PROBING STRIATAL SUBPOPULATIONS IN OUTCOME--DEPENDENT ACTION CONTROL IN RATS

JOANA FILIPA ANDRADE CATARINO

A dissertation submitted in partial fulfillment of the requirements for the Degree of Masters in Biomedical Research (Specialization Area: Neurosciences) at Faculdade de Ciências Médicas | NOVA Medical School of NOVA University Lisbon

December, 2021

PROBING STRIATAL SUBPOPULATIONS IN OUTCOME-DEPENDENT ACTION CONTROL IN RATS

Joana Filipa Andrade Catarino

Supervisors: Ingo Willuhn, PhD, Associate professor and Group leader at Netherlands Institute for Neuroscience;

Rita O. Teodoro, PhD, Group leader at Chronic Diseases Research Centre (CEDOC);

A dissertation submitted in partial fulfilment of the requirements for the Degree of Masters in Biomedical Research (Specialization Area: Neurosciences)

December, 2021

NETHERLANDS INSTITUTE FOR NEUROSCIENCE
Master the mind

Acknowledgments

" Those who pass by us, do not go alone, and do not leave us alone; they leave a bit of themselves, and they take a little of us."

Antoine de Saint-Exupery

Although this thesis work is mainly made of science, it is also made of all the people I encountered throughout this year. People that taught me so much, supported my dreams, and believed in me. Without you, this thesis work and this year would not be half as great.

First and foremost, I would like to thank my supervisor, Dr. Ingo Willuhn, for giving me the opportunity of being part of the neuromodulation and behavior group. To be part of your team was a great way to start my path as a neuroscientist and I could not be happier with all the opportunities and all the things I learned while there.

I would then like to sincerely thank Bastijn van den Boom. There are not enough words to express my gratitude for everything you taught me. Thanks for teaching and supervising me, but also for giving me the freedom to try and fail whenever needed and learn from it. It was great to share ideas and make good science next to you. I hope that our paths cross again in the future.

I would also like to express my gratitude to all the members of the neuromodulation and behavior group. Thank you for welcoming me so well and for always helping me when needed. It was great to spend this last year with all of you. A special thank you to Nicole Yee and Ralph Hamelink for teaching me a lot and always with a smile, and also to all the interns in the lab, who made this journey even more enjoyable. It was a pleasure to be a minion with you! I would also want to thank Sofia, who introduced me to the lab and since then has been always keen to help me with everything.

To Dr. Rita Teodoro, I want to acknowledge and thank all the support you gave me throughout the master's program. Besides being a great scientist, you are one of the kindest persons I met, and someone I will always remember and appreciate.

I would also like to thank my family with all my heart. Thank you for supporting me no matter what, for believing in me, and for making me chase my dreams, even when those dreams are in a different country and away from you. Without you, I would not be half of what I am today, and I hope to continue to make you proud in the future. All my victories will always be because of you.

To Bruno, I want to thank all the love and support. Thank you for being there for me every day and for celebrating my conquers and for helping me move over my defeats. With you by my side, everything is easier, and I hope we keep celebrating our achievements together for a long time.

To my NBR friends, I want to say thank you for always supporting my dreams and making them a bit yours too. Your conquers will always be my conquers and being able to celebrate them together will always be my favorite part. I hope you know that your friendship means the world to me, and I will always be your biggest fan. A "thank you" will never be enough.

Lastly, I want to thank Sara for being the best surprise of this Master's program and always having my back. Your friendship is something I will always cherish and want to keep for many years. Thank you for always being at the distance of one phone call and for making me see that I am capable of more than I think.

To all of you, a big thank you!

Financial Support

All the required financial support to perform all experimental and complementary work for the present project was provided by Grants to Dr. Ingo Willuhn (Principal Investigator):

1) H2020 European Research Council (ERC), Grant/Award Number: ERC-2014-STG 638013 2

2) Nederlandse Organisitie voor Watenschappelijk Onderzoek (NWO), Grant/Award Number: 864.14.010, 2015/06367/ALW

Abstract

A common characteristic of humans and animals lies in their ability to select the most appropriate action to achieve a desired goal. Voluntary action control (i.e., action initiation, and suppression), a process modulated by our internal needs and external factors, tends to adapt to the different situations we face and is biased to the outcome intended to achieve - to be rewarded or to avoid being punished. This process contributes to the achievement of optimal behavior by promoting the maximization of rewards and the minimization of punishments and is under influence of Pavlovian bias, a process that describes our tendency to activate actions in the face of reward (Go) and suppress them to avoid being punished (No-go).

Recent studies have associated activity in the striatum in outcome-dependent action control. More specifically, action control is facilitated by two subpopulations of medium spiny neurons (MSN): Dopamine 1 (D1) and Dopamine 2 (D2) receptor-expressing MSN. D1-MSN underlie the direct pathway of the basal ganglia while D2-MSN dominate the indirect pathway. The activity of these pathways generates controversy in the field, with different models trying to elucidate how they contribute to action control. A Classical model describes their activity as independent and antagonistic while a synergetic model argues that they act in a cooperative and synergetic way. The same controversy is seen regarding the way outcome-valence (reward and punishment) is encoded by these striatal subpopulations. Although a lot of effort has been made to better characterize outcome-dependent action control, it still remains poorly understood. Thus, with this thesis project, we intended to gain new insights into the way outcome-dependent action control is encoded by striatal subpopulations (D1-MSN and D2-MSN), and if their activity signals diverge between different striatal subregions (VMS, aDLS, pDMS).

To this end, we used a novel viral approach and optimized injection volumes to promote GCaMP expression in the rat' striatum. We then recorded activity in VMS, aDLS, and pDMS using *in vivo* fiber photometry while rats performed a novel Go/No-go reward and punishment task. Behavioral studies to better understand the subjective value of an aversive stimulus (loud white noise together with bright light) were also performed. We found that GCaMP expression in the rat striatum can be achieved by employing combined *Cre* x Flippase recombinases. In addition, we also showed that three seconds of continuous aversive stimulus formed by 90dB of white noise paired with bright light is aversive for the animals. More importantly, the subjective value was comparable to one sugar pellet, allowing us to compare reward and punishment conditions in the Go/No-go task. Furthermore, we found that animals can successfully learn and perform the Go/No-go task with Pavlovian bias influencing the performance between Go and No-go reward trials. Finally, neuronal data highlights that movement- and reward-related activity can be found in the three subregions but with different patterns of activity. Although activity in aDLS seems to be following the synergetic model, in VMS and pDMS the activity of D1- and D2-MSN seems to be less cooperative and synergetic. Differences were also found between trial types for some events and subregions.

The work presented in this thesis promotes the gain of knowledge about striatal contribution to outcome-dependent action control. Further investigation regarding the neuronal mechanisms underlying decision-making is needed to better understand which mechanisms might be impaired in psychiatric conditions and how we can create target-specific treatments for these patients.

Resumo

Uma característica comum entre humanos e animais é a capacidade de selecionar a ação mais apropriada para atingir um objetivo desejado. A isto chamamos controlo voluntário de ações, um processo modulado pelas nossas necessidades internas e fatores externos que tende a adaptar-se às diferentes situações com que nos deparamos sempre virado para o resultado que se pretende alcançar - ser recompensado ou evitar ser punido. Este processo contribui para a obtenção de um comportamento considerado otimizado através da maximização do número de recompensas e minimização do número de punições recebidas, encontrando-se sobre a influências do *Pavlovian bias*, um processo que descreve a propensão para a realização de ações para obter recompensas (*Go*) e de suprimi-las para evitar punições (*No-go*).

Estudos recentes associaram a atividade do estriado com o controlo de ações que são dependentes de um resultado, sugerindo que este é facilitado por duas subpopulações de neurónios espinhosos médios (MSN): neurónios que expressam o recetor de Dopamina 1 (D1) ou o recetor de Dopamina 2 (D2). D1-MSN constituem a via direta dos Gânglios da base, enquanto os D2-MSN constituem a via indireta. A forma como estas vias atuam ainda não é totalmente compreendida, existindo diferentes modelos que tentam explicar o seu modo de ação: o modelo clássico descreve a atividade destas vias como independente e antagonista, enquanto o modelo sinergético argumenta que elas atuam de forma cooperativa, estando simultaneamente ativas. A mesma controvérsia é gerada em relação à forma como as recompensas ou punições são codificadas por estas subpopulações. Ainda que muitos avanços tenham sido feito para melhor caracterizar o controlo de ações dependentes de um resultado, ainda existe muito por compreender. Com este projeto de tese, pretendemos obter novo conhecimento sobre a forma como o controlo de ações dependentes de um resultados é codificado por subpopulações do estriado (D1-MSN e D2-MSN), e se a sua atividade neuronal diverge entre as diferentes sub-regiões (regiões dorsais e ventrais do estriado - VMS, aDLS e pDMS).

Para este fim, usámos uma nova abordagem viral e volumes de injeção otimizados para promover a expressão de GCaMP no estriado de rato. Em seguida, registámos a atividade neuronal em VMS, aDLS e pDMS usando fotometria de fibra enquanto os animais realizavam uma tarefa de recompensa e punição ("Go/No-go"). Estudos comportamentais também foram realizados para elucidar o valor subjetivo de um estímulo aversivo (som de alta intensidade em combinação com luz forte). Com este projeto, descobrimos que a expressão de GCaMP no estriado de rato pode ser alcançada através da combinação das proteínas recombinantes *Cre* e *Flippases*. É também mostrado que três segundos de estímulo aversivo contínuo formado por 90dB de *white noise* em combinação com luz forte são aversivos para os animais. Mais importante ainda, mostrou-se que o seu valor subjetivo é semelhante a um *pellet* de açúcar, o que permite a comparação de condições de recompensa e punição na tarefa "Go/No-go". Finalmente, os dados de atividade neuronal destacam que atividade relacionada com o movimento e recompensa podem ser encontradas nas três sub-regiões estudadas, mas com diferentes padrões de atividade. Embora a atividade em aDLS pareça ir de acordo com o modelo sinergético, em VMS e pDMS, a atividade de D1- e D2-MSN parece ser menos cooperativa e sinergética. Diferenças entre os diferentes tipos de *trials* para alguns eventos e sub-regiões foram também encontradas.

O trabalho apresentado nesta dissertação promove a aquisição de conhecimento sobre a contribuição do estriado no controlo de ações dependentes de um resultado. Mais investigação sobre os mecanismos neuronais que estão na base do processo de tomada de decisões é necessária para melhor compreender quais os mecanismos que podem estar comprometidos em doenças psiquiátricas e como será possível, no futuro, criar tratamento mais especializados para este pacientes.

Table of Contents

Index of Figures

Introduction

Materials and Methods

Results

Appendix

Index of Tables

Introduction

Materials and Methods

Discussion

List of Abbreviations

1P	1 Pellet
1P+Comp	1 Pellet and Compound Stimulus
2P	2 Pellets
2P+Comp	2 Pellets and Compound Stimulus
A	Adenine
A2a	Adenosine 2a Receptor
AAV	Adeno-Associated Virus
AAV1	Adeno-Associated Virus Serotype 1
AAV9	Adeno-Associated Virus Serotype 9
aDLS	anterior Dorsal Lateral Striatum
AP	Anterior-Posterior
BG	Basal Ganglia
BSA	Bovine Serum Albumine
Ca^{2+}	Calcium
Comp	Compound Stimulus
D1	Dopamine 1 Receptor
D2	Dopamine 2 Receptor
DAPI	4', 6-diamidino-2-phenylindole
DIO	Doublefloxed Inverse Orf
DLS	Dorsal Lateral Striatum
DMS	Dorsal Medial Striatum
dMSNs	Direct Pathway Medium Spiny Neurons
DV	Dorsal-Ventral
EPC	Expression Pilot Cohort
FRT	Flippase Recognition Target
GCaMP	Genetically Encoded Calcium Indicator

GFP	Green Fluorescent Protein
GP	Globus Pallidus
hEF1α	Human Elongation Factor 1-Alpha
hSyn	Human Synapsin promotor
IC	Imaging Cohort
iMSNs	Indirect Pathway Medium Spiny Neurons
LED	Light-Emitting Diode
ML	Medial-Lateral
MSN	Medium Spiny Neurons
NAc	Nucleus Accumbens
NDS	Normal Donkey Serum
NP	Nose-poke
ns	Not Significant
PBS	Phosphate Buffered Saline
pDMS	posterior Dorsal Medial Striatum
PFA	paraformaldehyde
SEM	Standard Error of the Mean
SNr	Substantia Nigra pars Reticulata
STN	Subthalamic Nucleus
T	Thymine
VMS	Ventral Medial Striatum
vs	Versus
WN	White Noise

1. Introduction

Animals and humans can perform a variety of actions and, more importantly, are able to choose which action to perform, and when it should be performed. We can voluntarily choose a way to attain our goals, which takes into consideration our internal needs and the environment around us.

Throughout the years, philosophers, psychologists, and neuroscientists have argued about the definition of "voluntary actions" and their nature. Some thought that actions and the brain could be seen as dissociated subjects, while others knew that "actions" and "brain" were words that should never be used separately. Both contributed to what we know today: underlying the group of skills, that we so call actions, is a set of brain regions that form different neuronal circuits and work together to achieve successful behaviors. But how does the brain encode action control?

Despite all the work done in the field over the years, there is still a lack of information about the way outcome-dependent action control is encoded in the brain and what the contribution of striatal subpopulations' activity is in this process. Hence, this thesis aims to gather more knowledge regarding this topic by probing the role of D1- and D2-medium spiny neurons in outcome-dependent action control.

1.1. Voluntary action control and reinforcement learning

Our everyday actions are performed and adapted to the different situations we face in order to achieve desired outcomes. The ability to select the most appropriate action to maximize the outcome, taking into account internal and external factors (the environment around us that is in constant change), is called voluntary action control.[1,2] Independently of the action we perform, for it to be considered voluntary, the outcome should be anticipated, leading to the selection of the most suitable action.[2] But how do we acquire these action-outcome associations and how do we perform a voluntary action?

Several studies explain the acquisition of voluntary action control based on a two-stage model (**Figure 1**). In the first stage, a movement (motor pattern) is carried out, essentially driven by our internal demands and external inputs. From this action, a certain outcome is achieved, and the association between the action and the outcome is stored in the brain. This means that this association is registered and can be accessed in the future (**Figure 1A**). Moreover, in the second stage, the knowledge that was previously acquired (action-outcome association) is now used to activate specific behaviors and therefore perform an intentional and specific action to obtain the desired goal (**Figure 1B**).[2,3]

Voluntary actions are performed in order to achieve specific outcomes. By trial-and-error, actions are selected to optimize the outcome, and therefore sometimes fail our goals. Thus, to maximize the number of times the optimal outcome is achieved, one needs to learn which actions are correct and should be performed and which actions do not lead to the outcome (incorrect) and should be abandoned. The way this is done is through a process called reinforcement learning.[4,5]

According to reinforcement learning theory, we adapt our choices based on past experiences through an exploration and exploitation process.[4] During the exploration part, our behavior is shaped by the outcome of correct actions. Then, based on knowledge about previous outcome contingencies, certain actions will be performed with an increasing tendency (exploitation) to maximize the outcome. Both positive and negative reinforcements are capable of increasing the predisposition for an action to be selected by rewarding it with

something gratifying (positive reinforcement), or by promoting it through the avoidance of aversive outcomes (negative reinforcement)[4].

Moreover, voluntary action control was shown to be biased towards the outcome we intend to achieve. Depending on the outcome's valence, our actions can be selected in order to receive a reward (positive valence) or to avoid a punishment (negative valence).[1] Voluntary action control contributes to the achievement of optimal behaviors by maximizing the number of rewards one gets and minimizing the number of punishments. In addition, voluntary action control was also seen to be under the influence of Pavlovian bias, a process that promotes action to get rewards and inactivity to avoid punishments.[6–8]

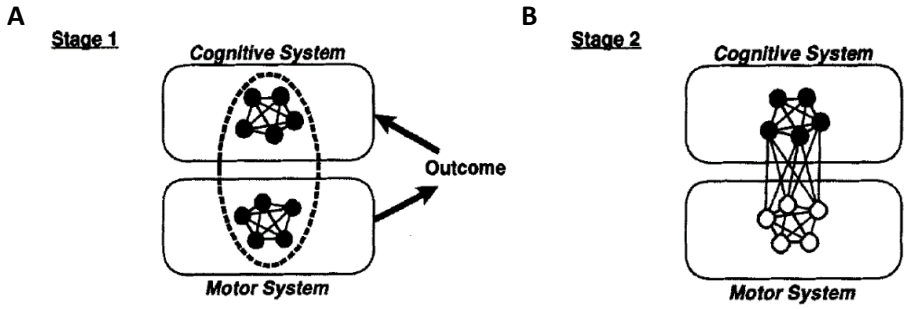

Figure 1: Two-stage model for the acquisition of voluntary action control. (A) In the first stage, the association between a random action and its outcome is formed and registered within cognitive codes. **(B)** In the second stage, specific cognitive codes are activated to perform an intentional action (motor pattern) and therefore achieve the anticipated outcome. Adapted from Eisner and Hommel, 2001.[2]

1.2. Pavlovian Bias

An important aspect of decision-making, specifically in action selection, is a process entitled as Pavlovian bias.[6] This process modulates behavior by working with two important variables: action and valence (outcome). Actions can involve effort (action initiation - Go) or inhibition (action suppression - No-go), and valence can entail the presence of reward or the absence of punishment.[7,8]

Studies performed in humans using decision-making tasks (e.g. Go/No-go tasks) showed that action selection is influenced by Pavlovian bias, which endorses the initiation of actions under rewarding conditions (Go to win) and the withholding of actions under punishing conditions (No-go to avoid lose).[6,7,9] Furthermore, the same behavioral studies elucidated that due to Pavlovian bias, humans find it more difficult to behave using the opposite action-valence associations: perform an action in the face of punishment (Go to avoid lose) or inhibit actions in the face of reward (No-go to win), leading to a suboptimal behavior.[6,8]

Behavioral studies performed in cohorts of patients suffering from psychiatric disorders, such as schizophrenia, depression, and impulsive control disorder found that the Pavlovian bias process was compromised in these patients. In disorders like schizophrenia, patients showed reduced bias towards action initiation (Go) in rewarding conditions. This can arise from both the fact that in this disorder the representation of reward values is likely to be degraded and that patients show difficulty in learning the adequate response to winning rewards (poorer Go learning).[10] Similarly to Schizophrenia, depression is characterized by reduced action initiation under rewarding conditions (under-responsivity) and accentuated

action suppression under aversive conditions (over-responsivity).[11,12] Additionally, and in opposition to the previous psychiatric conditions, in impulsive control disorder, there is enhanced action execution in rewarding conditions.[13,14] The lack of correct Pavlovian bias effect that is seen in these psychiatric conditions can be one of the factors contributing to the behavioral traits usually associated with these pathologies (e.g. impulsivity, anhedonia, etc).

To be able to find treatments that can target specific brain regions and help patients suffering from these conditions, we need to understand how these processes work in a healthy brain (absent of any pathology). Questions like: ”Where is outcome-dependent action control generated in the brain?”, “How can it be biased towards an outcome?” and “Which regions are involved in these processes?” need to be answered. As an attempt to answer this question, regions within the basal ganglia circuitry, with a major focus on the striatum, have been highlighted to play a significant role in action control and outcome-valence processing (i.e., outcome-dependent action control), thereby contributing to Pavlovian bias.[15]

1.3. Basal Ganglia Circuitry

“The human brain has 100 billion neurons, each neuron connected to 10 thousand other neurons. Sitting on your shoulders is the most complicated object in the known universe. “

Michio Kaku [16]

Neurons are one of the most important building blocks of our brain, and it is through the connections they make with each other that functional circuits arise.

The basal ganglia (BG) circuitry, a set of interconnected brain nuclei found in the base of the forebrain, plays an important role in cognitive and motor functions, as well as in the control of adaptive behavior driven by rewarding and punishing conditions.[17,18] The way information flow through the BG follows either the direct pathway via striatum - substantia nigra pars reticulata (SNr) - thalamus - cortex (the region that also sends the majority of the afferents found in BG) or the indirect pathway via striatum - globus pallidus (GP) - subthalamic nucleus (STN) - SNr - thalamus - cortex, thought to be involved in the execution or suppression of actions, respectively (**Figure 2)**.[17]

Moreover, it is through the cortico-basal ganglia-thalamo-cortical circuit that the majority of functions related to BG activity are implemented. The way this works is through projections sent by the cerebral cortex to the BG that are then sent back to the cortex via the thalamus.[19] Initially, it was thought that communication from BG mainly went to the motor cortex but more recently it was found that BG efferents go to different subdivisions of the thalamus, which then projects to a larger range of cortical and sub-cortical regions. This is why BG are involved in more functions than just motor functioning (e.g. cognitive and affective functions).[20]

Due to its connectivity, dysfunction of BG circuitry (either via cerebral cortex or cerebellum connections), can lead to diverse motor (neurologic) and non-motor (psychiatric) conditions.[19] Some well-known examples of BG dysfunctions are Parkinson’s and Huntington’s diseases, which affect motor control, but also schizophrenia, depression, and addiction (psychiatric conditions).[17,22]

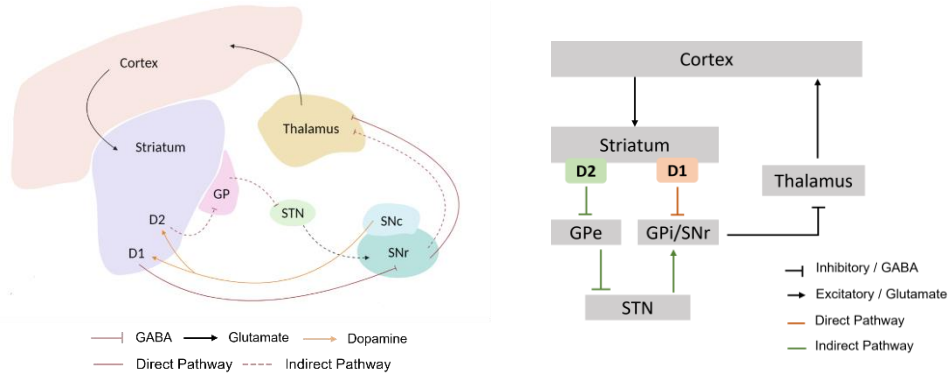

Figure 2: Representation of how information flows through the Basal Ganglia. Information flows through the basal ganglia employing the direct and indirect pathways that link regions like cortex, striatum, STN, SNr, GP and thalamus. On the left panel, the direct pathway is represented by a full line and the indirect pathway by a dashed line.

As we can see, the BG comprises different regions that receive and send projections to different areas in the brain, and therefore are related to different functions. All these regions are key players to the normal functioning of this circuit that allows a healthy state. Therefore, depending on the scientific questions that make us interested and motivated to investigate this circuit, we should pay more attention to certain regions to the detriment of others. Thus, in this thesis project, we intend to have a major focus on the role of striatum within this circuitry to disclose how it is involved in outcome-dependent action control.

1.3.1. An input nucleus called Striatum

The striatum is the primary input nucleus of the BG, receiving glutamatergic inputs from cortical, thalamic, and limbic areas, and dopaminergic inputs from the midbrain.[15,22,23] It is a region involved in diverse functions like motor control, learning, and decision-making upon reward and aversive conditions. Depending on the origin of the cortical input it receives, the striatum can be considered a sensorimotor, associative (cognitive), or limbic (emotional and motivational) center, processing different types of information.[19,24] The way the striatum helps in the action selection process (to learn which actions lead to reward and how they should be executed), is through the outputs it sends to the BG downstream regions after processing all the conveying information.[22] The way information goes from the striatum to the other output nucleus of the BG is through the direct and indirect pathways, which it gives rise to.

The rodent striatum can be divided into three subregions that are thought to be involved in learning and distinct forms of associations: the ventral medial striatum (VMS), dorsal medial striatum (DMS), and dorsal lateral striatum (DLS) - also known as limbic, associative and sensorimotor striatum, respectively.[15,25] Besides the different locations of these subregions within the striatum, they also present distinct roles. VMS has been shown to be important in outcome evaluation and motivation and involved in optimizing the selection of the most appropriate action.[15] The DMS is mainly involved in the acquisition of goal-directed behaviors, by establishing an association between an action and its outcome.[26] Finally, the DLS is involved in stimulus-response associations, which comprise habitual, non-flexible behavior.[15,25,27]

The striatum is formed by roughly 95% inhibitory (GABAergic) medium spiny neurons (MSN), which can be subdivided into two distinct populations based on their molecular properties. The remaining 5% are

composed of either GABAergic or cholinergic interneurons.[22] GABAergic interneurons can be divided into fast-spiking (FSI), low-threshold spiking (LTSI), and calretinin expressing (CR) interneurons, while cholinergic interneurons (CIN) are characterized as large and tonically active neurons (TAN), representing 1% of the cells found in the striatum.[17,22,28] Throughout this section we will be focusing our attention on the MSN since they embody the major group of neurons found in the striatum, giving rise to the direct and indirect pathway, and are thought to be related to action control.

Depending on the projection target and dopamine receptor that GABAergic MSN express, they can be classified into dopamine 1 (D1-MSN) or dopamine 2 (D2-MSN) subtype of neurons, both encompassing 50% of the MSN population. D1-MSN (striatonigral MSN) directly innervates the BG output nucleus SNr, giving rise to the direct pathway, while D2-MSN (striatopallidal MSN) indirectly innervates the SNr through the GP and STN, giving rise to the indirect pathway.[29] The net effect of the direct pathway on the thalamus is dis-inhibitory (due to inhibition of the SNr, which is an inhibitory nucleus), while the net effect of the indirect pathway is inhibitory on the thalamus.[23] Due to this, the direct pathway is thought to promote action initiation while the indirect pathway is hypothesized to inhibit action.[17,29,30]

Activity in striatal D1- and D2-MSN has been highlighted as a key player in the control of Go/No-go actions.[17] Hence, it is crucial to understand how their activity is coordinated and mediates outcome-dependent action control in both positive (reward) and negative (punishment) contingencies. In the next chapter (Introduction - Section 1.4.) we will explore the theories that attempt to explain the way these two pathways work.

1.4. Activity of direct and indirect pathways of the basal ganglia

Years of theoretical and experimental work tried to disclose how the activity in the striatum mediates action control and how these striatal signals vary depending on the action-outcome (reward versus punishment). Two theories try to describe how the direct and indirect pathways of the basal ganglia contribute to behavior. The classical model puts forward the "Go/No-go" model, where the direct and indirect pathways are described as antagonistic. The synergetic model, a more recent theory, describes the activity of these pathways as cooperative. In addition to action selection, direct and indirect pathways are also involved in outcome valence. Bridging the action selection and outcome valence theories urges the need for further investigation.

1.4.1. Action selection based on the classical and synergetic models

The classical model for the functioning of the basal ganglia circuitry describes the activity of the direct and indirect pathways, two primary projections of the striatum, as being independent and antagonistic. According to a model proposed by Albin *et al.* (1995)[31], the direct pathway is responsible for the facilitation of action initiation (Go), while the indirect pathway is hypothesized as responsible for the facilitation of action suppression (No-go).[30–32]

The opposite roles hypothesized by the classical model are supported by the structural organization and projections of the basal ganglia (anatomy) and by studies using lesions and behavioral tasks.[15] Regarding structural organization and basal ganglia projections, it was shown that D1-MSN (direct pathway) inhibit the GPi/SNr, promoting a disinhibitory net effect on the thalamus, thereby leading to action initiation (**Figure 3A**).[15,32] On the contrary, the D2-MSN (indirect pathway) promote inhibition of the GPe, which then

disinhibits the STN and ends up promoting the inhibition of the thalamus via inhibitory projections from the GPi/SNr (**Figure 3B**). Thus, the inhibition of the thalamic activity leads to a suppression of actions.[15,32]

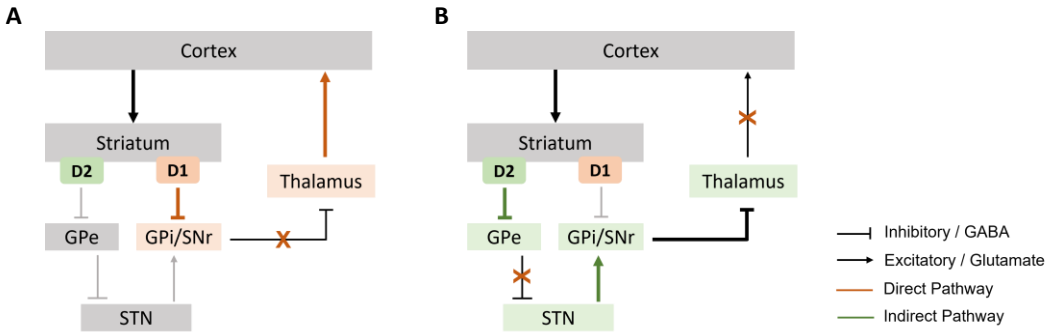

Figure 3: Differing effects of the direct and indirect pathways of the basal ganglia. (A) Model representing the functioning of the direct pathway of the basal ganglia. D1-MSN promote dis-inhibition of the thalamus by inhibiting basal ganglia output nucleus (GPi and SNr). **(B)** Model representing the functioning of the indirect pathway. D2-MSN promote inhibition of the thalamus by promoting the activation of the basal ganglia output nucleus (GPi and SNr). Orange represents the activity and regions involved in the direct pathway, while green represents the activity and regions involved in the indirect pathway.

Over the years, other models based on electrophysiological recordings and more advanced genetic tools tried to explain the activity of the two pathways and their involvement in action control. More recently, a second model - the synergetic model - was proposed by Mink (2003)[33] in an attempt to explain the functioning of these two important pathways.[33] In opposition to the classical model, here both striatal pathways are hypothesized to be involved in action selection, with the D1-MSN promoting the execution/initiation of the desired action and D2-MSN suppressing the execution of competing actions.[15,17,34] Overall, this model characterizes the activity of the direct and indirect pathways of the basal ganglia as cooperative and synergistic.[17,30,35]

The synergetic model was supported by different studies using cell-type-specific recording experiments that observed a transient increase in activity in both pathways upon action initiation (Cui *et al.*, 2013), as well as in lever pressing task in which both were active at the start and end of a lever press sequence (Jin *et al.*, 2014). These studies gathered evidence that supports the need for, and importance of, activation in both of these pathways during movement and action selection.[17,36–38]

Furthermore, **Table 1** summarizes findings achieved with optogenetic manipulation of D1-MSN and D2-MSN upon different behavioral paradigms and in different brain regions. This table also classifies the finding as following the classic model or the synergetic model.[15]

Table 1: Findings regarding optogenetic manipulation of D1-MSN and D2-MSN in decision making. Results of the optogenetic manipulation of this striatal subpopulations were compared to what is stated by the classic view model proposed by Albin for movement and action selection. Adapted from Cox et al. (2019).[15]

Classic view	Method	Result	Supports classic view?	Refs
D1R and D2R neurons oppositely modulate behaviour	Optogenetic manipulation	Activation of D1R MSNs increases whereas activation of D2R MSNs decreases spontaneous movement	Yes	39
		In value-based decision-making, activation of D1R and D2R MSNs oppositely biases choice	Yes	18
		Activation of D1R MSNs promotes performance of a stimulation-paired behaviour and activation of D2R MSNs decreases performance of a stimulation-paired behaviour (in the DMS and NAc)	Yes	44, 46
		Activation of D1R and D2R MSNs in the DLS promotes pressing of stimulation-paired lever, but activation of D2R MSNs in the DLS also increases pressing of an unpaired lever	No	40
		Activation of D1R and D2R MSNs promotes performance of a stimulation-paired behaviour, but only D1R activation increases time spent in a stimulation-paired location (in the NAc)	No	41
		Activation of D1R and D2R MSNs promotes go responses in a sensory go/no-go task	No	42
		Activation of D1R and D2R MSNs increases motivation	No	43
		D1R MSN inhibition slows action initiation but D2R MSN inhibition decreases task engagement	No	38

1.4.2. Striatal role in mediating action control dependent on the outcome (reward vs punishment)

The direct and indirect pathways of the basal ganglia are responsible for controlling motor output (movement and action selection) but they are also involved in mediating outcome valences: appetitive (reward) and aversive (punishment) stimuli. However, the way this mediation is done by striatal subpopulations is not completely understood.

Optogenetic studies performed by Kravitz *et al.* (2012)[44], selectively targeting these two distinct pathways, aimed to better understand the role of these sub-populations. Here, the authors tested the hypothesis that activation of the direct pathway would mimic appetitive stimuli, while the activation of the indirect pathway would instead mimic aversive stimuli.[44,45] Together, the results showed that activation of D1-MSN is sufficient to induce persistent reinforcement, while the activation of D2-MSN was sufficient to induce transient aversion, characterizing the activity of this subpopulations as antagonistic.[21,23,45,46]

Follow-up studies challenged the simplistic, dichotomy roles of direct and indirect pathways in appetitive and aversive encoding. Recent studies hypothesized that both striatal subpopulations can encode appetitive and aversive outcomes with different patterns of activity that can also vary between different striatal subregions.[23,47] Pharmacological approaches have shown that both D1 and D2-MSN in nucleus accumbens (NAc) are important for the motivational drive to work for rewards and that by using D1R and D2R antagonists in the same sub-region there is a decrease in the number of avoidance responses in aversive contexts. In addition, genetic studies showed that overexpression of D2R in NAc increases motivation for reward, which does not happen in dorsal regions of the striatum.[23] More studies using other approaches, such as optogenetics, were also performed to better understand the roles of these two pathways in appetitive and aversive encoding. However, the conflicting findings between experiments that still persist, underline the need to further investigate the role of these pathways.

Overall, the differences in models describing basal ganglia functioning, both in action selection and outcome valence, highlights the need for more elaborative behavioral studies combined with cell-type-specific recording techniques that allow the dissection of the role of each of the two pathways.

1.5. *In Vivo* Calcium Imaging

The importance and need to elucidate what is happening in our brain when we are awake and performing tasks that we normally do in our daily life has increased and become a major focus in neuroscience. Over the years, the development of new imaging techniques that allow the recording of specific neuronal populations in real-time in awake and behaving animals, has opened doors to gain new knowledge regarding the activity of diverse neural circuits. An example of this new recording methods is *in vivo* calcium imaging, which can be performed at the single-cell or population level, providing different readouts of brain activity.

Calcium imaging has been around to visualize neuronal activity in brain slices or even acute *in vivo* preparations. However, calcium imaging recordings in freely-behaving animals only started to be explored more recently. The way this technique works is by using calcium (Ca^{2+}) as a proxy for action potential activity. Given that during neuronal communication there is an influx of ions, among others Ca^{2+}, we can infer how neurons communicate by looking at their intracellular Ca^{2+} concentration.[48,49] This is possible by employing genetically encoded calcium indicators, fluorescent molecules that increase their fluorescence intensity when binding to Ca^{2+}.[48,50] Furthermore, the design of these genetically encoded calcium indicators (GECI) like GCaMP (which dynamics continue to be improved over the years) in combination with viral engineering and specific animal lines created using CRISPR technology, improved the versatility that this technique can offer.[51,52]

Additionally, with the advances of these techniques, new probes were also developed, allowing the targeting of deeper brain regions (e.g. gradient refractive index (GRIN) lenses), which was not possible before. This was accompanied by the development of new technology that allows the recordings to not require the animal to be immobile (e.g. miniaturized fluorescent microscopes that can be mounted on the animal's head and can be used to record while animals behave freely).[53,54] Overall, this constituted a great advance in science and allowed the gain of new knowledge regarding brain functioning.

Although this novel, state-of-the-art techniques are critical for nowadays neuroscience, it is also crucial for a researcher to choose the best method to answer the research question. Since there are several types of *in vivo* calcium imaging techniques using GECIs, we employed fiber photometry to best answer our research questions. Thus, we will be looking into more detail in the way this method works and also in its advantages and disadvantages.

1.5.1. Neuronal recordings using fiber photometry

Fiber photometry is an *in vivo* calcium imaging technique that uses chronically implanted optical fibers to record neuronal transients. As explained before, this is done through the expression of GECIs like GCaMP and allows the investigation of different neuronal populations' activity (**Figure 4**).[36] A characteristic of this technique lies in the fact that we can only record bulk activity instead of a single-cell activity. This means that we can explore what a certain population of neurons is doing but we cannot perform an analysis of

heterogeneous activity within that population (i.e., analyze how individual neurons behave within a group of neurons).[36,48]

Figure 4A shows a representation that explains how this technique works. Briefly, light emitted from fiber-coupled LEDS with excitation wavelengths of 465nm (Ca^{2+} dependent activity of the GECI used) and 405nm (isosbestic control channel for GCaMP) is combined into one path, due to dichroic mirrors, and delivered to the regions of interest via the fiber optic patch cable connected to the optical fiber implanted in the animal's head. When in the region of interest, the light will excite the cells that are located under the tip of the optical fiber (target region). If the cells are part of our target population (have GCaMP that was previously introduced through a specific viral construct), the GCaMP can bind to the intracellular Ca^{2+} present in the cell. The excitation of GCaMP will in return emit light of a different wavelength that will be captured back to the fiber patch cable and collected using a photo-sensor present in the setup. This photo-sensor will receive the light from both channels (GCaMP and isosbestic) and will be able to separate them into two distinct signal traces.[48,55,56]

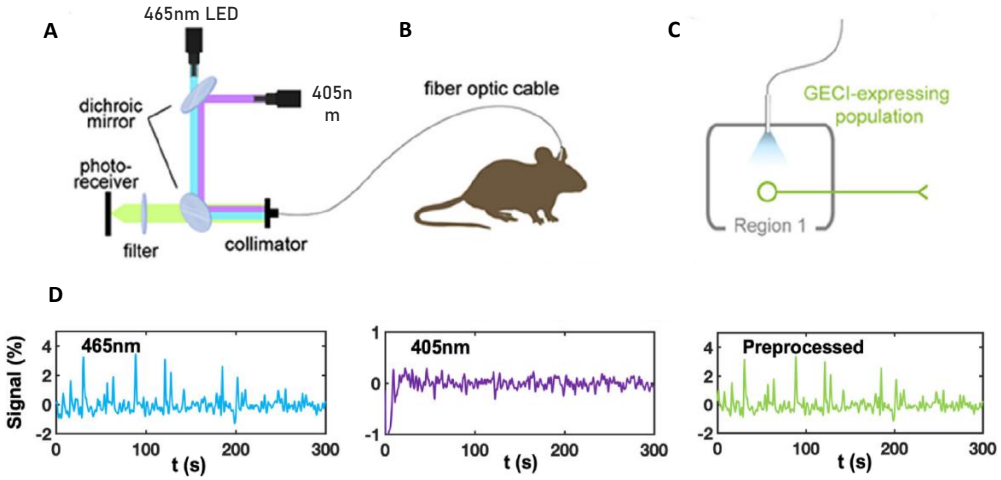

Figure 4: Schematic of *in vivo* calcium imaging using fiber photometry. (A) Schematic representation of how light travels withing the fiber photometry set up to be emitted and excite the brain regions implanted with an optical fiber and how light emitted by the cell is collected. Blue represents the wavelength of excitation for the Ca^{2+} dependent signal (465nm) and purple represents the wavelength of excitation for the isosbestic control signal (405nm). **(B)** Representation of how the optic patch cable is connected to the optical fibers implanted in the head of the animal. This connection is done in the beginning of the recording session and must allow the animal to freely behave within the operant chamber. **(C)** The optical fiber is implanted in the region of interest and GECI- expressing cells that are under the optical fiber will be recorded. This method allows the recording of either soma or axons. **(D)** Representation of the data obtained using fiber photometry. Left panel shows a representative trace of Ca^{2+}-dependent signal (465nm - blue trace). Middle panel shows a representative trace of the isosbestic channel, which is independent of Ca^{2+} (405nm – purple trace). Right panel shows a representative trace of the pre-processed signal, obtained by subtracting the isosbestic signal do the Ca^{2+}-dependent signal. Figure adapted from Siciliano and Tye, 2019, and Liang, Neuberger and Zhang, 2020.[48,57]

Advantages and Disadvantages of fiber photometry

Compared to other techniques used to record neuronal activity, this method has the advantage of being relatively easy to implement because we can record everything that is underneath the optical fiber, which requires less spatial resolution than other techniques (i.e. techniques that look at the single-cell activity and

therefore need to have a good number of individual neurons at an ideal distance from the tip of the lens).[48] Another advantage, and perhaps the most important, is the fact that we can implant more than one fiber per animal due to the small size of the optical fibers. This is valuable because it allows recording from different brain regions within the same animal.[58,59] Furthermore, the small size and weight of the optical fiber and the optical patch cable make it easier for the animals to perform a behavioral task.[59] Finally, another advantage of this method lies in the capability of having long imaging times due to low LED settings, thereby avoiding bleaching, which happens quickly with 1- and 2-photon calcium imaging.

Although there are several advantages to this technique, there are also some drawbacks. One of them is the inability to attend for heterogeneity within the population of neurons we are recording from.[48] Since we cannot look at single-cell activity, we cannot investigate the presence of different clusters of neurons within a population, ending up losing information that could be interesting and relevant for the study. Moreover, another drawback of this technique lies in the amount of noise that is injected into the data. Fiber photometry uses an optic patch cable connected to the optical fibers implanted in the head of the animal to excite and record light, the bending of the cable while the animal is moving in the operant chamber can introduce noise in our data, causing movement-related artifacts.[60] Fortunately, this can be minimized by using adequate controls to record movement artifacts, such as the isosbestic control channel.[60,61]

Isosbestic control channel

It is important to identify and minimize movement-related artifacts present in our data since they can easily be confused with activity-dependent signals (neuronal activity). To avoid these signals, an isosbestic channel can be used as a control. An isosbestic point is characterized as the wavelength at which physical or chemical reactions (e.g. Ca^{2+} reaction) occurring within a sample, do not change its photon absorbance (excitation light of 405nm).[48,61] This means, that when excited at the wavelength of its isosbestic point, the fluorescence resulting from the sample will not be Ca^{2+} dependent (not indicative of neuronal changes) and will only represent signals related to movement and other artifacts.[59] In addition, to improve the signal, the fitted isosbestic trace is usually used to subtract from the GCaMP trace and correct for motion.[62]

2. Research Questions and Aims

The experimental work done during this thesis project aimed to acquire new knowledge about the way outcome-dependent action control is encoded by striatal subpopulations (D1- and D2-MSN) and how their activity is coordinated. To this end, we proposed to answer the following research questions:

1. How do direct and indirect pathway medium spiny neurons contribute to outcome-dependent action control?
2. Do these signals diverge between different striatal subregions (VMS, aDLS, pDMS)?

To answer these questions, we established three different aims for this project:

Aim 1 - Optimization of *cre*-dependent GCaMP-expressing adeno-associated-virus to record striatal medium spiny neurons *in vivo*, by testing different volumes with a new dual-vector viral approach;

Aim 2 - Verify the subjective value of a compound stimulus (white noise paired with LED) in relation to a food reward (sugar pellet) to establish direct and trustful comparisons between neuronal activity during reward and punishment contingencies;

Aim 3 - Disclose the contribution of D1- and D2-MSNs in outcome-dependent action control in VMS, pDMS, and aDLS. To this end, we used *in vivo* fiber photometry to perform calcium imaging (i.e. measure the activity of genetically defined striatal neuronal populations) in new knock-in *Cre* rat lines during the performance of a complex Go/No-go reward and punishment task.

3. Materials and Methods

3.1. Experimental model

All animal procedures were performed in accordance with European and Dutch laws and approved by the Animal Experimentation Committee of the Royal Netherlands Academy of Arts and Sciences. Adult Long-Evans rats were single-housed and kept under a reversed 12-hour light-dark cycle (light on from 20:00 till 8:00) with controlled temperature (23^{o}C) and humidity. Experiments were conducted during the dark phase. Animals had *ad libitum* access to water and were food-restricted to 85-90% of their free-feeding body weight with daily monitorization. A total of 22 rats underwent surgery for viral injection and optical fiber implantation.

Transgenic Rat Lines

To record different neuronal subpopulations within the striatum, transgenic D1-i*Cre* and A2a-i*Cre* rat-lines were used (**Figure 20**, Appendix).[28] Transgenic rats with *Cre* protein under regulation of the dopamine D1 receptor gene were used to specifically target direct pathway medium spiny neurons (dMSNs), while animals with *Cre* protein under the A2a promoter were used to target the indirect pathway medium spiny neurons (iMSNs). The targeting selectivity of iMSNs was enhanced by using A2a receptors instead of D2. While A2a receptors are selectively expressed in the indirect pathway, D2 receptors can also be found on other striatal cells and synapses (e.g. Cholinergic interneurons, as well as, striatal glutamatergic and dopaminergic afferent projections), not being as specific for this pathway.[63] D1-i*Cre* and A2a-i*Cre* rats were bred in-house by crossing heterozygous parent lines. The animals used in the experiments were heterozygous for the relevant *Cre*-allele.

All animals used to conduct this study were selected and maintained according to specific parameters settled before the experimental work (**Table 2**). This allowed better control of our experiments and reduced discomfort for the animals.

Table 2: Description of experimental model's characteristics. All the animals used in this research project were selected and maintained according to all the different conditions present in the table below.

EXPERIMENTAL MODEL	
Specie	Rats
Developmental Stage	Adult (≈ 8 weeks)
***Cre* Lines**	D1-i*Cre* and A2a-i*Cre*
Strain	Long-Evans
Age	8-32 weeks
Weight	200-500g
Origin	Bred in-house

Experimental groups

For fiber photometry experiments, two cohorts of animals were used: The expression pilot cohort and the imaging cohort. The expression pilot cohort was a pilot group to optimize fiber photometry, the main technique employed in this project. Here, we used four female Long-Evans: two A2a-i*Cre* and two D1-i*Cre* to optimize virus expression. Animals in the imaging cohort were injected with optimized virus acquired using the pilot cohort and was constituted by 18 female Long-Evans - nine A2a-i*Cre* and nine D1-i*Cre*.

The expression pilot cohort was bilaterally implanted with three optical fibers in three different striatal subregions (ventral medial striatum (VMS), anterior dorsal lateral striatum (aDLS), and posterior dorsal medial striatum (pDMS)) and the imaging cohort was bilaterally implanted with only one optical fiber in one of the three striatal subregions.

Furthermore, a third experimental group, composed of D1-i*Cre* animals (two females and three males Long-Evans), was used to examine the subjective value of an aversive stimulus in relation to a reward delivery (Aversive/Appetitive Task). These animals were not submitted to any surgery and remained group-housed during all the experimental time.

Advantages of the model

Rodents are used as models for biomedical research due to their degree of similarity to humans (anatomically, physiologically, and genetically).[64] The use of rats as a model is an advantage for this project, not only because of their larger size, which facilitates surgical procedures (an important step in our experiments) but also due to their ability to perform more complex behavioral tasks.[28,64,65] This ability allows them to learn and carry out the Go/No-go task used in this project. Another advantage of using rats is the ability to implant several fibers in one animal, which reduces the total number of animals required in the experiments and allows us to follow the neuronal activity of different brain regions within animals.

3.2. Viral expression: Dual-vector approach

To achieve *Cre*-dependent expression of GCaMP6f in specific striatal subpopulations (D1-MSNs and A2a-MSNs), we used adeno-associated virus (AVV) in distinct transgenic rat lines expressing *Cre* recombinase (D1-i*Cre* and A2a-i*Cre* rat lines). Previous work from the lab showed low GCaMP expression in the striatum using standard *Cre*-dependent viral constructs (see Results section - **Figure 9A**). To overcome low expression, thereby hindering successful imaging *in vivo*, we employed a dual-vector viral approach (**Figure 5**), aiming to enhance the number of recombinase proteins, and consequently increase the amount of GCaMP expression.

This dual-vector approach consists of a mix of two different viral constructs in the same viral injection. The first virus was a *Cre*-dependent Flippase construct (AAV1.hEF1α.Dio.FLP) to increase the number of Flippase recombinase proteins relative to the number of *Cre* recombinase proteins. Briefly, this viral construct is formed by Flippase RNA that is inversed and flanked by two *loxP* sites with opposite directional orientations. Upon infection of neurons expressing *Cre* recombinase proteins, the inversion of the Flippase RNA will be mediated, making its translation into a functional protein possible.[66] Additionally, we used a second virus (AAV1.hSyn.FRT.GCaMP6f) with a Flippase-dependent GCaMP6f construct.[66] Here, we have GCaMP6f RNA inversed and flanked by two FRT (Flippase recombinase target) sites, with opposite orientations, that can be

recognized by the Flippase recombinases produced by the first virus. These will then revert GCaMP RNA to its correct orientation, allowing successful translation into functional proteins.

Together, this dual-vector approach increased the number of GCaMP6f proteins by enhancing the amount of Flippase recombinase proteins (compared to *Cre* proteins), which in turn can recombinase more GCaMP RNA.

All viruses used in these experiments were purchased from the Viral Vector facility - University of Zurich[UZH].

Importance of GCaMP expression

Inducing sufficient GCaMP expression is an important step since it is used as a proxy for neuronal activity. This is possible since Ca^{2+} binds to GCaMP present in the cell, which increases its fluorescence intensity and enables us to infer that activity is happening in the cell.[48,52]

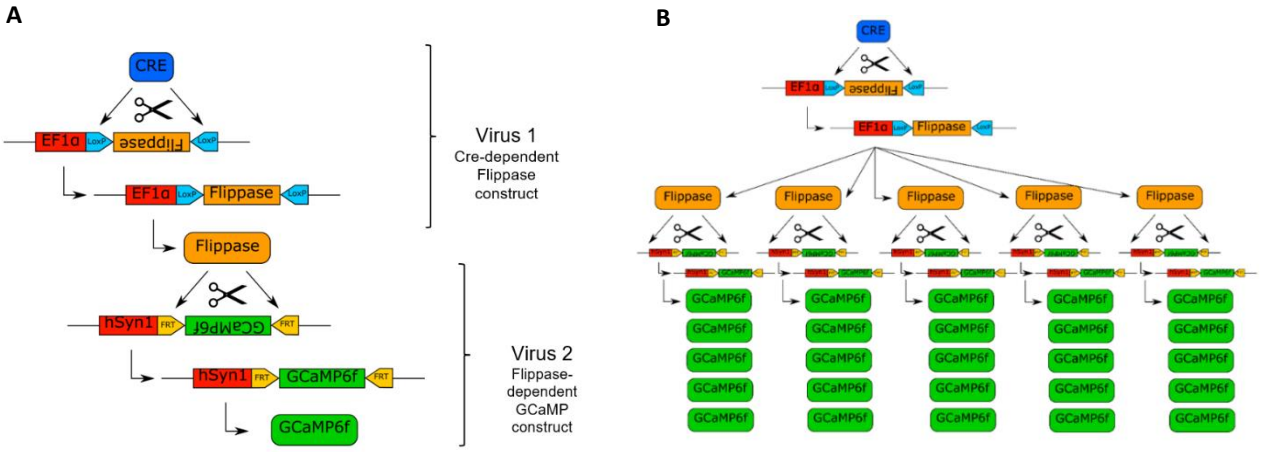

Figure 5: Dual-vector viral approach used in this project. (A) Schematic of the viral constructs used in the viral mix. The first virus (AAV1.hEF1α.Dio.FLP) is formed by a *Cre*-dependent Flippase construct and was used to increase the number of Flippase recombinases upon contact with cells expressing *Cre*-recombinase proteins. The second virus (AAV1.hSyn.FRT.GCaMP6f), constituted by a Flippase-dependent GCaMP construct, was used to improve GCaMP6f expression. The production of functional GCaMP proteins is only possible when the second virus contacts with the Flippase recombinases produced by the first virus. **(B)** Representation of the way the dual-vector approach increases the number of GCaMP6f proteins that will be available.

3.3. Surgical Procedures

To assure animals' safety and avoid infections, both the surgical area and tools were properly sterilized before surgery.

Prior to surgery, rats were taken from their home cage, weighted, and anesthetized using a small induction chamber (Petterson Scientific) filled with 3-5% isoflurane/oxygen. Upon loss of consciousness, the hairs on their head were shaved and disinfected with an alcohol-based solution (Sterillium med) to keep the surgical area as clean as possible.

All surgeries were performed in a stereotaxic frame (David Kopf Instruments, 962LS) under aseptic conditions and isoflurane anesthesia (1-2% isoflurane/oxygen with a flow rate of 0.6L/min – Burtons, Tec 850). Rats' body temperature was maintained at 37°C using an isothermal pad. Surgeries were conducted using the Zeiss microscope (Stemi 2000) and a light source (CL 15000 ECO).

Intracranial viral injection and optical fiber implantation

After being precisely positioned in the stereotaxic frame using the ear bars, animals were given subcutaneous injections of Metacam (Meloxicam, 1mg/Kg, Boehringer Ingelheim) and Dexamethasone (1mg/Kg) to decrease pain and prevent inflammation that can result from surgery, respectively. Following, depth of anesthesia was assessed by toe pinch and a small incision (from anterior to posterior) was made on the animal's head using a scalpel. The skin of the scalp was retracted to better expose the skull and Lidocaine (100 mg/mL, Xylocaine®) was applied as a local anaesthetic. Rats' eyes were covered with eye drops (TERRA-CORTRIL: Met-Polymyxin-B) and paper pads to keep them moist and avoid discomfort resultant from long light exposure.

Next, bregma and lambda coordinates were identified and used as reference points to align the skull. Based on the location of bregma, injection and implantation coordinates to specifically target VMS, pDMS and aDLS were calculated and marked on the skull of the animal (**Figure 6**). Using a hand-held surgical drill, small holes were drilled for every injection and optic fiber implantation site, as well as for the implantation of four stainless steel screws - aiming to increase head cap stability.

Animals were then subjected to bilateral viral injections with optimized viruses (AAV1.hEF1α.Dio.FLP and AAV1.hSyn.FRT.GCaMP6f - **Figure 5**). Volume and concentration of injection were different between experimental groups (**Table 3**). Injections were performed using Hamilton syringes (Hamilton, USA) and a nanoinjector (Micro4^{TM}) at a rate of 200nL/min. After infusion, the syringe was held in place for five to ten minutes to minimize virus leakage and was then slowly retracted.

Viral injections were followed by bilateral optical fiber implantation. Optical fibers, formerly cleaned with 70% ethanol, were implanted in the coordinates calculated previously (**Figure 6**). Fibers were slowly lowered to the target region to reduce tissue damage, as well as damage of the fiber. After lowering the fibers, the four screws were implanted: two more anterior and the other two more posterior. Optical fibers were fixed to the skull using super bond c&b (CTBA) and all the skull surface was covered with dental cement (Simplex Rapid Powder, Kemdent®). After this, the wound was treated with Prontonsan® (wound irrigation solution and wound gel, B|Braun), which cleans, decontaminates, and moisturizes acute chronic skin wounds. Lastly, the wound was sutured.

Following surgery, rats recovered from anesthesia in a temperature-controlled cabinet and were given water with Rimadyl (0.5/mL) for three days post-surgery to minimize pain. Animals were single-housed and given one to two weeks to fully recover with *ad libitum* access to water and food before initiating food restriction and behavioral experiments. Weight was monitored daily and parameters such as locomotion and food/water intake were checked.

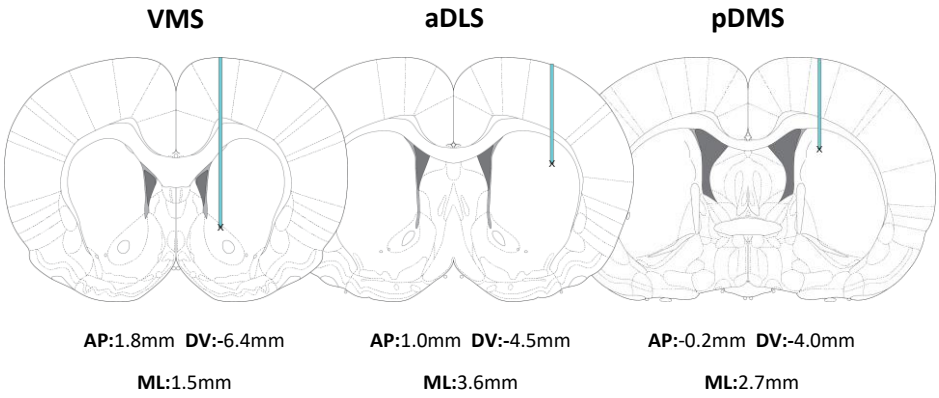

Figure 6: Brain coordinates used for viral injection and optical fiber implantation in the imaging cohort. Illustration of the coordinates used for viral injection and optical fiber implantation to target three regions of interest within the striatum (VMS, aDLS, and pDMS). Coordinates are given as Anterior-Posterior (AP), Dorsal-Ventral (DV), and Medial-Lateral (ML) and were calculated taking bregma as a reference point. Blue is a schematic representation of the track made by the optical fiber, with the black cross representing the location of the viral injection. In the imaging cohort, both optical fiber implantation and viral injection were done in the same coordinates. For the expression pilot cohort, the optical fiber was implanted in the same coordinates as in the imaging cohort but the viral injection was done 300µm below the viral location represented in these figures. Both viral injection and optical fiber implantation were done bilaterally. Illustrations of the coronal brain sections were purchased from Paxinos Rat Brain Atlas (2007, 6th edition).

Table 3: Viral constructs with respective volumes and concentrations for the expression pilot cohort and imaging cohort. The top table shows the two viral constructs used in the injections of the expression pilot cohort. Here, viral injections were done bilaterally with different volumes between hemispheres to optimize these variables. The bottom table shows the two viral constructs used in the imaging cohort injections, as well as the volume and concentration picked after optimization. In this cohort, both hemispheres were injected with the same volume and concentration of virus.

Expression pilot cohort

Viral construct	**Left Hemisphere** (High volume)	**Right Hemisphere** (Low volume)
1) AAV1.hEF1α.Dio.FLP	Volume = 3 x 1350 nL Concentration = 1:2	Volume = 3 x 900 nL Concentration = 1:2
2) AAV1.hSyn.FRT.GCaMP6f		

Imaging cohort

Viral construct	**Both Hemispheres**
1) AAV1.hEF1α.Dio.FLP	Volume = 1 x 1350 nL Concentration = 1:2
2) AAV1.hSyn.FRT.GCaMP6f	

3.4. Behavioral Assays

After recovering from surgery, animals were food restricted to 85-90% of their free-feeding body weight to increase their motivation to perform the behavioral task. Animals were habituated to the researcher before starting experiments.

Behavioral experiments were conducted in operant chambers (Med Associates Inc.) modified according to the behavioral task employed. In general, operant chambers were equipped with a food magazine linked to an automated food-pellet dispenser, a nose-poke hole, a house light, a metal grid floor (Med Associates Inc.), two levers, one white noise (WN) speaker, a tone generator and a magazine tray LED light (NIN mechatronics department). Operant chambers were placed inside metal Faraday cages revested with absorbing polyurethane foam to isolate sound. All boxes were equipped with a video camera for behavioral recordings.

For all behavioral tasks, before initiating the training period, rats were habituated to food pellets (Dustless Precision Pellets, Bio-Serv®) in their home cage. For this purpose, a few food pellets were given daily, together with chow (Mucedola S.R.L.), for three days.

3.4.1. Go/No-go Task

To study the role of distinct striatal subpopulations in outcome-dependent action control, we employed a novel and complex Go/No-go task (**Figure 7**) in which animals learned to work for different outcomes: food reward (positive reinforcement) or to avoid an aversive compound stimulus (WN paired with a magazine tray LED - negative reinforcement). Furthermore, to acquire these outcomes, animals have to either initiate an action (Go) or refrain from making a response (No-go). By associating both outcomes with both actions, we can compare neuronal activity within animal to these four conditions.

In this task, adapted from Syed *et al.* (2016)[67], animals are placed in an operant chamber as the one illustrated in **Figure 7A** and a trial starts when the nose poke (NP) hole light turns ON after a random interval of 5 to 15s. Animals are required to refrain from action and stay in the NP hole for 0.5s followed by the presentation of one out of four auditory tones, indicating one of the four trial types. In No-go trials, animals have to hold NP for 3s to either receive a reward (No-go reward) or avoid a punishment (No-go punishment). In Go trials, animals have to leave the NP and press a lever to either receive a reward (Go reward) or avoid a punishment (Go punishment). For all trials, the NP light and the auditory tone remain ON until the end of the trial. Sessions had a maximum duration of 120 minutes or 100 successful trials, whichever came first.

Task Progression

To habituate animals to the food magazine and pellet delivery, they were trained on two consecutive sessions of magazine training. In these sessions, animals were placed in the operant chamber and they received a total of 60 sugar pellets, timely spaced (60 seconds), without having to perform any action.

Afterward, animals started training in the Go/No-go task. Here, animals were trained in the four trial types in a specific order: No-go reward, Go reward, No-go punishment, and Go punishment. Progression to a new trial type was achieved when animals' performance (correct trials) was above 80% in two consecutive sessions.

No-go trials: In this first phase of training, animals learn to perform a NP to receive a food reward (NP short). The NP hold duration that animals need to perform to be successful starts with 10ms and increases overtraining until it reaches 0.5s. Next, we introduced the No-go reward training phase, in which, after performing a NP short of 0.5s, animals will hear a tone indicating that they are in a no-go reward trial type and therefore they have to hold the NP (No-go NP hold) for 3s to receive a reward. In an initial training phase, the No-go NP hold duration is 0.25s and it increases over time, in steps of 0.5s, until it reaches 3s.

Go trials: After acquiring a performance above 80% for No-go trials, animals were introduced to the Go reward training phase, which is divided into instructed Go training and non-instructed Go training. In instructed training, after a NP short of 0.5s, animals are presented with a tone that indicates a Go reward trial type, and a lever is presented in the operant chamber, delivering a reward upon pressing and consequently retracting. At the start of training, animals have 45s to press the lever but this time keeps decreasing until it reaches a maximum value of 3s. When the animals can perform well in instructed trials we move to non-instructed training. Here, the lever is extended throughout the session (not retraction upon pressing) and the animals need to pay attention to the tone presented to understand whether they have to press the lever or hold the NP to receive a reward. At the start, animals are trained to press only the right lever, and then only the left lever. However, when they are proficient in pressing each of the levers individually, we introduce block sessions, in which they have blocks of 5 trials per lever (right and left) within the same session.

For Go and No-go punishment, the training protocol is the same as described above but instead of working for a reward, animals are working to avoid the punishment. In this thesis report, we are only showing results for Go and No-go Reward since the punishment phase of training is still in progress.

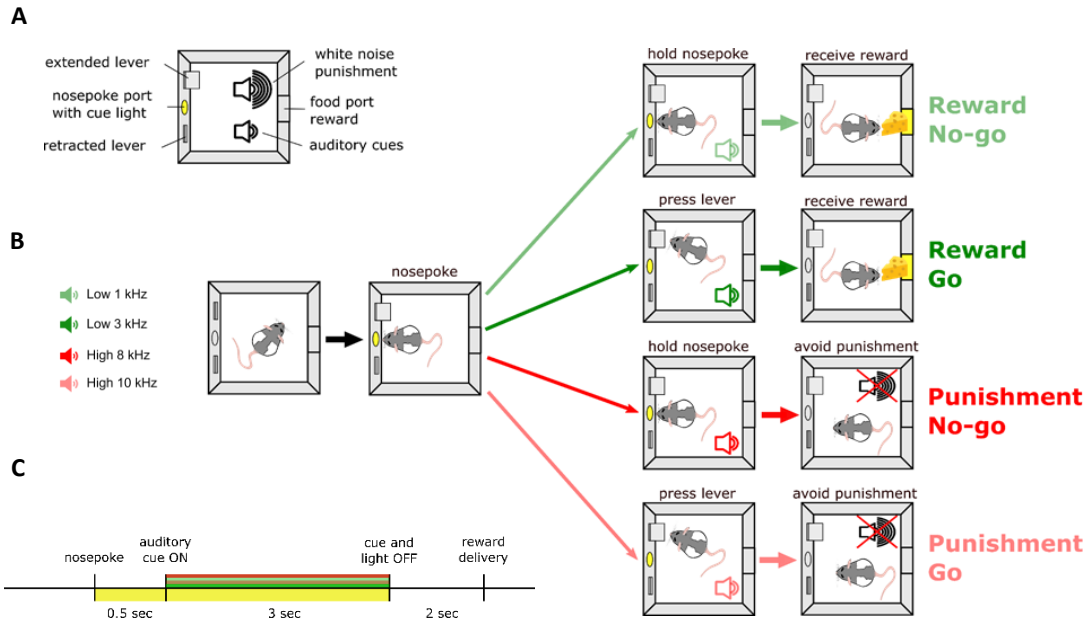

Figure 7: Schematic of the Go/No-go Task. (A) Representation of the experimental chamber used in this task. **(B)** Schematic of the four trial types: No-go Reward, Go Reward, No-go Punishment and Go Punishment. This task is based on four auditory cues with different frequencies. Each auditory cue is paired with a different condition as exemplified by the different tone and condition colors. Depending on the trial, rats need to hold nose-poke (No-go) or press a lever (Go) to either receive a reward or to avoid a punishment. **(C)** Timeline for the different task events.

3.4.2. Aversive/Appetitive Task

To better understand the subjective value of an aversive stimulus (WN paired with bright LED light, meant to be used as a punishment in our Go/No-go task) in relation to an appetitive stimulus (sugar pellet), we trained animals in an aversive/appetitive task.

We designed a lever press-based task (**Figure 8**) in which animals (n=5) were given the choice between two levers presenting different outcomes. The levers were identified as a "green lever" and a "red lever" for an easier distinction between lever outcomes throughout time. This task is divided into five stages aiming to evaluate how many sugar pellets are worth three seconds of a compound aversive stimulus (Comp) with continuous WN and LED and to confirm that the compound stimulus is considered aversive by the animal.

Briefly, at the beginning of the task, animals are placed in an operant chamber and a trial starts when a light in the NP hole turns ON, indicating that the animal needs to do a NP entry. After this, two levers, one on each side of the nose-poke hole, are introduced in the chamber to give the animal the choice of which one to press. In the first stage of the task, both levers deliver a single sugar pellet (1P vs 1P) but as training progresses, one of the levers starts being paired with the compound stimulus. Hence, the second stage is then characterized by one lever delivering a sugar pellet (green lever), and the other lever delivering a sugar pellet paired with the compound stimulus (red lever). Moreover, in the third stage of the task, we paired one lever with 1 pellet (green lever) and the other with 2 pellets (red lever) to confirm that animals choose the lever based on its outcome. Afterward, in stages four and five, we again paired one of the levers with the compound stimulus: the red lever delivers 2 pellets and the compound stimulus, while the green lever delivers only 1 pellet. The difference between stages four and five relies exclusively on the WN intensity. In stage four, we employed WN at 90dB, an intensity previously described to be effective in the promotion of aversiveness.[68] In stage five, we used WN at 70db to validate whether lower intensities have less subjective value for the animal.

Additionally, since we noticed difficulty of the animals to understand outcome contingencies upon lever press when we trained them in stages one and two, we introduced a reversal from stage 3 onwards to test their attention to the task and their capacity of making the lever-outcome relationship. Thus, after 30 consecutive presses in one of the levers, their contingency reversed: the lever that was on the right side becomes the one on the left side but delivers the same outcome as when in its initial position. This way, we confirmed that animals understand the outcome contingency of each lever.

Task Progression

As in previous behavioral essays, animals were habituated to the food magazine and food reward by undergoing two initial sessions of magazine training (60 sugar pellets, time spaced by 60 seconds). Next, they were trained for nose-poke entry and lever pressing to be able to perform the different stages of the task. Nose-poke training was composed of sessions with a maximum length of 120 minutes or 60 rewards, whichever came first, with the nose-poke entry being located between the two levers to assure that when a trial starts, the animal is at an equal distance to both of them. Like this, we can exclude the distance to the lever as a conditioning variable to lever choice.

After achieving a good performance in nose-poke entry, animals started to be trained for lever pressing. Thus, we started by introducing them to a lever that they could press without any retraction of the lever, ending up receiving one pellet per press. Next, we introduced a second lever in the operant chamber (inactive lever)

that would not retract upon pressing nor deliver a reward. Lastly, after having demonstrated a clear lever-outcome understanding, animals were trained in the different five stages of the task (**Figure 8**), explained above. It is important to notice that in stages where the compound stimulus is present, we tried to reduce the number of training sessions as much as possible to avoid potential habituation to the sound.

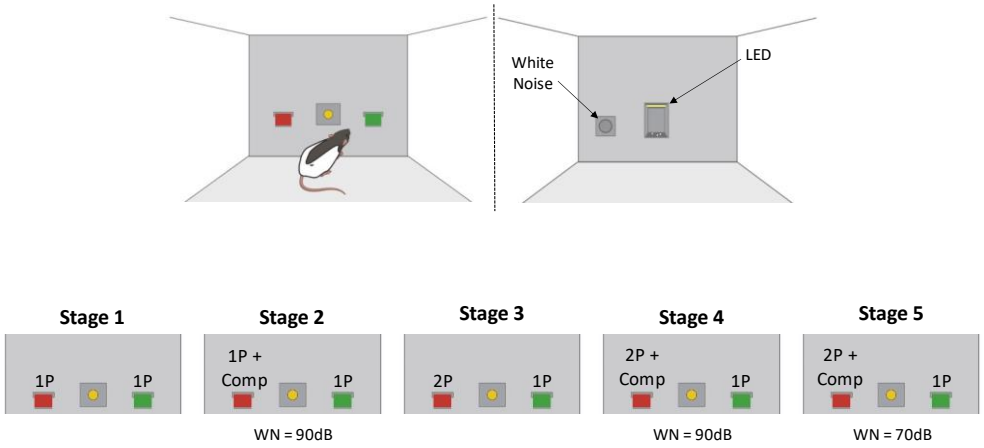

Figure 8: Schematic of the Aversive/Appetitive Task. Top panel - Configuration of the operant chamber used for this task. The left side shows the nose-poke hole placed in between the two levers, which are presented with different colors, aiming to an easier understanding of lever outcomes throughout the stages of the task. The right side shows the position of the WN speaker (right near the food magazine) and the LED. **Bottom panel –** Illustration of the different stages of the task. Red is associated with the lever presenting the bigger outcomes: more pellets (P) or pellets paired with the compound stimulus (Comp), except for stage 1, where both levers deliver the same outcome. Adapted from Biorender®.

3.5. Fiber Photometry

3.5.1. Optical Fiber Manufacturing

All optical fibers used in this project were made in-house. Different ceramic ferrules and optical fibers were used in the expression pilot cohort and imaging cohort to optimize the fiber photometry recordings. The specifications for the different materials can be found in **Table 4**.

To prepare optical fibers to be used in the fiber photometry experiments, we started by taking out the Tefzel® coating in which the optical fiber (Thorlabs) comes encased, using a fiber stripping tool (Thorlabs-T06S13). This coating is important to prevent any damage that can compromise the performance of the fiber (scratching or breakage) but upon fiber preparation, it needs to be removed. Following, we used either a diamond pen (Thorlabs Inc - 973-579-7227 – used for 200µm fibers) or a Shortix® tubing cutter (Supelco analytical - 21386-U – used for 400µm fibers) to cut the fiber in the desired length, accounting for the implantation length, ferrule size, and space for polishing and cementing. Thereafter, we placed the optical fiber inside a ceramic ferrule (Thor labs), and with the assistance of a microscope (Zeiss Stemi 200) to check the fiber surface, and a polishing disk (Thorlabs - D50-FC), we started polishing one side of the optical fiber until it becomes completely flat and shiny. This is an important step to maximize the transmission power of the fiber and decrease the scattering of light. To achieve this, we took advantage of several polishing steps using different polishing/lapping sheets (Thorlabs) with different size grits placed in a glass polishing plate

(Thorlabs – CTG913) with a rubber polishing pad (Thorlabs – NRS913A). Note that, to end up with a well-polished fiber, as polishing steps progress, a finer grit sheet was used.

Next, we took the fiber out of the ferrule, reversed it, and after adding a drop of cement (Precision Fiber Products, ref. PB199234 and PB199325) on top of the ferrule we again placed the fiber inside it. We then cleaned the excess cement in both ends of the ferrule and with a hot source, we dried it (**Figure 21**, Appendix). Once again, we polished this side of the fiber using a polishing disk and several polishing steps, as explained above.

Lastly, to confirm the performance of each optical fiber we measured the percentage of light that could pass through it using a blue laser (Laser century – Shanghai & optics century., LTD) and a digital optical power and energy meter (Thorlabs – PM100D). In our experiments, we only used optical fibers with a percentage above 80%, aiming to increase the detection of neuronal population activity in our recordings.

Table 4: Optical Fibers and Ceramic ferrules used in Fiber Photometry experiments for the expression pilot cohort and imaging cohort. The inner and outer diameter of the ceramic ferrule, as well as the diameter of the optical fiber, are different between cohorts. Images show optical fibers made in-house and prepared to be implanted.

	Ceramic Ferrule	Optical Fiber	
Expression pilot cohort	Inner diameter = 230µm Outer diameter = 1.25mm	Numerical Aperture = 0.5 Diameter = 200µm	
	Ceramic Ferrule	**Optical Fiber**	
Imaging cohort	Inner diameter = 440µm Outer diameter = 2.5mm	Numerical Aperture = 0.5 Diameter = 400µm	

3.5.2. Fiber Photometry Recordings

Rats used for fiber photometry experiments were trained in the Go/No-go task and recorded two times: first in a signal testing session and secondly when they were proficient in Go and No-go reward conditions. All sessions had a maximum duration of 120 minutes.

The recordings were done using a fiber photometry setup from Tucker-Davis Technologies (TDT).[69] The TDT's LUX RZ10X, with incorporated LED drivers, LEDs, and photosensors, was controlled by a Synapse software for recording and visualization in real-time of the fluorophore responses (GCaMP – Activity-dependent) with behavioral time-stamps. Additionally, by the end of the session, data can be easily exported to MATLAB to further analysis. This system comes with a Lux RZ10X processor that, in turn, permits the recording of two distinct animals simultaneously. Moreover, the system also has a lock-in amplifier which, not only makes it possible to subtract the fluorophore signal from the noisy raw signal but also the use of the isosbestic signal of GCaMP (a control for movement and other artifacts since this signal is not activity-dependent). This is possible due to its capacity of powering multiple LEDs at different frequencies. In our recordings, we made

use of this advantage and recorded from two channels: GCaMP (465nm) and the GFP isosbestic signal (405nm).

This system was used together with MedPC (through a MedPC SuperPort card), which sent timestamps of behavioral signals to the TDT system and allowed the visualization of identified behavioral events in real-time.

3.5.3. Signal test recording session

To test the quality of the signal of the animals used in this study, we performed a short recording session before initiating further experiments. In this recording session, animals were connected to the fiber photometry set up (through the optical fiber previously implanted) and placed in an operant chamber, where they were given five rewards with randomized inter-trial intervals.

Given the role of striatal regions in motivated behavior, we can confirm the presence of altered neuronal activity upon reward delivery. Additionally, the randomized inter-trial interval allows us to exclude any association of the signal with reward prediction instead of with the reward itself.

3.6. Histology and immunofluorescence

After finishing all behavioral experiments, rats were sacrificed and *ex vivo* assessment was used to confirm the correct location of the viral injections and optical fibers implantation.

Note that histology was only performed for the expression pilot cohort since animals from the imaging cohort are still alive and being trained in the Go/No-go task.

Preparation of brain tissue

Rats were anesthetized through an intraperitoneal injection of pentobarbital with the lack of reflexes being confirmed by toe pinching. Animals were then placed in the perfusion table and, using forceps and a scissor, an incision was made above the chest cavity to expose the heart. Next, a blunt needle was precisely placed within the heart of the animal in the ascending aorta and an incision was made in the right atrium. Following, using the perfusion pump, 20mL of phosphate buffered saline (PBS) 1x was pumped through the blunt needle to flush the vascular system (take all the blood out of the body). When no more blood was leaving the body, the tube connected to the needle was switched to 4% paraformaldehyde (PFA) to fixate brain tissue. The tube was maintained in PFA until the body became completely stiff.

After this process, animals were decapitated using a guillotine and their heads were stored in 4% PFA for post-fixation for 24 hours at 4^{o}C.

Afterward, the brains were removed from the head and placed in a solution composed of 30% sucrose and PBS 1x, in which they remained for a minimum period of two days. This is a crucial step since sucrose is used as a cryoprotectant to remove the water present in the brain, avoiding the formation of ice crystals upon brain freezing. The absence of this step could lead to cellular architecture damage, preventing the future use of this tissue.

Additionally, to properly section the brain without any tissue damage, we performed a snap-freezing protocol. Firstly, the brains were removed from the 30% sucrose/PBS 1x solution and placed in a petri dish containing Tissue-Tek® O.C.T.™ compound (Sakura). This step is once again important to reduce even more the amount of water around the tissue and to consequently avoid the formation of ice crystals.[70] The brain was then dropped for 6-8 seconds in a beaker containing 50mL of 2-methyl butane (Isopentane - Sigma Aldrich) and embedded in dry ice. After the designated time, the brain was removed to dry ice, moved into its original container, and stored at - 80^{o}C.

Next, coronal sections of 40µm were cut using a temperature-controlled cryostat (Leica SM3050), and slices containing the regions of interest (VMS, aDLS, and pDMS) were sequentially collected to a 24 well-plate previously filled with 0.02% PBS sodium-azide. This solution aims to prevent microbial growth, allowing us to store our brain tissue preserved for a longer period of time. The well-plate was then covered with aluminum foil to protect from light and stored at 4^{o}C.

Immunofluorescence staining and image acquisition

To confirm the correct location of the viral injection and optic fiber implantation, we performed an immunofluorescence staining for GFP (which amplifies GCaMP), and DAPI, a fluorescent marker that strongly binds to DNA regions that are rich in A-T, allowing the labeling of the cell's nucleus.[71] Briefly, we started this staining by washing the coronal brain sections three times with PBS 1x (10 minutes per wash) and by adding 0.3% Triton X-100 to the third wash for tissue permeabilization. Sections were then incubated in a blocking solution (5% bovine serum albumin (BSA), 5% normal donkey serum (NDS), 0.2% Triton X-100, and PBS 1x) for one hour at room temperature with agitation (shaker: Heidolph Instruments - Unimax 1010). Next, the primary antibody rabbit anti-GFP (Invitrogen by ThermoFisher, 1:1000) was added to the blocking solution, and the sections were incubated overnight at 4^{o}C with agitation. The following day, brain sections were washed four times with PBS 1x for 10 minutes and were then incubated in a blocking solution, which composition was previously described, with the secondary antibody donkey anti-rabbit A488 (Jackson Immuno Research Laboratories inc., 1:1000). Incubation was done for one hour at room temperature with agitation. Lastly, the sections were washed four times with PBS 1x, and DAPI (Sigma Aldrich,1:1000) was added to the second washing step. All the washes had a duration of 10 minutes.

Coronal brain slices were then organized from anterior to posterior in a dish containing PBS 1x and mounted in the right order in coated glass slides (Thermo Scientific). Afterward, they were cover-slipped (coverslips - Thermo Scientific) using Mowiol mounting medium (Calbiochem), and the edges of the slide were sealed with nail polish.

Lastly, images were collected for each coronal section using the Axio Scan.Z1 slide scanner (Zeiss). Images were acquired using two channels: DAPI for morphology (blue channel) and GFP for identification of correct viral injection location (green channel).

3.7. Data Analysis

Graphs generation and statistical analysis were performed using GraphPad Prism (GraphPad software version 8.1.1, California, USA). For all experiments, significance was set as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not significant.

Viral expression optimization

Area of viral expression and mean fluorescence intensity were calculated for the expression pilot cohort using Fiji-ImageJ (**Figure 9**). Both low volume and high volume injection sides were analyzed to establish a comparison and pick the best volume of the virus to use in further experiments (Imaging cohort). Therefore, raw figures collected with the slide scanner were opened in Fiji-ImageJ and after splitting channels, the green one was used for the analysis. A region with viral expression (region of interest) and a region without viral expression (background) was selected, and with the measurement tools of this program, the area in pixels and mean fluorescence intensity were calculated. For the latter, the values of the background were subtracted from the values of the region of interest to exclude brain autofluorescence. For every region studied, the analysis was done for two slices and the final value used was their mean.

Data were first tested for normality using the Shapiro-Wilk normality and lognormality test. Comparisons between volumes per region, regions, and volumes were done using a Two-way ANOVA test with multiple comparisons Sidak's test.

Aversive/appetitive task

The subjective value of the stimulus compound compared to a sugar pellet was evaluated through a red presses ratio (**Figure 10**). Since the red lever is associated either with more pellets or pellets paired with the compound stimulus we wanted to see, for each condition, whether the number of presses was bigger or smaller depending on its outcome. For this, we present our data as a red presses ratio, which was achieved by dividing the number of red presses by the number of total presses ([red presses/red presses + green presses]). Thus, in a ratio from 0 to 1, values closer to 1 indicate more red presses (preference for high reward value with compound stimulus), and values closer to zero indicate more green presses (low reward value with no compound stimulus). For conditions without compound stimulus, the ratio was built using the mean of the last two training sessions and for conditions in which the compound stimulus was introduced, we used the mean of the second and third sessions, to avoid any effect of habituation to the noise in our results.

As above, data were first tested for normality using the Shapiro-Wilk normality and lognormality test. A comparison between 1P vs 1P and 1P+Comp(90dB) vs 1P (two conditions tested in the same subjects) was done using a paired and parametric t-test (two-tailed; confidence interval of 95%). Furthermore, a comparison between 2P vs 1P, 2P+Comp(90dB) vs 1P, and 2P+Comp(70db) vs 1P (three different conditions) was done using a Friedman's test followed by a Dunn's multiple comparisons post-hoc analysis.

Behavioral performance in training sessions

Evaluation of animals' performance during training sessions in the Go/No-go task was done through the percentage of correct trials in each trial outcome (Go Reward and No-go Reward). Here, we first present the percentages for every training session until the first day of imaging and we then analyzed specific events, such as the percentage of correct trials when animals are proficient in the task, number of magazine entries, and latency to magazine entry after successful block trials (**Figure 13**). For all these events, graphs were generated using the mean of the three last sessions in which animals performed that specific trial type (i.e. Go right, Go left, Go blocks, and No-go).

Data were first tested using the Shapiro-Wilk normality and lognormality test. Comparisons between Go (Go right, Go left and Go blocks) and No-go conditions in the last three sessions of training were done using a

paired and non-parametric t-test (two-tailed; confidence interval of 95%): Wilcoxon matched-pairs signed-rank test. Furthermore, comparisons between Go blocks and No-go conditions regarding number and latency to magazine entry after successful trials were done using paired and parametric t-tests (two-tailed; confidence interval of 95%).

Behavioral performance in recording sessions

As previously described for training sessions, behavioral performance was also evaluated during the *in vivo* fiber photometry recording sessions. Graphs were generated using the average of two recording sessions per animal, and Go right, Go left, and No-go trials were considered (**Figure 14B**).

Data were first tested for normality using the Shapiro-Wilk normality and lognormality test. Comparisons between conditions (Go right, Go left, and No-go) were done using a repeated measure one-way ANOVA with a Silk post-hoc test.

***In vivo* Fiber photometry**

Fiber photometry data were pre-processed and plotted using custom-made MATLAB scripts (MathWorks, version R2020b).

All data were first pre-processed before conducting any analysis. Briefly, we started by correcting the raw data for motion artifacts. Here, we scaled the 405nm (Isosbestic) to the 465nm (GCaMP6f) channel using a polynomial fit function. After this, we obtained the ΔF/F by dividing the corrected trace by the polynomial fit. We then high pass filtered (30Hz) and low pass filtered (3Hz) the signal to attenuate high- and low-frequency oscillations. Lastly, we down-sampled the signal by a factor of 64 and used the resulting signal for further analysis.

In addition, imaging and behavioral data were aligned to allow inferring about changes in calcium transients related to specific task events (e.g. nose-poke entry, lever press, reward, etc.).

Signal test recording session

To test the signal obtained with different optical fibers and ferrules and therefore optimize the fiber photometry technique and guarantee that we could obtain a good neuronal signal and stability in the connections made between the ferrule and the fiber optic cable, we performed a test recording session with two cohorts of animals (expression pilot cohort and imaging cohort). Neuronal data was plotted as a heatmap and the resulting mean activity graph showing neuronal activity before and after reward delivery (**Figures 11** and **12**). These plots were built using 5 trials per animal (each one represented by a different color in the y axis), and the striatal subregion recorded from was the ventral medial striatum.

Values of ΔF/F in the heatmaps and mean activity graphs were normalized to the same range values across transgenic lines and cohorts (for heatmaps it ranges from -1.5 to +1.5 - represented in the color bar - and for the mean activity graph it ranges from -0.5 to +0.5). Like this, proper visual evaluation and comparisons can be done.

Go/No-go task recording sessions

To account for neuronal changes related to specific actions/events during the performance in the Go/No-go task, we assessed neuronal activity during five specific events related to movement and motivation: nose-poke entry, nose-poke exit, right lever press, left lever press, and reward delivery. For each of the events, a heatmap comprising all the trials for all the animals (divided by transgenic line and subregion) and the respective mean activity graph were plotted (**Figures 14 – 19**).

Values of ΔF/F for the heatmaps and mean activity graphs were normalized to the same range of values across transgenic lines and subregions (for heatmaps it ranges from -1 to +1 - represented in the color bar - and for the mean activity graph it ranges from -0.5 to 0.5). This allows an easier visual evaluation and a more accurate comparison between data.

Note that for all experiments using *in vivo* fiber photometry (signal test recording session and Go/No-go task recording sessions) no statistical analysis was performed due to the low number of animals. Given that not all the animals reached the level of behavior proficiency required to perform the recordings at the same time, only the ones that did were recorded and included in this study. Therefore, in the future, more data from the remaining animals will be added to the results, increasing the n and allowing a correct statistical analysis. Additionally, since these animals are still alive and performing the task, we do not have their histology, meaning that we cannot confirm that all viral injections and fiber implantation were done in the correct coordinates, which can influence the results. We also found that TDT can have random TTL signals, making the alignment of data really hard. This can be another factor that might affect the data.

Having this, for all neuronal data, we will be inferring based on a visual analysis of the heatmaps and mean activity graphs.

4. Results

In this project, we studied the role of striatal subpopulations (D1- and D2-MSN) in outcome-dependent action control, and how it differs between striatal subregions (VMS, aDLS, and pDMS). For this purpose, we recorded from these subpopulations using fiber photometry cell-type-specific calcium imaging while animals performed the Go/No-go task.

Before initiating the neuronal recording experiments, we started by optimizing important steps of the project that would contribute to the success of the recording sessions. Here, we tested aspects like viral approach and volumes, that should be used to increase expression, and ceramic ferrules and optical fibers that should be implanted to improve fiber photometry signal and cable connection. Additionally, we also developed an aversive/appetitive task to probe the subjective value of an aversive compound stimulus in relation to a sugar pellet and, therefore, be able to compare neuronal signals following reward and punishment in the same subject. After that, we trained animals in the Go/No-go task and evaluated their performance to assure that they perform above chance. When they were proficient in the task, we recorded them using *in vivo* fiber photometry. Hence, the results presented in this section, show the optimizations, behavioral performance, and neuronal recordings that were done throughout this project.

4.1. A dual-vector viral approach and a higher volume of virus improve GCaMP6f expression in VMS, aDLS, and pDMS.

Regular *Cre*-dependent viral constructs did not provide sufficient viral expression in striatal subregions (**Figure 9A** - AAV9.hSyn.DIO.GCaMP6f construct) in *cre* expressing rats. In an attempt to overcome suboptimal expression we used a new approach that combined two different viruses (AAV1.hEF1α.Dio.FLP and AAV1.hSyn.FRT.GCaMP6f). Thus, to confirm the efficacy of this approach and optimize the volume of virus that leads to a better expression we started by injecting an expression pilot cohort (EPC) of animals (composed of D1-i*Cre* and A2a-i*Cre* animals) with a high volume of virus (3x1350nL (1:2)) on one hemisphere and a low volume of virus (3x900nL (1:2)) in the other hemisphere. Injections were done in the three striatal subregions (VMS, aDLS, and pDMS), and mean fluorescence intensity and area of expression were evaluated.

Ex vivo histology results showed improved GCaMP6f expression using the dual-vector viral approach in relation to the use of a standard *Cre*-dependent viral construct (**Figure 9A**: EPC figure (left panel) versus figure with the AAV9.hSyn.DIO.GCaMP6f construct - bottom row, third panel).

Regarding optimization of the volume of the virus, the expression pilot cohort showed a significant increase in the area of expression when using higher volumes of virus compared to low volumes. No significant differences were found between distinct striatal subregions (**Figure 9C**, two-way ANOVA, volume effect: $F_{1,9}$=8.601, p=0.0167; region effect: F2,9=2.729, p=0.1185). Differences between the area of expression in hemispheres injected with high volume (left hemisphere) and low volume (right hemisphere) can also be seen in **Figure 9B**, in which the area of expression, per subregion, calculated from the *ex vivo* brain sections of four animals is plotted.

Moreover, fluorescence mean intensity, a measurement of signal strength, was also found to be greater in the hemisphere injected with a higher volume of the virus. Once again, no significant differences between striatal subregions were found (**Figure 9D**, two-way ANOVA, volume effect: F1,9=6.550, p=0.0307; region effect: F2,9=1.045, p=0.3907).

Furthermore, considering the results obtained, a new cohort of animals (imaging cohort) was bilaterally injected with the high volume of virus previously tested in the expression pilot cohort. *Ex vivo* sections of example animals are presented in **Figure 9A** (images labeled with IC, meaning imaging cohort) and show sufficient viral expression. Additionally, histology results for both the expression pilot cohort and imaging cohort, show a correct location of viral injections and optical fibers implantation.

Together these data suggest that increased GCaMP6f expression can be achieved in VMS, aDLS, and pDMS by using a dual vector viral approach with a high volume of virus (1350nL).

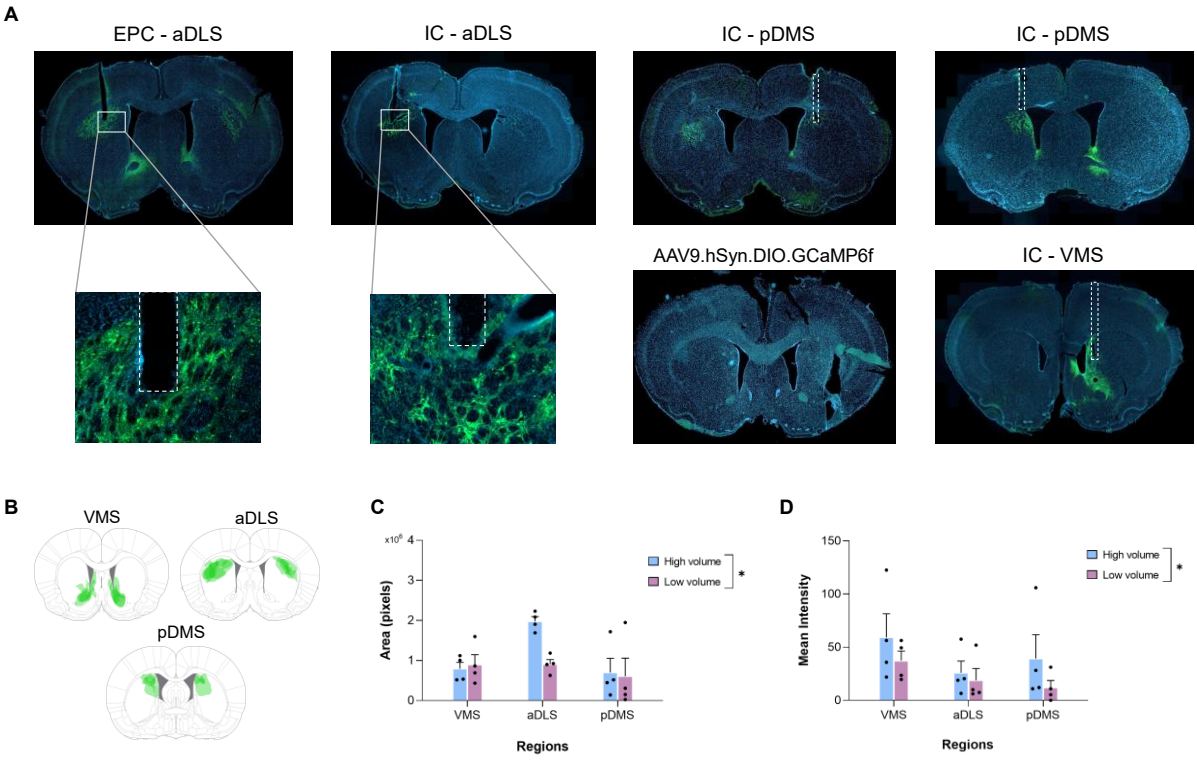

Figure 9: Viral expression in VMS, aDLS, and pDMS is significantly improved with a dual-vector viral approach and high volumes of virus injections. (A) *Ex vivo* representative examples of optical fiber placement and GCaMP6f expression using a dual-vector viral approach (AAV1.hEF1α.Dio.FLP and AAV1.hSyn.FRT.GCaMP6f) or a standard *Cre*-dependent viral construct (AAV9.hSyn.DIO.GCaMP6f). Green represents GCaMP6f expression, and the white dashed lines represent the track made by the optical fiber; **(B)** Representation of GCaMP6f expression location per animal (n=4) and per region (VMS, aDLS, pDMS). Area of expression was calculated per animal and overlapped in a Paxinos rat brain atlas section with the same degree of transparency. Darker areas indicate the location in which more animals got expression. Left hemisphere was injected with high volume of virus and right hemisphere was injected with low volume of virus; **(C)** Area of viral expression in pixels in VMS, aDLS, and pDMS for the expression pilot cohort. **(D)** Fluorescence mean intensity in VMS, aDLS, and pDMS for the expression pilot cohort. For panels **(C)** and **(D),** data are mean±SEM and for every region n=4, with each black dot representing an individual animal. Blue bars represent the hemisphere injected with a higher volume of virus and purple bars represent the hemisphere injected with a lower volume of virus. Statistical significance was determined by a two-way ANOVA test. *p<0.05. EPC, Expression pilot cohort; IC, Imaging cohort.

4.2. Three seconds of continuous compound stimulus with 90dB of WN is aversive for the animals and has similar subjective value as one sugar pellet.

To be able to compare neuronal activity data from reward and punishment conditions in the Go/No-go task, we needed to assure that we are evaluating outcomes with a similar subjective value. Thus, we need to know what the subjective value of a compound stimulus (reinforcer aimed to be used in punishment conditions) is in relation to a sugar pellet (reinforcer used in reward conditions).

We establish the aversiveness and subjective value of the continuous compound stimulus by training 5 animals (two females and three males) in an aversive/appetitive task (lever-based task) in which they were given the choice to press between two levers with different outcomes. In some sessions, both levers delivered sugar pellets only, while in other sessions one of the levers was paired with compound stimulus, as explained in section 3.4.2. of Materials and Methods.

In the first stage of the task (1P vs 1P - **Figures 10A** and **10B** (blue bar)), data showed a red press ratio of 0.6, meaning that animals press both levers with almost the same frequency. Additionally, in the second stage of training (red lever now paired with 1 pellet and 3s of continuous compound stimulus with white noise at 90dB), results exhibited a shift in animals lever preference towards the green lever (1P), with a red presses ratio value below chance (below 0.5) (**Figure 10B** - purple bar). Although animals increased pressing in the green lever in stage 2, this difference turned out not significant (**Figure 10B**, paired t-test, t=1.064, df=4, p=0.3472).

One explanation for reduced suppression of red lever presses during stage 2 could be insufficient lever-outcome associations. To induce stronger lever-outcome associations and prevent preferred lever-pressing, we introduce reversals in stages 3 and up (**Figures 10C** and **10D**). After 30 consecutive presses on one lever, the lever-contingency is reversed. We hypothesized that this would induce stronger lever-outcome associations.

During the third stage of the task (2P vs 1P), a red presses ratio of almost 0.8 was obtained, showing that animals developed a preference for the red lever. During these sessions, lever-outcome contingencies were reversed after 30 consecutive presses, meaning animals actively switched lever side during testing (**Figure 10D** - blue bar). Next, in stage 4 (2P+Comp. with WN at 90dB vs 1P), in which the compound stimulus was reintroduced, animals shifted their preference towards the green lever (no compound stimulus), showing red ratio presses above 0.4 (**Figure 10C and 10D** - purple bar). Overall, data shows that the red presses ratio is significantly lower in conditions with a compound stimulus with WN at 90dB compared with conditions without this stimulus (**Figure 10D**, Friedman's test with Dunn's post-hoc analysis, 2P+Comp(90dB) effect: Z=2.846, p=0.0089).

Finally, in the 5th stage of this task, in which the outcomes of the levers were the same as in stage 4 but with WN at 70dB instead of 90dB, the results showed a red presses ratio of 0.6, meaning that animals press more for the red lever than in the previous condition (stage 4). Additionally, when comparing results from stage 5 with conditions without compound stimulus (stage 3), no significant difference was found regarding the red presses ratio (**Figure 10D**, Friedman's test with Dunn's post-hoc analysis, 2P+Comp(70dB) effect: Z=0.949, p=0.6856, not significant).

Together, these results showed that animals prefer one pellet over two pellets coupled with a high WN compound stimulus. Importantly, the red lever ratio during 2P+Comp(90dB) vs 1P was 50% of the red lever

ratio lacking compound stimulus, which indicates that 3s of continuous compound stimulus with WN at 90dB is aversive for the animals and has the same subjective value as one sugar pellet. Therefore, the compound stimulus (90dB WN + LED) can be used as an aversive reinforcer (punishment contingencies) in the Go/No-go task to compare neuronal signals with reward.

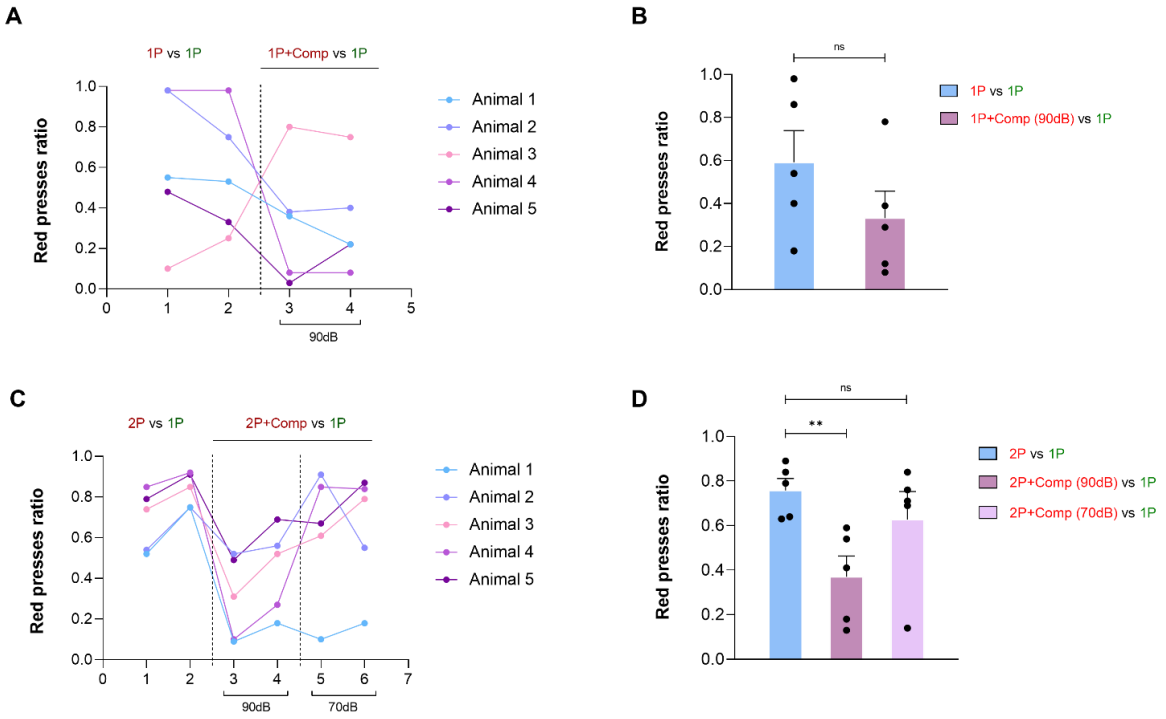

Figure 10: Three seconds of continuous compound stimulus with 90dB of WN is aversive for the animals and has the same subjective value as one sugar pellet. (A) Red presses ratio per single animal in 1P vs 1P and 1P+Comp(90dB) vs 1P conditions (n=5 animals). **(B)** Average of red presses ratio in sessions with 1P vs 1P (blue bar) and 1P+Comp(90) vs 1P (purple bar) conditions. **(C)** Red presses ratio per single animal in 2P vs 1P, 2P+Comp(90dB) vs 1P, and 2P+Comp(70dB) vs 1P conditions (n=5 animals). **(D)** Average of red presses ratio in sessions with 2P vs 1P (blue bar), 2P+Comp(90dB) vs 1P (purple bar) and 2P+Comp(70dB) vs 1P (pink bar) conditions. Red and Green in the legend indicates the outcome of each lever. For all panels red presses ratio was calculate as: [red presses/red presses + green presses], with results closer to 1 showing a preference for the red lever and results closer to 0 showing a preference for the green lever. For panels **(A)** and **(C)**, x axis represents the session number. For panels **(B)** and **(D)**, data are presented as mean±SEM with every dot representing the mean value of two sessions per individual animal. For conditions without compound stimulus, the ratio was built using the mean of the last two training sessions and for conditions in which the compound stimulus was introduced, we used the mean of the second and third sessions, to avoid any effect of a possible habituation to the noise in our results. Statistical significance was determined by a paired t-test (panel B) and Friedman's test followed by a Dunn's multiple comparison post-hoc analysis (panel D). **p<0.005, ns - not significant. Comp, Compound Stimulus; P, pellet; WN, White Noise.

4.3. *In vivo* fiber photometry signal test session

In this thesis project, we set up and consecutively tested *in vivo* fiber photometry in the lab. We started by testing the quality of the signal we could obtain with our imaging setup by performing a signal test session in which animals were given five rewards with random inter-trial intervals while VMS was being recorded (section 3.5.3. - Materials and Methods). To test and optimize both the signal and the connection (between

the ferrule in animals' head and the fiber optical cable from the set up) we used different ferrule and optical fibers between the expression pilot cohort and imaging cohort (see **Table 4** - Materials and Methods section 3.5.1.). The test signal recording session was performed for D1-i*Cre* and A2a-i*Cre* animals in both cohorts.

Due to the low number of animals, we did not perform statistical analysis on the neuronal data obtained from this session. Conclusions for all the results presented in this section will be inferred based on a visual analysis of the heatmaps and mean activity graphs that were obtained. It is important to remind that there is no histology data for the majority of the animals from the imaging cohort because they are still alive and being trained in the Go/No-go task. This means that we cannot guarantee that the optical fibers of all the animals included in the study are in the correct location.

Moreover, regarding the neuronal data obtained from the signal test recording session for D1-MSN, we started by evaluating the quality of the signal by looking at the GCaMP6f and isosbestic channel traces from an example animal. For both cohorts, the traces showed clear differences, with peaks of activity present in the GCaMP6f channels that were absent in the isosbestic channel (**Figure 11A** and **Figure 11B**). Furthermore, D1-MSN from animals present in the expression pilot cohort showed increased activity in the period post reward delivery (**Figure 11C**). In the heatmap, this is seen through a slight change to a more yellow color after the reward is delivered, which is indicative of increased neuronal activity. In the mean activity graph, this increase is seen through the peak of activity in the post-period (blue square). Additionally, in the imaging cohort, we see the opposite. The heatmap and mean activity plot showed an accentuated suppression of activity after reward delivery – blue color in the heatmap indicating a decrease in ΔF/F and a decrease in the trace of mean activity (**Figure 11D**).

Regarding the neuronal data obtained for the D2-MSN, different patterns of activity were found between the GCaMP6f and isosbestic channels in both cohorts, with activity peaks in the GCaMP6f channel that were not present in the isosbestic channel (**Figure 12A** and **Figure 12B**). Additionally, the activity of this neuronal subpopulation was seen to be increased in the expression pilot cohort after reward delivery, which is reflected by lighter yellow colors in the heatmap (increased ΔF/F) and an activity peak in the post-event period in the mean activity graph (**Figure 12C**). On contrary, and as seen before, D2-MSN from animals in the imaging cohort show suppression of activity after reward delivery (**Figure 12D**).

Next, during the signal test recording session, we also accounted for the stability of the connections established between the fiber implanted in the animals' heads and the fiber optic cable to which the animals are connected during the recordings. Although there is no data figure regarding this matter, based on what we observed during the signal test recording session, a decrease in the number of disconnections in the imaging cohort occurred, most likely due to the bigger diameter of the ferrules providing increased stability. The number of optical fibers being pulled out of the headcap during disconnection was also decreased in this cohort.

Overall, these results show that we have signal in both cohorts with higher connection stability when using the type of optical fibers that were implanted in the imaging cohort (bigger ferrules + optical fiber). The area we can record from it is also larger in this cohort (double the diameter of the optical fiber). Lastly, no definite conclusions can be drawn from the neuronal data of the imaging cohort due to the lack of histology.

Dopamine 1 Medium Spiny Neurons

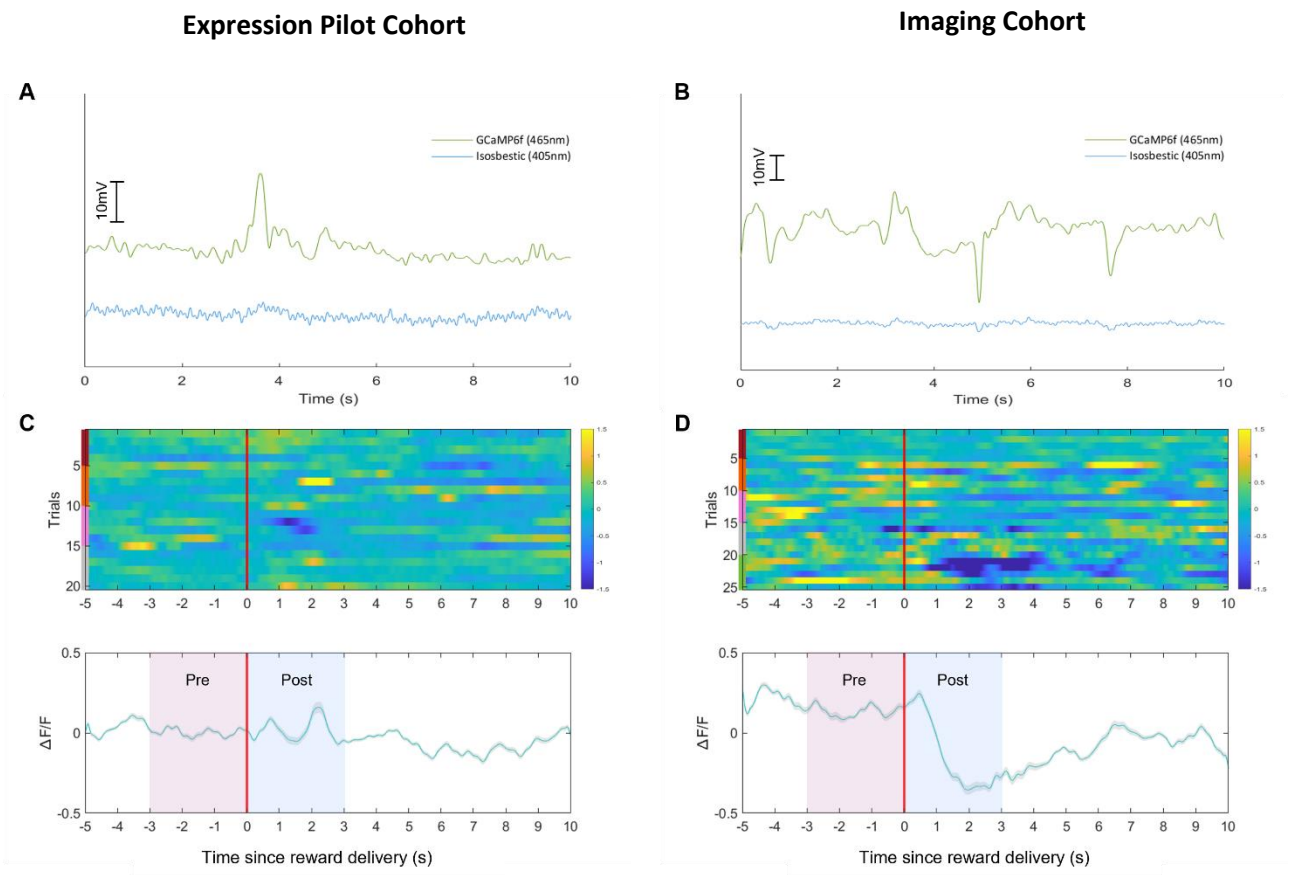

Figure 11: Neuronal activity of D1-MSN in VMS during the signal test recording session for expression pilot cohort and imaging cohort. (A) Example traces of isosbestic (blue - 405nm) and GCaMP6f (green - 465nm) channels from one D1-iCre animal of the expression pilot cohort. **(B)** Example traces of isosbestic (blue - 405nm) and GCaMP6f (green - 465nm) channels from one D1-iCre animal of the imaging cohort. **(C)** Top panel shows GCaMP6f fluorescence around reward delivery for four animals from the expression pilot cohort - five trials per animal. Bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of all the animals in the expression pilot cohort. **(D)** Top panel shows GCaMP6f fluorescence around reward delivery for five animals from the imaging cohort - five trials per animal. Bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of all the animals in the imaging cohort. For top panels of figures **(C)** and **(D)**, each color in the y axis represents a single animal and for the bottom panels, the two boxes indicate the time windows before (pre-purple) and after (post-blue) the reward delivery (t=0s). Pre- and post-reward delivery time-windows are both composed of 3 seconds. Red line represents the reward delivery with all trials being aligned to this event (t=0s). ΔF/F colorbar ranges from a maximum value of +1.5 (yellow) to a minimum value of -1.5 (dark blue).

Dopamine 2 Medium Spiny Neurons

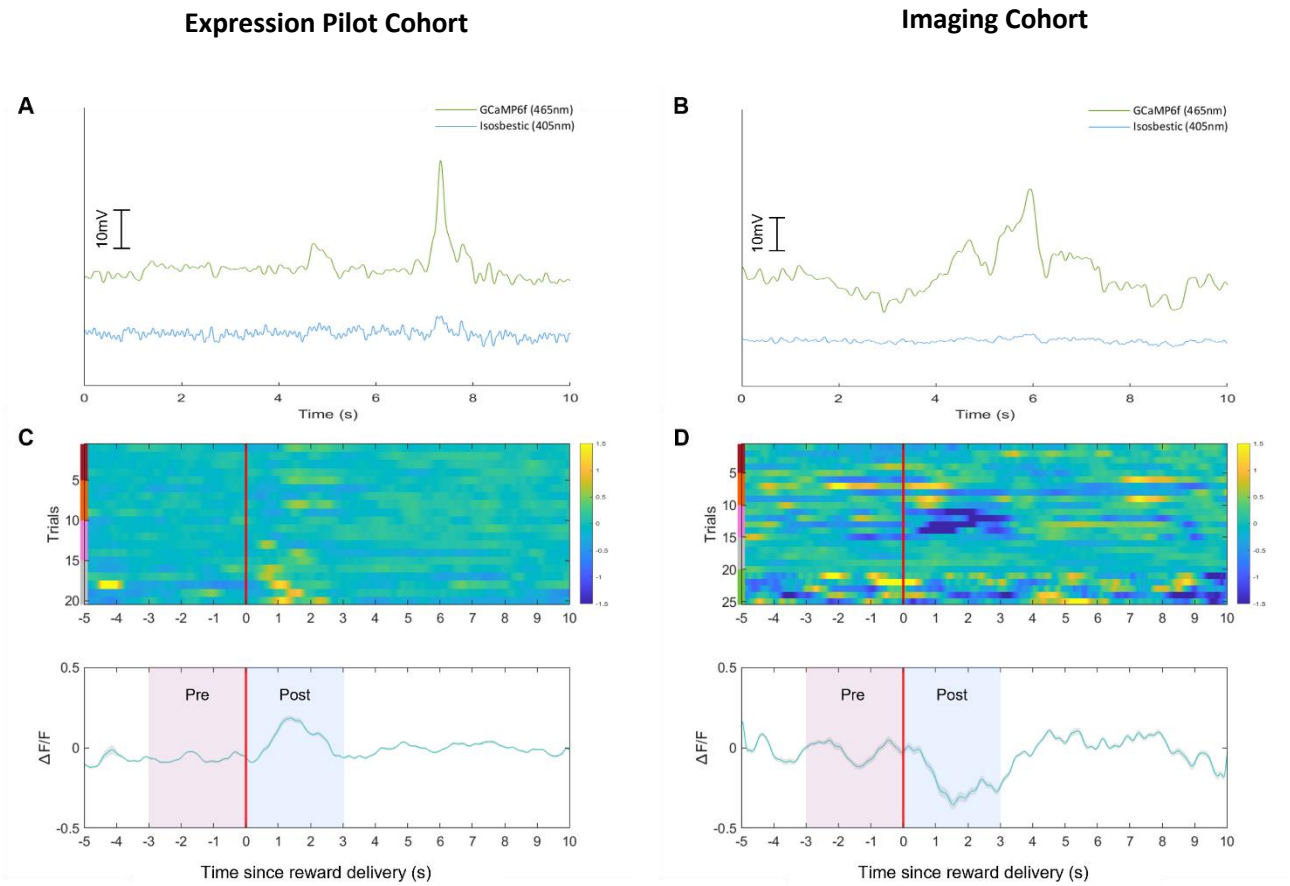

Figure 12: Neuronal activity of D2-MSN in VMS during the signal test recording session for expression pilot cohort and imaging cohort. (A) Example traces of isosbestic (blue - 405nm) and GCaMP6f (green - 465nm) channels from one D2-iCre animal of the expression pilot cohort. **(B)** Example traces of isosbestic (blue - 405nm) and GCaMP6f (green - 465nm) channels from one D2-iCre animal of the imaging cohort. **(C)** Top panel shows GCaMP6f fluorescence around reward delivery for four animals from the expression pilot cohort - five trials per animal. Bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of all the animals in the expression pilot cohort. **(D)** Top panel shows GCaMP6f fluorescence around reward delivery for five animals from the imaging cohort - five trials per animal. Bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of all the animals in the imaging cohort. For top panels of figures **(C)** and **(D)**, each color in the y axis represents a single animal and for the bottom panels, the two boxes indicate the time windows before (pre-purple) and after (post-blue) the reward delivery (t=0s). Pre- and post-reward delivery time-windows are both composed of 3 seconds. Red line represents the reward delivery with all trials being aligned to this event (t=0s). ΔF/F colorbar ranges from a maximum value of +1.5 (yellow) to a minimum value of -1.5 (dark blue).

4.4. Animals can successfully learn and perform the Go/No-go task for both Go and No-go reward contingencies.

To disclose how striatal subpopulations encode outcome-dependent action control, and how their activity is coordinated in different striatal subregions, we trained 18 animals in the Go/No-go task. We started by testing the ability of animals to successfully perform the Go/No-go task above chance. To this end, we assessed behavioral performance during the training (38 sessions), until the day before the recording session.

Animals started to be trained with No-go reward trials only, in which they performed above chance (50%) (**Figure 13A**). However, a slight decrease in their performance is notorious when Go reward trials were introduced in the training (**Figure 13A**). All Go reward trials were also performed above chance, except for the session in which uninstructed Go trials were introduced for the first time.

Moreover, when looking at the last three sessions for each Go trial type (Go right, Go left, and Go blocks - **Figure 13B**), the data provides that, independently of the trial type, the percentage of correct trials is significantly higher in Go reward conditions than in No-go reward conditions (**Figure 13B**, Wilcoxon matched-pairs signed-rank test, Go Right effect: $W=-121$, $p=0.0026$; Go Left effect: $W=-136$, $p<0.0001$; Go Blocks effect: $W=-126$, $p=0.0003$). For all the three conditions, average performance is above chance (more than 50% of correct trials), with the majority of the animals having a percentage of correct trials higher than 70%.

Next, to have more readouts of the influence of Pavlovian bias, we assessed the number of times animals made a magazine entry after Go blocks and No-go successful trials, as well as the latency for these magazine entries (**Figures 13C** and **13D**). Results regarding the number of magazine entries after successful trials show that animals perform significantly more magazine entries whenever a trial was a Go reward than when it was a No-go reward (**Figure 13C**, paired t-test, $t=2.353$, $df=15$, $p=0.0327$). Regarding the latency for magazine entry after performing a successful trial, data showed that for both trial types (Go blocks and No-go) the difficulty between making the action (lever press or hold nose-poke) and going to the food magazine to receive a reward is approximately the same, with animals showing similar performance in both conditions (**Figure 13D**, paired t-test, $t=0.663$, $df=15$, $p=0.5176$).

Together, these results show that animals can successfully learn and perform the Go/No-go task for both Go and No-go reward contingencies. Data also confirm that animals are equally motivated to perform both contingencies and show the presence of Pavlovian Bias towards Go reward trials. Overall, this means that behaviorally, we can use the task designed by our group to study outcome-dependent action control.

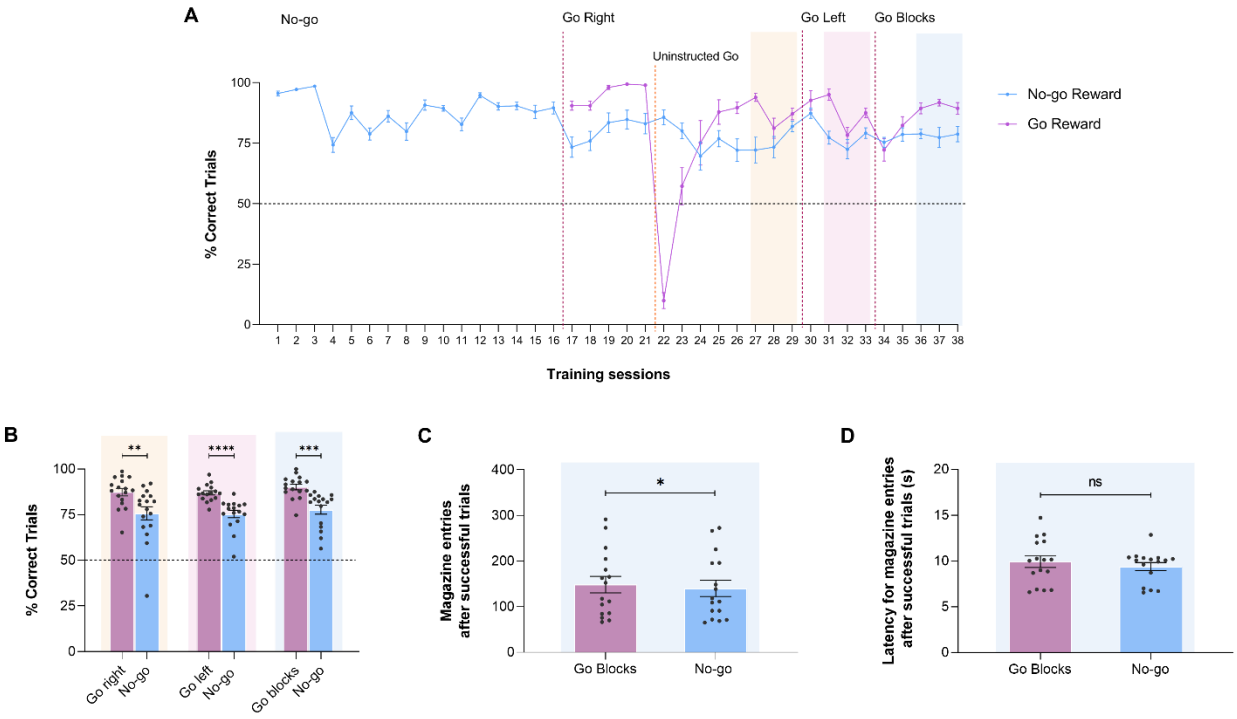

Figure 13: Animals can successfully learn and perform the Go/No-go task for both Go and No-go reward contingencies. (A) Animals' performance in the Go/No-go task during training sessions (n=38 sessions). Data is presented as percentage of correct trails, with each data set being the mean±SEM of all the animals in that stage of the task. Blue represents No-go reward trials and purple represents Go reward trials. Red dashed lines show the first session in which a new condition was introduced (No-go, Go Right, Go Left, and Go Blocks). Orange dashed line shows the first day of training in the uninstructed Go Right condition. Shaded background represents the three last sessions for specific conditions that were used in further analysis (yellow: Go Right, pink: Go Left, and blue: Go Blocks). **(B)** Percentage of correct trials in the last three sessions of training for Go Right (n=17), Go Left (n=16), Go Blocks (n=16) and No-go conditions. **(C)** Number of magazine entries after successful trials in the last three sessions of training in Go Blocks (n=16). **(D)** Latency for magazine entries after successful trials in the last three sessions of training in Go Blocks (n=16). For panels **(B)**, **(C)**, and **(D)**, data are presented as mean±SEM. Each black dot represents an individual animal. Purple bars represent Go conditions and blue bars represent No-go conditions. Statistical significance was determined by a Wilcoxon matched-pairs signed-rank test (panel B) and paired t-tests (panel C and D).*p<0.05, **p<0.005, ***p<0.0005, ****p<0.0001, ns - not significant.

4.5. *In vivo* Fiber photometry recordings during the performance of the Go/No-go task

To study how outcome-dependent action control is encoded by distinct striatal subpopulations (D1- and D2-MSN) and better understand if their signals differ between striatal subregions (VMS, aDLS, and pDMS), we recorded animals using *in vivo* fiber photometry while they performed a Go/No-go task. In this section, we will present behavioral and neuronal data from the recording sessions (Go reward and No-go reward conditions). Out of a total of 18 animals, 11 were included in this study due to their degree of proficiency in the task.

As explained in the data analysis section of this thesis (section 3.7. - Materials and Methods), we analyzed the data by taking into consideration five specific task events related to movement and motivation: nose-poke entry (when the auditory tone turns ON, indicating the trial type), nose-poke exit (when the auditory tone turns OFF and the trial was successful), right lever press, left lever press, and reward delivery. Data is

divided by subregions, and within subregions by transgenic animal line (D1-MSN and D2-MSN). For each event, Go and No-go trials are presented separately.

It is important to note that due to the low number of animals, we did not perform statistical analysis on neuronal data. All the results presented in this section (except behavioral data) will be inferred based on a visual analysis of the heatmaps and mean activity graphs that were obtained. It is important to remind that there is no histology data for the majority of the animals included in this section because they are still alive and being trained in the Go/No-go task. This means that we cannot guarantee that their optical fibers are in the correct location. In addition, we found that TDT can have unexpected TTL signals, making the alignment of data less reliable. This can be another factor that influenced the interpretation of the data.

A schematic representation of the task for Go reward and No-go reward conditions is presented in **Figure 14A**. Animals were recorded throughout the events represented in this figure and only the successful trials are included in this section.

We started by evaluating the behavioral performance of all animals during the recording sessions. The majority of the animals showed a performance above change (50%) for all trial types (Go right, Go left, and No-go)(**Figure 14B**). The percentage of correct trials was significantly higher for Go left when compared with No-go trials (Repeated measure one-way ANOVA with Silk's post-hoc test; Go left vs No-go: p=0.0051, t=4,247, dF=10). Furthermore, no significant difference was seen between both Go trial types (Repeated measure one-way ANOVA with Silk's post-hoc test; Go right vs Go left: p=0.1564, t=2.170, dF=10) or between Go right and No-go (Repeated measure one-way ANOVA with Silk's post-hoc test; Go right vs No-go: p=0.0865, t=2.533, dF=10).

Before starting the analysis of the neuronal data, we assessed the quality of our signal. Data obtained show that traces from an example animal have distinct activity patterns between GCaMP6f and isosbestic channels (**Figure 14C**), with activity peaks present in the GCaMP6f channel that are absent in the isosbestic one.

4.5.1. D1-MSN in VMS show movement- and reward-related activity for Go conditions

We evaluated how D1-MSN in VMS encode action control by recording from animals with the optical fiber implanted in this region. For nose-poke entry, animals show no activity in both Go and No-go conditions (no change in ΔF/F in the heatmaps or peaks of activity in the mean activity graphs, **Figure 14D**). Moreover for nose-poke exit, D1-MSN showed a peak of activity right after the event in Go conditions, which was not seen in No-go conditions (**Figure 14E**). Furthermore, regarding lever pressing, we presented data for the right lever press (**Figure 14F**) and left lever press (**Figure 14G**) in Go conditions. Here, increased activity for both right and left levers was observed, with a higher peak when animals performed a left lever press. Also, upon reward delivery, increased activity was, again, present in Go conditions but not in No-go conditions (**Figure 14H**).

Additionally, for task events that showed activity in VMS (nose-poke exit, lever pressing, and reward delivery), heatmaps showed an accentuated increase in activity around t=2s. This time point represents the moment in which the animals are in the food magazine eating the reward. It is important to note that this increase in activity is also stronger for Go conditions than No-go conditions.

Although VMS is a region more involved in outcome evaluation and motivation, results showed increased activity of D1-MSN, not upon reward delivery only, but also for task events that involve movement (nose-poke exit, right lever press, and left lever)

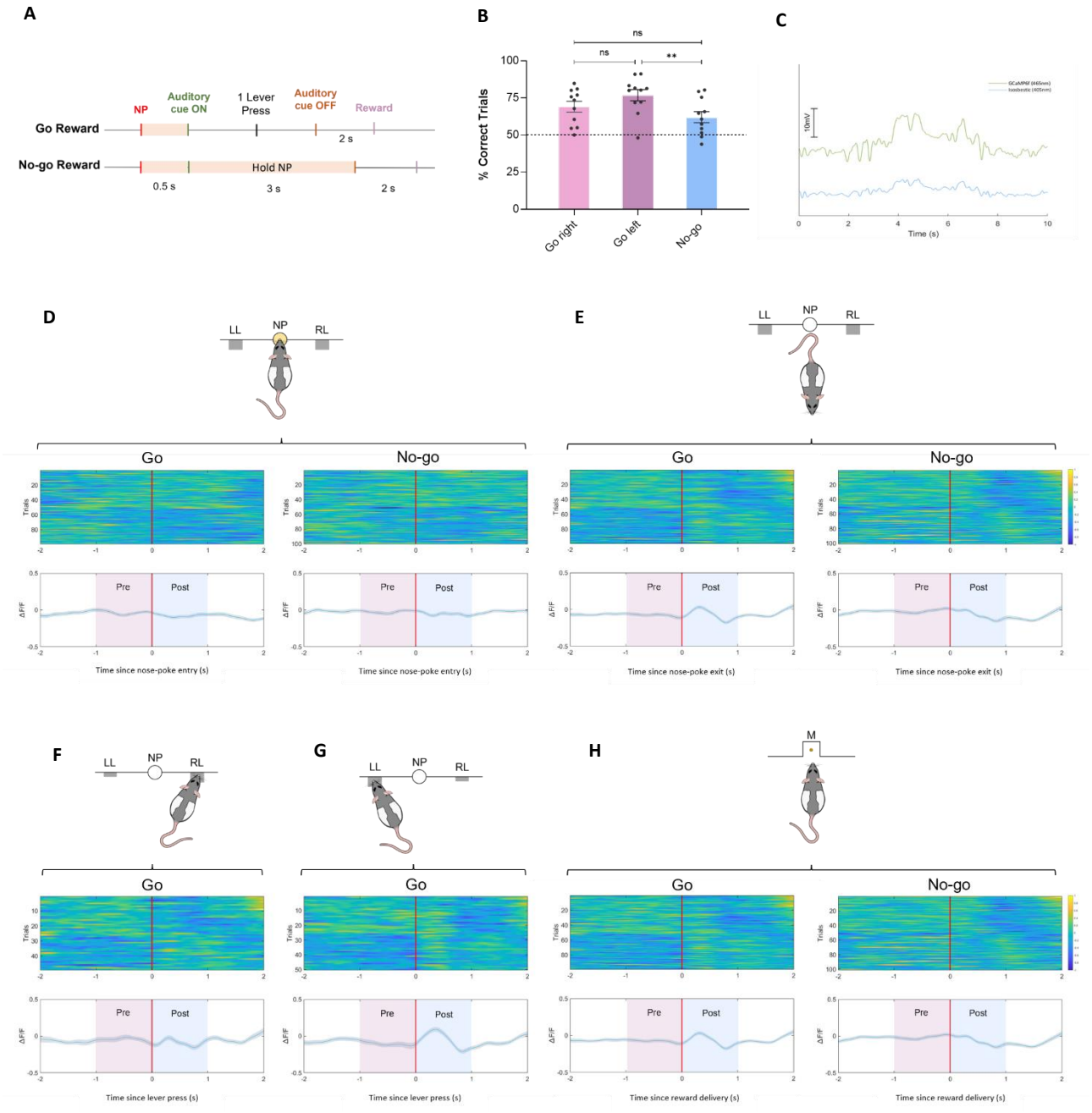

Figure 14: D1-MSN in VMS show movement- and reward-related activity in Go conditions. (A) Schematic representation of the Go/No-go task with the conditions recorded using *in vivo* fiber photometry (Go reward and No-go reward). **(B)** Percentage of correct trials during the recording sessions for Go right, Go left, and No-go conditions (Average of two sessions). Data are presented as mean±SEM and each black dot represents an individual animal. Statistical significance was determined by a repeated measure One-way ANOVA with Silk's post hoc test. **p<0.005, ns - not significant. **(C)** Example traces of isosbestic (blue - 405nm) and GCaMP6f (green - 465nm) channels from one D1-i*Cre* animal. Figures **(D)** to **(H)** show the neuronal activity of D1-MSN in Ventral Medial Striatum during five different events of the Go/No-go task: Nose-poke entry **(D)**, Nose-poke exit **(E)**, Right lever press **(F)**, Left lever press **(G)**, and Reward delivery **(H)**. For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for two animals (n=2), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the two D1-i*Cre* animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(D)**, **(E)**, and **(H)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

4.5.2. D2-MSN in VMS showed a slight increase in activity around movement-related task events and reward delivery in Go and No-go conditions

Upon nose-poke entry, D2-MSN showed no activity in both Go and No-go conditions, as depicted in the heatmaps and mean activity figure (**Figure 15A**). Furthermore, in the nose-poke exit event, we see a slight increase in activity in both Go and No-go trials (**Figure 15B**). A peak in activity is also notorious in t=2s when the animal is at the food magazine eating the reward (also seen in both trial types). Additionally, and as seen for the previous event, data show a slight increase in activity for the right lever press (**Figure 15C**) and left lever press (**Figure 15D**). This increased activity is not different between right presses and left presses. In addition, data show strong reward-related activity while animals are eating the reward around t=2s in these two events. Lastly, and coherent to these past results, reward delivery is encoded by a slight increase in activity of D2-MSN in both Go and No-go trials (**Figure 15E**). As expected, a bigger increase in activity can also be found around t=2s in both trial types.

Together, data showed D2-MSN in VMS responding to movement-related task events and reward delivery in Go and No-go trials.

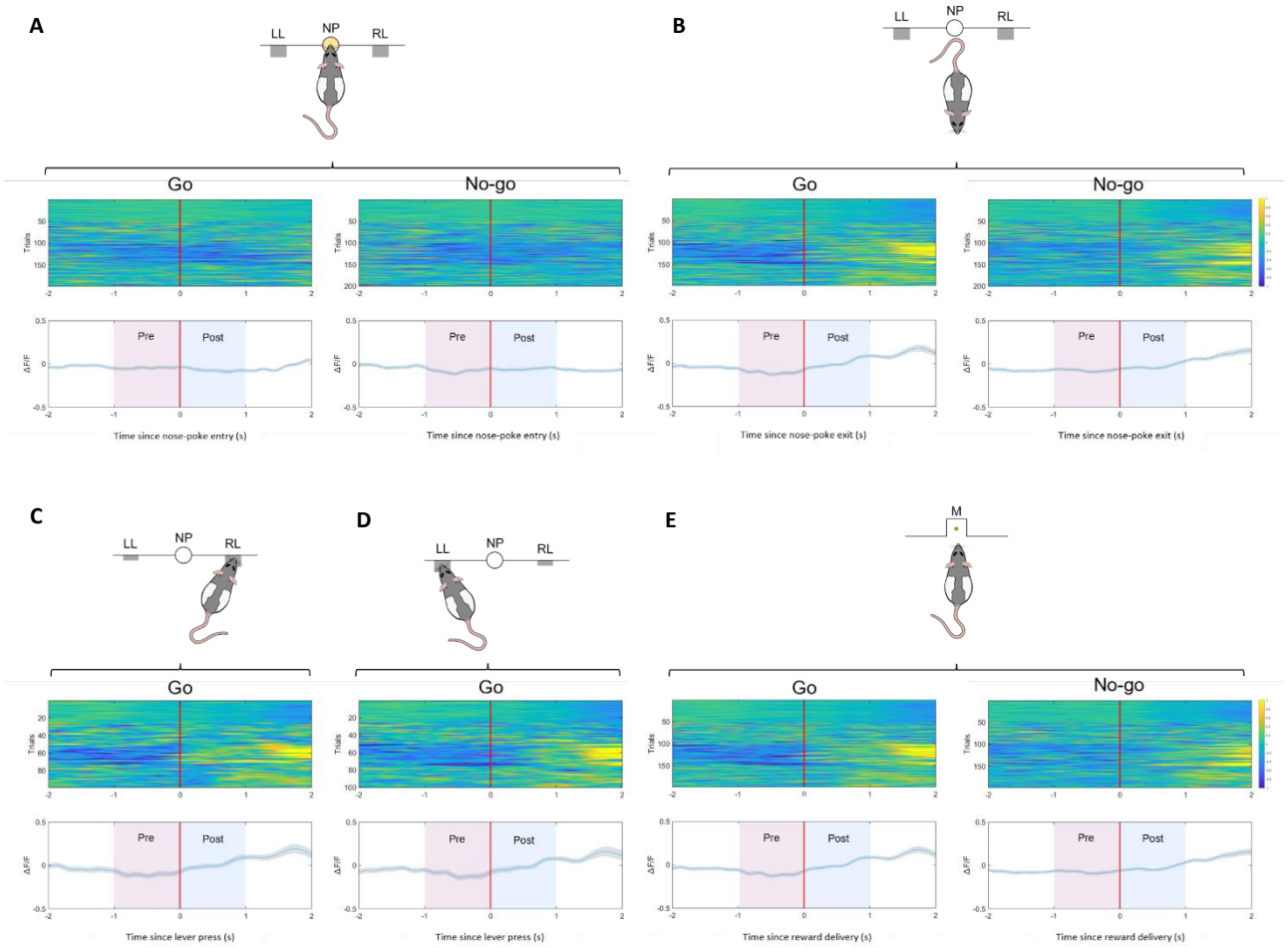

Figure 15: D2-MSN in VMS showed a slight increase in activity around movement-related task events (nose-poke exit and lever press) and reward delivery in Go and No-go conditions. Neuronal activity of D2-MSN in Ventral Medial Striatum during five different events of the Go/No-go task: Nose-poke entry **(A)**, Nose-poke exit **(B)**, Right lever press **(C)**, Left lever press **(D)**, and Reward delivery **(E).** For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for four animals (n=4), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the four D2-iCre animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(A)**, **(B)**, and **(E)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

4.5.3. D1-MSN in DLS responded to movement-related task events and reward delivery in Go and No-go conditions

D1-MSN in DLS do not show activity upon nose-poke entry for both Go and No-go trials (**Figure 16A**). Moreover, in both Go and No-go trials, there is a subtle increase in activity after nose-poke exit (**Figure 16B**). In addition, for both conditions in this event, there was an increase in ΔF/F at t=2s, which is around reward eating. Also, this effect is observed in both Go and No-go trials in the other three task events: right lever press (**Figure 16C**), left lever press (**Figure 16D**), and reward delivery (**Figure 16E**). Regarding lever pressing, no differences were found between the right lever and the left lever.

Together, these results indicate that D1-MSN in DLS responded to movement-related task events and reward delivery in both Go and No-go trials.

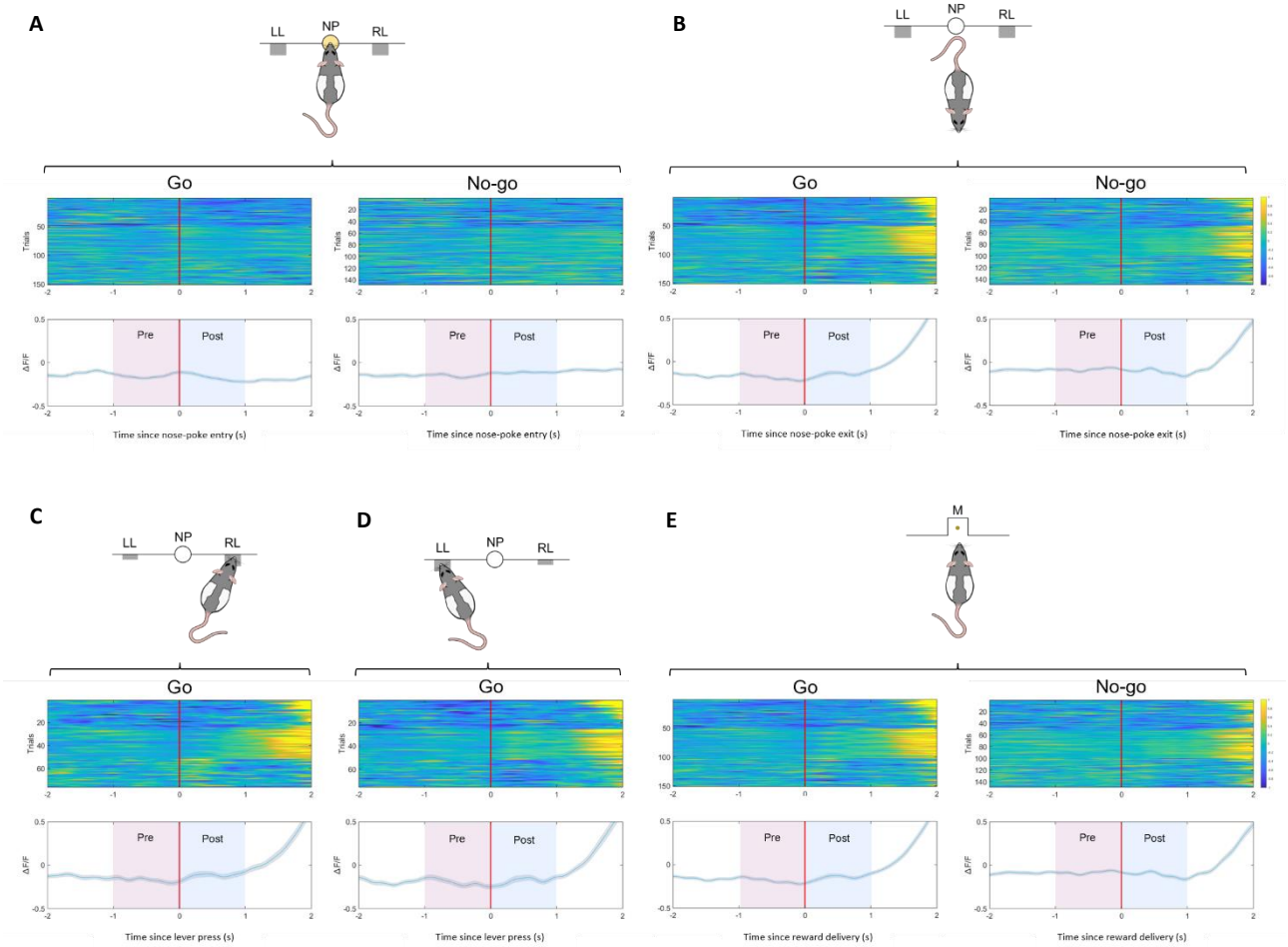

Figure 16: D1-MSN in DLS responded to movement related task events (nose-poke exit and lever press) and reward delivery in Go and No-go conditions. Neuronal activity of D1-MSN in Dorsal Lateral Striatum during five different events of the Go/No-go task: Nose-poke entry **(A)**, Nose-poke exit **(B)**, Right lever press **(C)**, Left lever press **(D)**, and Reward delivery **(E).** For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for three animals (n=3), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the three D1-i*Cre* animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(A)**, **(B)**, and **(E)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

4.5.4. D2-MSN in DLS respond to movement-related task events and reward delivery in Go and No-go conditions

D2-MSN in DLS did not show activity upon nose-poke entry in both Go and No-go trials (**Figure 17A**). Moreover, upon nose-poke exit, D2-MSN presented a small increase in ΔF/F for Go and No-go trials (**Figure 17B**). Furthermore, regarding lever pressing events, the increase in activity is seen for both of them in Go and No-go trials but it seems to be slightly higher in the right lever press (**Figure 17C**) in relation to left lever press

(**Figure 17D**). Lastly, in the period post-reward delivery, the same pattern of activity is seen: a slight increase in both Go and No-go trial types (**Figure 17E**).

Contrarily to what was seen before, no increase in activity was seen near the timepoint (t=2s) in which animals are eating the reward.

Overall, these results suggest that D2-MSN in DLS respond to movement-related task events and reward delivery in Go and No-go trials.

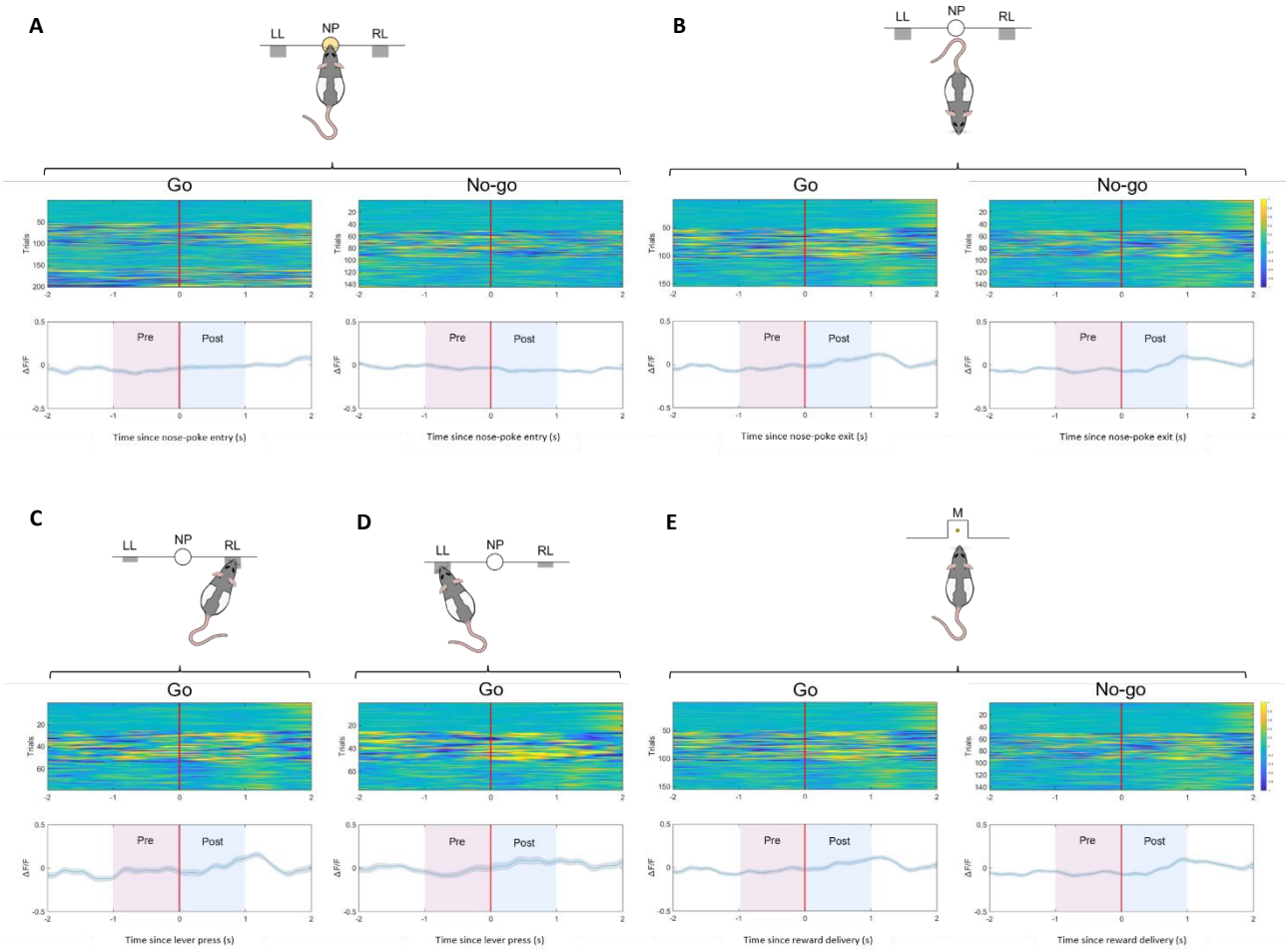

Figure 17: D2-MSN in DLS respond to movement-related task events (nose-poke exit and lever press) and reward delivery in Go and No-go conditions. Neuronal activity of D2-MSN in Dorsal Lateral Striatum during five different events of the Go/No-go task: Nose-poke entry **(A)**, Nose-poke exit (**B**), Right lever press **(C)**, Left lever press **(D)**, and Reward delivery **(E)**. For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for four animals (n=4), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the four D2-i*Cre* animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(A)**, **(B)**, and **(E)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

4.5.5. D1-MSN in DMS show a slight increase in activity around nose-poke entry and minor suppressions after nose-poke exit, lever pressing, and reward delivery in Go trials

D1-MSN in DMS showed increased activity in Go trials upon nose-poke entry (**Figure 18A**), which was absent for No-go trials. Moreover, during nose-poke exit, D1-MSN showed a small suppression of activity for Go trials. Contrary to Go, during No-go no difference in activity was found (**Figure 18B**). Regarding lever press events, the right lever press showed minor suppression in the period post-event (**Figure 18C**), while during the left lever press figure (**Figure 18D**) no alteration of activity can be found. Finally, when looking at data obtained around the reward delivery task event, D1-MSN activity decreased slightly after the reward being delivered in Go trials, while in No-go trials no alterations of activity were detected.

Contrary to what was seen for this neuronal subpopulation in other striatal subregions, there is no increase in activity related to the time point in which the animals are eating the reward for any of the trial types in any of the task events.

Overall, these data show that D1-MSN in DMS produced a slight increase in activity after nose-poke entry and minor suppressions after nose-poke exit, lever pressing, and reward delivery in Go trials.

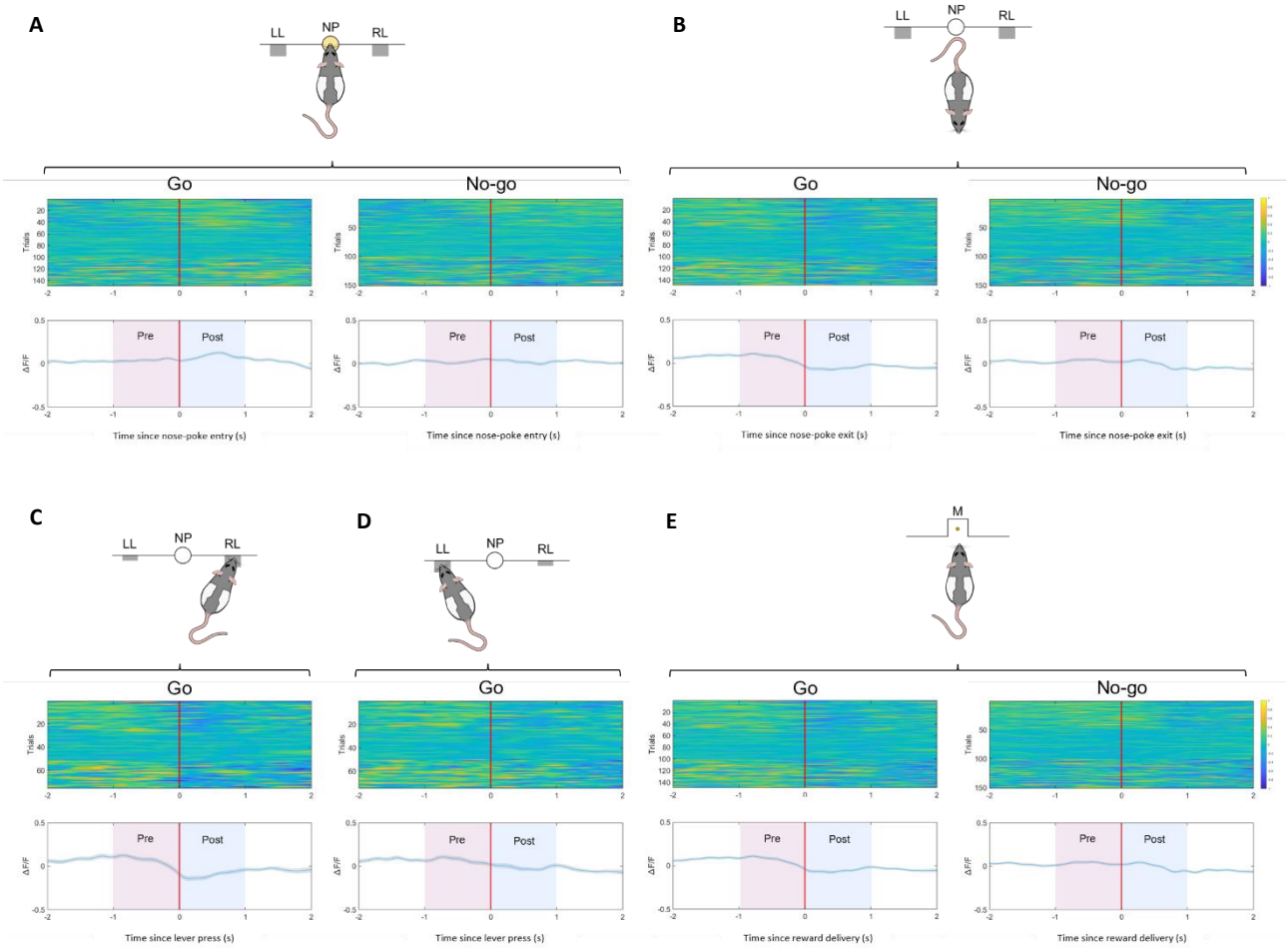

Figure 18: D1-MSN in DMS show a slight increase in activity around nose-poke entry and minor suppressions after nose-poke exit, lever pressing, and reward delivery in Go trials. Neuronal activity of D1-MSN in Dorsal Medial Striatum during five different events of the Go/No-go task: Nose-poke entry **(A)**, Nose-poke exit **(B)**, Right lever press **(C)**, Left lever press **(D)**, and Reward delivery **(E).** For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for three animals (n=3), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the three D1-i*Cre* animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(A)**, **(B)**, and **(E)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

4.5.6. D2-MSN in DMS responded to movement-related task events and reward delivery in Go and No-go conditions

D2-MSN in DMS showed a small suppression of activity for nose-poke entry in Go trials with no difference in activity observed in No-go trials (**Figure 19A**). Regarding nose-poke exit, there is an increase in activity for both Go and No-go trials (**Figure 19B**). Increased activity was found around t=2s (animals eating the reward) with a higher rise in No-go conditions. Moreover, in lever pressing task events, there is also a slight increase in activity for both the right lever press (**Figure 19C**) and left lever press (**Figure 19D**). Additionally, no differences were found between the left and right sides. Lastly, upon reward delivery, this neuronal subpopulation shows increased activity in Go and No-go trials with the rise being slightly bigger in No-go conditions (the bright yellow color in the heatmap for t=1s illustrates this) (**Figure 19E**).

Together, these results indicate that D2-MSN in DLS responded to movement-related task events and reward delivery in Go and No-go trials.

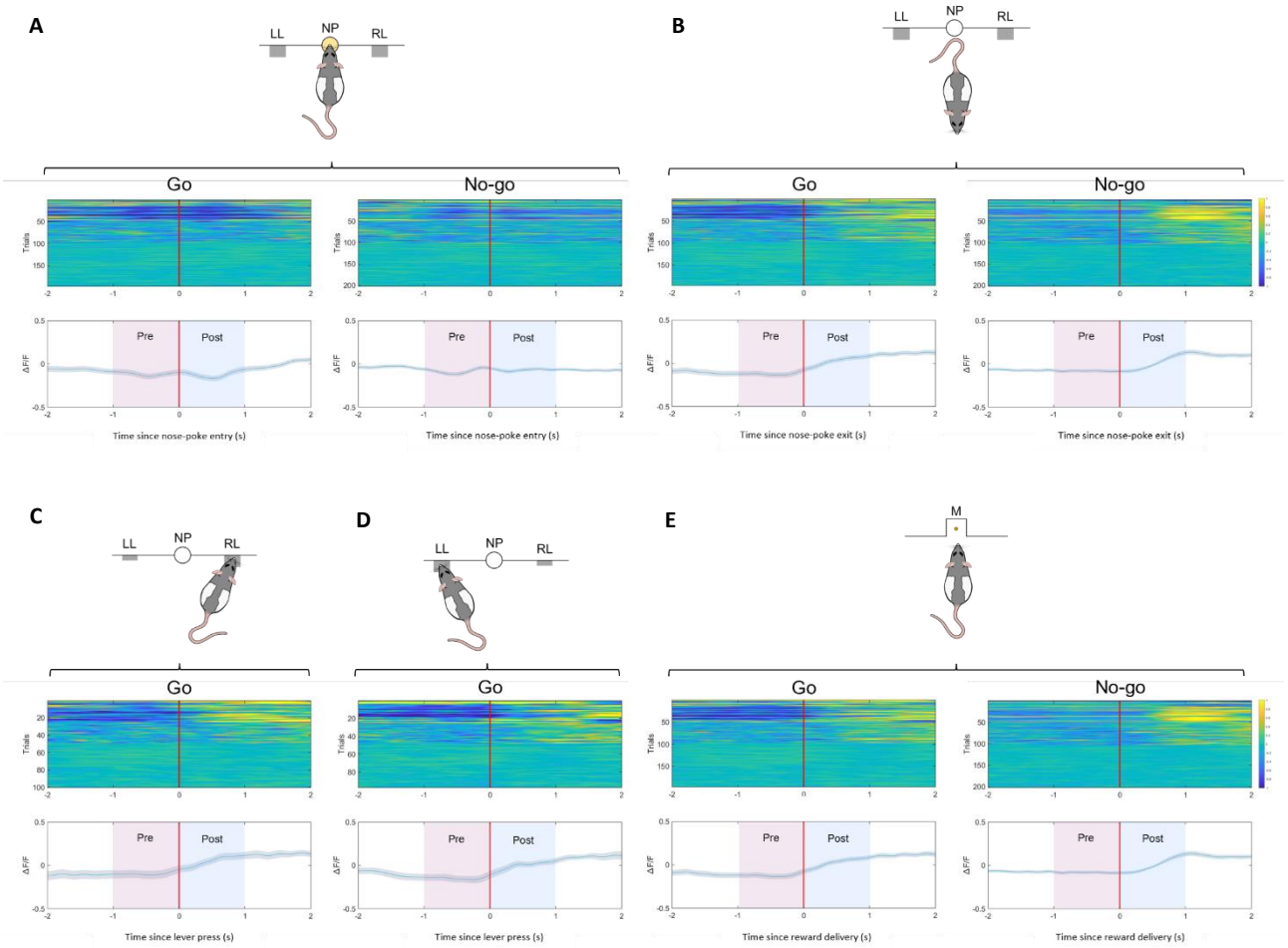

Figure 19: D2-MSN DMS responded to movement-related task events (nose-poke exit and lever press) and reward delivery in Go and No-go conditions. Neuronal activity of D2-MSN in Dorsal Medial Striatum during five different events of the Go/No-go task: Nose-poke entry **(A)**, Nose-poke exit **(B)**, Right lever press **(C)**, Left lever press **(D)**, and Reward delivery **(E)**. For all figures, the top panel shows GCaMP6f fluorescence around the designated task event for four animals (n=4), and the bottom panel shows the mean±SEM of GCaMP6f fluorescence for all the trials of the four D2-i*Cre* animals; The two boxes indicate the time windows before (pre-purple) and after (post-blue) the designated task event. Pre- and Post- task event time windows are both composed of 1 second. For figures **(A)**, **(B)**, and **(E)**, the right panels correspond to No-go conditions while left panels correspond to Go conditions. For all figures, the red line represents when the task event happened with all trials being aligned to it (t=0s). ΔF/F colorbar ranges from a maximum value of +1 (yellow) to a minimum value of -1 (dark blue).

5. Discussion

Voluntary action control is a decision-making process important in our daily life. It is through this process that we take into consideration internal and external information to guide our behavior and attain desired goals. Outcome-dependent action control is thought to recruit striatal subregions, but precise underlying mechanisms remain unclear. Therefore, the dissection of the neuronal circuits that guide outcome-dependent action control is of high importance to better understand the nature of psychiatric conditions, in which it seems to be impaired, and how we can develop more target-specific treatments to help these patients.

With this thesis project, we intended to gain new knowledge about the way outcome-dependent action control is encoded by neuronal subpopulations (D1-MSN and D2-MSN) found in the striatum, a brain region thought to be involved in this process. Moreover, this project also proposed to disclose if the activity signals of these neuronal subpopulations diverge between different striatal subregions (VMS, aDLS, pDMS).

- **A dual-vector viral approach and a higher volume of virus improve GCaMP6f expression in VMS, aDLS, and pDMS**

In the first aim of this project, we proposed to optimize *cre*-dependent GCaMP-expressing adeno-associated-virus to record striatal medium spiny neurons *in vivo*. For that, we tested a new dual-vector viral approach and different volumes of injection. We found that high volume of virus (1350nL) produces robust expression that is sufficient for imaging purposes, in VMS, aDLS, and pDMS.

Although previous studies showed robust GCaMP expression in mouse striatum, this has not been accomplished in D1- and A2a-*cre* rat striatum.[36] One potential reason for the lack of expression in rats might be that viruses are optimized for mice and have lesser binding/expression affinity in rats. In addition, we used transgenic rat lines that were recently developed and in which GCaMP expression was never tested before.[28] The insufficient expression found using standard viral constructs (example shown in **Figure 9A** - bottom row, third panel - AAV9.hSyn.DIO.GCaMP6f construct) led to the need of employing a new approach. Our experiments, for the first time, show that virus expression is possible by employing a combined *cre* x Flippase system with high volume of viral injection. This results in sufficient expression, which can be used to measure cell-type specific signals in diverse striatal subregions during a Go/No-go task, but also in other behavioral tasks.

In our work, we tested different volumes of viral injection, but we were not able to measure over-expression or apoptosis. The optimization of the volume of injection is crucial for calcium imaging experiments. On one hand, using a volume of virus that is too small may lead to inadequate expression of GCaMP (low number of infected cells), while on the other hand, a volume that is too high might promote GCaMP over-expression. This can be an issue since GCaMP is known to disrupt calcium signaling when available in high quantities, which can lead to cell death by apoptosis.[50] However, in our histological data, we did not find outliers of bright neurons, which would indicate over-expression.[50] Future analysis should test more injection volumes and quantify cell death by performing immunostainings with specific markers (e.g., Annexin A5 + PI).[72]

Another factor that should be taken into consideration to achieve good recordings is the area of expression and the location of that expression. With a larger area being infected by the virus, the probability of having more neurons expressing GCaMP under the tip of the optical fiber, and whose signal is being recorded from,

is also higher. This is of extreme importance when recording neuronal activity since we always want to record as many neurons as we can in order to have a big and statistically reliable n. Although we saw a big spread using a high volume of injection in aDLS and pDMS, this spread was smaller in VMS. This smaller spread with high injection volume might be in part explained by the way our data is plotted. We only considered the viral expression found right below the tip of the optical fiber, discarding other regions infected with the virus. In VMS it is common to find viral expression around the ventricles due to their proximity to the coordinates injected to target this subregion. This means that we could also have a bigger spread with high volumes of virus in VMS, but it was probably more off-target than what is seen in the other two subregions. Viral spread near the ventricles would not result in a better signal since our fibers are targeted further away from the ventricles. A way to optimize this is by testing different injection coordinates, which we also started doing when performing viral infection in the imaging cohort. Here, we injected the virus in the same coordinates as fiber implantation, instead of 300µm below (see section 3.3. – Materials and Methods). Since most of the animals are still alive and being trained in our behavioral task, there are only histology results for two out of eighteen animals. Hence, we do not know yet if this optimization is working.

It is important to consider that although our results for the expression pilot cohort have shown a significant increase in fluorescence mean intensity and area of expression using a high volume of virus, they are based on a n of four animals. More studies should be performed in order to increase this n and have a more robust analysis.

Overall, the fact that we were able to promote viral expression in rat's striatum using the dual-vector viral approach, giving highlight on which volumes of injection can contribute to successful experiments, will help other scientists in the field that need to use these novel rat lines in their experiments and want to perform calcium imaging techniques in striatal subregions.

- **Three seconds of continuous compound stimulus with 90dB of WN is aversive for the animals and has the similar subjective value as one sugar pellet**

In the second aim of this project, we intended to better understand the subjective value of an aversive stimulus in relation to a sugar pellet. For that, we employed a lever press-based aversive/appetitive task. In the presented work, we report that three seconds of continuous compound stimulus with 90dB of white noise and bright light is aversive for the animals and has a similar subjective value as one sugar pellet. The aversiveness of the compound stimulus seen with these results confirms that it can be used as a punishment in the Go and No-go punishment stages of our Go/No-go task. In addition, the similar subjective value of the negative stimulus in relation to the positive stimulus is also important for the establishment of more truthful comparisons between the different contingencies in our task.[73]

Over the years, literature has described continuous white noise as an aversive stimulus to the animals when used within ranges of moderate-intensity (80-90dB).[74] Additionally, the use of WN has also been referred to as advantageous compared to the use of shocks, the most common form of punishment employed in behavioral studies. Besides being better tolerated by the animals, WN is considered aversive without inducing freezing behaviors (when used in the correct intensities) and does not introduce artifacts during the neuronal activity recordings (this might happen when using shocks), an important factor to be considered in these type of experiments.[75,76]

Previous studies from our group (Goedhoop *et al*. 2021, *bioRxiv*) showed that white noise at 90dB promotes aversion in the animals and therefore can be used as a punishment in our behavioral tasks.[68] In our studies, we used three seconds of continuous white noise at 90dB but we paired it with LED to strengthen its effect. This was done because rodents prefer dark environments over light, tending to hide to avoid it. The LED was placed in the magazine tray to directly shine light on the animal when it goes to the food magazine to collect the reward.

In our aversive/appetitive task, animals had to choose between two levers based on the outcome they delivered. When presented with equal value in both levers, animals were able to understand that by pressing any of the levers they would receive the same amount of reward. When we first introduced the compound stimulus to the animals by pairing it with 1 pellet, animals shifted their preference towards the lever without the compound stimulus. Although they press more the green lever in relation to the red lever, the difference to conditions without the compound stimulus was not significant, suggesting that the animals do not mind sometimes receiving the compound stimulus even though they could get the same amount of reward by pressing only the green lever. However, when we introduce reversals in our study, they seem to react differently to the levers, being more flexible and paying more attention to the outcome paired with their choice. From stage 3 onwards, we noticed that in conditions without the aversive compound stimulus (third stage of the task) animals press more for levers with higher outcomes (more sugar pellets), but when the compound stimulus is introduced (fourth stage of the task), they shift their preference to the lever without the compound even though that means to receive less reward. Overall, it seems that two pellets are not worth it when compared to the compound stimulus, meaning that 1) the compound stimulus is indeed considered aversive by the animals, and 2) its subjective value is approximately the same as 1 sugar pellet.

The conclusions taken from the fourth stage of the task were done using a compound stimulus with WN at 90dB, which was previously shown to work as an aversion stimulus in rats. In our experiments, we also wanted to test the effectiveness of WN at 90dB comparatively to lower intensities (70dB) by validating whether lower intensities have less subjective value for the animal. As stated by Mollenauer *et al.* (1992), using 70-75dB of WN does not induce significant differences between lever pressing compared to a control group (condition without the compound stimulus), which shows that using a compound stimulus with WN at 70dB is not as aversive as with WN at 90dB.[74] Our results corroborate these previous affirmations by showing very similar red presses ratios between conditions without compound stimulus (stage 3) and conditions with a compound stimulus with WN at 70dB. Since the effect of a compound stimulus with 70dB of WN is similar to conditions with the absence of the compound stimulus, animals still prefer to do some red lever presses (more than half of the time) and get double the pellets.

Although literature describes WN as not inducing freezing behaviors when used within a certain range of intensities, this was something we did not test in our experiments. This should and could be easily done by tracking animals' performance in the different sessions through the recording videos. This should be accounted for in the future as a way to support the use of the compound stimulus over shocks in this type of behavioral studies.

In general, the experiments performed in this section clarified what the subjective value of a compound stimulus is in relation to a sugar pellet. This will be very important to other studies in the field that want to establish a straight comparison between rewarding and punishing conditions or that want to test how outcome value is encoded in the brain. In addition, these experiments also showed that it is possible to use WN paired with LED to promote aversion in rats. This could be used as an alternative to foot shocks in diverse

behavioral studies that need to induce an aversive context in their task. The fact that we can use stimuli that are less harmful to the animals is also an important factor to be considered in future studies.

- ***In vivo* fiber photometry signal test session**

In this thesis project, we aimed to record the bulk activity of distinct neuronal populations in the striatum.[48] For that, we set up and consecutively tested *in vivo* fiber photometry for the first time in the lab, accounting for the quality of the signal obtained with the setup and the stability of the connections between the ferrules and the fiber optical cable. In our experiments, we found that the recorded signals were reflecting neuronal activity, excluding the presence of movement artifacts by comparison with traces obtained in the isosbestic channel. We also found that using larger ceramic ferrules and optical fibers we can achieve higher stability of the connections and increased area of recording, which constitutes an advantage for this project.

In our signal test session, we were able to confirm that the recorded signal was activity-dependent (dependent on GCaMP), reflecting the response of neurons to certain events. The importance of assuring that the recorded signal is activity-dependent lies in the fact that artifacts related to the movement of the animals on the operant chamber or the bending of the fiber optical cable can be easily present in signal traces, being misunderstood as neuronal activity. The comparison between the signal obtained with the GCaMP6f channel and the isosbestic channel (control for the presence of artifacts) allows us to infer if we are indeed recording the activity of neurons. The way this inference is done is by verifying if activity peaks present in the GCaMP6f channel are absent in the isosbestic channel. If so, it means that those peaks were only caused by the response of neurons to a certain event, but if the peaks are present in both channels, they are due to an artifact. Besides confirming the quality of the signal, our results also confirmed the quality of the viral expression, since without GCaMP expression we would not be able to detect any neuronal signal.

In the experiments of this section, we also tested the quality of the signal by recording from VMS while animals received rewards in an operant chamber. Due to the importance of this striatal subregion to outcome evaluation and motivated behaviors, we were expecting to find a reward-related signal.[77] The results we obtained were opposite between cohorts. In the expression pilot cohort, increased activity after reward delivery was seen in both D1- and D2-MSN, while in the imaging cohort both subpopulations showed suppressed activity for the same event. The suppression of activity, which was not expected, might be explained by the time-point at which animals were submitted to the signal test session. Unlike the animals from the expression pilot cohort that never engaged in behavioral tasks before the signal test session, animals from the imaging cohort were already being trained in the first stages of the Go/No-go task when they were recorded. The fact that they were used to receiving rewards when placed in the operant chamber (due to what they already knew from the Go/No-go task) could explain the lack of increased activity when the reward was delivered. To confirm this, a cohort with the same characteristics as the imaging cohort should be submitted to the signal test session before engaging in any behavioral task.

In addition, we were also able to infer that by using larger ceramic ferrules and optical fibers, besides having an increased area of recording that allows us to capture signals from more neurons, we have improved stability of the connections between the ferrules on the animals' head and the fiber optical cable. This connection is vastly important and should be as tight as possible, otherwise, the movement of the animal during the task can promote a disconnection and the recording will be interrupted, leading to the loss of important data. By using larger ferrules, we noticed that upon connection and disconnection of the cable

fewer fibers were pulled out of the animal's head. Likewise the stability of the connections, It seems that we can also achieve more stability of the fiber implantation by using larger optical fibers.

Although we were able to test the quality of the signal through the traces obtained for the GCaMP6f and isosbestic channels and the stability of the connections, we were not able to take proper conclusions regarding the response of neurons to reward delivery in VMS. The fact that most of the animals in the imaging cohort are still alive and performing the Go/No-go task means that we do not have their histological results yet. Therefore, we do not know if the optical fiber is placed in VMS and if we have viral expression targeting the region implanted, which makes it difficult to conclude about the signal observed. After having the histological data, a better evaluation of the signals should be done.

Besides the lack of histological data, we also had some problems with the pre-processing of the data that might slightly influence the results. Given that we analyzed this type of data for the first time in this thesis project, we are still trying to figure out the best way to pre-process it, which is a crucial step to the achievement of results that indeed reflect the neuronal activity of the animals. We know that there is still space for improvement in our pre-processing scripts and that it is something that needs to be worked on in the future.

- **Animals can successfully learn and perform the Go/No-go task for both Go and No-go reward contingencies.**

In the third aim of this thesis, we intended to disclose how D1- and D2-MSN encode outcome-dependent action control in VMS, aDLS, and pDMS and if the activity of these striatal subpopulations varies depending on the subregion where they are located. For that, we recorded animals using *in vivo* fiber photometry while they perform in a Go/No-go task designed by our group. Regarding behavioral performance, we found that animals can successfully learn and perform the Go/No-go task for both Go and No-go reward contingencies with Pavlovian bias influencing the behavior.

In our experiments we showed that animals easily achieved a good degree of proficiency in No-go trials, meaning that animals can easily learn how to hold the nose-poke for 3 seconds to receive a reward. However, this performance slightly drops when Go trials are introduced. The fact that the percentage of correct trials is significantly higher for Go trials in respect to No-go trials for all conditions (Go right, Go left and Go blocks) highlights the difference seen between executing an action (Go trials) and refraining from action (No-go trials) in the face of reward. This difference seen in our results can be influenced by two different things: 1) the presence of Pavlovian bias and 2) the difference in the duration between Go and No-go trials. The literature describes action selection as being influenced by a process entitled as Pavlovian Bias. According to this process, upon rewarding conditions, we tend to activate actions (Go) to achieve the reward, while in punishing conditions we tend to withhold an action (No-go) to avoid being punished.[6,7,78] Additionally, studies performed in humans using decision-making tasks, showed that we can perform using the opposite action-outcomes (Go to avoid punishment and No-go to receive a reward) but this is usually accompanied by a higher level of difficulty.[7] Since in this thesis project we are evaluating Go and No-go reward contingencies, the data obtained is in coherence with what is stated by the Pavlovian bias, showing a better performance (higher percentage of correct trials) when it is a Go trial.

Although our data fits with the concept of Pavlovian Bias that has been described by literature over the years, it is also important to take into consideration aspects of the task that might have a pivotal role in the

discrepancy between Go and No-go performance. The way our task is designed makes that in Go trials, after hearing the auditory tone that indicates the trial type, the animals have to press the lever and 2 seconds later they receive a reward. However, in No-go trials, after hearing the auditory tone, animals most hold the nose-poke for 3 seconds, and 2 seconds after the end of the nose-poke hold they receive the reward. As we can see, the time between hearing the auditory tone until receiving the reward is shorter in Go trials (in case animals are proficient in the task) compared to No-go trials. Overall, this means that when the animals receive the tone that indicates the trial type they already know if they will receive the reward closer or farther in time and might perform better in Go trials due to that. A way to overcome this issue and improve the duration of Go trials is by increasing the number of presses animals have to do to receive a reward. The idea is to make animals do 2 to 3 presses that will last for 3 seconds, which is the same amount of time that they need to hold the nose-poke in No-go trials. Like this, the reward would be delivered at the same time point of the task (in relation to the moment the auditory tone is heard) for both trial types, allowing better comparisons between them for both behavioral and neuronal data. Moreover, we started training our animals using this approach (data not available yet) and we observed that they can do up to three presses within the time window of 3 seconds, meaning that in the future this should be the way to go since they seem to successfully perform in this new task version.

Furthermore, an important feature of this task is that we trained and recorded animals performing Go reward trials in both the right and the left levers (Go right, Go left and Go blocks of left and right). This was done to attend the striatal lateralization of movement signals that have been described in the literature.[25,36] Since studies performed in rodents have shown stronger neuronal activity signal in striatal regions upon contralateral movements to the site of recording, we decided to also account for this lateralization by performing recordings when the animal is pressing the right and the left lever (Go blocks – 5 trials on the right lever, than 5 trials on the left lever and so on until the total number of Go trials is achieved (50% of total trials)). To account for this, we implanted animals with optical fibers in both hemispheres for the same striatal subregion.[36]

In our experiments, we also confirmed that animals were equally motivated to perform both Go and No-go reward trial types by looking at the number of magazine entries made after a successful trial and also the latency for those magazine entries. The data we obtained showed that animals tend to do significantly more magazine entries after successful Go trials, which might be explained by either the Pavlovian bias or the shorter duration of this trial type. However, when looking at the latency for magazine entries after successful trials, no difference is found between these two trial types, which indicated that they can easily execute an action or withhold it and they were equally motivated to perform any of the trials to receive the reward.

Overall, the results obtained with these experiments showed that we can use this task (with some alterations) to study outcome-dependent action control in rats and perform *in vivo* fiber photometry recordings. This is also good for the field since it presents a new task that allows the study of actions and outcomes and that it is feasible for the animals to learn and perform. Besides this, the results also highlighted the influence of Pavlovian bias in behavior, showing that this is a process that is present and needs further investigation. Since we know that Pavlovian bias interferes with our action control, we now need to focus on the impact that its dysfunction could have on behavior and how to reverse those possible effects.

- ***In vivo* Fiber photometry recordings during the performance of the Go/No-go task**

Over the years activity in the striatum has been highlighted as playing an important role in action control, especially the control of actions that are dependent on the outcome.[15] A lot of research has been done to better understand how D1-MSN and D2-MSN encode actions and their outcomes with the classical and synergetic model oppositely trying to explain that way this is done. Although a lot of effort has been done, outcome-dependent action control is still poorly understood and needs further investigation. In the third aim of this project, we intended to contribute to this gap in knowledge by performing *in vivo* fiber photometry in different striatal subregions using D1-i*Cre* and A2a-i*Cre* rat lines and by also accounting for their behavior throughout the task.

By looking at the behavioral performance of the animals during recording sessions, we found that almost all animals performed above chance, but with a degree of performance (percentage of correct trials) somewhat lower than the one obtained during the training sessions. This decrease in performance might be explained by the fact that the recording day was the first time that animals were connected to the fiber optical cable for a longer period. This connection ends up constraining their movements a bit and might influence the overall performance of the task. This data also showed that animals perform Go right and Go left trials without significant differences and that they also perform Go left trials with more success than No-go trials. This could be in part explained by the presence of Pavlovian bias towards the Go left reward conditions. A way to improve behavioral performance during the training session is by starting to habituate the animals to the fiber optic cable during the training sessions by connecting them to the cable without shining light.

In our experiments, we also evaluated neuronal activity. We confirmed the quality of the signal by comparing traces between the GCaMP6f and isosbestic channels and concluded that our signal was activity-dependent, excluding the presence of any artifacts.

In addition, when looking at the neuronal data presented in this dissertation, it is important to remember that no statistical analysis was performed due to the low number of animals included in each group. Therefore, this discussion is based on the results inferred from a visual analysis of the heatmaps and mean activity graphs obtained for each task event. This analysis was done for D1-MSN and D2-MSN separately in each of the three striatal subregions focused on in this work.

Table 5 presents a summary of the neuronal activity observed in our data for D1- and D2-MSN in VMS, aDLS, and pDMS during the five task events (nose-poke entry, nose-poke exit, right lever press, left lever press, and reward delivery.

Table 5: Summary of the neuronal activity observed for D1-MSN and D2-MSN in VMS, aDLS, and pDMS for the five task events investigated. "**X**" represents no alteration in activity between the periods pre- and post-event; "↑" represents increased activity in the post-event period; and "↓" represents slightly suppressed activity in the post-event period.

Region	Neuronal type	NP entry		NP exit		Right Press	Left Press	Reward	
		Go	No-go	Go	No-go	Go	Go	Go	No-go
VMS	D1-MSN	X	X	↑	X	↑	↑	↑	X
	D2-MSN	X	X	↑	↑	↑	↑	↑	↑
aDLS	D1-MSN	X	X	↑	↑	↑	↑	↑	↑
	D2-MSN	X	X	↑	↑	↑	↑	↑	↑
pDMS	D1-MSN	↑	X	↓	X	↓	X	↓	X
	D2-MSN	↓	X	↑	↑	↑	↑	↑	↑

Activity of D1- and D2-MSN in Ventral Medial Striatum

The ventral medial striatum has been shown to be involved in motivation and reward processing.[25] Studies performed in this subregion have shown that both MSN subtypes are capable of encoding rewarding conditions, which is in line with the results we obtained showing that both D1-MSN and D2-MSN have increased activity after reward delivery, being necessary to drive normal motivated behaviors.[22,77] Moreover, studies performed by Burton *et al.* (2014) have also highlighted that neurons in VMS fired more strongly for larger rewards and shorter delays.[25] Because our Go trials are shorter than No-go trials, meaning that animals receive the reward closer in time when performing a Go trial, this might explain our data showing that D1-MSN increase activity in Go but not in No-go trials upon reward delivery. For D2-MSN we saw a rise in activity for both trial types.

Besides increased activity around reward delivery, which was already expected to be found in this striatal subregion, we also observe a rise in activity around movement-related events (nose-poke exit and lever press) for both D1- and D2-MSN. This is probably related to the motivational aspect of VMS, in which the motivation for the reward is driving the action (in the task, animals receive the reward after lever pressing and nose-poke exiting). This motivational and reward processing role can also be noted by the increased activity when animals are at the food magazine eating the reward (around the time-point t=2s).[25] In D1-MSN this increase is more notorious in Go trials, while in D2-MSN it is similar for Go and No-go trials.

It is also important to note that peaks for reward delivery are higher in VMS compared to aDLS and pDMS, which is coherent with previous studies showing that there are twice as many neurons increasing fire to reward in both Go and No-go conditions in the ventral striatum than in the dorsal striatum regions.[79] This also means that activity upon reward delivery should not be dependent on action execution or withholding.[79] Our data goes in agreement with this previous statement only for D2-MSN, in which the increased activity is seen for both trial types.

Although we found increased activity around the same task events for both neuronal subpopulations, this increase is noticeably higher (bigger peaks) in D1-MSN compared to D2-MSN, which might indicate that the cooperative and synergistic activity of these two subpopulations in this striatal subregion might not be as strong as in other regions.

Lastly, further neuronal activity data from the conditions of our task presenting the negative stimuli (Go and No-go punishment) are necessary to better understand and make conclusions regarding which model better represents the way VMS neurons encode reward versus punishment. Although we do not have these results yet, recent evidence leans strongly towards the possibility of both subpopulations being able to encode reward and punishment, depending on the striatal sub-region they are located on.[23]

Activity of D1- and D2-MSN in Dorsal regions

Dorsal regions of the striatum have been gaining more attention in the field due to their pivotal role in reward-related actions, namely goal-directed (DMS) and habitual behaviors (DLS).[15,25,26]

Dorsal striatal regions encode locomotion information, with both direct and indirect pathways being important for the control of behavior, in which D1-MSN and D2-MSN show similar patterns of activity.[15,34,36,38] These studies showed that neurons located in these regions were modulated by movement type for both Go and No-go conditions, not depending on action execution or withholding.[25,79] The data we obtained in this project for DLS is in accordance with previous studies, with both subpopulations of MSN showing increased activity around movement-related events (nose-poke exit and lever press) for both Go and No-go trials. However, the same is not completely observed in data obtained from neurons located in DMS. Here, we see this increased activity around movement-related events only in D2-MSN for both trial types, with D1-MSN presenting different patterns of activity.

D1-MSN activity in DMS showed increased activity after NP entry for Go trials, and suppression after lever press for Go and No-go trials, magazine entry (reward delivery) for Go trials, and nose-poke exit for Go trials. This suppression in activity might be caused by a prior period in which the animal moves to a different location inside the operant chamber (e.g. right or left lever, food magazine, etc). Although this activity is not specific to most of the task events we are recording, it seems to be enhanced around movement in the operant chamber. Given the fact that DMS has been described as modulated by locomotor speed, one way to better understand this increase in activity would be to use DeepLabCut to track the animals during the task and look at D1-MSN activity around high versus low movement.[35,80] If with this approach we find increased activity during periods of high movement than we could correlate to our first results and conclude that activity in D1-MSN located in DMS would be enhanced when the animals are moving inside the operant chamber to go to different locations required for their successful performance in the task.

Moreover, considering what is known about these dorsal regions, activity in DLS should be stronger during action selection when animals are already proficient (overtrained) in the task.[25] This stronger activity should be noticed in both subpopulations since they seem to be necessary for the initiation and execution of actions that were previously learned, promoting the ongoing behavior and avoiding eventual shift towards distinct actions.[25,37] Contrarily, activity in DMS should be more selective early in training. The fact that one region could be more involved early in training and the other after extended training, might in part explain why we find stronger activity after task events in DLS compared to DMS.

Furthermore, studies performed with a focus on DLS have shown that this striatal region is modulated by response direction and consequently has a strong effect of lateralization - movements that are contralateral to the recording site encode stronger activity.[25,36] Therefore, when data from more animals are added to this study, they should be subdivided by recording site (left or right hemisphere) to attend possible lateralization effects during task events in which there is a distinction between ipsi- and contralateral movements (e.g.,

lever pressing). This analysis of lateralization should be accounted for in all the three striatal subregions included in this thesis project.

Although these dorsal regions are described to vigorously encode locomotion, they also encode reward although with weaker signals.[15] In coherence with this, our data show that in DLS both D1- and D2-MSN have increased activity around reward delivery in both Go and No-go trials. The same happens in DMS but only for D2-MSN. Moreover, regarding increased activity around the time point in which the animals are at the food magazine eating the reward, it was only seen in D1-MSN in DLS, and in D2-MSN in DMS. In the two cases, this was seen for both trial types.

When comparing activity in D1-MSN upon reward delivery in aDLS and pDMS to VMS, one can see that the rise in activity is markedly bigger in VMS, which goes in agreement with the main role of this subregion comparatively to the role of more dorsal regions of the striatum.

Overall, our neuronal results show evidence that explains the activity in DLS according to what is stated by the synergetic model of the basal ganglia: there is a co-activation of D1-MSN and D2-MSN, in which D1-MSN might be promoting the execution of the desired action and D2-MSN the suppression of possible competing actions.[22,35,38] Although this model seems to capture functioning of the direct and indirect pathways, is thought that basal ganglia circuits might be more difficult and complex than this.[30,38] Additionally, activity in DMS does not seem to reflect a synergy between the direct and indirect pathways since we have different patterns of activity for D1- and D2-MSN during the same task events (suppressed activity versus increased activity).

Regarding the encoding of the outcome (reward vs punishment), we saw that both dorsal regions encoded reward, but we cannot make claims regarding the model that best describes the way activity is modulated by positive and negative stimuli. Since we do not have the data for the punishing conditions of our task (Go and No-go punishment) with do not have evidence to explain how neurons behave in these contingencies and how this is correlated with rewarding conditions.

To sum up, some factors should be considered when discussing the neuronal data presented in this thesis project. First of all, it is important to remember that because we were using this technique (fiber photometry) for the first time in our group we still need to improve the way our data is pre-processed. There are still some issues that should be taken into consideration when writing the MATLAB scripts. By solving these issues we might be able to improve the results. Another factor that might also end up changing our overview of the results is the histological data. Since our animals are still alive and being trained in the punishment conditions of the task, we do not have histology results which means that we cannot confirm if 1) we have a good viral expression that translates into a good signal that we will be able to record, and 2) if all the fibers are implanted in the correct coordinates. Depending on these histology results (that should be added in the future), the interpretation of the data might change.

Furthermore, it is also important to note that the way neuronal and behavioral data are aligned might be optimized in the future. For events like nose-poke entry and nose-poke exit, we aligned our data to the moment the auditory cue was turned ON for nose-poke entry and turned Off for nose-poke exit. To be more precise in this analysis we should align the recorded videos to the TTL and find the exact time frame in which animals perform the first nose-poke entry or exit.

By accounting for all these factors in the future we will have data that is robust enough to take proper conclusions and represent the way striatal subpopulations are encoding outcome-dependent action control in VMS, aDLS, and pDMS.

6. Conclusion and Future Perspectives

This thesis project was built in order to probe striatal subpopulations in outcome-dependent action controls in rats by performing *in vivo* fiber photometry while animals perform a Go/No-go reward and punishment task.

In summary, our findings report that it is possible to achieve GCaMP expression in rat's striatum by using a combined Cre x Flippase system. This viral expression was enhanced with high volumes of injection (1350nL), showing stronger fluorescence mean intensity and area of infection. The work developed in this project also shows that three seconds of continuous compounds stimulus formed by WN at 90dB paired with bright light is aversive for the animals and has a similar subjective value as one sugar pellet. This means that the aversive compound stimulus can be used as a punishment in the Go/No-go task, allowing direct and accurate comparisons between neuronal data from rewarding and punishing contingencies. In addition, our study demonstrates that rats can successfully learn and perform this novel and complex task, with Pavlovian bias influencing their performance between Go and No-go trials. Finally, in respect to neuronal data obtained using *in vivo* fiber photometry, we show that movement- and reward-related activity can be found in VMS, aDLS, and pDMS for both D1- and D2-MSN but with different patterns of activity.

The overall goal of this thesis project was to better investigate two different research questions: 1) How do direct and indirect pathway medium spiny neurons contribute to outcome-dependent action control? and 2) Do these signals diverge between different striatal subregions (VMS, aDLS, and pDMS)? The results we obtained with this work indicate that the way the direct and indirect pathways contribute to outcome-dependent action control seems to differ between striatal subpopulations and in some cases between action initiation or suppression (Go and No-go trials in our task, respectively). Although activity in aDLS seems to be following the synergetic model, in which both subpopulations are active and cooperate in order to execute the correct action and suppress competing actions, in VMS and pDMS the activity of D1- and D2-MSN seems to be less cooperative and synergetic. Here, we saw that for the same event we can have increased activity in one pathway and suppressed activity in the other. In general, we can infer that the direct and indirect pathways may present different patterns of activity, but both seem to be crucial for the attainment of desired and optimal behaviors.

Although in this thesis project we gathered a lot of data regarding the activity of these pathways, it is important to note that are still some optimizations that need to be done in order to achieve robust results and be able to perform vigorous statistical analysis. Thus, in the future, we should add data from animals that are still in the training phase of the task (including data of punishing contingencies) to the data that was here presented in order to increase our n and statistically probe our results. We should also optimize the MATLAB scripts we used to pre-process our data and consider histological data to assure viral expression and correct positioning of the optical fiber. By combining all these steps, we will be more prone to make conclusions regarding the way the direct and indirect pathways contribute to outcome-dependent action control.

The importance of the work here developed lies in the attempt of better understanding the neuronal mechanisms underlying decision-making, a process that seems to be impaired in several psychiatric conditions. Given that studies that purpose the dissection of these mechanisms are not feasible in humans due to their invasive component, we are still far from understanding how to manipulate these circuits to find target-specific treatment for patients suffering from these conditions. The use of animals models offers

solutions to provide better insights into the field. Thus, the use of novel transgenic rat lines that allow pathway-specific analysis, allowed us to ask more complex questions that intended to place us one step closer in the fight against psychiatric conditions. Although considered advances have been done regarding this matter, more studies are still needed to better understand how outcome-dependent action control is encoded by the brain, generating optimized behaviors.

Futures studies, focusing on single-cell activity using these transgenic animals lines combined with more specific imaging techniques might add new knowledge to what is known so far. Furthermore, the dissection of these mechanisms in animal lines that mimic certain psychiatric conditions, combined with behavioral tasks designed to test specific behavioral features seen in patients might also be taken into consideration as a way to go in the future.

7. References

(1) Guitart-masip, M.; Duzel, E.; Dolan, R.; Dayan, P. Action versus Valence in Decision Making. *Trends Cogn. Sci.* **2014**, *18* (4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003.

(2) Eisner, B.; Hommel, B. Effect Anticipation and Action Control. *J. Exp. Psychol. Hum. Percept. Perform.* **2001**, *27* (1), 229–240. https://doi.org/10.1037/AM96-1523.27.1.229.

(3) Balleine, B. W.; Doherty, J. P. O. Human and Rodent Homologies in Action Control : Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacol. Rev.* **2010**, 48–69. https://doi.org/10.1038/npp.2009.131.

(4) Sutton, R. S.; Barto, A. G. *Reinforcement Learning : An Introduction*, 2nd ed.; The MIT Press, **2017**.

(5) Bloem, B.; Huda, R.; Amemori, K.; Abate, A.; Krishna, G.; Wilson, A.; Carter, C. W.; Sur, M.; Graybiel, A. M. Multiplexed Action-Outcome Representation by Striatal Striosome-Matrix Compartments Detected with a Novel Cost-Benefit Foraging Task. *bioRxiv* **2021**, 1–56. https://doi.org/10.1101/2021.08.17.456542.

(6) Chen, X.; Holland, P.; Galea, J. M. The Effects of Reward and Punishment on Motor Skill Learning. *Curr. Opin. Behav. Sci.* **2018**, *20*, 83–88. https://doi.org/10.1016/j.cobeha.2017.11.011.

(7) Guitart-masip, M.; Huys, Q. J. M.; Fuentemilla, L.; Dayan, P.; Duzel, E.; Dolan, R. J. Go and No-Go Learning in Reward and Punishment : Interactions between Affect and Effect. *Neuroimage* **2012**, *62* (1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024.

(8) Guitart-masip, M.; Economides, M.; Huys, Q. J. M.; Dayan, P.; Dolan, R. J. Differential, but Not Opponent, Effects of L -DOPA and Citalopram on Action Learning with Reward and Punishment. *Psychopharmacology (Berl).* **2014**, *231*, 955–966. https://doi.org/10.1007/s00213-013-3313-4.

(9) Cavanagh, J. F.; Eisenberg, I.; Guitart-masip, M.; Huys, Q.; Frank, M. J. Frontal Theta Overrides Pavlovian Learning Biases. *J. Neurosci.* **2013**, *33* (19), 8541–8548. https://doi.org/10.1523/JNEUROSCI.5754-12.2013.

(10) Albrecht, M. A.; Waltz, J. A.; Cavanagh, J. F.; Frank, M. J.; Gold, J. M. Reduction of Pavlovian Bias in Schizophrenia : Enhanced Effects in Clozapine-Administered Patients. *PLoS One* **2016**, 1–23. https://doi.org/10.1371/journal.pone.0152781.

(11) Huys, Q. J. M.; Gölzer, M.; Friedel, E.; Heinz, A.; Cools, R.; Dayan, P.; Dolan, R. J. The Specificity of Pavlovian Regulation Is Associated with Recovery from Depression. *Psychol. Med.* **2016**, *46*, 1027–1035. https://doi.org/10.1017/S0033291715002597.

(12) Nord, C. L.; Lawson, R. P.; Huys, Q. J. M.; Pilling, S.; Roiser, J. P. Depression Is Associated with Enhanced Aversive Pavlovian Control over Instrumental Behaviour. *Sci. Rep.* **2018**, No. August, 1–10. https://doi.org/10.1038/s41598-018-30828-5.

(13) Weidacker, K.; Whiteford, S.; Boy, F.; Johnston, S. J. Response Inhibition in the Parametric Go / No-Go Task and Its Relation to Impulsivity and Subclinical Psychopathy. *Q. J. Exp. Psychol.* **2016**, *70* (3), 1–52. https://doi.org/10.1080/17470218.2015.1135350.

(14) Lopez, A. M.; Weintraub, D.; Claassen, D. O. Impulse Control Disorders and Related Complications of Parkinson's Disease Therapy. *Semin. Neurol.* **2017**, *37*, 186–192. https://doi.org/https://doi.org/ 10.1055/s-0037-1601887.

(15) Cox, J.; Witten, I. B. Striatal Circuits for Reward Learning and Decision-Making. *Nat. Rev. Neurosci.* **2019**. https://doi.org/10.1038/s41583-019-0189-2.

(16) Kaku, M. *The Future of Mind*, First Edit.; New York Times, **2014**.

(17) Bahuguna, J.; Aertsen, A.; Kumar, A. Existence and Control of Go / No-Go Decision Transition Threshold in the Striatum. *PLOS Comput. Biol.* **2015**, 1–36. https://doi.org/10.1371/journal.pcbi.1004233.

(18) Nonomura, S.; Nishizawa, K.; Sakai, Y.; Kawaguchi, Y.; Kato, S.; Uchigashima, M. Monitoring and Updating of Action Selection for Goal-Directed Behavior through the Striatal Direct and Indirect Pathways Article. *Neuron* **2018**, *99* (6), 1302-1314.e5. https://doi.org/10.1016/j.neuron.2018.08.002.

(19) Bostan, A. C.; Dum, P.; Strick, L. Functional Anatomy of Basal Ganglia Circuits with the Cerebral Cortex And. *Curr. Concepts Mov. Disord. Manag.* **2018**, *33*, 50–61. https://doi.org/10.1159/000480748.

(20) Alexander, G. E.; DeLong, M. R.; Strick, P. L. Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. *Annu. Rev. Neurosci.* **1986**, *9*, 357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041.

(21) Kravitz, A. V; Kreitzer, A. C.; Barker, J. M.; Taylor, J. R.; Chandler, L. J.; Maguire, E. P.; Macpherson, T.; Swinny, J. D.; Dixon, C. I.; Herd, M. B.; et al. Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology* **2012**, *27*, 167–177. https://doi.org/10.1152/physiol.00004.2012.

(22) Burke, D. A.; Rotstein, H. G.; Alvarez, V. A. Perspective Striatal Local Circuitry : A New Framework for Lateral Inhibition. *Neuron* **2017**, *96* (2), 267–284. https://doi.org/10.1016/j.neuron.2017.09.019.

(23) Soares-cunha, C.; Coimbra, B.; Sousa, N.; Rodrigues, A. J. Reappraising Striatal D1- and D2-Neurons in Reward and Aversion. *Neurosci. Biobehav. Rev.* **2016**, 370–386. https://doi.org/10.1016/j.neubiorev.2016.05.021.

(24) The, I.; Parent, A.; Hazrati, L. Functional Anatomy of the Basal Ganglia. The Cortico-Basal Ganglia-Thalamo-Cortical Loop. *Brain Res. Rev. - Elsevier* **1995**, *20*, 91–127. https://doi.org/10.1016/0165-0173(94)00007-c.

(25) Burton, A. C.; Nakamura, K.; Roesch, M. R. From Ventral-Medial to Dorsal-Lateral Striatum: Neural Correlates of Reward-Guided Decision-Making. *Neurobiol. Learn. Mem.* **2015**, *117*, 51–59. https://doi.org/10.1016/j.nlm.2014.05.003.

(26) Yin, H. H.; Ostlund, S. B.; Knowlton, B. J.; Balleine, B. W. The Role of the Dorsomedial Striatum in Instrumental Conditioning. *Eur. J. Neurosci.* **2005**, *22* (March), 513–523. https://doi.org/10.1111/j.1460-9568.2005.04218.x.

(27) Yin, H. H.; Knowlton, B. J.; Balleine, B. W. Inactivation of Dorsolateral Striatum Enhances Sensitivity to Changes in the Action – Outcome Contingency in Instrumental Conditioning. *Behav. Brain Res.* **2006**, *166*, 189–196.

https://doi.org/10.1016/j.bbr.2005.07.012.

(28) Jeffrey R. Pettibone, Jai Y. Yu, Rifka C. Derman, Thomas W. Faust, Elizabeth D. Hughes, Wanda E. Filipiak, Thomas L. Saunders, C. R. F. and J. D. B. Knock-in Rat Lines with Cre Recombinase at the Dopamine D1 and Adenosine 2a Receptor Loci. *eNeuro* **2019**.
https://doi.org/10.1523/ENEURO.0163-19.2019.

(29) Klaus, A.; Alves, J.; Costa, R. M. What , If , and When to Move : Basal Ganglia Circuits and Self-Paced Action Initiation. *Annu. Rev. Neurosci.* **2019**, *42*, 459–483.
https://doi.org/10.1146/annurev-neuro-072116-031033.

(30) Klaus, A.; Martins, G. J.; Paixao, V. B.; Zhou, P.; Paninski, L.; Costa, R. M.; Klaus, A.; Martins, G. J.; Paixao, V. B.; Zhou, P.; et al. The Spatiotemporal Organization of the Striatum Encodes Action Space. *Neuron* **2017**, *95* (5), 1171-1180.e7.
https://doi.org/10.1016/j.neuron.2017.08.015.

(31) Albin, R. L.; Young, A. B.; Penney, J. B. The Functional Anatomy of Basal Ganglia Disorders. *Perspect. Dis. - Elsevier* **1989**, *12* (10), 366–375.
https://doi.org/10.1016/0166-2236(89)90074-x.

(32) Bahuguna, J.; Weidel, P.; Morrison, A. Exploring the Role of Striatal D1 and D2 Medium Spiny Neurons in Action Selection Using a Virtual Robotic Framework. *Eur. J. Neurosci.* **2019**, *49* (May 2018), 737–753. https://doi.org/10.1111/ejn.14021.

(33) Mink, J. W. The Basal Ganglia and Involuntary Movements: Impaired Inhibition of Competing Motor Patterns. *Neurol. Rev.* **2003**, *60*, 1365–1368.
https://doi.org/10.1001/archneur.60.10.1365.

(34) Barbera, G.; Liang, B.; Zhang, L.; Chen, R.; Li, Y.; Lin, D. Spatially Compact Neural Clusters in the Dorsal Striatum Encode Locomotion Relevant Information. *Neuron* **2016**, No. 92, 1–12. https://doi.org/10.1016/j.neuron.2016.08.037.

(35) Parker, J. G.; Marshall, J. D.; Ahanonu, B.; Wu, Y.; Kim, T. H.; Grewe, B. F. Diametric Neural Ensemble Dynamics in Parkinsonian and Dyskinetic States. *Nature* **2018**, *557*, 177–182.
https://doi.org/10.1038/s41586-018-0090-6.

(36) Cui, G.; Jun, S. B.; Jin, X.; Pham, M. D.; Vogel, S. S.; Lovinger, D. M.; Costa, R. M. Concurrent Activation of Striatal Direct and Indirect Pathways during Action Initiation. *Nature* **2013**, *494* (7436), 238–242. https://doi.org/10.1038/nature11846.

(37) Jin, X.; Tecuapetla, F.; Costa, R. M. Basal Ganglia Subcircuits Distinctively Encode the Parsing and Concatenation of Action Sequences. *Nat. Neurosci.* **2014**, *17* (3), 423–430.
https://doi.org/10.1038/nn.3632.

(38) Tecuapetla, F.; Jin, X.; Lima, S. Q.; Costa, R. M. Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. *Cell* **2016**, *166* (3), 703–715.
https://doi.org/10.1016/j.cell.2016.06.032.

(39) Kravitz, A. V; Freeze, B. S.; Parker, P. R. L.; Kay, K.; Thwin, M. T.; Deisseroth, K.; Kreitzer, A. C. Regulation of Parkinsonian Motor Behaviours by Optogenetic Control of Basal Ganglia Circuitry. *Nature* **2010**, *466* (7306), 622–626.
https://doi.org/10.1038/nature09159.

(40) Vicente, A. M.; Galvão-ferreira, P.; Tecuapetla, F.; Costa, R. M. Direct and Indirect Dorsolateral Striatum Pathways Reinforce Different Action Strategies. *Curr. Biol.* **2016**, *26* (7), R267–R269. https://doi.org/10.1016/j.cub.2016.02.036.

(41) Cole, S. L.; Robinson, M. J. F.; Berridge, K. C. Optogenetic Self-Stimulation in the Nucleus Accumbens : D1 Reward versus D2 Ambivalence. *PLoS One* **2018**, *63649*, 1–29. https://doi.org/10.1371/journal.pone.0207694.

(42) Wang, L.; Rangarajan, K. V; Gerfen, C. R.; Krauzlis, R. J. Activation of Striatal Neurons Causes a Perceptual Decision Bias during Visual Change Detection in Mice. *Neuron* **2018**, 1–13. https://doi.org/10.1016/j.neuron.2018.01.049.

(43) Soares-cunha, C.; Coimbra, B.; David-pereira, A.; Borges, S.; Pinto, L.; Costa, P.; Sousa, N.; Rodrigues, A. J. Activation of D2 Dopamine Receptor-Expressing Neurons in the Nucleus Accumbens Increases Motivation. *Nat. Commun.* **2016**, *7* (May), 1–11. https://doi.org/10.1038/ncomms11829.

(44) Kravitz, A. V; Tye, L. D.; Kreitzer, A. C. Distinct Roles for Direct and Indirect Pathway Striatal Neurons in Reinforcement. *Nat. Neurosci.* **2012**, 1–3. https://doi.org/10.1038/nn.3100.

(45) Paton, J. J.; Louie, K. Reward and Punishment Illuminated. *Nat. Neurosci.* **2012**, *15* (6), 807–809. https://doi.org/10.1038/nn.3122.

(46) Lobo, M. K.; III, H. E. C.; Chaudhury, D.; Friedman, A. K.; Sun, H.; Damez-Werno, D.; Dietz, D. M.; Zaman, S.; Koo, J. W.; Kennedy, P. J.; et al. Cell-Type Specific Loss of BDNF Signaling Mimics Optogenetic Control of Cocaine Reward. **2010**, *330* (October), 385–390. https://doi.org/10.1126/science.1188472 Salmonella.

(47) Kupchik, Y. M.; Brown, R. M.; Heinsbroek, J. A.; Lobo, M. K.; Schwartz, D. J.; Kalivas, P. W. Coding the Direct / Indirect Pathways by D1 and D2 Receptors Is Not Valid for Accumbens Projections. *Nat. Neurosci.* **2015**, No. July, 1–4. https://doi.org/10.1038/nn.4068.

(48) Siciliano, C. A.; Tye, K. M. Leveraging Calcium Imaging to Illuminate Circuit Dysfunction in Addiction. *Alcohol* **2019**, *74*, 47–63. https://doi.org/10.1016/j.alcohol.2018.05.013.

(49) Hodgkin, L.; Ridgway, E. B. Depolarization and Calcium Entry in Squid Giant Axons. *J. Physiol.* **1971**, *218*, 709–755. https://doi.org/10.1113/jphysiol.1971.sp009641.

(50) Resendez, S. L.; Jennings, J. H.; Ung, R. L.; Namboodiri, V. M. K.; Zhou, Z. C.; Otis, J. M.; Nomura, H.; Mchenry, J. A.; Kosyk, O.; Stuber, G. D. Visualization of Cortical, Subcortical and Deep Brain Neural Circuit Dynamics during Naturalistic Mammalian Behavior with Head-Mounted Microscopes and Chronically Implanted Lenses. *Nat. Protoc.* **2016**, *11* (3), 566–597. https://doi.org/10.1038/nprot.2016.021.

(51) Daigle, T. L.; Madisen, L.; Hage, T. A.; Li, L.; Tasic, B.; Walker, M.; Graybuck, L. T.; Yao, Z.; Fong, O.; Nguyen, T. N.; et al. A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* **2018**, No. 174, 465–480. https://doi.org/10.1016/j.cell.2018.06.035.

(52) Chen, T.; Wardill, T. J.; Sun, Y.; Pulver, S. R.; Renninger, S. L.; Baohan, A.; Schreiter, E. R.; Kerr, R. A.; Orger, M. B.; Jayaraman, V.; et al. Ultra-Sensitive Fluorescent Proteins for Imaging Neuronal Activity. *Nature* **2013**, *499* (7458), 295–300. https://doi.org/10.1038/nature12354.

(53) Flusberg, B. A.; Nimmerjahn, A.; Cocker, E. D.; Mukamel, E. A.; Barretto, R. P. J.; Ko, T. H.; Burns, L. D.; Jung, J. C.; Schnitzer, M. J. High-Speed, Miniaturized Fluorescence Microscopy in Freely Moving Mice. *Br. Commun.* **2008**, *5* (11), 3–6. https://doi.org/10.1038/NMETH.1256.

(54) Groot, A. De; Boom, B. J. G. Van Den; Genderen, R. M. Van. NINscope, a Versatile Miniscope for Multi- Region Circuit Investigations. *Elife* **2020**, 1–24. https://doi.org/10.7554/eLife.49987.

(55) Kloet, S. F. De; Bruinsma, B.; Terra, H.; Heistek, T. S.; Passchier, E. M. J.; Berg, A. R. Van Den; Luchicchi, A.; Min, R.; Pattij, T.; Mansvelder, H. D. Bi-Directional Regulation of Cognitive Control by Distinct Prefrontal Cortical Output Neurons to Thalamus and Striatum. *Nat. Commun.* **2021**, *12* (1994). https://doi.org/10.1038/s41467-021-22260-7.

(56) Mansy, M. M.; Member, I. S.; Kim, H.; Oweiss, K. G.; Member, I. S. Spatial Detection Characteristics of a Single Photon Fiber Photometry System for Imaging Neural Ensembles. *9th Int. IEEE/EMBS Conf. Neural Eng.* **2019**, 969–972. https://doi.org/10.1109/NER.2019.8717005.

(57) Liang, Z.; Neuberger, T.; Zhang, N. Global Brain Signal in Awake Rats. *Brain Struct Funct.* **2020**, *225* (January), 227–240. https://doi.org/10.1007/s00429-019-01996-5.

(58) Guo, Q.; Zhou, J.; Feng, Q.; Lin, R.; Gong, H.; Luo, Q.; Zeng, S.; Luo, M.; Fu, L. Multi-Channel Fiber Photometry for Population Neuronal Activity Recording. *Biomed. Opt. Express* **2015**, *6* (10), 7319–7324. https://doi.org/10.1364/BOE.6.003919.

(59) Kim, C. K.; Yang, S. J.; Young, N. P.; Kauvar, I.; Jennings, J. H.; Lerner, T. N.; Berndt, A.; Lee, S. Y.; Ramakrishnan, C.; Davidson, T. J.; et al. Simultaneous Fast Measurement of Circuit Dynamics at Multiple Sites across the Mammalian Brain. *Nat. Methods* **2016**, *13* (4), 325–328. https://doi.org/10.1038/nmeth.3770.

(60) Lu, L.; Gutruf, P.; Xia, L.; Bhatti, D. L.; Wang, X.; Vazquez-guardado, A. Wireless Optoelectronic Photometers for Monitoring Neuronal Dynamics in the Deep Brain. *PNAS* **2018**. https://doi.org/10.1073/pnas.1718721115.

(61) Barnett, L. M.; Hughes, T. E.; Drobizhev, M. Deciphering the Molecular Mechanism Responsible for GCaMP6m's Ca 2+ -Dependent Change in Fluorescence. *PLoS One* **2017**, *12* (2), 1–24. https://doi.org/10.1371/journal.pone.0170934.

(62) Lerner, T. N.; Shilyansky, C.; Davidson, J.; Luo, L.; Tomer, R.; Deisseroth, K. Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell* **2015**, *162* (3), 635–647. https://doi.org/10.1016/j.cell.2015.07.014.

(63) Alcantara, A. A.; Chen, V.; Herring, B. E.; Mendenhall, J. M.; Berlanga, M. L. Localization of Dopamine D2 Receptors on Cholinergic Interneurons of the Dorsal Striatum and Nucleus Accumbens of the Rat. *Elsevier* **2003**, *986*, 22–29. https://doi.org/10.1016/S0006-8993(03)03165-2.

(64) Bryda, E. C. The Mighty Mouse: The Impact of Rodents on Advances in Biomedical Research. *Mo. Med.* **2013**, *110* (3), 207–211.

(65) Iannaccone, P. M.; Jacob, H. J. Rats ! *Dis. Model. Mech.* **2009**, *2*, 206–210. https://doi.org/10.1242/dmm.002733.

(66) The Jackson Laboratory. The Cre-LOX and FLP-FRT systems https://www.jax.org/news-and-insights/2006/may/the-cre-lox-and-flp-frt-systems. (Acessed: june 2021)

(67) Syed, E. C. J.; Grima, L. L.; Magill, P. J.; Bogacz, R.; Brown, P.; Walton, M. E. Action Initiation Shapes Mesolimbic Dopamine Encoding of Future Rewards. *Nat. Neurosci.* **2016**, *19* (1), 4–9. https://doi.org/10.1038/nn.4187.

(68) Goedhoop, J. N.; van den Boom, B. J. G.; Arbab, T.; Willuhn, I. Nucleus-Accumbens Dopamine Tracks Aversive Stimulus Duration and Prediction but Not Value or Prediction Error. *bioRxiv* **2021**, 1–29. https://doi.org/10.1101/2021.01.16.426967.

(69) Technologies, T.-D. Fiber Photometry: Record Real-Time Responses with the RZ10X https://www.tdt.com/system/fiber-photometry-system/. (Acessed: june 2021)

(70) Goding, J. W. Monoclonal Antibodies - Principles and Practice. In *Elsevier*; **1996**; Vol. 3, pp 400–423. https://doi.org/10.1016/B978-0-12-287023-1.50061-4.

(71) Kapuscinski, J. DAPI : A DNA-Specific Fluorescent Probe. *Biotech. Histochem.* **2016**, *0295* (October), 220–233. https://doi.org/10.3109/10520299509108199.

(72) Bahmani, P.; Schellenberger, E.; Klohs, J.; Steinbrink, J.; Cordell, R.; Zille, M.; Mu, J.; Harhausen, D.; Hofstra, L.; Reutelingsperger, C.; et al. Visualization of Cell Death in Mice with Focal Cerebral Ischemia Using Fluorescent Annexin A5 , Propidium Iodide , and TUNEL Staining. *J. Cereb. Blood Flow Metab.* **2011**, *31*, 1311–1320. https://doi.org/10.1038/jcbfm.2010.233.

(73) Florillo, C. D. Two Dimensions of Value : Dopamine Neurons Represent Reward but Not Aversiveness. *Science (80-.).* **2013**, *341* (6145), 546–549. https://doi.org/10.1126/science.1238699.

(74) Mollenauer, S.; Bryson, R.; Robison, M.; Phillips, C. Noise Avoidance in the C57BL/6J Mouse. *Anim. Learn. Behav.* **1992**, *1* (20), 25–32. https://doi.org/10.3758/BF03199943.

(75) Hughes, R. A.; Bardo, M. T. Shuttlebox Avoidance by Rats Using White Noise Intensities from 90-120dB SPL as the UCS. *J. Aud. Res.* **1981**, No. 21, 109–118.

(76) Campbell, B. A.; Bloom, J. M. Relative Aversiveness of Noise and Shock. *J. Comp. Physiol. Psychol.* **1965**, *60* (3), 440–442. https://doi.org/10.1037/h0022572.

(77) Soares-cunha, C.; Vasconcelos, N. A. P. De; Coimbra, B.; Domingues, A. V.; Silva, J. M.; Loureiro-campos, E.; Gaspar, R.; Sotiropoulos, I.; Sousa, N.; Rodrigues, A. J. Nucleus Accumbens Medium Spiny Neurons Subtypes Signal Both Reward and Aversion. *Mol. Psychiatry* **2020**, *25*, 3241–3255. https://doi.org/10.1038/s41380-019-0484-3.

(78) Chen, X.; Rutledge, R. B.; Brown, H. R.; Dolan, R. J.; Bestmann, S.; Galea, J. M. Age-Dependent Pavlovian Biases Influence Motor Decision-Making. *PLOS Comput. Biol.* **2018**, *14* (7), 1–22. https://doi.org/10.1371/journal.pcbi.1006304.

(79) Apicella, P.; Ljungberg, T.; Scarnati, E.; Schultz, W. Responses to Reward in Monkey Dorsal and Ventral Striatum. *Exp. Brain Res.* **1991**, *85*, 491–500. https://doi.org/10.1007/BF00231732.

(80) Mathis, A.; Mamidanna, P.; Cury, K. M.; Abe, T.; Murthy, V. N.; Mathis, M. W.; Bethge, M. DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning. *Nat. Neurosci.* **2018**, *21* (September), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y.

8. Appendix

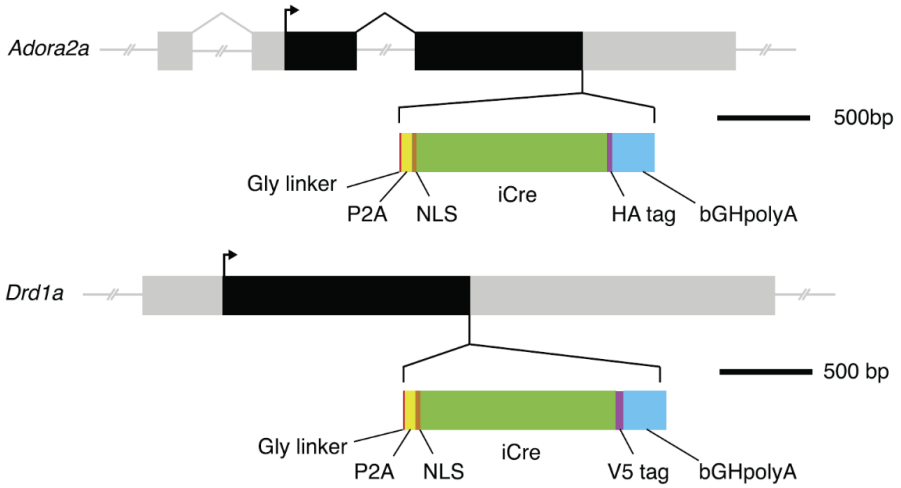

Figure 20: Novel transgenic D1-i*Cre* and A2a-i*Cre* rat-lines used in this thesis project. Schematic representation of insertion cassettes into A2a (Adora2a - Top panel) and D1 (Drd1a - bottom panel) genes. Abbreviations used: P2A = porcine teschovirus-1 self-cleaving peptide; NLS = nuclear localization sequence; HA = influenza hemagglutinin protein tag YPYDVPDYA; V5 = peptide tag GKPIPNPLLGLDST; bGH = bovine growth hormone polyadenylation sequence. Adapted from Pettibone *et al.* (2019).[28]

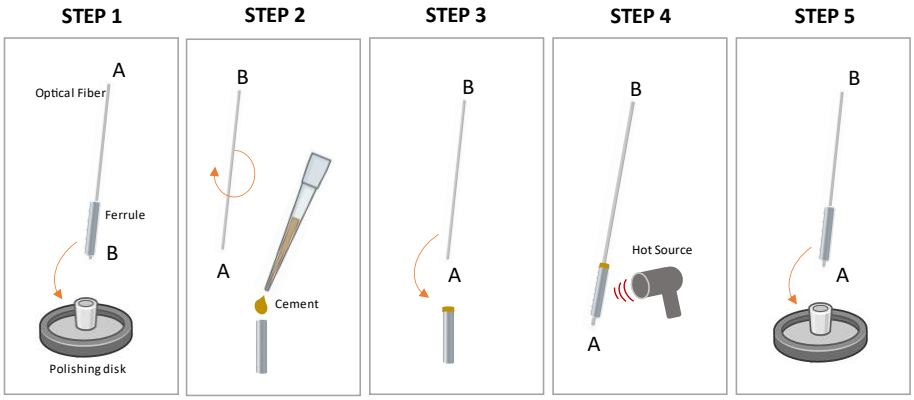

Figure 21: Schematic representation of optical fiber manufacturing. To prepare optical fiber to use in *in vivo* fiber photometry experiments we followed steps one to five. After cutting optical fibers to the desired length and insert them inside a ceramic ferrule, we started by polishing one extremity (side B) of the optical fiber using a polishing disk (Step 1). Once the extremity B was completely shiny and flat, we took the fiber out of the ferrule, reverse it, and after adding a drop of cement on top of the ferrule (Step 2), we again placed the fiber inside it (Step 3). Next, we cleaned the excess of cement in both ends of the ferrule and with a hot source we dried it (Step 4). Lastly, we polished the extremity A of the optical fiber using the polishing disk until it becomes shiny and flat (Step 5). After this process is finished, the performance of the optical fiber (percentage of light that passes through the fiber) should be assessed using a blue laser and a digital optical power and energy meter.