Mesolimbic dopamine and circuit level mechanisms in action initiation and restraint

George Jenkins

Hertford College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity term 2023

Mesolimbic dopamine and circuit level mechanisms in action initiation and restraint

George Jenkins
Hertford College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Abstract

Beneficial decision making does not just involve selecting which action to make based on anticipated future reward but also being able to withhold responding at appropriate times. Mesolimbic dopamine is associated with both reward processing and the activation of reward seeking action. Yet an understanding of how phasic dopamine signals may influence reward guided action, and how extended circuits may act in concert with dopamine when action must be controlled, remains lacking. To probe these questions, we applied different causal manipulations within a common behavioural paradigm – a reward-guided Go/No-Go task – where animals must either restrain or enact action to obtain different sizes of reward. In Chapter 3 we demonstrate that transient activation of VTA dopamine neurons can promote action over inaction, in a manner dependent on the magnitude of activation and its timing relative to task relevant cues. Though effects on behaviour were observed most strongly on trials receiving stimulation, we also found stimulation influenced within-session behavioural strategy, biasing animals towards premature action as the session progressed. In Chapter 4, we show that the ability to successfully restrain action depends on prelimbic and infralimbic, but not the medial orbital, subregions of the medial prefrontal cortex. However, while pharmacological inactivation of prelimbic and infralimbic cortex both increased premature responses on No-Go trials, analysis of the patterns of errors showed distinct contributions of these areas to action restraint. Conversely all three subregions are shown to be crucial for executing the correct action sequence for reward. In Chapter 5 we determined that potentiating endocannabinoid signalling did not boost motivation as had been observed in simpler behavioural paradigms, but rather selectively increased the likelihood of entering into an unfocused state when needing to complete the Go trial action sequence to gain reward. Together, this thesis demonstrates the capacity for phasic dopamine signals to causally influence reward guided action, and highlights how medial prefrontal circuits and the endocannabinoid system contribute to the initiation and restraint of action, key information for progressing towards an integrated understanding of the systems governing appropriate action control.

Acknowledgments

Before beginning a DPhil I probably though two things about a thesis. First that it was a common experience - which every student must endure in a similar fashion, and second that it was an individual venture and must be faced alone. I now believe that neither of these statements are true, while every DPhil student I know has written a thesis its seems that each has had their individual highs and individual horrors. And this document, as I'm sure is the common case, would not exist without the support, assistance and friendship of many people other than myself.

I would first like to thank Mark, my primary supervisor, for being a great teacher and mentor and an excellent guiding voice of reason. I would also like to extend thanks to the Walton lab in general – science is not a solo endeavour! I would like to particularly highlight Lauren Burgeno, Hironiri Ishii and Mason Silveria for scientific assistance and technical aid. I would like to extend thanks also to Marta Blanco-Pozo, Merima Šabanović and Raquel Pinacho – for statistical wisdom and being a general positive force during the slog that is dedicated thesis writing. I would also like to acknowledge Freya Marijatta, for being a dopaminergic comrade-at-arms and for helpful discussions – about science and in general.

Finally I would like to thanks friend, in Oxford and further afield, and my parents without whom nothing would be possible. Finally (finally) I would like to thank Maddy Wyburd, without whom my time in Oxford would have been unthinkable (!)

Covid-19 impact statement

My DPhil spanned the global COVID-19 pandemic, which commenced 15 months into my project (as an iCASE-funded student, my lab work started in January 2019). For 5 months from mid-March until mid-August laboratory facilities were completely closed, preventing data collection. The nature of my experimental work required significant training of animals, with each cohort taking several months to complete. Therefore, when facilities closed in March, the cohort I had been working with for ~2 months had to be culled – meaning I lost at least 7 months of effective data collection time.

I received a 3.5-month extension from the BBSRC, which enabled key experiments to be completed. However, the course of my DPhil remained markedly altered due to Covid. Moreover, my ability to work efficiently even when the lab was reopened was substantially affected. First, the pandemic exacerbated an existing anxiety disorder. Second, I lost my day-to-day supervisor, a Japanese post-doctoral researcher who was not granted a fellowship extension to mitigate the lab closure and therefore had to return to Japan early. Given the long-lasting uncertainties during the pandemic about whether experimental work would remain possible or would again be halted by additional lockdowns, this eventually led to a decision to switch away from additional planned experiments to focus on analyses of existing data sets.

Table of Contents

1. General Introduction 1

1.1 – An introduction to reward guided action 2

1.2 - Anatomy and physiology of the midbrain dopamine system 7

1.3 - Theories of mesolimbic dopamine function 12

1.4 - Mesolimbic dopamine and the initiation of reward guided action 23

1.5 - Regulation of action and regulation of dopamine 29

1.6 - Medial prefrontal cortex, dopamine and action 33

1.7 - Endocannabinoids, Dopamine and action 40

1.8 - Thesis Aims 44

2. General Methods 46

2.1 - Go/No-Go Task 47

2.2 -Optogenetics experiments 54

2.3 -Infusion experiments 64

2.4 - Video tracking 68

3. How do phasic dopamine transients influence action initiation and restraint? 72

3.1 - Introduction 73

3.2 - Methods 76

3.3 - Results 82

3.4 - Discussion 99

4. Do mPFC subregions make separable contributions to reward guided action control? 107

4.1 - Introduction 108

4.2 - Methods 112

4.3 - Results 116

4.4 - Discussion 144

5. How does potentiated endocannabinoid signalling effect action initiation and restraint? 151

5.1 - Introduction 152

5.2 - Methods 155

5.3 - Results 158

5.4 - Discussion 171

6. General Discussion 178

6.1 - Introduction 179

6.2 - Summary and implications of key findings 181

6.3 - Limitations, technical considerations, and future directions 187

6.4 - Future directions 193

References 195

1. General Introduction

This thesis aims to investigate processes underlying appropriate control of action, with a particular focus on mesolimbic dopamine and complementary circuits and neurotransmitters. In this introduction, we will first discuss theories of dopaminergic function that relate to reward guided action, before examining two particular systems that may regulate both aspects of action and mesolimbic dopamine: the medial prefrontal cortices and the endocannabinoid system.

1.1 – An introduction to reward guided action

Why do anything? – a question pondered by countless philosophers and angst-ridden teenagers. For most biological organisms, the answer lies in maximising outcomes that are likely to promote the success of the individual or species, such food and sex, while minimizing those that have an opposing effect. In addition to being advantageous to personal and collective survival, these kinds of outcomes are often reinforcing, making the pattern of behaviour that led to their obtainment more likely to be repeated, as well as tending to induce subjective feelings of pleasure. These positive experiences are often referred to as 'rewarding' and behaviour organised by actual or potential rewards termed 'reward guided'. The category of reward is often expanded to include any stimuli, such as drugs of abuse or rewarding brain stimulations, which are mediated by the same underlying neural circuity and effect the same changes in behaviour as their naturally occurring counterparts (Ikemoto & Bonci, 2014).

While the term 'reward' is often employed, the effects of rewards on behaviour are not a singular entity, but rather reflect the collective properties of many components. Rewards can induce a motivational drive, which is distinct from the hedonic 'pleasure' of anticipating or consuming a reward, both psychologically and in terms of the underlying neural circuitry. In addition to effects on moment-by-moment behaviour, rewarding outcomes can influence future behaviour as organisms learn the value of different states and actions – and this process may also be psychologically and biologically separable from both motivation and pleasure (Berridge et al., 2009; Cogliati Dezza et al., 2022; Ikemoto, 2010; Scholl et al., 2015). The influence of natural rewards on behaviour is therefore multifaceted.

Reward-guided behaviour is by nature a multi-step process (Rangel et al., 2008). First, information relating to the availability of possible rewards arising from the environmental context and cues must be processed. Second, the value of those rewards and the costs of obtaining them must be calculated according to factors such as the internal state of the agent – a stale sandwich is not very valuable when sated but might be rather more attractive when starving. Finally, a pattern of behaviour, or action, must be selected with goal of obtaining available rewards, and finally executed. Following the execution of reward guided behaviour, the outcome of the action is evaluated, and learning from these outcomes updates the processes of valuation and action selection in order to optimize future behaviour.

While reducing these processes to a schematic gives the illusion of simplicity, in reality dissecting reward guided behaviour presents a complex problem. The identity and probability of rewards, as well as their subjective value according to the current internal state, can vary greatly. Actions are how organisms interact with their environment to obtain potential rewards, and the term 'action' can refer to anything from the complete pattern of motor behaviour used to achieve an outcome or a single movement component produced by an individual muscle contraction to progress towards that goal.

As with rewards, action requirements too can differ depending on the current state of the world. Cues and contextual information often signal which action strategy would be best implemented under the current conditions. For example, something in the environment that looks like reward, or smells like a reward, can often elicit simple actions such as approach and consumption (known as Pavlovian responses).

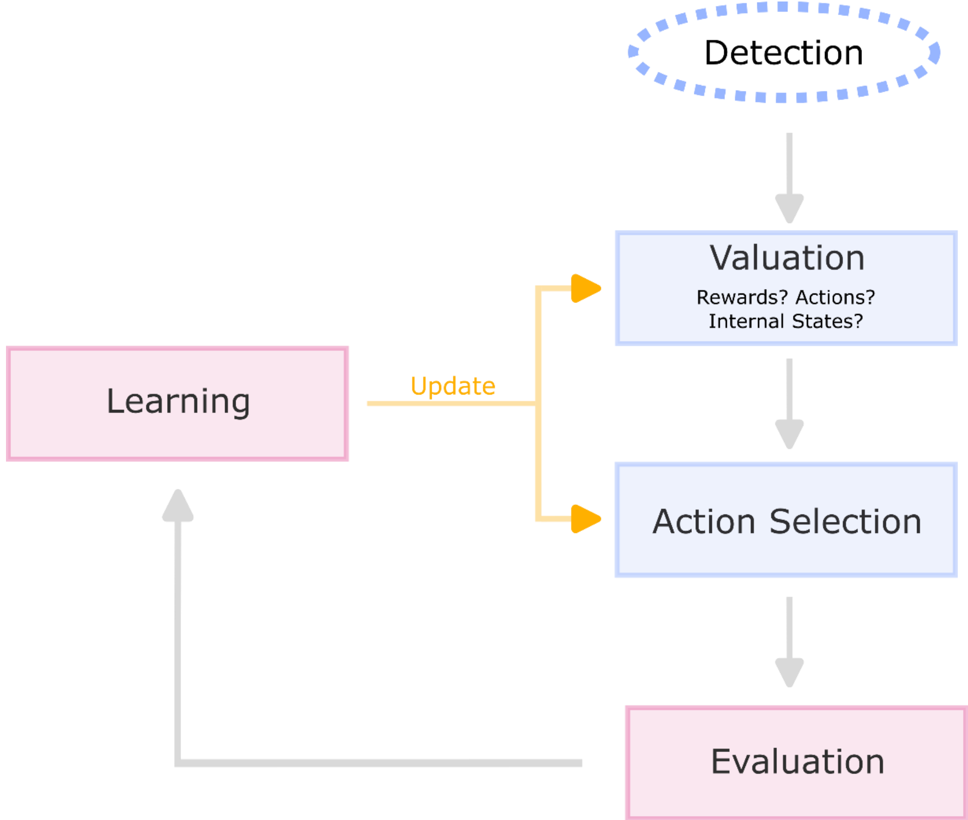


Figure 1. Basic stages in reward-guided behaviour. Adapted from Rangel et al 2008.

However, there are many situations where simple Pavlovian approach behaviour is *not* the most optimal. For example, a predator that immediately engages in pursuit as soon as prey is detected may miss out on the vital hunting advantage that a stealthy approach may provide. For prey animals, sighting a predator may indicate that potential costs of approaching a currently available reward far outweigh the benefits. Even in the absence of predation risk, pursuing reward can involve paying time and effort costs. Sometimes restraining from approaching a reward in the present may mean being able to take advantage of richer rewarding opportunities in the future. Such situations potentially generate conflict between the simple approach behaviour governed by Pavlovian systems,

and instrumental responses, where actions can be performed to cause a desired outcome. Such situations of conflict require engagement of instrumental control to regulate appropriate responding (Dayan et al., 2006; Dickinson, 1980; Gershman et al., 2021; Mackintosh, 1983).

On a moment-by-moment basis, the behavioural output of reward-guided decision processes can be defined as either action initiation or action restraint. *Initiation* is the expression of motor behaviour to obtain reward. However, behavioural output can be shaped by different processes. For example, reward-associated cues can elicit Pavlovian approach behaviours, which can be considered as pre-specified responses to environmental stimuli such as rewards (e.g., approach response, as outlined above). There are then instrumental responses, which can extend from a single operant response to complex sequences that chain many actions together. Instrumental responses can be executed under the control of different systems. Goal-directed instrumental action selection chooses a behavioural strategy in order to achieve a particular outcome. Habitual behaviour, on the other hand, selects actions not with a particular result in mind, but rather because repeated experience has shown that executing a particular action after a particular environmental stimuli is advantageous. As such, goal directed action selection is rapidly sensitive to changes in outcome, whether or not the desired aim was achieved or whether internal needs have changed, while actions selected under habitual control tend to be more reflexive in nature, but highly efficient in stable environments. Both failing to act swiftly, or failing to act with the appropriate goal in mind, can lead to missed opportunities, meaning these systems of action selection and initiation must be balanced according to environmental demands (Dayan 2006).

Alternatively, in some situations the drive to emit action must be withheld, delayed or cancelled in order obtain greater rewarding outcomes. Just as the reasons for inhibiting action can be many and varied, such as to avoid danger or meet instrumental action requirements, so can the underlying psychological processes. The taxonomy of inhibitory control as proposed by Bari and Robbins (2013) distinguishes between the restraint, postponement and cancellation of action as separable forms of response inhibition (Figure 2). For example, the process of *action cancellation* requires ongoing action to be arrested. Conversely *action postponement* describes the ability to withhold from acting while waiting, and *action restraint* is defined as the inhibition of planned responding under certain conditions. These process are thought be subserved by different neural circuits (Bari & Robbins, 2013; Dalley & Roiser, 2012), and can interact with the systems that promote action initiation in complex ways. Should the ability to withhold action misfunction, organisms may gain less reward overall or even expose themselves to danger.

Optimal control of action therefore must be situationally appropriate, using available information to determine the best action strategy (whether making or withholding action) within the current state. This process likely involves the interaction of many systems that carry sensory information about rewards, that can compute the values of possible actions, and that index motor circuits to shape behaviour. A full dissection of the neural circuits that contribute to the appropriate utilisation of initiation and restraint is beyond the scope of this thesis. Instead, as will be described in sections below, discussion will focus on midbrain dopamine, which is well placed to shape reward guided action, before expanding to other relevant systems that might interact and regulate dopamine release.

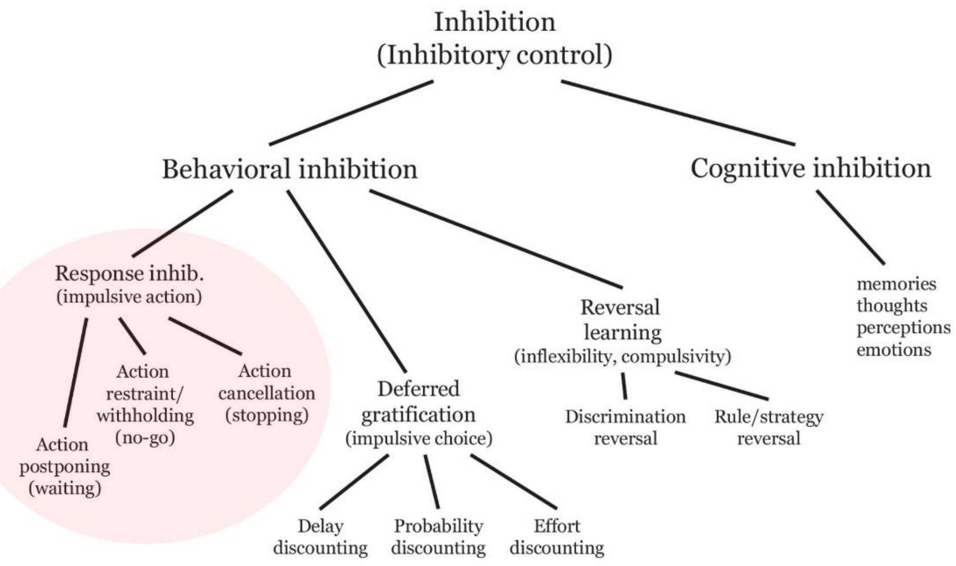


Figure 2. Taxonomy of behavioural inhibition proposed by Bari and Robbins (2013). Components of behavioural inhibition that involve action (response inhibition) highlighted in red. (Adapted from Bari and Robbins, 2013)

1.2 - Anatomy and physiology of the midbrain dopamine system

The application of immunohistochemical techniques in the 1960s first defined the localisation of dopamine cell bodies in the midbrain to three nuclei, the Ventral Tegmental Area (A10, VTA), Substantia Nigra pars compacta (A9, SNc), and the retrorubral field (A8) (Dahlstrom, 1964; German & Manaye, 1993). The projections of these nuclei are organised into three major pathways. In rodents, the nigrostriatal pathway contains afferents arising from the SNc and terminating in the dorsomedial and dorsolateral regions of the striatum. Dopaminergic cells located in the more medial VTA instead largely innervate the more ventral and medial portion of the striatum, also known as the nucleus accumbens (NAc). These projections contribute to the mesolimbic pathway, which also sends afferents to other limbic structures including the amygdala and parts of the hippocampus. The VTA also sends a more diffuse projection throughout several regions of the cortex, thought to be

composed of a population of neurons largely separable from those targeting the striatum (Margolis et al., 2006). These dopamine neurons form the mesocortical pathway and innervate the entorhinal, prefrontal and anterior cingulate cortex in particular.

The striatal-projecting dopamine neurons of the VTA/SNc complex form the densest dopaminergic projections. These pathways have been the most widely studied and have been implicated in both movement control and reward processing. For these reasons further discussions of dopamine circuitry, physiology and function will focus on these subpopulations of dopaminergic neurons.

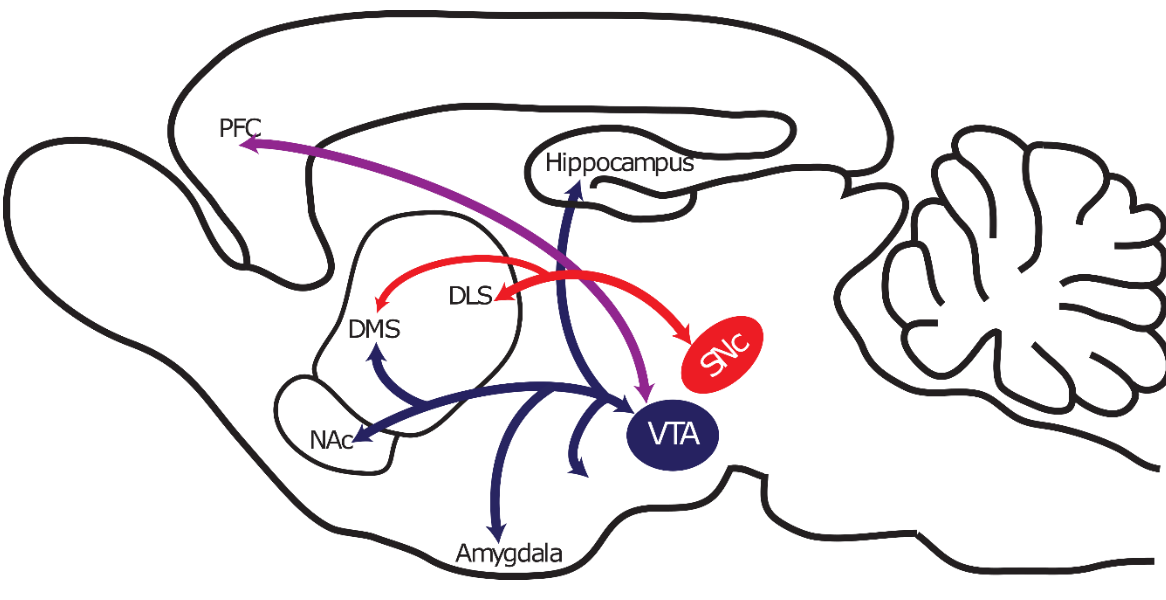


Figure 3: Midbrain dopamine system
Dopamine neurons in the midbrain are located within three nuclei, the substantia nigra pars compacta (SNc), the ventral tegmental area (VTA) and the retrourbal field. Axons of midbrain dopamine neurons form three main pathways, the mesolimbic pathway, the mesocortical pathway and the nigrostriatal pathway. NAc = Nucleus accumbens, PFC = prefrontal cortex, DLS = dorsolateral striatum, DMS = dorsomedial striatum.

The striatum is the main nexus of input to the basal ganglia, an interconnected network of nuclei thought to contribute to many functions including motor learning and movement control. The striatum itself is delineated into ventral and dorsal portions according to cell distribution and patterns of anatomical connectivity (Voorn et al., 2004) and receives distributed input across areas of motor, sensory and associative cortex as well as the thalamus (Bolam et al., 2000; Doig et al., 2010; Smith et al., 2004)

The majority of neurons within the striatum are spiny projection neurons (SPNs), long range GABAergic projection neurons (Gerfen & Surmeier, 2011). Dopamine release within the striatum can modulate excitability of SPNs to either facilitate or suppress activity via action at dopamine receptors. Dopamine receptors are a group of G-protein coupled receptors and can be divided into two subfamilies: D1-like receptors (D1R, D5R) and D2-like receptors (D2R, D3R, D4R). D1-like receptors stimulate the G-proteins $G\alpha_{s/olf}$, which boosts production of cyclic AMP (cAMP) and activates protein kinase A (PKA) (Beaulieu & Gainetdinov, 2011) resulting in a series of events that heighten the receptivity of the neuron to glutamatergic input. D2-like receptors are instead coupled to $G\alpha_{i/o}$ proteins, which inhibit cAMP and limit activation of PKA (Beaulieu & Gainetdinov, 2011), making the intracellular signalling cascades induced by the two sub-families of dopamine receptors diametrically opposed.

Midbrain dopamine neurons are usually described as having two modes of activity, which are posited to contribute to different functions and have distinct effects upon D1-like and D2-like receptors (Grace, 1991; Grace et al., 2007). Dopamine neurons fire at a steady 'pacemaker' rate of 2-10Hz, providing a baseline supply of dopamine to afferent structures. This mode of firing has been termed tonic, and tonic levels of dopamine have been

suggested to play a role in in calculations of reward rate and motivational state (Niv et al., 2007; Ostlund et al., 2011) (though note there are debates about the relative contribution of tonic dopamine firing to tonic dopamine levels in terminal regions in the striatum: e.g ., Owesson-White et al., 2012). Ligand sensitivity differs between the dopamine receptor subfamily, with D2-like receptors having greater affinity than D1-like receptors. For this reason, it has been classically thought that the D2-like and D1-like families are better positioned to track smaller and greater changes in dopamine concentration respectively, though recent studies have suggested this may not be a strict functional division (Yapo et al., 2017).

In addition to changes in baseline tonic firing rate, dopamine cells exhibit short periods of elevated firing. These phasic ‘bursts’ induce rapid increases in dopamine concentration in downstream regions such as the NAc (Sombers et al., 2009) and are thought to be linked to behaviourally relevant events including salient cues and actions (Bromberg-Martin et al., 2010; Coddington & Dudman, 2019; Schultz, 2019b; Watabe-Uchida & Uchida, 2018). The lower affinity D1R are less likely to be occupied at baseline dopamine concentrations and are therefore more able to track these swift increases in available dopamine.

Dopamine receptor expression is organized across striatal SPNs. D1R-expressing SPNs are classically associated with the “direct” pathway, which, from dorsal striatum, has a monosynaptic connection to the output nuclei of the basal ganglia, the substantia nigra pars reticula, globus pallidus internal and VTA. D2R-expressing SPNs form the “indirect” pathway from the dorsal striatum (Gerfen et al., 1990; Surmeier et al., 1996), which reach these output nuclei indirectly by first projecting to the globus pallidus external and subthalamic nuclei (Soares-Cunha, Coimbra, Sousa, et al., 2016). In the dorsal striatum

activation of the direct pathway has been proposed to promote movement, while the indirect pathway promotes inhibition (Albin et al., 1989; DeLong, 1990; Gerfen & Surmeier, 2011; Kravitz et al., 2010). Differential dopamine receptor expression between the two pathways would allow the level of dopamine release to bias activation towards one pathway over the other. However, recent work has challenged both the specified circuitry and functional opponency laid out by the canonical direct-indirect pathway model (Cazorla et al., 2014; Cui et al., 2013; A. Saunders et al., 2015; Soares-Cunha, Coimbra, David-Pereira, et al., 2016; Vicente et al., 2016). In addition, this division maps less neatly to the ventral striatum, where the indirect pathway projecting to the ventral palladium is composed of both D1R and D2R SPNs (Kupchik et al., 2015).

Dopamine receptors are also present on cortical afferents to the striatum where they can regulate release (Bamford et al., 2004, 2018; Wang & Pickel, 2002) as well as being expressed on cholinergic and GABAergic interneurons (Ibáñez-Sandoval et al., 2010; Yan et al., 1997). Thus, fluctuating dopamine concentrations within the striatum can shape the responsiveness of SPNs via multiple routes. In addition to mediating moment by moment receptiveness of striatal neurons to excitatory drive, dopamine has the capacity to aid the formation of long-term plastic changes at cortico-striatal synapses (Bamford et al., 2018; Reynolds et al., 2001; Yagishita et al., 2014).

In summary, dopaminergic inputs are one of many that converge in the striatum, which receives additional input from a wide range of afferent structures. Differential receptor expression within the striatum makes dopamine well placed to modulate the sensitivity of subpopulations of striatal SPNs to incoming information. Additionally, the capacity of dopaminergic transmission to both facilitate plasticity and change momentary excitability

enable changes in both long term processes such as learning and short term online behaviour. The potential for this neurotransmitter system to evoke an array of effects in circuits involved in many different processes has led to the development of multiple theories of dopamine's function.

1.3 - Theories of mesolimbic dopamine function

Initial investigations linking dopamine and reward

The first associations between the midbrain dopamine system and reward came from intracranial self-stimulation (ICSS) studies, where rodents would perform an instrumental action such as lever pressing in order to obtain electrical brain stimulation (Olds & Milner, 1954). This behaviour was distinguished from the instrumental pursuit of natural reinforcers such as food by the fact that animals did not satiate, and was the first indication of the existence of 'reward circuitry'.

The reinforcing effects of stimulation are by no means equally distributed across the brain, and an early hotspot for reinforcement was found be just above the medial forebrain bundle, a collection of axons passing from the midbrain to the forebrain that include dopaminergic projections arising from the VTA and SNc (Olds & Olds, 1958). Further explorations found ICSS to also be effective when targeting sites in the midbrain that held dopaminergic nuclei (Corbett & Wise, 1980). This finding arose in conjunction with studies demonstrating that if animals were pre-treated with dopamine receptor antagonists, they would attenuate lever pressing or reward pursuit on a linear track over the course of a

session (Fouriezos et al., 1978; Fouriezos & Wise, 1976; Wise, 1978). The decrease in responding following dopamine receptor blockade matched how animals would behave if the rewarding effects of a lever press were suddenly decreased or removed entirely, indicating that dopamine transmission was necessary to maintain responding. Subsequent studies confirmed that blocking dopamine receptors had similar effects across reward modalities, such as in animals pressing for food rewards (Wise & Schwartz, 1981) or drugs of abuse such as amphetamine (Yokel & Wise, 1976).

Together these results lead to the initial proposal of an anhedonia hypothesis of dopaminergic function, suggesting that blocking or depressing dopamine transmission prevented 'hedonic' or pleasurable aspect of rewards (Wise, 1982, 2008), and blunted the ability of rewarding experiences to reinforce actions. However, subsequent studies demonstrated that dopamine depletion with 6-OHDA does not impair pleasurable reactions to food consumption (Berridge et al., 1989), nor does it block reward-motivated behaviour when the effort required to obtain the reward is minimal (Salamone et al., 1997). In addition, electrophysiological recordings demonstrated that in well trained animals that have learnt an association between a cue and consequent reward, dopamine neurons will respond to the cue *and not* to delivery of the predicted reward (Mirenowicz & Schultz, 1994; Schultz et al., 1997). Through these experiments it became understood that dopamine could not be mediating the subjective positive experience of reward consumption, but rather contributed in a different way to process of reward-driven reinforcement. Despite the average lay person still attributing to dopamine the status of the 'feel good' neurochemical, critical reappraisal of the anhedonia hypothesis has led the

research community to conclude that dopamine's function in shaping behaviour must lay elsewhere (Wise, 2008).

Dopamine correlates with a reward prediction error signal.

The rise of one prominent theory relating the activity of dopamine neurons to reward, and particularly reward learning, was sparked by series of seminal electrophysiological studies in the 1980s and 90s from Wolfram Schultz's laboratory. This group performed a number of experiments recording firing rates of putative midbrain dopamine cells in behaving nonhuman primates, for example in cued reaching tasks (Schultz & Romo, 1990a). Phasic bursts recorded in these tasks in response to cue were *initially* posited to relate the capacity of cues to promote behavioural reactions (Romo & Schultz, 1990; Schultz, 1986) or motivational states (Schultz & Romo, 1990a).

However, reappraisal of these phasic dopamine bursts in the 1990s found that the activity of dopamine neurons correlated with a reward prediction error (RPE), a key teaching signal in reinforcement learning models (Montague et al., 1996; Schultz et al., 1997). This meant that in early training, when rewards were still relatively unpredictable, the activity of dopamine neurons increased when a reward was presented. As the reward became reliably preceded by a cue, firing would instead propagate back to the time of cue presentation. Should a cue-predicted reward be omitted, then dopamine neurons would exhibit a pause in firing at the time of expected reward. This pattern of activity meant that dopamine neurons tracked the difference between actual and rewarding outcomes within the task.

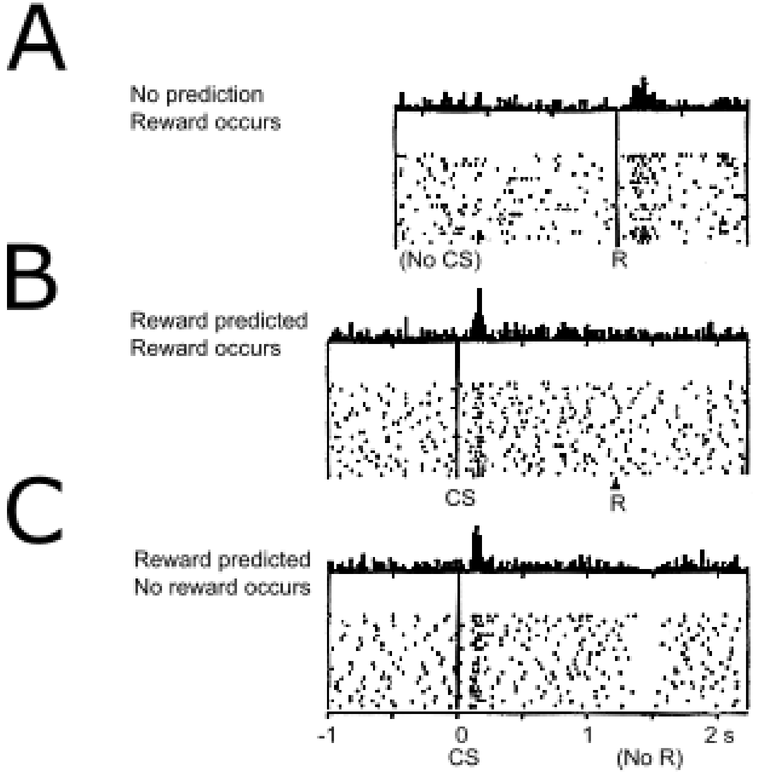


Figure 4. Dopamine neuron firing correlates with a reward prediction error.
Electrophysiological recordings from putative dopamine neurons in the midbrain of non-human primates after A) presentation of unexpected reward B) a conditioned stimulus (CS) that has been learnt to predict upcoming reward C) reward omission. Adapted from Schultz et all (1997)

The conceptualization of RPE theory has been hugely influential. Numerous studies have followed seeking to further define the properties of phasic dopamine signalling in a RPE context. Dopamine neurons have been shown to encode the relative value of rewards and cues across dimensions including reward size (Roesch et al., 2007), likelihood (Tobler et al., 2005), and subjective preference (Lak et al., 2014). Reward predictive changes in dopamine concentration have also been recorded in sites downstream of midbrain dopamine neurons such as the NAc (Day et al., 2007; Gan et al., 2010; Hart et al., 2014; Papageorgiou et al., 2016), though it remains an open and fiercely debated issue how well dopamine release across terminal regions reports an RPE-like signal (Blanco-Pozo et al., 2023; H. Kim

et al., 2020; Mohebi et al., 2019; Tsutsui-Kimura et al., 2020; van Elzelingen et al., 2022). While earlier studies relied on basal firing rates and electrophysiological properties to identify neurons as dopaminergic, more recent studies in rodents have utilised optogenetic manipulations to confirm the dopaminergic identity of RPE encoding neurones (Cohen et al., 2012; Eshel et al., 2016).

Reward prediction error theory positions dopamine as a teaching signal that uses information about the present to shape future behaviour. This causal role for dopamine transmission ties in mechanistically with the ability of dopamine to modulate synaptic plasticity and generate long term changes in connectivity at the cortico-striatal synapse (Reynolds et al., 2001; Yagishita et al., 2014).

The advent and more widespread utilisation of optogenetics has also enabled the link between phasic dopamine responses and reward learning to be moved from correlative to causal. ‘Blocking’ describes the phenomenon by which learning an association between a cue and reward is prevented (or blocked) if the cue is presented concurrently with another stimulus that has already been associated with reward. This effect is theorised to occur due the lack of an RPE at the time of the fully predicted reward (Kamin, 1968; Rescorla & Wagner, 1972; Waelti et al., 2001). Steinberg and colleagues found that transient optical stimulation of dopamine neurons at the time of reward during compound conditioning, in essence artificially inducing an RPE even when reward was fully predicted, was sufficient to ‘unblock’ learning about the previously blocked cue-reward pairing (Steinberg et al., 2013). The authors of this study also showed that by stimulating at the time of reward omissions, when dopamine neurones should be pausing to encode a negative RPE (Schultz, 1998), they could attenuate the reduction of responding in an extinction paradigm.

The results of this experiment suggest stimulating dopamine neurons was sufficient to causally drive learning, providing evidence that the phasic dopamine signals recorded in numerous previous studies could be being used to update animals' future behaviour. Though later studies have produced complementary and aligned results (Chang et al., 2015; Keiflin et al., 2019), it is worth noting that when activating dopamine neurons in these associative learning paradigms, stimulation is often applied for 1 second or more. Given this time period exceeds the typical window of endogenous dopamine burst firing, it raises the question of what effect more physiologically-aligned patterns of stimulation may have had in these behavioural settings.

Dopamine motivates reward pursuit behaviour.

RPE theory is based on the often tight correlation between phasic dopamine neuron activity and a prediction error signal. However, evidence suggests information beyond reward expectation may be encoded in dopamine neuron firing rates. Satoh et al recorded from midbrain dopamine neurons in a behavioural task where cues predicted upcoming reward. In line with RPE theory, firing rates from putative dopamine neurons scaled with the volume and likelihood of reward predicted by different cues (Satoh et al., 2003). However, when comparing the evoked activity of cues that predict identical rewarding outcomes, and therefore should have identical reward expectations, differences in firing rates were observed. These discrepancies in activity were found to correlate with the speed at which animals initiated movement following cue presentation. Correlations with the vigour or speed of movements have also been recorded in downstream dopamine release (Collins, Greenfield, et al., 2016; Ko & Wanat, 2016) and dopamine-dependent

activity in striatal SPNs (Panigrahi et al., 2015). These studies present a case for dopamine not only representing sensory information relating to reward, but also reflecting a dimension of the motivation with which animals move to obtain those rewards.

Ideas around dopamine acting to invigorate behaviour employed to obtain reward are tied together in theories of behavioural activation (Robbins & Everitt, 2007) and incentive salience (Berridge, 2007). These frameworks position dopamine as mediating the motivational effect that reward associated cues have on pursuit behaviour, and are largely evidenced by pharmacological manipulations that chronically disrupt dopaminergic signalling, particularly in the NAc. Presentation of a cue previously associated with reward can invigorate instrumental responding, a process called Pavlovian instrumental transfer (Colwill & Rescorla, 1988; Lovibond, 1983). Dopamine receptor blockade within the NAc reduces this effect, suggesting that intact NAc dopamine transmission is required for reward-paired cues to have a motivating effect on behaviour (Lex & Hauber, 2008). NAc dopamine depletions or receptor blockade can also bias behaviours away from exerting effort to obtain rewards, causing rodents to preferentially consume freely available lab chow rather than press levers for more palatable sucrose pellets (Koch et al., 2000; Salamone et al., 2001, 2007).

In accordance with these theories, release of dopamine in the NAc core has been shown to 'ramp' upward as animals approach reward, either through physical space or through 'state space' as animals proceed through an action sequence (Hamid et al., 2016; Howe et al., 2013; Syed et al., 2016; Wassum et al., 2012). This ramping signal provides a continuous estimate of proximity to reward and has been theorised to maintain motivational drive during prolonged reward pursuit behaviours.

Strong evidence exists for both dopamine encoding an RPE teaching signal and motivating behaviour. It may be possible to reconcile these two theories of dopamine function computationally in certain situations (McClure et al., 2003). However, an alternative way dopamine may seemingly play two different roles come from a series of experiments recording and manipulating dopamine in rats engaged in a two-armed bandit task. Mohebi et al found a dissociation between electrophysiological recordings of VTA cell firing and fibre photometry recordings downstream in the NAc. VTA neurons responded to the presentation of cues and rewards in a manner consistent with RPE (Mohebi et al., 2019). NAc dopamine release also contained cue responses but demonstrated ramping activity upon reward approach and modulation according to background reward rate. Both of these features are consistent with a value signal that could be used to motivate ongoing behaviour, and both were absent from VTA cell recordings.

This study presented the possibility for both value and RPE to be represented across the mesolimbic dopamine system and, more generally, fits in with an abundance of evidence that release at dopamine terminals can be locally regulated (Cachope & Cheer, 2014; Sulzer et al., 2016; Threlfell et al., 2012). However, other studies combining electrophysiological recordings of VTA cell bodies with photometry of dopaminergic axon calcium activity and release in the NAc have instead argued for homogenous RPE encoding in VTA and NAc (Kim et al., 2020). Further experiments will be required to fully describe situations in which terminal release can disassociate from VTA cell body firing.

In addition to demonstrating that what dopamine might be encoding may differ depending on *where* one records, a study using the same task featured in Mohebi et al also demonstrated a difference in the causal influence stimulation had on behaviour depending

on *when* that simulation was applied. Optogenetic activation of the VTA at the time of choice made it more likely for that choice to be repeated, as if stimulation were acting as teaching signal in order to reinforce the selected action (Hamid et al., 2016). However, stimulating concurrently with the cue that signalled a trial was available resulted in faster trial initiation, as if mediating increased motivation to work for reward - but did not change choice behaviour on the following or subsequent trial. This study presents the possibility that a phasic dopamine signal may be read out differently in downstream regions depending on the behavioural state of the animal, an attractive prospect to anyone seeking to reconcile the diverse functional roles that dopamine seems to play.

Dopamine in motor control

A considerable body of work relates dopamine transmission to the capacity for rewards to guide behaviour. Yet from a neurological perspective, dopamine is associated strongly with movement, as Parkinson's disease is classically thought to affect SNc dopamine neurons and produces a plethora of motor symptoms (Jankovic, 2008). While movement related activity has been recorded since early electrophysiological studies (Schultz et al., 1983), the relationship between dopamine activity and motor behaviour has appeared weaker and more inconsistent when compared to reward related activity (Schultz, 2019b). Indeed, in lateral parts of the VTA, the activity of nearly every identified dopamine cell was reported to conform to an RPE (Eshel et al., 2016), making prediction error seem like an attractive unifying theory for dopamine function.

However, a series of recent studies recording from dopamine neurons have observed changes in activity around the time of action initiation, even in environments devoid of explicit rewards (Da Silva et al., 2018; Dodson et al., 2016). Dodson et al employed juxtacellular recordings, which allow the precise recovery of the anatomical location and molecular identity of target neurons, in head fixed mice. They found significant modulations in activity around the timepoint that postural adjustments were initiated. Despite differences in the patterns of movement related activity and the proportion of responsive cells, changes in firing rates were observed in both the SNc and VTA.

Movement signalling in the absence of explicit reward has also been observed in freely moving mice. In animals exploring an open field, SNc dopamine neurons displayed phasic changes before movement initiation, with the majority of recorded neurons showing an increase in activity (Da Silva et al., 2018). This modulation of dopamine neurons was not specific to the type of movement being initiated, although in a subpopulation of recorded neurons the magnitude of the increase in activity did scale with the vigour of initiated movement. Utilising optogenetics, this experiment also demonstrated that phasic activation of dopamine neurons in the absence of rewards can induce movement, though other studies have only replicated this effect with supraphysiological levels of stimulation (Coddington & Dudman, 2018). Regardless, when taken together results from these experiments make it difficult for any framework that deals solely with reward to explain dopaminergic signalling.

Evidence of movement-related activity in the absence of reward brings forth the question of how these two relatively distinct functional roles that dopamine is thought to play can be reconciled. One study recording in the VTA found that dopamine neurons universally

encoded a reward response, but also could be clustered into specialized groups according to how they encoded a selection of secondary variables, which include movement-related kinematics (Engelhard et al., 2019). This study presented an attractive possible avenue by which movement activity could be accounted for in a way that enabled common reward-based algorithm of dopamine function to be maintained (Schultz, 2019b).

However other studies have presented opposing evidence to the proposal of dopamine neurons multiplexing movement and reward. Studies in both freely moving and head-fixed contexts have shown that dopamine neurons responding to reward and those responding to movement belonged to largely separable populations (Da Silva et al., 2018; Howe & Dombeck, 2016). In the case of head-fixed example, this separation is correlated with anatomical location, with VTA neurons more strongly responding to reward while the activity of those in the SNc was dominated by movement signals. Though assigning dopamine neurons a role in either movement or reward along a clean anatomical line is an attractively simple prospect, several studies have shown spontaneous movement related modulations in VTA neurons activity to be present (Coddington & Dudman, 2018; Dodson et al., 2016). When combined with the fact that foundational RPE recordings took place in primate SNc (Romo & Schultz, 1990; Schultz & Romo, 1990b), a purely anatomical explanation for the seemingly multiple functions of dopamine neurons seems unlikely to capture the full picture.

1.4 - Mesolimbic dopamine and the initiation of reward guided action

While VTA neurons can be modulated by spontaneous movement, though less prominently and consistently than those in the SNc, studies manipulating mesolimbic dopamine transmission associate VTA dopamine particularly with movements or actions that aim to obtain reward. Nicola and colleagues trained rats in a discriminative stimuli task, where one cue (the DS) indicated that an operant response could be made for reward, while another cue (the NS) was unrewarded. This paradigm provides a situation where reward seeking action should be initiated dependent on the identity of the cue presented.

Nicola et al found that cue-evoked activity in the NAc core discriminated between reward opportunity and likelihood of action. NAc neurons displayed a stronger increase in firing rate for DS presentation over NS, and for cues that were followed by reward seeking actions (Nicola et al., 2004). This cue-evoked action initiation was found to depend on dopaminergic input, with both cue-evoked changes in firing rate and the speed to initiate reward seeking decreasing following VTA inactivation (Yun, Wakabayashi, et al., 2004) or NAc D1R blockade (Yun, Nicola, et al., 2004).

In a more complex version of the task, responding following a third cue (PS) was only probabilistically rewarded (Nicola et al., 2005). Animals adjusted their behaviour accordingly, nearly always seeking reward following a DS, occasionally following a PS, and rarely following a NS. Boosting NAc dopamine increased the probability of reward pursuit after PS, but not NS presentation. These results link increased mesolimbic dopamine transmission specifically with an increase in *reward-seeking* actions, and not an indiscriminate increase in movement or action more generally.

While the DS task provides an environment where cues guide reward seeking, in other purely instrumental tasks operant responding for reward is self-paced. One such example is the progressive ratio task traditionally employed to assess motivation, where the requisite number of lever presses to obtain reward increases within a session. Chemogenetically boosting mesolimbic dopamine transmission within this task increases the total number of lever presses animals will make to obtain reward (Boekhoudt et al., 2018). This effect is not mediated by increasing the number of presses animals will make in a bout, but rather the number of bouts that animals initiate within a session. Crucially these effects were specific to activating VTA over SNc dopamine neurons, highlighting the role for the mesolimbic projection over the nigrostriatal pathway in management of reward directed action. Taken together, these results indicate that elevated VTA activity increases the likelihood that reward seeking actions will be initiated. However due to chronic nature of the manipulations utilised in these experiments, it not possible to determine the dependence of effects mediated by the mesolimbic dopamine transmission on the timing of external cues or internal states.

Though elevating mesolimbic dopamine transmission can promote reward directed action, many studies have recorded increases in VTA cell activity within tasks such as Pavlovian conditioning, where animals are not required to make an action in order to obtain reward. These experiments frequently occur in head fixed preparations, where the assumption is that movement is restricted, and therefore motor responses outside of reward collection are often not recorded. However restricting movement of an animal does not necessarily equate to no movement being attempted. Hughes et al recorded VTA dopamine neurons whilst animals received rewards on a fixed time schedule, and simultaneously recorded the

directional force produced by the mice (Hughes et al., 2020). In this study the activity of dopamine neurons was tightly correlated to a seemingly prominently motor variable, the impulse vector, or the integral of force of time. The authors of the study presented a striking proposal that RPE encoding across the body of work already performed could instead reflect patterns of force produced by animals in response to cues and rewards.

To investigate causality, the same study applied optogenetic manipulations of VTA neurons. In trained animals' stimulation produced patterns of force production consistent with how directional force was encoded across the population. However, the production of movement was dependent on reward expectation; stimulation prior to training when no rewards had been received in the context did not result in the same pattern of force. In addition, activation is often accompanied by anticipatory licking behaviour indicating some association with reward experience or memory. This aligns with previous findings that some degree of contextual reward experience is necessary for VTA activation to drive motor behaviour (Coddington & Dudman, 2018), and positions mesolimbic dopamine as operating at the intersection of reward and movement, rather than in a purely motor space.

What the previously discussed study highlights well is the importance of measuring and understanding the movements produced by animals in rewarded tasks (see Jenkins & Walton, 2020). Separating out movement and reward in a non-trivial exercise; discerning when dopamine neurons may be encoding reward information and when they may be encoding movement is complicated by their entwined nature. Animals often make actions in order to improve the state they are in, and the presentation of rewards often prompts actions to obtain them. Even in Pavlovian conditioning paradigms where no action must be

made to obtain reward, anticipatory movements such as movements towards and interacting with a lever cue ('sign tracking': Flagel et al., 2010) or licking or lipping (in primates) are observed and often used as a measure of reward anticipation to index learning. These anticipatory movements are executed with a short latency following cue presentation in well trained animals, complicating the interpretation of any dopaminergic signals occurring at the same time. This is highlighted in cued instrumental tasks where, due to slower responding on a subset of trials, the timepoint of cue and action become disassociated. For example, in a study where a cue signalled that a lever could be pressed for reward, animals often reacted swiftly to cue presentation, meaning NAc dopamine could not be independently related to either cue or action. However, on slow latency trials, cue and action signals became dissociable, and peak dopamine release was shown to correlate best with the time of lever press, not cue presentation (Roitman et al., 2004).

Other studies have designed tasks to dissect the influence of action initiation and reward expectation more systematically. In a symmetrically rewarded Go/No-Go task, Syed et al recorded NAc dopamine while animals were instructed to either initiate (Go) or restrain (No-Go) action depending on the identity of the cue presented. While phasic signals did carry information relating to RPE, increases in dopamine were also tightly aligned to the initiation of a correct reward-seeking action. NAc release did not elevate when animals initiated an incorrect action on Go trials and was instead *suppressed* when action instead had to be withheld for reward following No-Go cues. This study indicated that that dopamine release in the NAc of well-trained animals may promote the selection of reward seeking actions only when it is correct to do so. These results dovetail with findings in a

human Go/No-Go task that show striatal BOLD responses are higher for Go trials relative to No-Go trials, irrespective of trial valence (Guitart-Masip et al., 2011, 2014).

In line with evidence associating elevations in mesolimbic dopamine with the *initiation* of reward seeking actions, manipulations that affect dopamine transmission have also been associated with the impairment the ability of animals to *inhibit* action when it would be premature or inappropriate. However, the influence of these manipulations has been shown to depend on the nature of the behavioural inhibition. The taxonomy of inhibitory control as proposed by Bari and Robbins (2013) distinguishes between the restraint, postponement and cancellation of action as separable forms of response inhibition, which are differentially influenced by mesolimbic dopamine.

For example, the process of *action cancellation* – or the capacity of animals to cancel an ongoing action, appears independent of NAc dopamine receptor activation instead being influenced by manipulations of the dorsal striatum (Eagle et al., 2011). Conversely *action postponement*, or the ability to withhold from acting while waiting, is influenced by mesolimbic dopamine transmission. The 5-choice serial reaction time task (5CSRTT) is mostly commonly used to assess action postponement, where animals must wait animals must wait for a light to appear briefly in one of 5 noseports. Nose poking in the illuminated port results in reward delivery, but responses made before the light is displayed are counted as premature. Both injections of amphetamine, which induces terminal dopamine release and blocks re-uptake amongst other effects, and D1R agonists into the NAc increase premature responding(Pattij, Janssen, Vanderschuren, et al., 2007; Pezze et al., 2007; Van Gaalen et al., 2006).

Whilst pharmacological investigation relates NAc dopamine receptor activity to successful action postponement, investigation utilising other methodologies have produced mixed results. Both chemogenetic and optogenetic activation of the VTA do not produce impulsive responding in the 5CSRTT(Boekhoudt et al., 2016; Flores-Dourojeanni et al., 2021). One possible explanation for these negative results is that opposing pro-restraint and pro-impulsivity effects mediated by increased dopaminergic tone in different sites downstream of the VTA result in a 'zero-sum' effect, a similar lack of impact follows systemic D1R agonist administration (Winstanley et al., 2010). In line with this thinking, projection specific activation of VTA-NAc shell, and not VTA-NAc core dopamine neurones did increase premature responding (Flores-Dourojeanni et al., 2021). However, this effect is conditional on stimulation being delivered at the end and not the start of the period of waiting. It is important to note that to date no recordings of mesolimbic dopamine release have been performed in the 5CSRTT, so it is difficult to relate effects of chronic pharmacological manipulation, chemogenetic potentiation of phasic activation and optogenetic induction of phasic transients to *endogenous* DA activity within the task.

Finally, *action restraint* is defined as the inhibition of planned responding on a subset of trials and is usually assessed in Go/No-Go paradigms. As previously mentioned, it is indicated that NAc dopamine release is suppressed when a cue instructs that action should be withheld(Syed et al., 2016). Activating D1Rs impairs cued action restraint, both when applied systemically and infused into the NAc core, as does intra-NAc core infusions of d-amphetamine to potentiate dopamine levels (Grima et al., 2022; Härmson et al., 2022). However, unlike the studies examining action postponement in 5CSRTT, experiments thus

far are limited to tonic stimulation of receptors, and the effect of potentiating or artificially inducing phasic transients upon action restraint remain unknown.

1.5 - Regulation of action and regulation of dopamine

A wealth of evidence suggests that dopamine dependent activity in the NAc contributes to the initiation of reward seeking action, and the impairment of restraint when behaviour must be inhibited. These studies align well with the classical proposal from Mogenson et al (1980) situating the NAc as an interface between limbic and motor systems. Mogensons proposal was based upon two strands of evidence. Firstly, the involvement of VTA afferents arising from limbic areas such as the lateral hypothalamus in behavioural responses such as attack. Secondly that locomotion mediated by Nac dopamine injections could be attenuated by blocking NAc output to the globus pallidus. Since Mogensons original paper, further theories have built upon the capacity for NAc circuits to modulate behavioural output, positioning the NAc as mediating action selection(Khamassi & Humphries, 2012), and shaping reward or goal directed behaviour (Mannella et al., 2013).

Of note is that in Mogensons proposal the NAc is not designated as a motor area itself, but rather as a relay between limbic and motor systems. In line with this theory inhibiting NAc activity does not prevent movement, and the NAc is rather thought to bias the direction or intensity of behaviour (Floresco, 2015) - via projecting to regions that go on to shape motor output. One proposal for how NAc activity may influence behaviour in this way is via altering activity in PFC regions. NAc neurones project either directly, or indirectly via the ventral pallidum, to the SNr (Humphries & Prescott, 2010). Increased activity of NAc

neurons therefore has the capacity to bidirectionally modulate SNr firing, which in turn can either inhibit or disinhibit portions of the thalamus that input to regions of the PFC, including the PL, IL, and AIC(Voorn et al., 2004). Mannella et al suggest that prefrontal circuits hold representations of situationally relevant outcomes, and via disinhibition of thalamic subregions the NAc can bias selection between possible outcomes or goals (Mannella et al., 2013).The PFC would then engage the relevant actions associated with the selected outcomes via a mutlisynaptic projection to premotor cortex, via supplementary motor cortex (Nachev et al., 2008)or parietal cortex(Fogassi et al., 2005). Of note, this proposed mechanism could explain why the effects of some transient manipulations of NAc dopamine depend on a particular context or timepoint within the task – as in order for NAc activity to alter behaviour, an outcome representation would have to be engaged in PFC circuits.

Within this model, mesolimbic dopamine would be able to shift the balance in activity between populations of NAc MSNs that express D1 or D2R, thereby bias the selection of actions to achieve a particular outcome. For example, phasic increases in NAc dopamine would facilitate excitatory drive to D1 expressing MSNs – which may promote the activation of a particular goal representation in PFC circuits. An alternative path by which NAc activity may arrive at altering motor output involves subcortical projections. Striatal circuits are not independent from one and other, but rather have a spiralling organisation whereby more ventral regions able to influence dopamine cells that input to more dorsal sections (Haber et al., 2000). As such the NAc is able to alter dopamine release, and therefore the balance of D1R to D2R expressing MSNs activity, in dorsal regions such as DLS that can influence sensorimotor regions of cortex(Yin & Knowlton, 2006). Both of these

pathway's position NAc as generally driving or invigorating behaviour – aligning with studies that find NAc activity increases before onset of reward pursuit but does not correlate with specific movement kinematics (McGinty et al., 2013).

Evidence associates elevations of mesolimbic dopamine transmission with an increase in reward seeking action, even when it advantageous to exercise restraint. Modulating activity in VTA dopamine neurons offers a route by which external structures could bias behaviour towards or away from the initiation of reward seeking actions. The mesolimbic dopamine pathway receives functional input from a range of limbic and cortical regions, which may seek to influence dopamine release. It has been demonstrated that the hippocampus (Legault et al., 2000; Luo et al., 2007) and prefrontal cortex (Jackson et al., 2001; Overton et al., 1996; Tong et al., 1996) can modulate VTA activity and alter downstream NAc dopamine fluctuations. Other areas such as the amygdala send glutamatergic projections to the NAc and can mediate local dopamine release (Floresco et al., 1998; Jones et al., 2010), demonstrating another route by which mesolimbic dopamine output may be shaped.

In addition to long range projections altering VTA activity or NAc release, there are local circuits and systems that act to sculpt mesolimbic dopamine under endogenous conditions. These sources of local regulation could offer an alternate method by which dopamine release, and the associated behaviour, could be shaped, as well as providing possible therapeutic targets in disorders of impulsivity or amotivation. One such example of local circuitry are striatal cholinergic interneurons, which comprise a mere 2% of the cells in the striatum (Ookschot, 1996), but make strong and numerous synaptic connections with dopamine terminals (Zhou et al., 2003). Activity of cholinergic interneurons can regulate

striatal release of dopamine without activation of the midbrain cell bodies (Threlfell et al., 2012), and activity in within this network is thought contribute to regulating cue-evoked behaviour (Collins, Aitken, et al., 2016).

In addition to the interneuron-mediated circuits within the striatum, local regulatory systems also shape dopamine cell activity at the level of the soma in the VTA. Endocannabinoids are synthesised and released from VTA neurons following periods of sustained activity (Oleson et al., 2021; Peters et al., 2020). Inputs to the VTA that express receptors for cannabinoids are supressed following their release, allowing selective modulation of afferent drive when dopamine neurons are highly active. This regulation is thought to be important for shaping the dopamine release patterns that guide reward seeking behaviour(Covey et al., 2018; Oleson et al., 2012), and can result in long term plastic changes at terminals that synapse onto VTA neurons (Pan et al., 2008).

Though evidence points to a dopaminergic contribution to reward guided action, appropriate action control is a complex process that likely relies on a wide network of systems and neural structures, some which interact with mesolimbic dopamine and others which act independently. The serotonin system is an established contributor to action control, particularly to components of waiting or postponing action (Eagle et al., 2008). Depletions of serotonin can increase premature responding in the 5CSRTT, and also impair performance when animals must restrain action on a subset of trials in a go/no-go task (Harrison et al., 1997, 1999). Though serotonin can influence striatal dopamine release(Alex & Pehek, 2007), manipulations of serotonin receptors in a symmetric Go/No-Go produces a distinct set of impairments to those following dopamine manipulations

(Grima et al., 2022; Härmson et al., 2022) – suggesting the two neuromodulator systems influence action control via different mechanisms.

In addition to neuromodulator systems there are a number of cortical and subcortical structures that are thought to play a role in regulation of appropriate action. Amongst them are the ventral hippocampus and medial prefrontal cortex, whose functional connectivity is required for impulse control within the 5CSRTT (Chudasama et al., 2012). Disrupting medial prefrontal regions can also affect appropriate responding on a range of other behaviour tasks (Hardung et al., 2017a; Kamigaki & Dan, 2017; Narayanan et al., 2006; Risterucci et al., 2003), and neurons in these regions display changes in activity that reflect the execution of actions (de Haan et al., 2018; Halladay & Blair, 2015; Horst & Laubach, 2013), indicating these regions may coordinate and play an important role in general action control.

A thorough investigation of every component in the circuits that regulate action control, both in concert with and outside of the mesolimbic dopamine system, is outside the remit of this thesis. Further discussion will focus on two complementary systems: the medial prefrontal cortex and the endocannabinoid system.

1.6 - Medial prefrontal cortex, dopamine and action

The rodent medial frontal cortex (mPFC) is thought to correspond to parts of human posterior agranular medial prefrontal cortices (Roberts, 2020; Wise, 2008), and has been implicated in a range of cognitive and behavioural processes (Dalley et al., 2004; Eichenbaum, 2017; Euston et al., 2012; Vertes, 2006). Of particular interest to this thesis,

however, mPFC has been implicated in functions relevant to reward guided action selection and inhibition. For example, mPFC function has been implicated in restraining impulsive behaviours (Bari & Robbins, 2013; Kim & Lee, 2011) and using contextual cues to guide behaviour flexibly (Sharpe et al., 2019; Sharpe & Killcross, 2015), both key components of appropriate reward-guided action. The mPFC not only projects to midbrain dopamine nuclei (Carr & Sesack, 2000; Geisler & Wise, 2008), but is also well placed to alter behaviour via extensive connections with the striatum, thalamus and limbic system (Ferenczi et al., 2016; Gabbott et al., 2005; Sesack et al., 1989). For the sake of brevity, the rest of this section will consider the subsection of the mPFC literature that relates to the control of action.

The mPFC has classically been divided into subregions along a dorso-ventral axis, including of the prelimbic cortex (PL), the infralimbic cortex (IL), the dorsopeduncular cortex (DP) and the medial orbital cortex (MO) (Öngür & Price, 2000), though some suggest that these designations may need re-evaluating (van Heukelum et al., 2020). These delineations are based on differences in both function and anatomical connectivity. Within the mPFC, subregions target generally non-overlapping projection targets (Vertes, 2004). For example, anatomical tracing studies in rats have shown that the IL innervates the shell region of the NAc most densely, while the PL instead preferentially projects to the NAc core (Brog et al., 1993; Öngür & Price, 2000). These differences in connectively between the IL and PL are reflected in the effects of selective inactivation or lesion experiments, where subregion specific interference can result in distinct or even opponent effects on behaviour (Capriles et al., 2003; Capuzzo & Floresco, 2020; Gourley & Taylor, 2016; Hardung et al., 2017a; Sierra-Mercado et al., 2010).

Appropriately selecting whether to act or whether to exercise restraint involves matching behaviour to the situation at hand. This process involves using external information to recognise the state of the world, so that the most advantageous strategy can be executed. Normal function in prefrontal circuits has been shown to differentially affect the ability to use contextual cues to shape behaviour. In an experiment by Sharpe et al, rats were presented with two stimuli in two distinct contexts (Sharpe & Killcross, 2015). In the first context, cue A was paired with a footshock while cue B was not. In the second context these contingencies were reversed – such that cue B now predicted imminent danger. Presentation of stimuli paired with footshock elicits defensive freezing behaviour in rats, and control animals were able to use the contextual information to shape behaviour – i.e., they expressed defensive behaviour only to the shock-predicting cue in each context. Inactivation of PL impaired this process and animals were no longer able to use contextual cues to modulate responding appropriately.

While an intact PL contributes to aligning behaviour to the contextual requirement, the IL has been suggested to have an oppositional affect, instead promoting the generalisation of behaviour. In ABA renewal paradigms, animals first learn an association in context A and then have it extinguished in context B. Returning animals to the acquisition context (A) promotes a transient increase in conditioned behaviour relative to being returned to the extinction context (B). IL lesions make animals more sensitive to the contextual cue, increasing the renewal of condition behaviour both when the behaviour is appetitive (ie approaching a food magazine) and aversive (eg freezing) (Rhodes & Killcross, 2007; Zelikowsky et al., 2013). This suggests that under normal condition the IL acts to make

behaviour *less* sensitive to the demands of the current environment, placing the two mPFC subregions in opposing roles.

In addition to matching appropriate behaviour to different environments, the presentation of sensory information in the form of cues can signal a change of state within the same environment, and therefore a change in which action strategy may be appropriate. As previously discussed, discriminative stimuli tasks present two cued states, one where action results in reward (following the DS) and one where it does not (following the NS). A series of studies have shown that inactivating dorsal regions of the mPFC, including the PL and neighbouring anterior cingulate, reduced both the responses made during a DS and the latency with which animals reacted to DS presentation (Ishikawa et al., 2008a, 2008b). Conversely, inactivations targeted to the ventromedial PFC, consisting largely of the IL, instead increased responding to the NS. These experiments again suggest opposing roles for these two mPFC subregions, with PL activity contributing to the activation of appropriate reward guided action, while the IL appears to be necessary in order to inhibit responses that are inappropriate.

Historical theories of descending control position the prefrontal cortex as biasing activity patterns in lower structures in order to produce the appropriate output for a given task or situation (Miller & Cohen, 2003). This aligns with a role for the mPFC in selecting action appropriate behaviours, and also ties in with studies demonstrating that disrupting mPFC activity can cause an impulsive phenotype. When considering the process of action postponement in the 5CSRTT, early lesion experiments indicated that damage which spared the ventral portion of the mPFC left action restraint intact (Chudasama & Muir, 2001; Passetti et al., 2002a). However later studies using reversible inactivation

demonstrated that the more dorsal PL also contributed to postponing action successfully (Paine et al., 2011; Pezze et al., 2014), suggesting that the null effect of PL lesions may be due to compensatory changes in downstream regions.

While the 5SCRTT provides an environment to test inhibitory control, it is also designed to track other psychological process including attention and impulsivity. Other studies have used more simplified reaction time (RT) tasks to test contributions of the mPFC to action control. In RT tasks animals must make a specific action, such as release a lever, within a specific time window signalled by a cue. Selective lesioning of PL or IL, but not the more dorsal anterior cingulate, has been shown to increase the number of premature responses occurring before cue presentation in a RT task (Risterucci et al., 2003). Hardung et al took a more extensive approach, optogenetically inhibiting the mPFC and orbitofrontal cortex in a subregion specific manner on a subset of trials. In line with previous reports, inhibiting the PL increased early errors (Hardung et al., 2017a). However, inhibition of the IL had the opposing effect *reducing* premature responses. When the more ventral MO was targeted no effects on premature responding were detected, instead inhibition resulted in a slower reaction time to the cue.

Approach avoidance tasks present complex situations where action must be controlled appropriately. At times, action must be restrained to avoid aversive outcomes and at other time it should initiated to obtain maximal reward. Verharen et al utilised a task where during the presentation of an audio-visual cue magazine entry would be punished with a shock, while reward could be collected freely after the terminus of the cue (2019). This experiment found a homogenous role for mPFC subregions in the restraint of action, as

inactivation of either the PL, IL, or MO all resulted in an increase in the number of shocks received via inappropriate action initiation.

While the aforementioned study used a single cue to guide behaviour, Capuzzo et al employed a more complex task where animals had to either make (active) or refrain from making (inhibitory) a specific action in order to avoid aversive footshock, depending on the cue presented (Capuzzo & Floresco, 2020). Here inactivation resulted in differing effects according to the subregion targeted. PL inactivation impaired the ability of animals to make an action to avoid shock, while inactivation of IL disrupted both active and inhibitory avoidance. The authors of this study also tested animals in an appetitive vision of this task, where inactivation of both regions selectively impaired *inhibitory* reward seeking. Results from these studies suggest that involvement of the mPFC subregions in action control differs between situations of pure avoidance, where the PL and IL subserve different functions, and situations featuring appetitive reward, whereas all regions contribute to successful restraint.

mPFC activity contributes to both restraint of premature responding and appropriately shaping behaviour according to external cues, making it well placed to shape action strategies. The varied patterns of connectivity between the mPFC and other structures enable many possible routes by which the mPFC may influence reward-guided action control. The mPFC, particularly PL and MO, projects to the dorsal striatum (Gabbott et al., 2005) and activity in this pathway is important for postponing action in the 5CSRTT (Terra et al., 2020). The mPFC also sends afferents to the NAc that have been shown to play a role in the control of reward seeking behaviour, as inhibition or activation of this projection can bidirectionally regulate reward-seeking (Kim et al., 2017). ~~Note however that in addition to~~

~~striatal connectivity, mPFC is functionally (though not directly anatomically) connected to motor cortex, and inactivation of the mPFC can disrupt activity in this region (Narayanan & Laubach, 2006), highlighting another possible way mPFC circuits could shape motor behaviour.~~

In addition to projections to the striatum, the mPFC is functionally connected to motor cortex - providing an alternative route by which the mPFC could shape motor output and action selection. Spiking in primary motor cortex can be decoded from mPFC activity in reward guided tasks(Wu et al., 2022), and inactivation of the mPFC can disrupt activity in this region (Narayanan & Laubach, 2006). Though the mPFC does not directly project to motor cortex, multisynaptic pathways link the two regions via parietal (Fogassi et al., 2005) and supplementary motor cortex (Nachev et al., 2008).

In addition to these diverse connections with motor regions, the mPFC could act though the mesolimbic dopamine system to index reward guided action. Electrical stimulation of the mPFC can elicit either excitation or inhibition (Lodge, 2011; Overton et al., 1996; Tong et al., 1996) of VTA neurons, the latter presumed via increased drive to GABAergic interneurons (Omelchenko et al., 2009). Electrical stimulation of the mPFC at physiological frequencies can shape output of the mesolimbic pathway, reducing dopamine release in the NAc. Jo et al have demonstrated that this functional connection regulates mesolimbic dopamine in awake behaving animals. Reversible inactivation of the mPFC results in elevated phasic responses to both reward-predictive cues and rewards themselves (Ferenczi et al., 2016; Jo et al., 2013; Jo & Mizumori, 2016). This body of work demonstrates the capacity for mPFC to influence VTA firing and NAc release, and suggests that mPFC

inputs to the mesolimbic dopamine system form part of the regulatory system shaping release and associated behaviour according to environmental demands.

1.7 - Endocannabinoids, Dopamine and action

Physiology of the endocannabinoid system

The endocannabinoid system refers to a group of G-Protein coupled receptors that can modulate neuronal function, their ligands (endocannabinoids) and the set of enzymes that synthesise and degrade these ligands. There are two endocannabinoid receptors, both associated with $G_{i/o}$ G-proteins which induce inhibition of a neuron via cAMP inactivation upon ligand binding (Howlett & Mukhopadhyay, 2000). CB1Rs are the most numerous receptors in the mammalian brain (Burns et al., 2007; Glass et al., 1997), while expression of CB2Rs in the CNS is mostly restricted to immune cells, with much denser expression recorded in peripheral tissue (Atwood & MacKie, 2010).

The two most prominently studied ligands for these receptors are small lipid molecules, 2-arachidonoylglycerol (2-AG) and N-arachidonoylethanolamide (anandamide). These endocannabinoids are synthesised by separate pathways and degraded by separate molecules, with the monoacylglycerol lipase (MAGL) being mainly responsible for 2-AG while fatty acid amidohydrolase (FAAH) breaks down anandamide (Parsons & Hurd, 2015). Both molecules can activate either CB1Rs or CB2Rs (Moreira et al., 2015; Parsons & Hurd, 2015). Due to their vastly greater prominence in the CNS, the rest of this section will focus on CB1Rs, which are primarily found on presynaptic GABAergic and glutamatergic terminals.

Unlike ‘classical’ neurotransmitters such as dopamine, endocannabinoids are not synthesised and stored in synaptic vesicles ready for release. Instead, endocannabinoid signalling is activity dependent (Alger & Kim, 2011). Persistent neuronal depolarisation triggers the synthesis and subsequent release of endocannabinoids. Once released endocannabinoids can diffuse and bind to presynaptic CB1R, mediating inhibition of synaptic inputs. This process enables either depolarisation-induced suppression of inhibition (DSI) or excitation (DSE) of the postsynaptic neuron, depending on whether CB1R are expressed on presynaptic GABAergic or glutamatergic terminals respectively (Ohno-Shosaku & Kano, 2014; Wilson & Nicoll, 2001). These signalling processes have been best described for 2-AG, with the role of anandamide less certain. MAGL, responsible for 2-AG degradation, is located on presynaptic axons and under most conditions’ high expression levels of this enzyme ensure that the signalling effects of 2-AG are transient (S.-H. Lee et al., 2015).

Effects of endocannabinoids on dopamine release and action control

Much evidence suggests that regulation by endocannabinoids plays a vital role in normal functioning of the mesolimbic dopamine system and in appropriate reward seeking behaviour (Fattore et al., 2010; Moreira et al., 2015; Oleson et al., 2021; Parsons & Hurd, 2015; Peters et al., 2020). Endocannabinoid signalling is thought to be an important component in the endogenous regulation of VTA neuron activity. Dopaminergic cells in the VTA express enzymes for the synthesis of 2-AG and receive input from CB1R-expressing glutamatergic and GABAergic axons (Mátyás et al., 2008). In addition, the application of CB1R agonists can depress postsynaptic inhibitory and excitatory currents in VTA neurons

(Melis et al., 2004; Szabo et al., 2002), making DSI or DSE mediated by retrograde endocannabinoid signalling a likely contributor to VTA activity under normal conditions. If expression of CB1R is more prominent on GABAergic inputs than glutamatergic, then treatment with CB1R agonists should reduce net inhibition and increase VTA firing. This upshift in VTA basal firing rates, due to prominence of DSI over DSE, has been observed both in vitro (Cheer et al., 2000) and in vivo (Gessa et al., 1998). In addition to these short-term effects of CB1R activation on VTA activity, endocannabinoids can also mediate long term depressions of inhibitory inputs to VTA (Pan et al., 2008). Via this mechanism, endocannabinoids are involved in long term regulation of VTA firing as they shape the strength of synaptic inputs.

Further studies have confirmed that the CB1R-mediated increase in VTA cell firing is reflected in NAc dopamine release in awake behaving animals. The frequency of NAc dopamine transients recorded with FCV is markedly increased following treatment with a CB1R agonist in freely moving animals (Cheer et al., 2004). Similar studies have demonstrated that endogenous endocannabinoid tone contributes to phasic dopamine release evoked by primary reinforcers. Disrupting endocannabinoid signalling with a CB1R antagonist attenuates the phasic fluctuations in NAc dopamine release elicited by cocaine (Cheer et al., 2007), and novel food rewards (Melis et al., 2007).

Fluctuations in NAc dopamine, such as those attenuated by CB1R blockade in the previous discussed studies, are thought to shape reward pursuit for both food and drugs (Phillips et al., 2003; Roitman et al., 2004). In line with this thinking, animals will respond less to obtain both nicotine (Simonnet et al., 2013), and ethanol (Alvarez-Jaimes et al., 2009) following CB1R antagonist infusions directly into the VTA. In addition to the frequency with which

reward seeking actions are mediated, endocannabinoids can bidirectionally modulate the number of lever presses animal are willing to make for food rewards (Solinas & Goldberg, 2005). Together these results present a picture where intact endocannabinoid signalling enables VTA dopamine release to shape reward pursuit behaviour.

Further investigations employing FCV with cannabinoid pharmacology have provided information about the relationship between cannabinoid signalling, NAc dopamine and reward. Oleson et al recorded NAc dopamine release in rats during an ICSS paradigm, where a cue signalled that a lever could be pressed for rewarding brain stimulation (2012). Blocking CB1R in the VTA markedly reduced cue-evoked dopamine, an effect accompanied by an equally dramatic increase in the latency to initiate reward seeking actions. The suppression of cued-reward seeking with CB1R antagonism is also observed with food-based reinforcement (Ward et al., 2007). These studies indicate that endogenous endocannabinoid tone contributes to both normal dopamine release and the associated reward seeking behaviour.

In addition to the contributions that endocannabinoids make to reward guided action, some investigation has been made into their role in restraining or withholding action. Blockade of CB1R reduces premature responding in the 5CSRTT, suggesting that endocannabinoid tone contributes to baseline impulsivity (de Bruin et al., 2011; Pattij, Janssen, Schepers, et al., 2007). Interestingly, a systemic CB1 agonist had no effect in the task (Pattij, Janssen, Schepers, et al., 2007). This could possibly be mediated by maximal occupancy of CB1R during the task, at least during the pre-cue period where animals must restrain action. CB1Rs can regulate the release of a vast array of neurotransmitters (Egerton et al., 2006; Schlicker & Kathmann, 2001), which presents the possibility that a

CB1 agonist may promote pro-impulsive and anti-impulsive action via its influence distinct systems. In line with this concept, it is difficult to ascertain if CB1R-mediated effects on withholding action in the 5CSRTT occur via dopamine interactions. CB1R blockade does attenuate the increased 5SCRTT impulsivity following systemic administration of the stimulant amphetamine (Wiskerke et al., 2011), which evokes terminal dopamine release. However the promiscuous distribution of the endocannabinoids system makes it difficult it localise this effect. Further experiments are needed to ascertain the how endocannabinoids may contribute to dopamine-influenced dimensions of reward-guided action – particularly in more complex tasks where restraint must be balanced with appropriate initiation.

1.8 - Thesis Aims

Dopamine transmission has been purported to mediate reward learning, action, and motivation, but complete understanding of when dopamine may play a role in each of these functions is unclear. Prior work has implicated mesolimbic dopamine in prompting the initiation of reward-guided action, placing dopaminergic signalling in a core position within a wider network responsible for appropriate action control. However, a critical understanding of the role that phasic dopamine transients, such as those so often recorded in behavioural tasks, may play in reward guided action initiation and restraint remains to be determined. Furthermore, a wider understanding of the circuits and systems that enable appropriate reward seeking, both independently of and in conjunction with the mesolimbic dopamine pathway, remains elusive.

To examine this question, the thesis will apply an established behavioural paradigm – the Go/No-Go task first used by Syed et al. (2016) – to examine the effects of different causal manipulations (see Experimental Chapters for specific hypotheses). This task provides a balanced symmetric design where the influence that changing reward sizes and action requirements have on behaviour can be dissected.

This thesis will first examine the effect of optogenetically evoked dopamine release upon reward guided behaviour within the task to address the causal role of dopamine release at different points in the task and in different magnitudes has on behaviour (Chapter 3). Then, to broaden the perspective on this question, manipulations of two other systems thought to be important for the regulation of endogenous dopamine release and flexible reward guided behaviour, subregions of mPFC (Chapter 4) and the endocannabinoid system (Chapter 5), will be examined.

2. General Methods

Here the general methods used in the following experimental chapters are described, including details of the behavioural paradigm, surgical procedures, and statistical analysis. The approach for optogenetic experiments (Chapter 3) and infusion experiments (Chapter 4) will also be outlined in detail.

2.1 - Go/No-Go Task

The Go/No-Go task utilises a four-way factorial design to enable behaviour to be tracked in conditions with differing reward expectations and differing action requirements. This allows the effects of manipulations on behaviour to be compared across situations where an action had to either be *made* or *withheld* to obtain a reward, and across trials with different amounts of reward on offer.

The task was implemented in operant chambers, arranged such that one wall was equipped with a central noseport port flanked by two retractable levers and the opposing wall held a food magazine (Figure 1A). A trial was initiated by a voluntary nose-poke in the central noseport, which triggered the presentation of one of four possible auditory cues. The identity of the presented cue conveyed information about both the amount of reward on offer for that trial and the action required to obtain that reward.

On trials requiring an action (Go Trials) animals had to make two responses on either the left or right extended lever, dependent on cue identity. On the other half of trials where action had to be withheld (No-Go Trials), animals instead had to remain in the central port for a further 1.5-1.7s. The amount of reward that could be obtained varied. On Small

reward trials the correct action would result in delivery of one 45mg sucrose pellet. Conversely on Large reward trials successful performance resulted in a reward of two 45mg pellets. The combination of two different action requirements and two different reward amounts gave a total of four possible trials, No-Go Small, No-Go Large, Go Small and Go Large (Figure 1B), indicated by different auditory cues (see below for details).

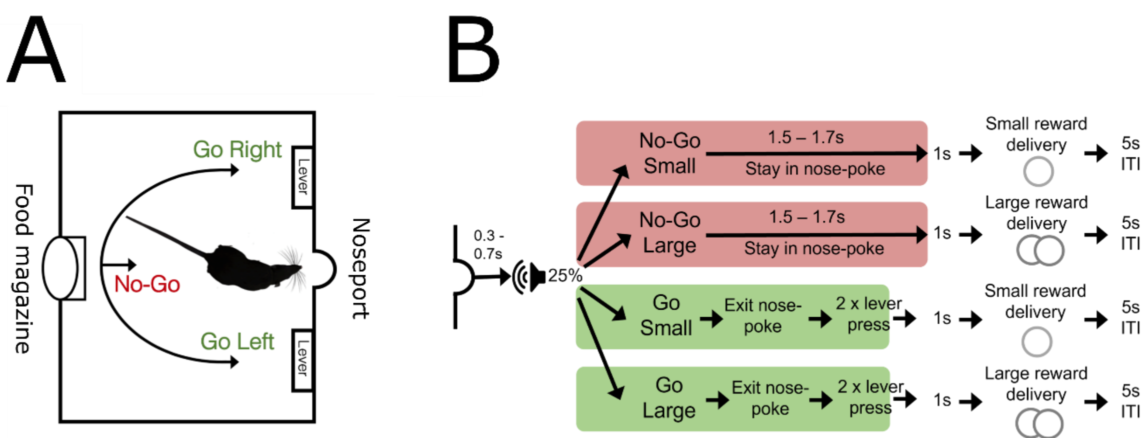


Figure 1. Task schematic.

A) Four-way factorial design of the Go/No-Go task. B) Operant box layout

Behavioural apparatus

Behavioural testing occurred in 30.5 x 24.1 x 29.2cm operant chambers (Med Associates VT, USA). Each chamber was installed in a custom sound-attenuating cabinet. An infrared beam inside the noseport enabled the detection of noseport entry and exit. The opposite wall housed a central food magazine where 45mg sucrose pellets (Test Diet, Sandown Scientific, UK) were delivered. Operant chambers were also equipped with a house light and a speaker for the delivery of auditory cues.

Behavioural training

Auditory cues were one of the following: white noise, tone, buzz or click, all delivered at ~70dB. Cues were pseudo-randomly assigned to trial types (No-Go Small, No-Go Large, Go Small and Go Large) before training began in order to be counterbalanced across animals.

Animals were first habituated to the operant chamber for two 30 minute sessions on consecutive days, followed by an initial session of magazine training, where a sucrose pellet was delivered on a 90 second random interval schedule until 21 pellets had been delivered. Upon successful collection of at least 18 pellets animals progressed to No-Go Trial training. During habituation and magazine training the central noseport was covered.

Each training session lasted until either 100 rewards had been delivered or 60 minutes had elapsed. Training began with No-Go trials. To successfully receive reward in No-Go training, animals had to maintain a given nose-poke hold duration, which consisted of a pre-cue period and a cue period. Across training, the pre-cue period was gradually increased to 0.3-0.7s and the cue period to 1.5-1.7s, where animals had to hold for incrementally longer duration on the current training session if they had achieved < 60% success rate on their previous session. To account for infrared beam breaks resulting from small postural adjustments, a 0.1s buffer period was allowed. If animals exited the noseport and did not enter within the buffer period, then that trial was counted as incorrect. If animals successfully completed the hold duration, then that trial was counted as correct, and reward was delivered. Correct trials were followed by a 5s intertrial interval (ITI) where the nose-poke was inactive. The end of the ITI was uncued. Incorrect trials results in an error period where the houselight turned on for 5s, followed by the standard 5s ITI. During No-Go training both levers were retracted. On No-Go trials the cue was

presented until the end of the holding period on correct trials or until premature head exit on incorrect trials.

After animals had achieved a < 60% success rate on the full No-Go hold duration, a second No-Go trial was introduced, with the same hold duration but a different cue, which resulted in the delivery of two sucrose pellets (No-Go Large).

After reaching criterion for both No-Go trial types, training for Go trials commenced. The first session of Go training consisted of a fixed ratio 1 schedule with the nosepoke blocked, where animals progressed if they had made at least 20 presses on both levers. In the following session, both levers were extended, and nosepoke entry that was sustained for the pre-cue period would result in the presentation of one of two auditory cues. Animals had to complete a single lever press on either the left or right lever dependent on cue identity. Cue presentation lasted until the rat pressed the correct lever, the incorrect lever or until 60s had elapsed with no lever press made. A 5s ITI followed correct Go trials and incorrect Go trials resulted in an error period where the houselight turned on for 5s, followed by the standard 5s ITI.

To aid error correction, following an incorrect Go trial the subsequent trial would be of the same cue and trial type, but with the incorrect lever retracted to prevent the rat pressing it. Achieving > 60% success rate resulted in increasing the lever press requirement to two, and the cue duration (and therefore maximum possible response time) to be incrementally reduced until it lasted for 5s.

Following completion of Go training, No-Go trials were reintroduced. Each of the four trial types were interleaved and had equal probability of being selected (Figure 2). Once animals

had > 60% on all four trial types Go correction trials ceased, and animals progressed to the full version of the task.

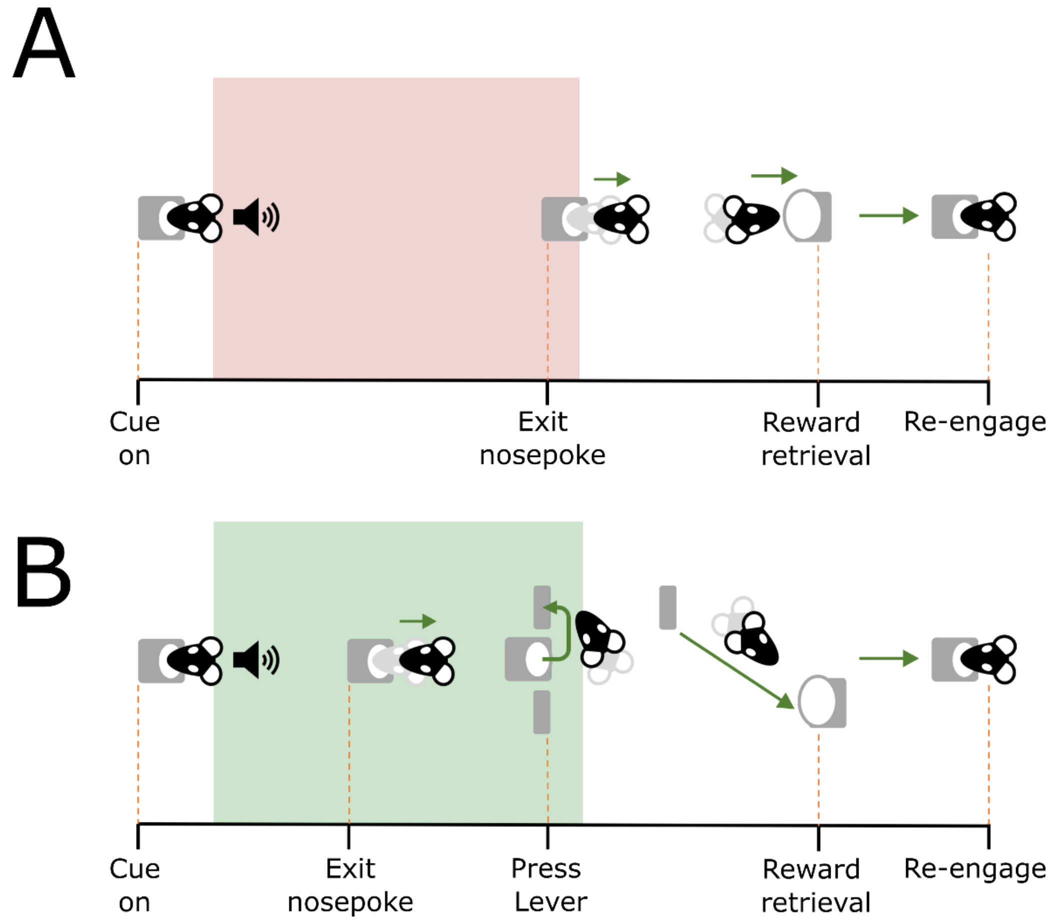


Figure 2. Trial events

A) Schematic detailing the series of events on a successful No-Go Trial. B) same as in A but for Go Trials

---

Measures of performance

In addition to defining trials as ‘correct’ or ‘incorrect’ a more detailed examination of how animals completed, or failed trials can be obtained in the Go/No-Go task. On Go trials animals could fail by either selecting the wrong lever (*Go Wrong Lever Press*) or by failing

to press any lever in the 5s response time (*Go Response Omission*). On No-Go trials only one type of error could be made - premature head exit before the end of the holding period. However, erroneous head exits could be reasoned to occur from two separable processes, a failure to inhibit cue-driven responses or a failure to wait for sufficient time before making a response. As the former would be expected to result in errors clustered near cue presentation, while the latter would result in errors clustered towards the end of the holding period(Härmson et al., 2022), we quantified No-Go errors as either being early (>800ms) or late (<800ms).

Additionally, there are two other methods for completing trials that could signal different behaviours. On Go trials animals could maintain a nosepoke hold for longer than 1.7s – the maximum hold time required to successfully complete a No-Go trial. These trials were labelled as *Invalid Go* trials. Conversely on No-Go trials, after completing the hold duration successfully, animals go then make a lever press. These trials were labelled *Invalid No-Go* trials. While these trials were recorded, as animals successfully completed the requirements of the trial type they were presented with, both types of invalid trials were included as 'correct' when calculating overall success rates.

Finally, animals could leave the noseport before the precue period had been completed and cue presentation triggered. These trials were recorded as *Aborted trials*.

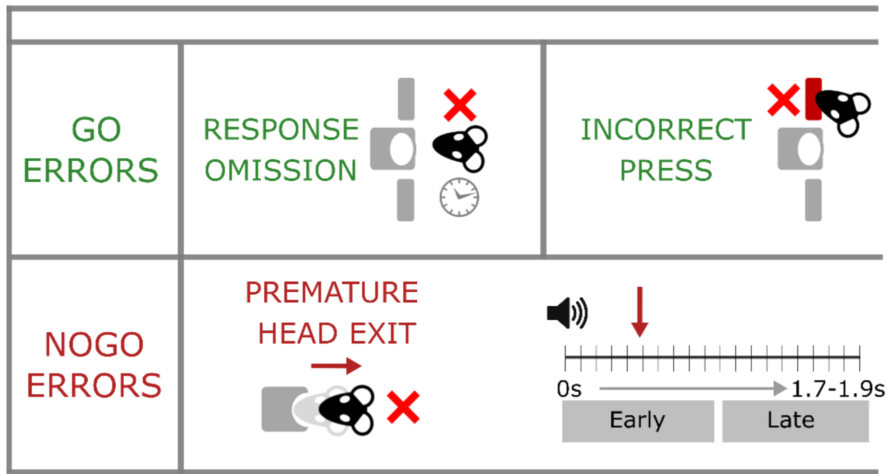
Measures of latency

In addition to recording whether and how animals completed or failed trials, behavioural latencies within the Go/No-Go task were also recorded. Measures of how quickly or slowly

animals performed different action sequences within the tasks contains potential information on action dependent features such as vigour or motivation. The latencies monitored in the Go/No-Go task is listed below:

Time in nose-poke: Noseport exit times aligned to cue onset for both successful and failed trials.

Travel time: on Go trials, the time taken from nose-poke exit to first lever press.



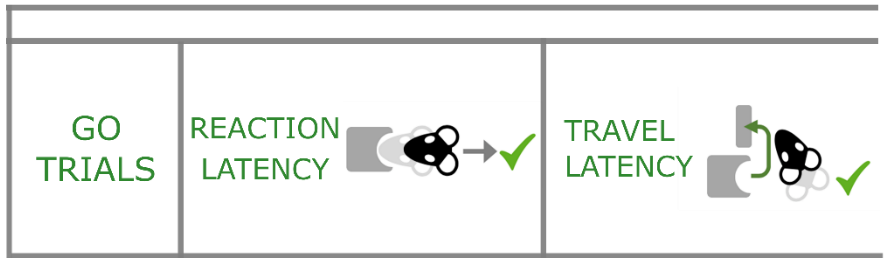


Table 1. Summary of performance and latency measures within the Go/No-Go task

Statistical approaches

For analysis of behaviour in Chapter 3 (optogenetic manipulation), mixed effect models were used (see Section below for details). For data in Chapters 4 (regional mPFC inactivation) and 5 (MAGL inhibition), repeated measures ANOVAs were used. A mixture of within subject and mixed effects ANOVAs were employed depending on the experimental design in question. Any significant effects were further investigated post-hoc for significant pairwise comparison. If the assumption of sphericity was violated, Greenhouse-Geiser corrected p-values were used. Post-hoc tests were corrected for multiple comparisons using a Sidak correction. Effect sizes were reported alongside significance testing.

2.2 -Optogenetics experiments

This section outlines the procedures necessary for the optogenetic activation of VTA dopamine neurons in the Go/No-Go task, results presented in Chapter 3.

Optogenetics

Optogenetics is a technique that enables fast temporal control over neuronal activity with cell type specificity. The technique involves expression of light activated proteins in neurons, which enable depolarisation, hyperpolarisation, or modulation of the excitability of the target cell depending on which construct is expressed. The use of cre-dependent viruses that encode optogenetic tools in conjunction with transgenic animals that express the enzyme Cre-recombinase in a specific subtype of neurons allows the activity of a select population to be manipulated.

For these experiments, a Channelrhodopsin (ChR2) variant was used. ChR2s are cation channels that undergo a conformational change when exposed to blue light of ~470nm, allowing positive ions to flow into the cell and induce depolarisation. In this experiment, the E123T/T159C (ET/TC) variant was used, which carries two mutations relative to the wild type ChR2 (Berndt et al., 2011). ChR2-(ET/TC) has rapid on/off kinetics and generates large photocurrents upon activation, enabling depolarisation of neurons without requiring excessive expression, in addition to reliably inducing depolarisation even at high frequencies of stimulation. Th-Cre rats (SD-*THem1(IRES-Cre)Sage*) which express Cre-recombinase in tyrosine hydroxylase-expressing neurons, needed to synthesise catecholamines including dopamine, were used in these experiments. The combination of this transgenic strain, the coordinates of viral injection, and the cre-dependence of the viral vectors, ensured that expression was restricted to dopaminergic neurons.

Optimization of transfection

In order to tie effects of stimulation specifically to activation of the mesolimbic pathway we aimed to engineer expression of channelrhodopsin within dopamine neurons of the VTA, which project predominantly to ventral and medial striatum including the nucleus accumbens, while sparing the neighbouring SNc, which projects more strongly to dorsal and lateral striatum.

We initially trialled 3 different approaches: two different viral vector serotypes, AAV2 and AAV5, targeted to the VTA, and one of these – AAV5 – targeted directly into NAc.

Delivering AAV5 to the VTA gave good transfection of dopamine neurons, but expression often spread to include both the VTA and the SNc resulting in the whole striatum being transfected (data not shown).

One recent study reported using AAV5 to *retrogradely* transfect dopamine axons(B. T. Saunders et al., 2018), so we tested whether an analogous approach could enable us to reliably transfect NAc-projecting dopamine neurons. However, in animals that received AAV5 injections to the NAc, little to no dopamine release in NAc could be evoked through optogenetic stimulation of VTA in anesthetised FCV recordings, and histology confirmed highly sparse transfection of the VTA.

In contrast to AAV5 delivered to VTA, we found that using an AAV2 virus to deliver the viral construct to the VTA enabled more restricted expression that was largely localised to the VTA. Moreover, combined optogenetic-FCV recording experiments in anaesthetised rats suggested that dopamine could be reliably evoked in the NAc using this method.

We then proceeded to train animals in the Go/No-Go task and use AAV2 to transfect the VTA with either ChR2 (n=10) or EYFP (n = 4). Once performance had stabilised, we applied optogenetic stimulation and found no effect upon any behavioural measures (data not shown). As a positive control, we treated these animals with a systemic D1R agonist. This produced the same pattern of behavioural effects, namely a reduction in No-Go performance, as had been previously observed in wild animals (Grima et al., 2022) – indicating the null effect of optogenetic stimulation was not due to some behavioural aberrance in the transgenic Th-Cre animals.

To confirm the validity of the lack of effect in the Go/No-Go task we next investigated how stimulation may affect behaviour in AAV2 transfected animals in an intracranial self-stimulation paradigm. Here animals will repeat an action such as a lever press or nosepoke to obtain optogenetic stimulation of dopamine neurons. However, stimulation failed to maintain responding in the majority of animals tested. Moreover, in a subset of animals that had FCV electrodes implanted (n=3) and in which reward-evoked dopamine could be observed, optical stimulation failed to evoke reliable dopamine release. Histology confirmed that expression in AAV2 transfected animals was often unreliable and patchy, and also that our optic fibre implantations tended to be too ventral, particularly in females.

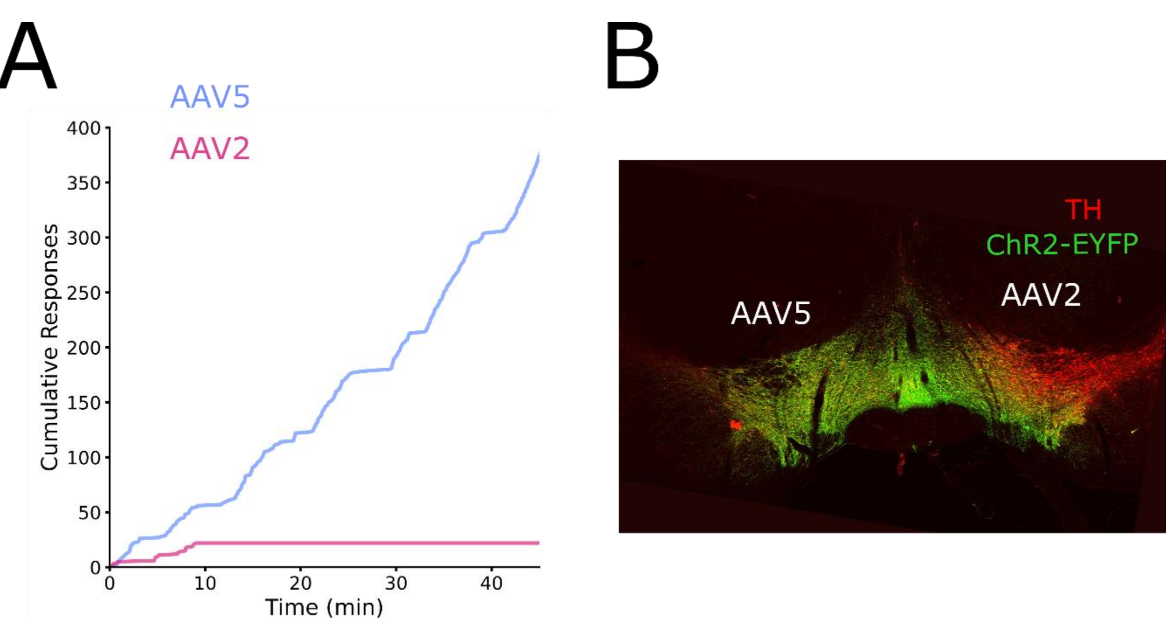


Figure 3. Optimization of viral transfection strategy

A) Cumulative responses in an ICSS paradigm from an example animal transfected with an AVV2 virus (pink) compared to an example animal transfected with an AAV5 virus (blue). B) Example histology from an animal receiving injections of an AAV5 virus in one hemisphere and an AAV2 virus in the other hemisphere.

To address these issues, we performed a number of optimization experiments. First, we adjusted coordinates so that dorso-ventral measurements were taken from the surface of the brain, as opposed to the surface of the skull. We confirmed that these adjusted coordinates enable consistent targeting of the optic fibre terminus to just dorsal to the VTA in both male and female animals. Second, to resolve transfection issues, we returned to using the AAV5 serotype but adjusted the location and reduced the volume of the virus injection to 400nl / site, aiming to achieve VTA specific transfection with targeted injections. We found that this approach allowed for restricted transfection covering, but largely limited to, the VTA. Moreover, optical stimulation induced consistent reinforcement behaviour in ICSS paradigms.

Thus, for the main Go/No-Go experiments reported here, we used this last approach: an AAV5 serotype viral vector and optic fibres targeted to the VTA.

Surgical procedures

Adeno-associated virus (AAVs) were obtained from the University of North Carolina Viral Core. Experimental animals received a cre-dependent virus encoding Channelrhodopsin (AAV-ChR2/ AAV5-EF1a-DIO-hChR2(E123T/T159C)-EYFP) (titre: 5.2×10^{12} vg/ml), while control animals received a virus that only induced expression of EYFP (AAV-EYFP/ AAV5-EF1a-DIO-EYFP-WPRE-pA) (titre: 4.5×10^{12} vg/ml). Viruses were delivered in saline as a 1:1 dilution.

Surgeries took place after rats had reached stable performance in the Go/No-Go task. Anaesthesia was induced in the rats using inhaled isoflurane (4% vol/vol in O2 for

induction, 1-1.5% delivered via facemask for maintenance). Meloxicam (Metacam, 2mg/kg, subcutaneous), buprenorphine (Vetergesic, 0.03 mg/kg subcutaneous) and 2ml of saline (0.9%, Aquapharm, subcutaneous) were given at the start of surgery with a further 1-2ml of saline given during surgery dependent on the length of the procedure. A homothermic heating blanket was used to maintain body temperature at 37 ± 0.5c. Prior to surgery the animal's scalp was shaved and cleaned with diluted hibiscrub and 70% ethanol.

Animals were secured in a stereotactic frame (Kopf Instruments) and the scalp was treated with local anaesthetic (bupivicaine, 2 mg/kg). Eyes were protected with a layer of eye gel (Lacrilube, Allergan) which was reapplied midway through surgery. An incision was made to expose the skull and the bregma and lambda were taken. The tilt of the head was then adjusted, if necessary, such than the dorsal/ventral difference between bregma and lambda was less than 0.1mm.

Small craniotomies were drilled for the injection of viral vectors, the implantation of optic fibres, and the attachment of anchoring screws to ensure the headcap stayed firmly attached to the skull. Screws were secured and a custom pulled glass pipette fitted to an injector (Nanoject, Drummond Scientific) was filled with either AAV-ChR2 or AAV-EYFP depending on whether the animals was assigned to the control or experimental group.

The four craniotomies for delivery of viral vectors were made at the coordinates (-5.2, -6.2 AP; ± 0.75 ML) and two injections were made in each site (-7.2/-7.6 DV to surface of the brain) for a total of eight injections across both hemispheres. All injections were of 400nl volume and injected at a speed of 2nl/second. Following the first injection, the pipette was raised, and the second injection made. Following the second injection in each site the

pipette was left in place for 5 minutes before being removed to avoid pulling virus into the pipette tract.

The craniotomies for the implantation of optic fibres (200um diameter, 0.5NA, Thorlabs) were located at the coordinates (-5.7 AP; $\pm$ 1.825 ML). After injections were completed, fibres were implanted at a depth of -7.6mm D/V to the surface of the brain at a 10° angle. Optic fibres and a headpost were then secured to the surface of the skull using dental cement.

Following surgery animals were given buprenorphine and palatable food. Meloxicam was given in an oral jelly suspension where possible, or via sub-cutaneous injection, for three days following surgery. On average, animals had 2 weeks recovery with ad libituim access to food and water before restarting food restriction and behavioural training.

Anaesthetised fast-scan voltametric recordings.

NAc core dopamine release was recorded using fast-scan cyclic voltammetry in anesthetised animals while applying optogenetic stimulation of varied parameters to the VTA. For recordings, craniotomies were drilled over the VTA, NAc core and posterior cortex for the implantation of optic fibres, carbon-fibre voltammetry recording electrode and Ag/AgCl reference electrode respectively.

Voltammetry recordings were performed as previously described (Gan et al., 2010). Voltametric scans were performed at a frequency of 10hz. Prior to each scan the carbon fibre voltammetry electrode was held at potential of -0.4V compared to the reference

electrode. During a scan, the recording electrodes potential was increased to +1.3V and then returned to – 0.4v at a speed of 400v s^{-1}. Applying this waveform to the carbon fibre results in redox reactions in electrochemically active species, such as dopamine, which can then be recorded over time as fluctuations in current.

Recordings were passed through a 2khz low pass filter. A described in (Flagel et al., 2010; Wanat et al., 2010), principal component analysis was used to extract fluctuations in the concentration of dopamine from other unrelated electrochemical changes such as variation in pH. Recordings where the dopamine current could not be successfully extracted on > 50% of data points were excluded. Data was smoothed using a 0.5s rolling window and visualised using custom python scripts.

Behavioural analysis

In the optogenetic stimulation experiments, individual animals underwent many sessions of data collection, resulting in a large number of trials collected. Analysing this data with an ANOVA results in averaging over all trials to obtain an aggregate measure for each individual. In this experiment, the average number of trials completed per session by each animal varied, as did the number of sessions completed, resulting in an unequal total number of trials completed by each animal. Obtaining an aggregate measure gives equal weight to the data provided by each individual, meaning that animals that have completed fewer trials, and are therefore more likely to suffer from random sampling variation, contribute equally to the final conclusion as animals that have completed significantly more trials. In addition, due to two animals not completing the 'late' stimulation condition, this dataset is unbalanced, and due to exclusions from the ChR2 group the group sizes are

unequal. Though ANOVAs account well for deviations in the normality of underlying datasets, they are less able to handle missing data and unequal groups.

The large number of data points obtained for this experiment enables a regression to be fit as an alternative measure to ANOVAs, which does not suffer from the dependence of ANOVAs on equal group sizes. However, a key assumption of a regression is the independence of individual data points. When individual observations, e.g., trials, can be meaningfully grouped based on which individual completed those trials this assumption is violated. For example, trials from an individual that has a fast average reaction time are more likely to be similar to each other in terms of recorded action initiation latency, compared to trials from a slower individual.

In order to account for non-independence and prevent statistically overconfident results, mixed effects modelling was employed, utilising random intercepts and slopes to account for individual variability.

For all variables where mixed effects models were applied, the formula was as follows:

$$
\begin{aligned}
&Dependent\ variable \\
&\qquad = reward * timing * magnitude * group \\
&\qquad + (random\ effects\ ||\ subject)
\end{aligned}
$$

Variables were coded as follows:

- *Reward:* 2 level ordinal catergorical variable indicating whether the trial could result in a small or large reward. (coded as small = 1, large = -1)

- *Timing:* 2 level categorical variable indicating whether the trial took place in the early or late block. (coded as early = 1, late = -1)
- *Magnitude:* 3 level categorical variable indicating whether stimulation on the trial was not present (No stimulation), consisted of 5 pulses (5p) or consisted of 10 pulses (10p). Regressor 1 ('10p stimulation:' coded as 10p = 1, 5p = 0, no stimulation = -1), Regressor 2 ('5p stimulation': coded as 10p = 0, 5p = 1, no stimulation = -1)
- *Group:* 2 level categorical variable indicating whether the trial is from a control (EYFP) or experimental (ChR2) subject, (coded as ChR2 = 1, EYFP = -1).

In all cases the maximal random effect structure was initially fitted. If the maximal model was singular, the random effect structure was simplified until the model became non-singular (Bates et al., 2015; Matuschek et al., 2017; Singmann & Kellen, 2019).

Mixed models were implemented using the afex package in R programming language. Orthogonal sum-to-zero contrasts and likelihood ratio tests were used to calculate p-values and construct an ANOVA table by comparing the full model to a reduced model where the parameters corresponding to the effect being tested are set to zero. Any interactions including group that reached significance of $p < 0.05$ were investigated further by fitting smaller models post-hoc that split the dataset by group. Any significant effects ($p < 0.05$) were investigated for pairwise comparisons using estimated marginal means.

Histology

Following completion of behavioural experiments animals were terminally anesthetised with sodium pentobarbitone (200mg/kg, i.p). Animals where then transcardially perfused

with 0.9% saline followed with 10% formalin solution (vol/vol). Brains were kept in 10% formalin for a minimum of 2 days before being transferred to 30% sucrose solution in sodium azide. A vibrating blade microtome was used to cut brains into 50um coronal sections. For immunohistochemistry, sections were washed three times in PBST (0.3% Triton, source) before being blocked in 10% normal donkey serum for 1 hour. Primary antibodies for GFP (chicken anti-GFP, 1:50 dilution, Aves Labs) and Tyrosine hydroxylase (rabbit anti-TH, 1:1000 dilution, Millipore) were then applied in a 2% normal donkey serum solution and left overnight. The next day sections were washed three times in PBST and secondary antibodies (Alexa-488 anti-chicken, 1:1000 dilution, Jackson ImmunoRes; Cy3 anti-rabbit, 1:1000 dilution, Jackson ImmunoRes) were applied in a 1% normal donkey serum solution and left for 4 hours. Sections were finally washed 3 times in PBST and mounted onto 2% gelatin-coated glass slides in a DAPI-containing mounting medium (Vectasheild) and cover slipped.

2.3 -Infusion experiments

This section outlines the procedures necessary for the inactivation of mPFC subregion with baclofen/muscimol infusions in the Go/No-Go task, results presented in Chapter 4.

Surgical procedures

As with the optogenetic experiments, once animals had reached stable baseline performance in the Go/No-Go task they underwent surgery. Surgical procedure for the infusion experiments was identical to that for the optogenetic experiments in terms of anaesthesia, analgesia, use of a stereotaxic frame, incision, and levelling of the skull.

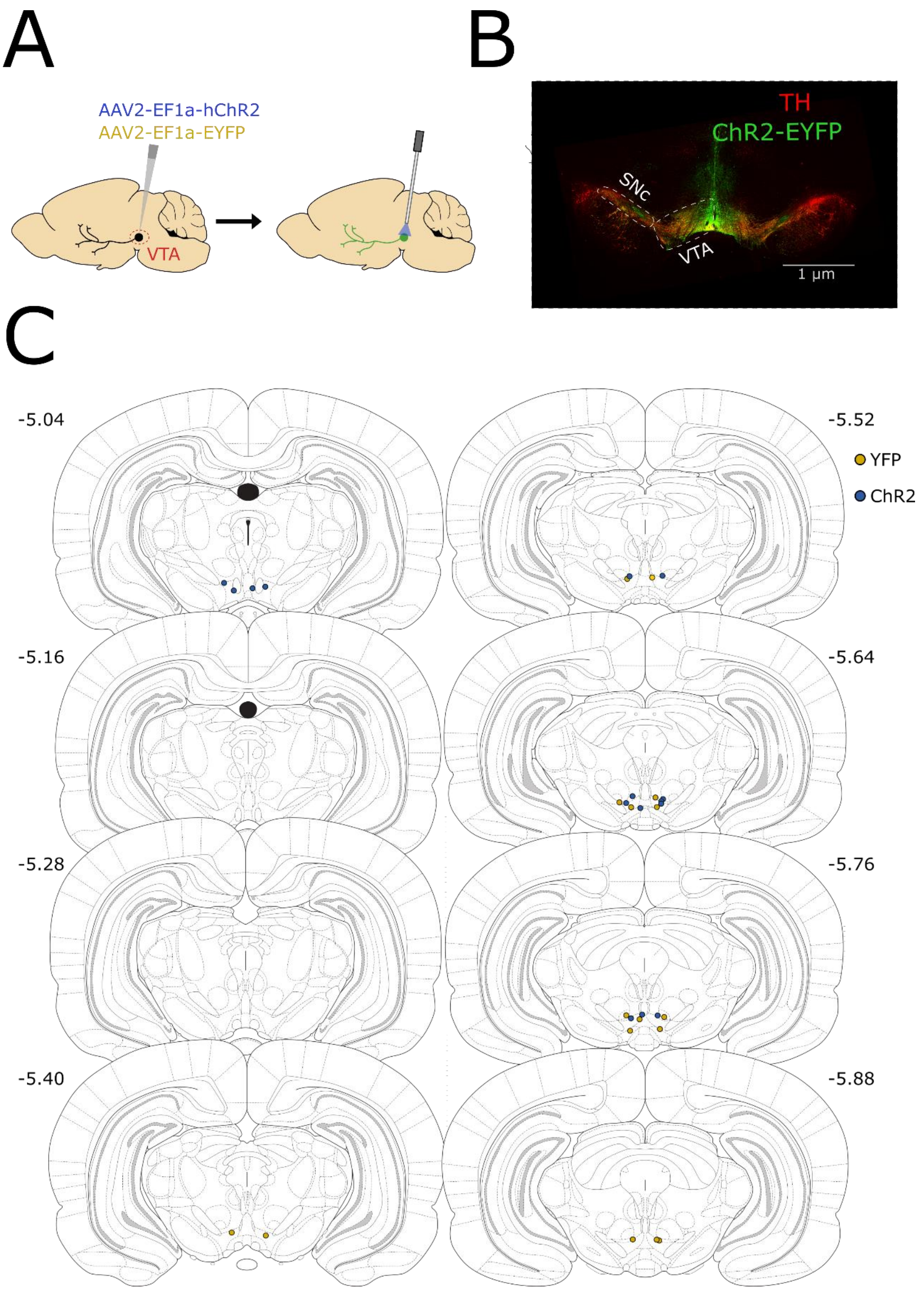


Figure 4. A) Schematic of viral injection and optic fibre placement over the VTA. B) Example coronal section stained for TH and EFYP showing transfection with AVV5-EF1a-ChR2-EYP restricted largely to the VTA. C) Fibre terminus in the VTA for ChR2 animals (blue) and EYFP animals (yellow).

Adapted from Paxinos and Watson (2009)

For infusion experiments, 6 small craniotomies were made in the skull: one for each cannula and four for the dental screws that affixed the headcap to the skull. Animals were implanted with 22-gauge stainless steel guide cannula (Plastics One) bilaterally into the medial prefrontal cortex at the following coordinates: prelimbic cortex (PL) : AP = +2.7 mm, ML = ±.75 mm from bregma, DV = −1.5 mm from dura; infralimbic cortex (IL) : AP = +2.7 mm, ML = ±.75 mm from bregma, DV = −2.5 mm from dura; medial orbital cortex (MO): AP = +3.9mm, ML = ±.75 mm from bregma, DV = −2.0 mm from dura (Paxinos and Watson 1998), as shown in Figure 5. Dental acrylic (Associated Dental Products Ltd.) was then applied to secure the cannula to the skull and screws and any loose skin around the headcap was brought together with stitches. Obdurators flush with the cannula tip were inserted to prevent clogging prior to testing, and dust caps were screwed on.

Post surgery recovery occurred in an identical manner to the optogenetic surgeries.

Infusion procedures

Following stable post-operative performance on the Go/No-Go task, infusions began. Infusions were delivered via two 10 µl glass Hamilton syringes loaded onto an infusion pump (Cole Parmer). The syringes were connected to the intracranial injectors with polyethylene tubing filled with sterile saline. Initially rats received two mock infusions, in which the dust caps and obdurators were removed, and injectors were lowered bilaterally into the cannula. During these sessions no infusions were performed, but instead were carried out to habituate rats to the experience of being handled prior to testing.

The mPFC was inactivated by a mixture of the $GABA_B$ agonist baclofen and the $GABA_A$ agonist muscimol (Sigma–Aldrich), prepared separately at 0.5 µg/µL in 0.9% saline, and mixed together in equal volumes to form a 0.25 µg/µL solution. A stock solution of this GABA agonist cocktail was prepared and frozen. Individual aliquots for each rat were thawed at least one hour prior to infusion. Infusions adhered to a 3-day cycle, starting with a baseline session, followed by drug or saline infusion session, and ending with a washout day. On infusions days, dust caps and obdurators were removed, and 28-gauge injectors were lowered into the cannula. 0.5 µL per hemisphere injections of saline or baclofen/muscimol were administered bilaterally at a rate of 0.4 µL/min, and injectors were left in place for an additional minute to allow diffusion. Once completed, injectors were removed, obdurators replaced, and animals were returned to their home cages for 10 minutes before being placed in the operant chambers and starting the task.

Histology

Once data collection had been completed, animals were deeply anesthetised and transcardially perfused. Extracted brains were submerged in 10% formalin and later 10% sucrose before sectioning. 50 µm coronal sections using a vibratome (Leica) and stained with cresyl violet. The projected locations of the injector tips protruding from the guide cannula were mapped onto coronal section images (Paxinos and Watson, 1998).

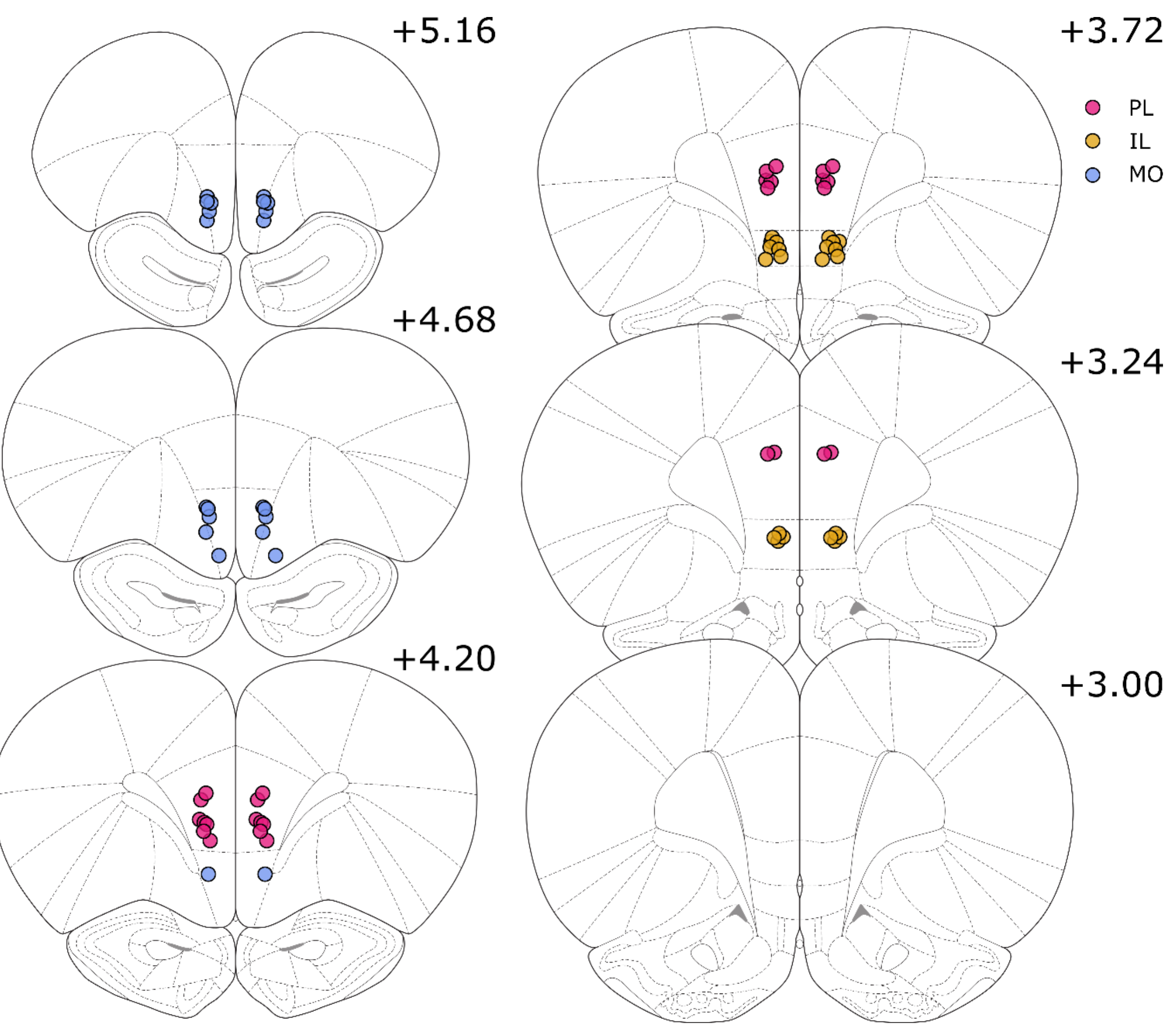


Figure 5. Cannulae placement in mPFC for the PL (pink) Il, (yellow) and MO (blue) groups.

Adapted from Paxinos and Watson (2009)

2.4 - Video tracking

For experiments presented in Chapters 4 and 5, video tracking was applied using DeepLabCut (Nath et al., 2019). DeepLabCut is an open-source toolbox that adapts an algorithm designed for human pose-estimation to track the position of animals in laboratory experiment settings. By annotating a relatively small number of video frames,

the experimenter can train a network to label a number of videos with near-human accuracy.

The model was trained on a subset of 12 videos that gave equal representation to behavioural boxes used and different levels of the experimental manipulation being made. All videos were captured at 25fps. Key body parts of the animals, including the nose, head, shoulders, and tail were labelled – in addition to features of the box such as the four corners of the floor, the levers, and the food magazine (Figure 6A). 15 frames from each video were initially labelled, and a further 5 frames from each video were labelled after the initial output from the model was examined. The model was trained for over 1,030,000 iterations until the average error between test results containing labels from the model, and the training data containing labels from the experimenter was less than 5 pixels.

The model produces estimations for the x and y coordinates of each tracked body-part for each frame in every video to which tracking was applied. The coordinates are supplied along with a confidence score corresponding how certain the model was that tracked body-part in question occupied those coordinates in that frame. Any data points that had a confidence score below a threshold of 0.9 were discarded. Coordinates for a tracked body-part were also discarded if they jumped a distance greater than 20 standard deviations of the frame-by-frame change for that body part, or if they were short sections of valid data (< 5 consecutive points) isolated between unconfident coordinates. All discarded data was interpolated across. Only files with more that 75% of points meeting these criteria were included in analysis.

After an error was made in the Go/No-Go task – the houselight turned on for 5s. The average brightness of each frame of each video was calculated, and the onset of the sharp

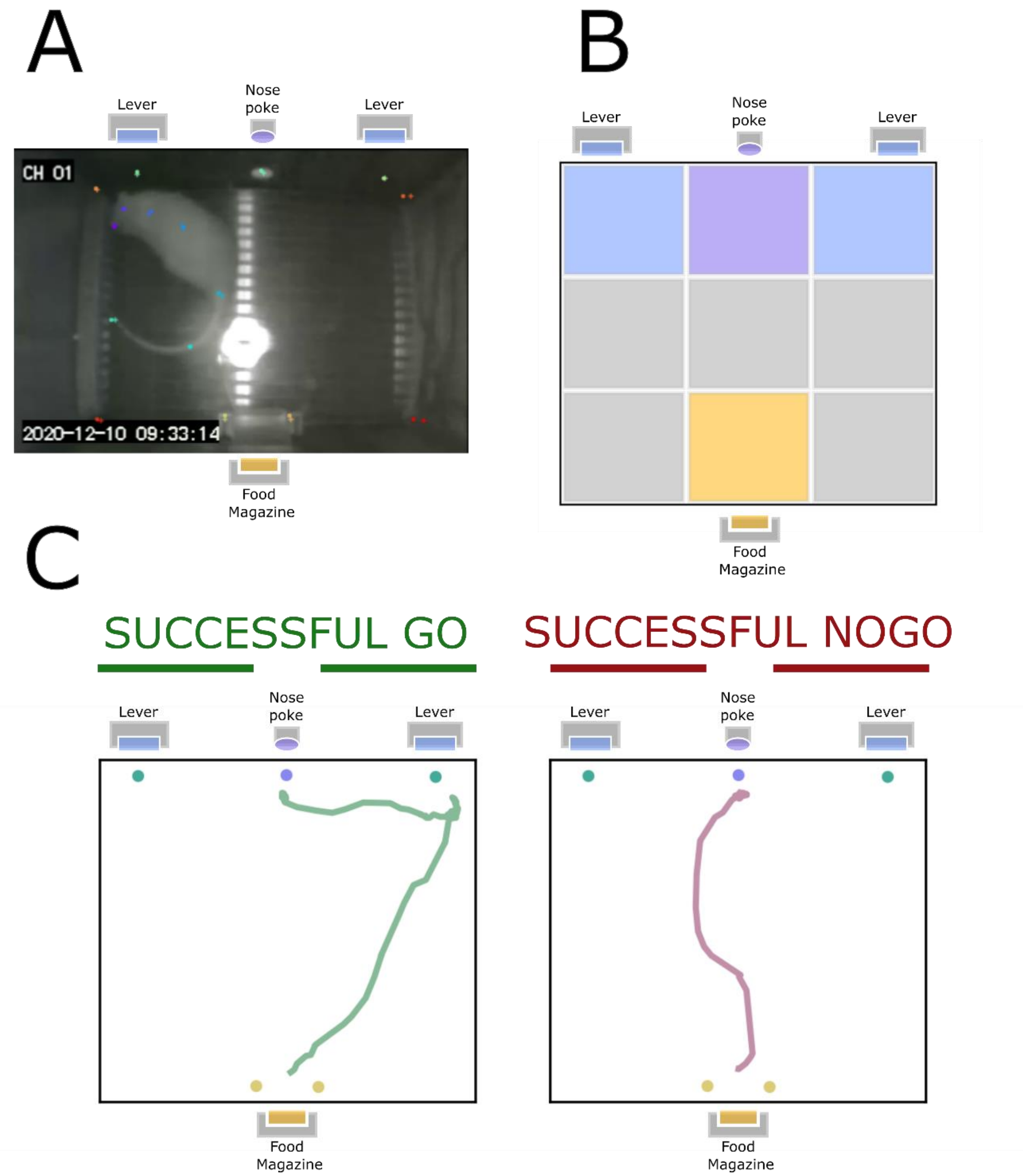


Figure 6. Video Tracking

A) Example labelled frame. Labelled points of the animal (coloured purple through to blue) were as follows: nose, head, shoulders, tail base, mid tail, tail tip. Labelled points of the operant box (coloured green through to red) were as follows: nosepoke, left lever, right lever, left food magazine, right food magazine, box top left, box top right, box bottom left, box bottom right. Crosses indicate the labelling assigned by the experimenter and circles indicate model predictions. B) schematic displaying division of the operant box into zones. C) Example trajectory on a successful Go trial (green, right) and No-Go trial (red, left)

increases of brightness as the house light turned on following an error were used to align the videos with the behavioural data. Any video where the onsets of houselight illumination could not be clearly extracted was discarded. Any file where the range of timestamp differences between behavioural and video error onsets greater than 500ms were discarded. Only animals with video data passing inclusion criteria in every experimental condition were included in analysis (Chapter 4: n= 31 out of 32 total; Chapter 5: n= 10 out of 15 total).

Trajectories were split into trials starting from cue onset. On successful trials trajectories were tracked until 0.75s after entry into a region defined around the food magazine, or 5s after cue presentation, whichever came first. On failed Go trials, tracking lasted from cue onset until 5s after cue presentation and on failed No-Go trials tracking lasted from error onset until 5s after error onset. Trajectories were normalised by subtracting the median value of the nose-poke from all other coordinates. For analysis of zone visitation, the behavioural box was divided into a 3x3 grid, consisting of two lever zones, one food magazine zone, one nosepoke zone and 5 neutral zones that did not contain any task relevant box features (Figure 6B).

3. How do phasic dopamine transients influence action initiation and restraint?

Mesolimbic dopamine is thought to be involved in both learning about rewarding outcomes and activating reward guided behaviour. However, the causal influence that phasic dopamine transients exert on reward guided action is uncertain. In this chapter we optogenetically activated VTA dopamine neurons while animals made or withheld action for reward. We found that VTA dopamine stimulation can promote transitions to action in a manner that depends on the magnitude and timing of stimulation. Moreover, effects on behaviour were not only restricted to the trial on which stimulation was applied, but also caused sustained modifications to behavioural strategy.

Contributions

For the experiments presented in this chapter, George Jenkins, Hironori Ishii, Emilie Werlen and Mark Walton designed and planned the experiments, Emilie Werlen collected anesthetised FCV data, George Jenkins and Lauren Burgeno conducted surgeries, George Jenkins collected all the behavioural data, and George Jenkins analysed and presented all data with input from Mark Walton.

3.1 - Introduction

An extensive body of work describes how dopamine neuron activity and release in regions of the striatum correlates with a reward prediction error signal (Schultz et al., 1997). RPEs have been reported to be encoded in changes in dopamine neuron firing rates across reward dimensions including reward size (Roesch et al., 2007), likelihood (Tobler et al., 2005), and subjective preference (Lak et al., 2014), making RPE theory a prominent framework that proposes to capture the role of dopamine neurones in reward-guided behaviour.

However, dopamine is also closely associated with movement. From a neurological perspective, a myriad of motor symptoms are presented following the death of dopaminergic neurones in Parkinson's disease (Jankovic, 2008). While Parkinsonian pathology has been largely ascribed to SNc neurons, recent studies have demonstrated that dopamine neurons in the SNc *and* VTA are modulated by movement in the absence of explicit rewards (Coddington & Dudman, 2018; Dodson et al., 2016), challenging the proposition that prediction error may provide a unifying theory of dopamine function.

Mesolimbic dopamine in particular is associated with promoting or invigorating movements that aim to obtain rewards, such as the actions taken following the presentation of a reward-associated cue (Robbins & Everitt, 2007). Disrupting dopamine within the NAc can reduce the likelihood that reward seeking actions are initiated (Lex & Hauber, 2008; Yun, Nicola, et al., 2004), while pharmacological elevation of NAc dopamine transmission can drive cued-reward seeking (du Hoffmann & Nicola, 2014), even when it is inappropriate to do so (Grima et al., 2022). Moreover, fluctuations in mesolimbic dopamine have been shown to closely track the initiation of reward seeking actions (Phillips et al.,

2003; Roitman et al., 2004; Syed et al., 2016), with NAc dopamine being suppressed if a cue instructs that action must instead be *withheld* to obtain reward (Syed et al., 2016).

However, it remains unclear how to integrate these different conceptions of dopamine function and predict what causal role transient phasic dopamine signals may exert over reward driven action initiation and restraint. If transient increases in dopamine are used to update the estimated value of associated states and actions to optimize future behaviour (Schultz, 2019a) or as a moment-to-moment estimate of current value (Hamid et al., 2016), stimulating dopamine should act to improve ongoing behaviour and motivate responding required to obtain reward as the value of the current state will have been boosted. If this mimics an RPE signal, then stimulation should also have an effect on performance on *future* trials. Alternatively, should the relationship between dopamine release and cued-promoted reward seeking action be causal, then dopamine transients should primarily act to promote action over inaction.

Moreover, there is increasing evidence that the effect of dopamine neuron stimulation on motor output depends on a number of factors, including an animal's experience within a rewarding context (Hughes et al., 2020), the behavioural state of the animal at the time of stimulation (Da Silva et al., 2018), and when within the framework of a behavioural task stimulation is applied (Hamid et al., 2016). Moreover, while movement can be elicited by driving dopamine release in 'resting' animals, this is only consistently observed with levels of stimulation outside of physiological ranges (Coddington & Dudman, 2018). Together, these results suggest that the effects of dopamine neuron stimulation on reward-driven action may depend on the magnitude and timing of the stimulation with respect to internal states or external cues.

Aims

Therefore, this study aimed to resolve questions regarding the capacity of phasic dopamine events to causally promote action over inaction. In order to separate action requirements from the reward for performing them correctly, we tested animals in a cued Go/No-Go task. Here the cue presented at trial initiation both instructed animals whether to make or withhold action, and informed whether the reward contingent on successful performance would be small or large.

In order to transiently activate dopamine neurons, an optogenetic approach was employed, using light-activated cation channels selectively expressed in VTA dopamine cells. Dopamine neurons can be depolarised with pulses of light, enabling a tight temporal control of activity that mimics endogenous release events more closely than manipulations such as chronic pharmacology. To determine how the effect of dopamine neuron stimulation depends on A) proximity to task-relevant cues B) magnitude of stimulation applied, the timepoint during a trial when stimulation is applied, as well as the number of light pulses delivered in a stimulation train, were manipulated.

Based on the effects of stimulating D1 receptors in the NAc we observed in a previous study (Grima et al., 2022), we predicted that contrary to a pure RPE perspective, activation of dopamine neurons following a No-Go cue when dopamine release is suppressed endogenously should increase the likelihood that inappropriate action is initiated, and impair performance on No-Go trials. In line with these results, we also predicted that stimulation should impair No-Go performance to a greater extent when it is delivered

shortly following cue presentation. Moreover, if dopamine is promoting action over inaction, we also hypothesised that boosting dopamine release should speed action execution on Go trials.

3.2 - Methods

All procedures were performed in line with the UK Animal (Scientific Procedure) Act 1986 and in accordance with the University of Oxford animal use guidelines and were approved by the local ethical review panel at the Department of Experimental Psychology, University of Oxford.

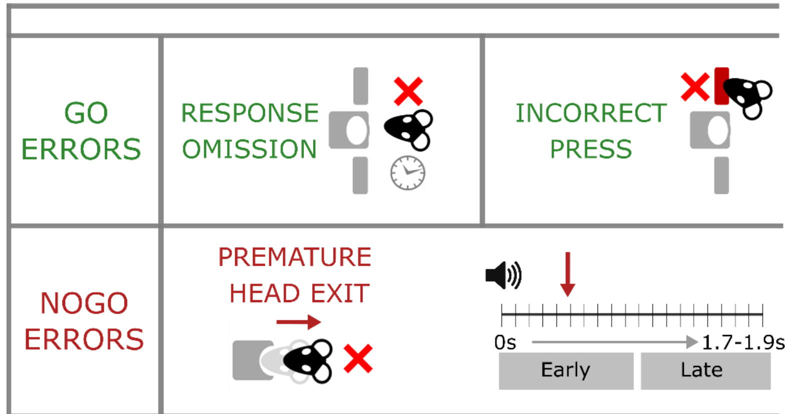
For the Go/No-Go experiments, male and female adult heterozygous TH-Cre rats (SD-*THem1(IRES-Cre)Sage*, n=16), approximately 8 weeks old at the start of training, were trained in the Go/No-Go task as described in Chapter 2. Following initial training, these were divided into an experimental group receiving a Cre-dependent virus expressing ChR2 (n=8) and a control group receiving a Cre-dependent virus expressing just a fluorescent EYFP (n=8). Two animals were excluded from the experimental group either for lack of expression of ChR2 or for expression not being restricted to VTA but also transfecting the neighbouring SNc. This resulting in final group sizes of n=6 for the ChR2 group and n=8 for the EYFP group. For the anesthetised voltammetry experiments male and female adult heterozygous TH-Cre rats (SD-*THem1(IRES-Cre)Sage*, n =5) received a Cre-dependent virus expressing ChR2. A minimum of 5 weeks was allowed between viral injection surgery and either behavioural testing or anesthetised recording to allow for sufficient transfection of dopamine neurons with ChR2.

Animals were maintained on a twelve-hour light/dark cycle. All testing occurred in the light phase. Animals were singly housed post-surgery to prevent damage to optic fibre canula. Food was restricted such that animal bodyweights were kept between 85%-90% of their free feeding weight. Water was available ad libitum.

Details of the behavioural task, training and apparatus are described in Chapter 2 (Methods).

Behavioural measures

Measures of performance and latency are outlined in Chapter 2 (Methods) but briefly reiterated in the table below.



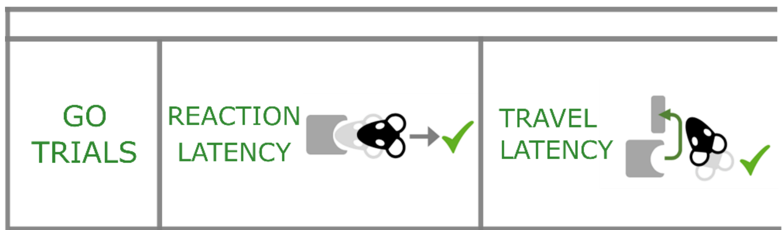


Table 1. Summary of performance and latency measures within the Go/No-Go task

Surgical procedures

Adeno-associated virus (AAVs) were obtained from the University of North Carolina Viral Core. Experimental animals received a cre-dependent virus encoding Channelrhodopsin along with a fluorescent tag (AAV-ChR2/ AAV5-EF1a-DIO-hChR2(E123T/T159C)-EYFP), while control animals received a virus that only induced expression of EYFP (AAV-EYFP/ AAV5-EF1a-DIO-EYFP-WPRE-pA).

Details of the surgical procedures are outlined in Chapter 2 (Methods) but to briefly reiterate, eight injections of AAV-ChR2 or AAV- EYFP were targeted to the VTA, and optic fibres were then implanted over the VTA at a 10° angle.

Anaesthetised fast-scan voltametric recordings.

Details of the methods for recording and analysing FCV data are outlined in Chapter 2 (General Methods). To give a brief overview, in anesthetised animals carbon fibre electrodes were lowered into the NAc core. While optical stimulation was applied, a triangular waveform was applied at 10Hz, driving the potential from -0.4V to +1.3V and back to -0.4V against an Ag/AgCl reference.

Optogenetic manipulations.

Optogenetic stimulation during the Go/No-Go task was delivered bilaterally to VTA dopamine neurons using two 465 nm LEDs (Plexon, Plexbright) connected to two 200um 0.5NA patch cords, with a light power of 8-11mWs at the fibre stimulation. Based on the

anaesthetised recordings, altering the number of pulses in a stimulation train allowed for the cleanest titration of dopamine release downstream in the NAc. Therefore, frequency and pulse width were fixed (20hz, 5ms pulse width), while the number of pulses was varied across sessions.

To examine the effect of two different magnitudes of stimulation: 5-pulse (5p, 'short') or 10-pulse stimulation (10p, 'long') were used. To examine the effect of stimulation timing relative to salient events such as cues, stimulation was delivered at two timepoints on a trial either 50ms ('early') or 800ms ('late') after cue presentation (corresponding to the early or late half of the No-Go holding period: see Härmson et al., 2022).

In each behavioural session, animals only received one type of stimulation, i.e., one combination of 5p or 10p pulse either 50ms ('early') or 800ms ('late') after the cue. Stimulation was delivered pseudorandomly on 33% of all trials in a session. Animals first underwent testing in the early stimulation block, and 5 pulse and 10 pulse sessions were pseudo randomly interleaved across sessions. Following completion of the early block animals then underwent a late stimulation block (Figure 1A). Animals had one session per day, which ended when either 300 rewards had been obtained or one hour had elapsed, whichever came first. Two animals only contributed data to the early block due to being culled due to illness.

On average animals completed 219 $\pm$ 51 trials per session, and underwent 7-16 sessions in the early block and 4-10 sessions in the late block for a total of 57,854 trials across all animals and conditions.

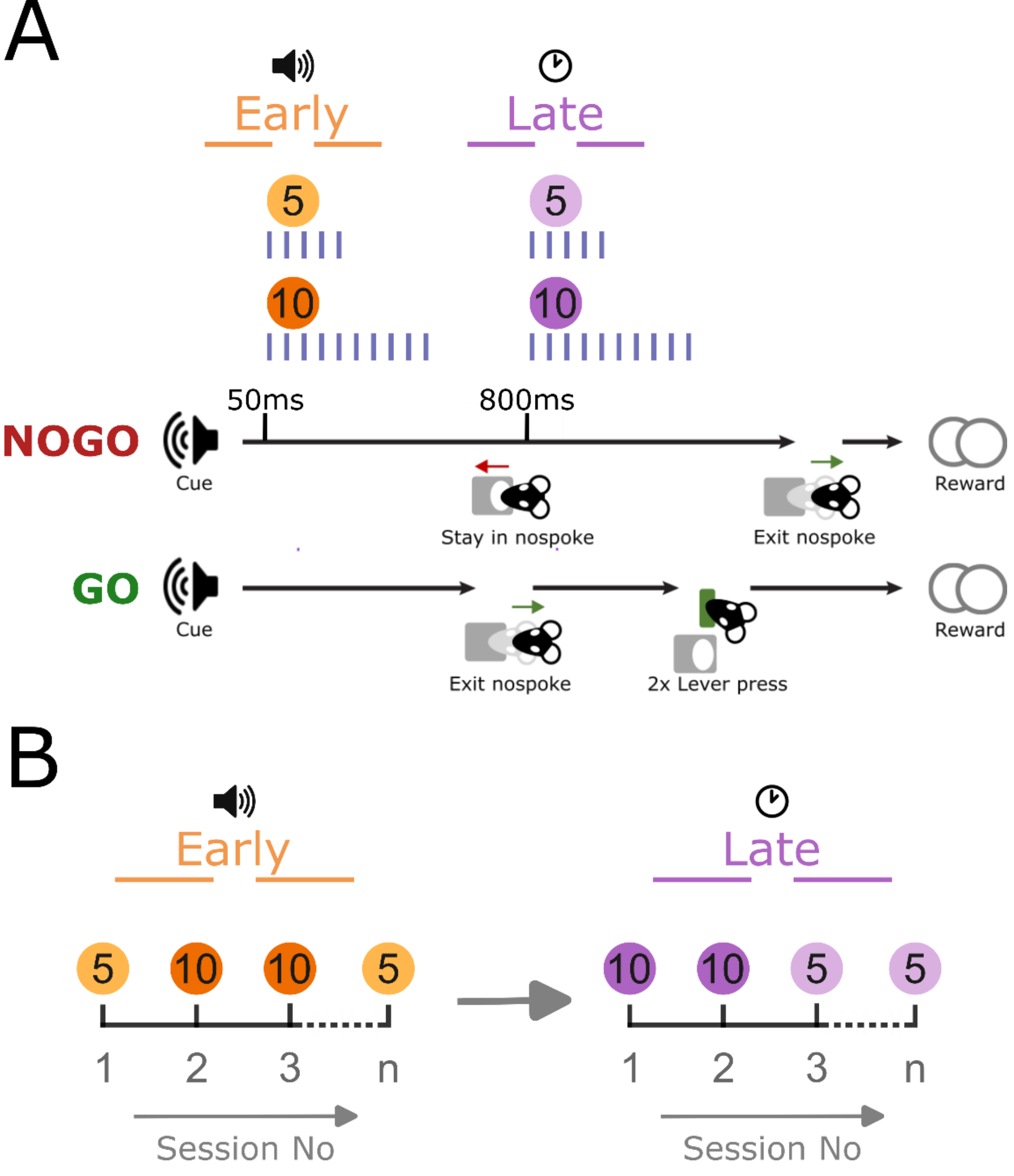


Figure 1. Experimental design

A) Stimulation schematic showing placement of 5-pulse and 10-pulse stimulations relative to task events on Go and No-Go trials when delivered either early (50ms after cue presentation) or late (800ms after cue presentation). B) Timing block schematic. Animals completed 5-pulse and 10-pulses session in a pseudorandom order until they had completed the early block before moving onto the late block.

Statistical analysis of behaviour

A full discussion of the statistical approaches used in the analysis of data from the optogenetics experiments can be found in Chapter 2 (Methods).

To briefly reiterate mixed effect modelling was employed, utilising random intercepts and slopes to account for individual variability.

For all variables where mixed effects models were applied, the formula was as follows:

$$
\begin{aligned}
&Dependent\ variable \\
&\qquad = reward * timing * magnitude * group \\
&\qquad + (random\ effects\ ||\ subject)
\end{aligned}
$$

Variables were coded as follows:

- *Reward:* 2 level ordinal catergorical variable indicating whether the trial could result in a small or large reward. (coded as small = 1, large = -1)
- *Timing:* 2 level categorical variable indicating whether the trial took place in the early or late block. (coded as early = 1, late = -1)
- *Magnitude:* 3 level categorical variable indicating whether stimulation on the trial was not present (No stimulation), consisted of 5 pulses (5p) or consisted of 10 pulses (10p). Regressor 1 ('10p stimulation:' coded as 10p = 1, 5p = 0, no stimulation = -1), Regressor 2 ('5p stimulation': coded as 10p = 0, 5p = 1, no stimulation = -1)
- *Group:* 2 level categorical variable indicating whether trial is from a control (EYFP) or experimental (ChR2) subject, (coded as ChR2 = 1, EYFP = -1).

3.3 - Results

Before applying optogenetics in a behavioural setting, anesthetised recordings were performed in order to understand the relationship between different stimulation parameters applied to the VTA and downstream NAc dopamine release. The parameters varied included the number of pulses in a simulation train (5, 10, 20), the power of the laser (1mW 5mW, 10mW, 20mW), the frequency with which light pulses were delivered (4Hz, 20Hz, 50Hz), and the width of individual pulses of light (0.5ms, 2ms 5ms 10ms, 20ms). For each stimulation only one parameter was varied, and all others were kept fixed.

Varying different parameters produced different effects upon downstream dopamine release. Changing the frequency of the stimulation did not give a clear gradation of dopamine release, as the low frequency 4hz stimulation did not reliably evoke dopamine and stimulating at 50hz produced traces that were highly variable (Figure 2C). Increasing the laser power initially increased evoked dopamine, but this effect quickly saturated at higher values. Varying either pulse width or the number of pulses enabled titrated levels of evoked dopamine, however increasing pulse number gave a cleaner and more linear gradation of peak dopamine at the values tested. On this basis, for use in behavioural experiments two different stimulations were selected that differed in the number of pulses but were of consistent frequency, power, and pulse width.

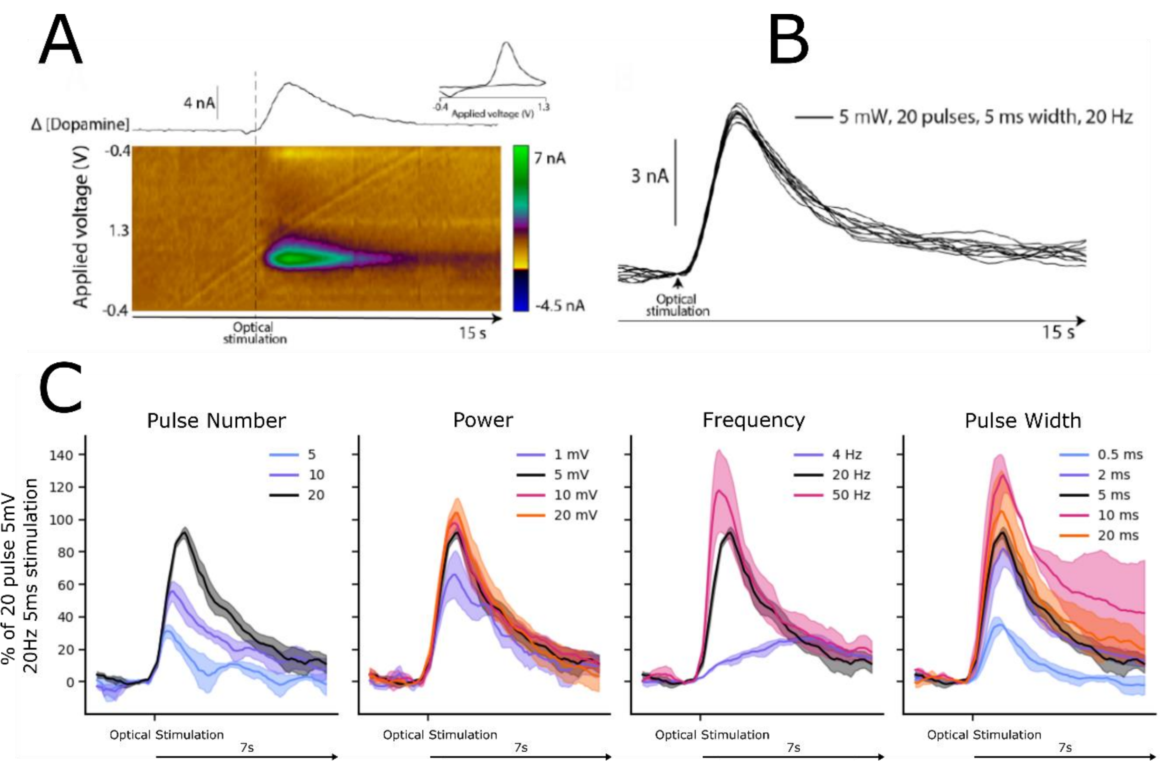


Figure **2**. Optically evoked **NAc** dopamine can be titrated.

A Representative example of dopamine release in the nucleus accumbens after optical stimulation of the VTA. Color plot shows the background-subtracted current as a function of the applied-voltage over time; inset example cyclic voltammogram identifying the detected current as dopamine. B) Repeated stimulation using the same parameters elicits stable dopamine release over many minutes. C) Average dopamine traces for each level of varied stimulation parameter, expressed as percentage of average dopamine evoked by a 20hz 20 pulse stimulation at 5mW with 5ms pulses (black trace in each subplot).

VTA stimulation conditionally impairs action restraint.

We first examined the effect of optical activation of dopamine neurons on No-Go trials. To determine the effect of the stimulation on No-Go performance we employed logistic mixed effects modelling, initially fit to data from both groups (EYFP, ChR2). The model included information on the magnitude of stimulation applied on each trial (i.e., unstimulated, 5p, 10p), the timepoint within the trial at which the stimulation was applied (early or late) and

the size of the reward on offer (small, large). In particular, we wanted to determine if there was a significant effect of applying either a 5p or a 10p stimulation relative to the unstimulated condition on No-Go performance in the ChR2 group, and if so whether that effect interacted with other factors such as the timing at which stimulation was applied.

As can be seen in Figure 3, optical stimulation of VTA dopamine neurons selectively changed No-Go success rates in the ChR2 group, but in a way that depended on the timing and magnitude of the stimulation. This was reflected in a stimulation magnitude * stimulation timing * group effect ($p=0.011$), indicating the effect of stimulation differed depending on both stimulation features (magnitude/timing) and experimental group, as well as a reward * group interaction ($p=0.046$).

To understand what was driving these effects, separate mixed effect logistic regressions were then fit to data from the ChR2 and EYFP groups. This showed that the effect of optogenetic stimulation was modulated both by stimulation timing and stimulation magnitude in the ChR2 group [significant stimulation magnitude * timing block interaction, $p <0.001$]. In the early block both the 5p and the 10p stimulation impaired No-Go performance, inducing a significant increase in the probability of premature responses [pairwise comparisons: early stimulation block, no stimulation vs 5 pulses $p < 0.0001$; no stimulation vs 10 pulses $p < 0.0001$, Figure3A]. When stimulation was delivered late in the No-Go period, the 10p stimulation again impaired performance [late stimulation block, no stimulation vs 10 pulses $p = 0.0002$]. By contrast, when the 5p stimulation was delivered at the late time point, it had an *opposing* effect on behaviour, *increasing* the likelihood animals would successfully restrain action [late stimulation block, no stimulation vs 5 pulses $p = 0.0001$]. These effects were specific to dopamine neuron activation as there

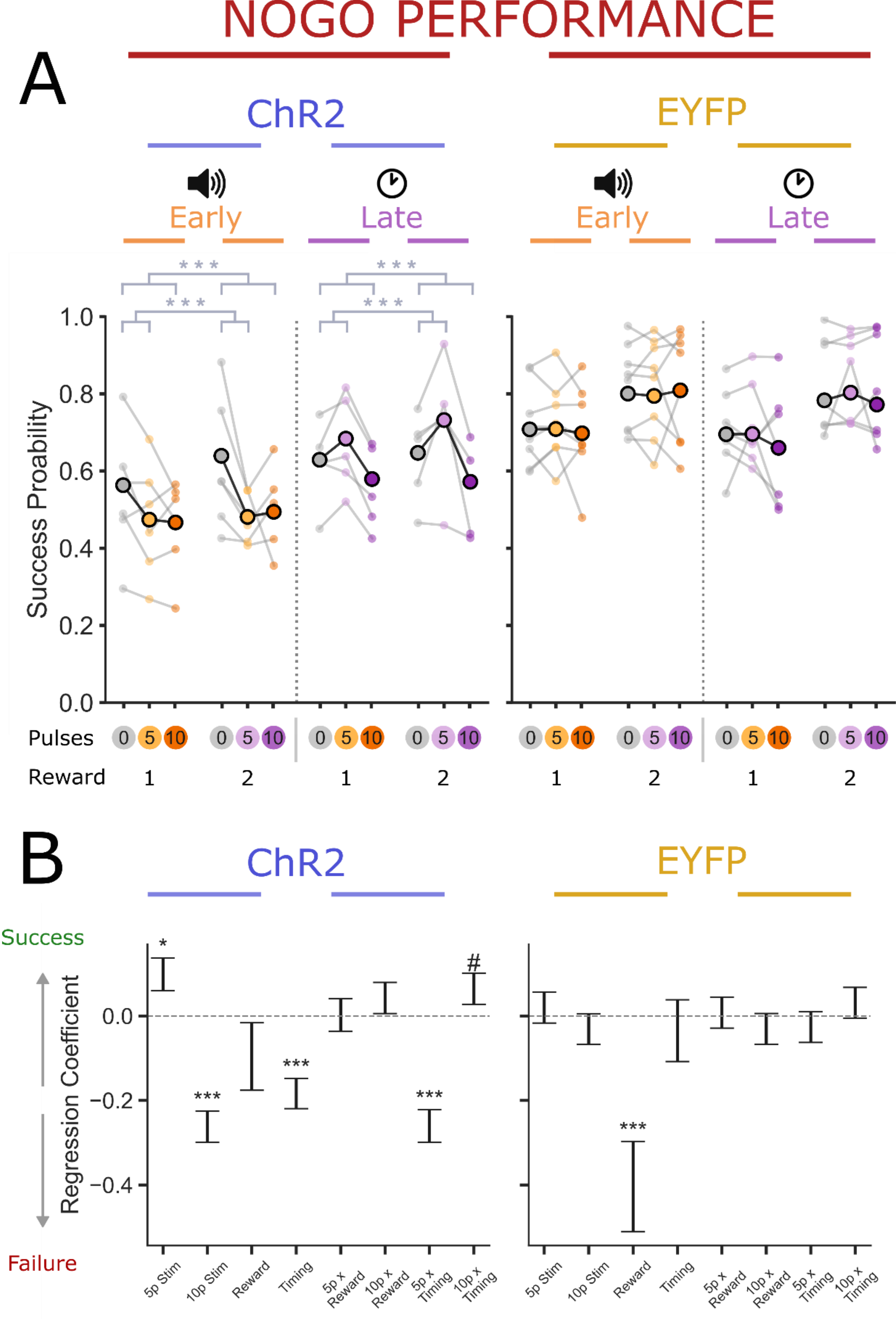


Figure 3. The effect of stimulation on No-Go performance depends on magnitude and the timepoint of delivery.

A) Probability of No-Go Success across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent success probability over all trials. Faded points represent success probability for individual animals, data from the same individuals joined by a faded line. B) Mixed-

effects logistic regression on data from the ChR2 group (left) or the EYFP group predicting probability of success based upon stimulation, time of delivery, size of reward and their interactions. Three-way interaction coefficients not shown. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

---

were no analogous effects in the control EYFP group. However, the EYFP group, but not the ChR2 group, had an increased probability of success when a large reward was on offer [significant pairwise comparisons in the EYFP group: small reward vs large reward $p = 0.0002$].

In summary, the effect of activating VTA neurons on action restraint depends on the magnitude of stimulation and the timepoint of its delivery. The longer 10p stimulation consistently promoted premature action, while the effect of the shorter 5p stimulation had either a disruptive or beneficial effect on restraint depending on whether it is delivered early or late respectively.

Effects of VTA stimulation extend past a single trial

Stimulation of VTA dopamine neurons modulated No-Go performance on trials receiving stimulation. However, as can be seen in Figure 3, differences between ChR2 and EYFP groups appeared not just restricted to trials where stimulation was applied, as the ChR2 group also were less successful at withholding action even on *unstimulated* trials [stimulation magnitude * timing block * group interaction, $p = 0.0011$; pairwise comparisons, unstimulated trials, early and late stimulation blocks: ChR2 vs EYFP, all $p < 0.0022$].

If VTA activation merely perturbed online behaviour, then any impairments of No-Go performance would be restricted to trials where stimulation was applied. The fact that the ChR2 and EYFP groups differ in No-Go performance even on *unstimulated* trials presents one of two possibilities. First, it could be that animals selected for the experimental group were generally less able to restrain action as required to successfully complete No-Go trials. To examine for pre-existing differences in baseline No-Go performance we analysed the last session of training for each animal, where performance was stable, but no stimulation was applied. ChR2 animals were not found to differ significantly on baseline No-Go performance relative to EYFP animals [$t(12) = 1.028$, $p = 0.325$].

Alternatively, VTA activation may not only influence online action restraint, but could also have a sustained effect that reduces the probability of successful action. If this effect were persistent and cumulative, we may expect No-Go performance to be similar between groups early in testing, but for ChR2 animals to deteriorate *across* behavioural sessions when compared to EYFP animals, as more simulations are received.

To investigate whether impairment on No-Go trials accumulated as more days of testing were experienced, we took unstimulated trials from across the first sessions animals underwent. As animals experienced the early stimulation block first, stimulation in these sessions was always delivered 50ms after cue presentation, which impaired performance in ChR2 animals irrespective of the stimulation magnitude applied. The minimum number of early block sessions completed by an individual animal was 8, so for all animals we examined the first 8 sessions. Unstimulated trials from these sessions were pooled regardless of whether the stimulation for that session was 5 or 10 pulses in magnitude. As can be seen in Figure 4A, performance of the ChR2 group is lower relative to EYFP

[significant main effect of group $F_{1,14} = 6.689$, $\eta_p^2 = 0.357$, $p = 0.023$, Figure 4A], but importantly, neither group display changes across testing days [no significant main effect of session number or interaction, all $Fs < 0.73$ all $ps > 0.64$]. To understand whether there was a general change in motivation, the same analysis was run on Go trials from the same block of sessions. Here both groups remain at a stable performance and this time overlay each other [Figure 4A, no main effects or interactions, all $Fs < 1$ all $ps > 0.44$].

While performance remained stable *across* sessions, a sustained effect of stimulation may still accumulate and effect No-Go performance in ChR2 animals to an increasing extent *within* a session. To investigate whether ChR2 performance degraded within a session relative to EYFP controls, performance on all unstimulated No-Go trials from sessions with the same stimulation timing and magnitude were grouped, and then binned into quartiles. To account for differences in individual baseline performance, the success rate for the first quartile was subtracted from every subsequent quartile for each animal. This revealed No-Go performance on unstimulated trials *within a session* was significantly influenced by both group and stimulation timing [significant timing block * quartile * group interaction $F_{3,30} = 3.191$, $\eta_p^2 = 0.242$, $p = 0.038$].

To further understand this effect, separate ANOVAs were run on the data from the early block and the late block. In the early timing block, we detected a quartile by group interaction [$F_{3,30} = 5.424$, $\eta_p^2 = 0.311$, $p = 0.003$]. ChR2 animals' performance degraded within a session, such that they were significantly worse in quartile 3 and 4 when compared to the control group [Figure 4B, quartile 2, ChR2 vs EYFP $p = 0.054$, quartile 3, ChR2 vs EYFP $p = 0.014$, quartile 4 ChR2 vs EYFP $p = 0.018$]. This was specific to the group receiving dopamine activation as EYFP animals' performance improved between the 1st and 2nd

quartiles and then remained stable for the rest of the session [pairwise comparisons, EYFP, quartile 1 vs quartile 2 $p = 0.037$, all other $ps > 0.299$].

In the late timing block, changes in unstimulated No-Go performance within a session did not differ between the ChR2 and EYFP groups [Figure 4C, no significant main effect of group or interaction, all $Fs < 1.565$, all $ps > 0.218$]. Animals did show greater improvement in No-Go performance in 5p sessions compared to 10p sessions, but this effect did not depend on group [significant stimulation magnitude * quartile interaction $F_{3,30} = 3.857$, $\eta_p^2 = 0.278$, $p = 0.019$: pairwise comparison: quartile 2, 5 pulse session vs 10 pulse session, $p = 0.025$; quartile 3, 5 pulse session vs 10 pulse session $p = 0.008$; quartile , 5 pulse session vs 10 pulse session $p = 0.1$]. As the frequency at which stimulation occurs does not differ between the early and late block, the total concentration of optically evoked dopamine should be similar. The lack of accumulating impairment in the late block suggests that the effect is not due to rising extracellular dopamine.

In summary, the difference in No-Go performance between the ChR2 and EYFP control group on unstimulated trials is not caused by baseline differences between the groups or an accumulated deterioration in performance across sessions. Instead, it primarily reflects a within-session increase in premature responding on No-Go trials such that optical stimulation of dopamine release has effects that last beyond the stimulated trial itself. This effect was only significantly detected in early block sessions, and not late block sessions.

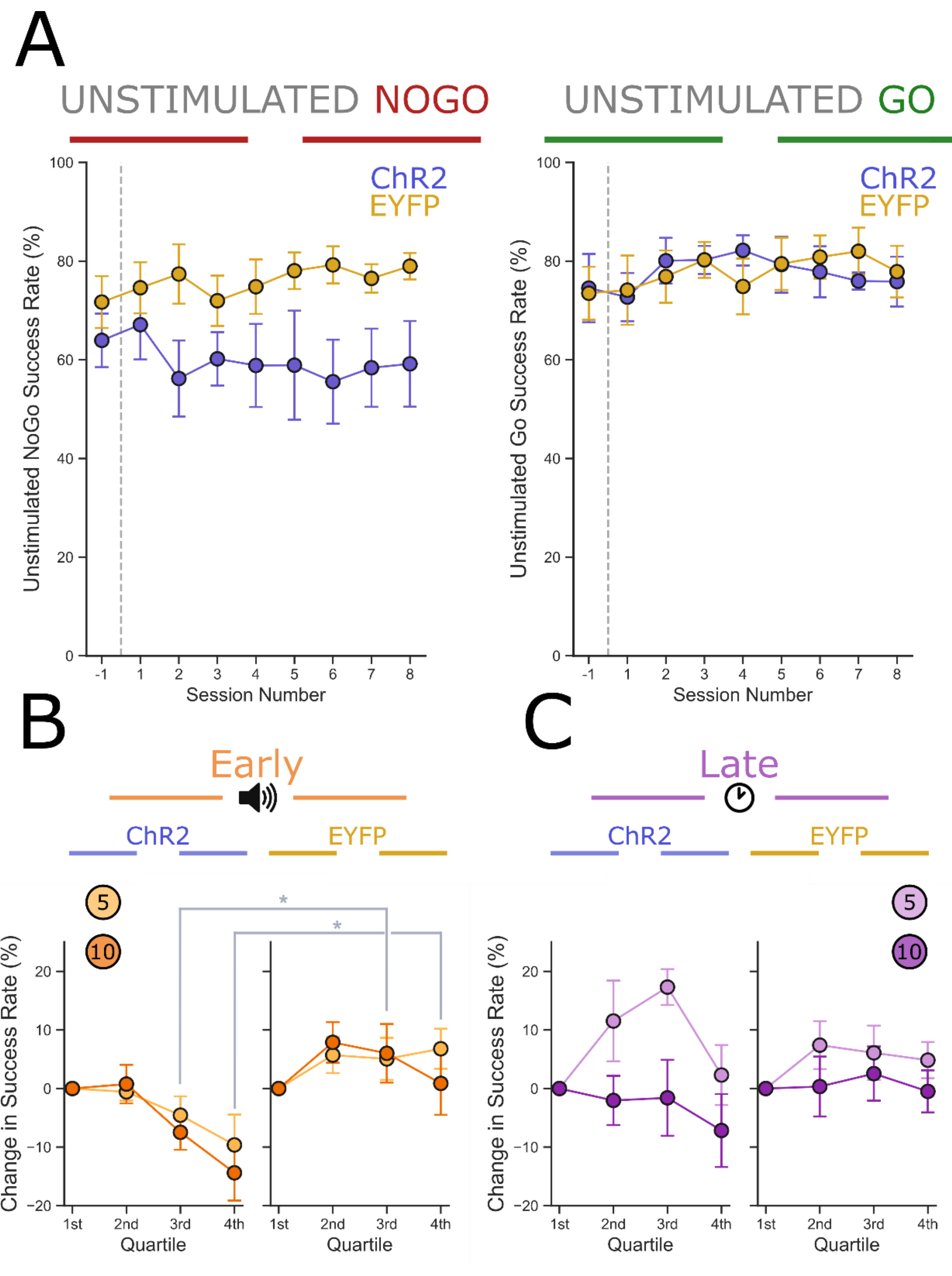


Figure 4. Unstimulated No-Go performance degraded within but not across early block sessions.

A) Left panel, average No-Go performance on unstimulated trials in the ChR2 group (blue) and the EYFP group (gold) in the last session of training (session -1) and across the first 8 sessions in the early block (sessions 1-8). Right panel same as in left panel but for unstimulated Go performance. B) Average change in unstimulated No-Go performance within a session relative to performance on the 1st quartile of a session for the ChR2 group (left) and the EYFP group (right) in session in the early block (5p sessions: light orange, 10p sessions: dark orange). C) same as in B but for the light

block. (5p sessions: light purple, 10p sessions: dark purple). Filled points represent means. Bars show standard error.

VTA activation alters aspects of action execution

We next investigated how VTA dopamine stimulation affects action selection on Go trials. We employed the same analytical approach as on No-Go trials, and fit a logistic mixed regression to the data. As observed previously (Grima et al., 2022), performance on Go trials was modulated overall by the size of reward, with better performance on high reward trials [main effect of reward $p < 0.001$]. In addition, there was a main effect of timing, with Go success being higher in the late stimulation block (which also occurred later in the testing schedule) [Main effect of timing, $p = 0.01$, early block vs late block). Importantly, however, there was no reliable change following optical stimulation [Figure 5A, No main effect of stimulation or any interactions, all $ps > 0.19$], nor did performance differ between control and experimental groups [No main effect of group or any interactions, all $ps > 0.147$].

On Go trials, errors can occur either if an animal responds on the wrong lever (Incorrect Press error) or if they fail to complete a sequence of two lever presses within a 5s time limit after cue presentation (Response Omission error). In order to determine whether stimulation might have induced a specifc *type* of error, without signifcantly altering overall Go performance, logistic regressions were also used to assess changes in the probability of each type of error occuring.

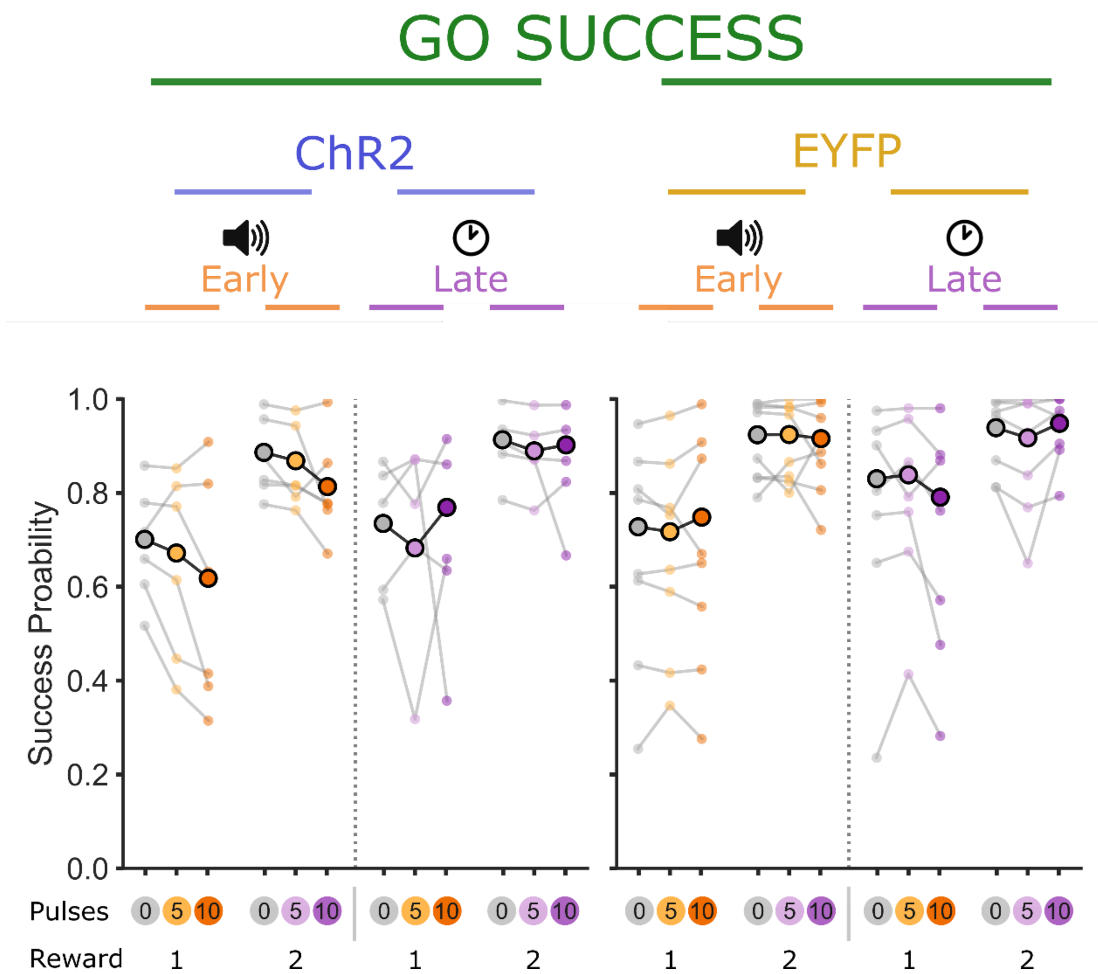


Figure 5. Stimulation does not significantly impair Go performance.

A) Probability of Go Success across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent success probability over all trials. Faded points represent success probability for individual animals, data from the same individuals joined by a faded line. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

We first examined effects on the probability of making a Response Omission error. Fitting a model to all of the data revealed a stimulation magnitude * timing block * reward interaction [p = 0.041]. This is effect was driven by an increase in omissions on large trials receiving a 5p stimulation in the 800ms block relative to unstimulated trials [pairwise comparisons: no stimulation vs 5p p = 0.0172], but as it was observed across animals in

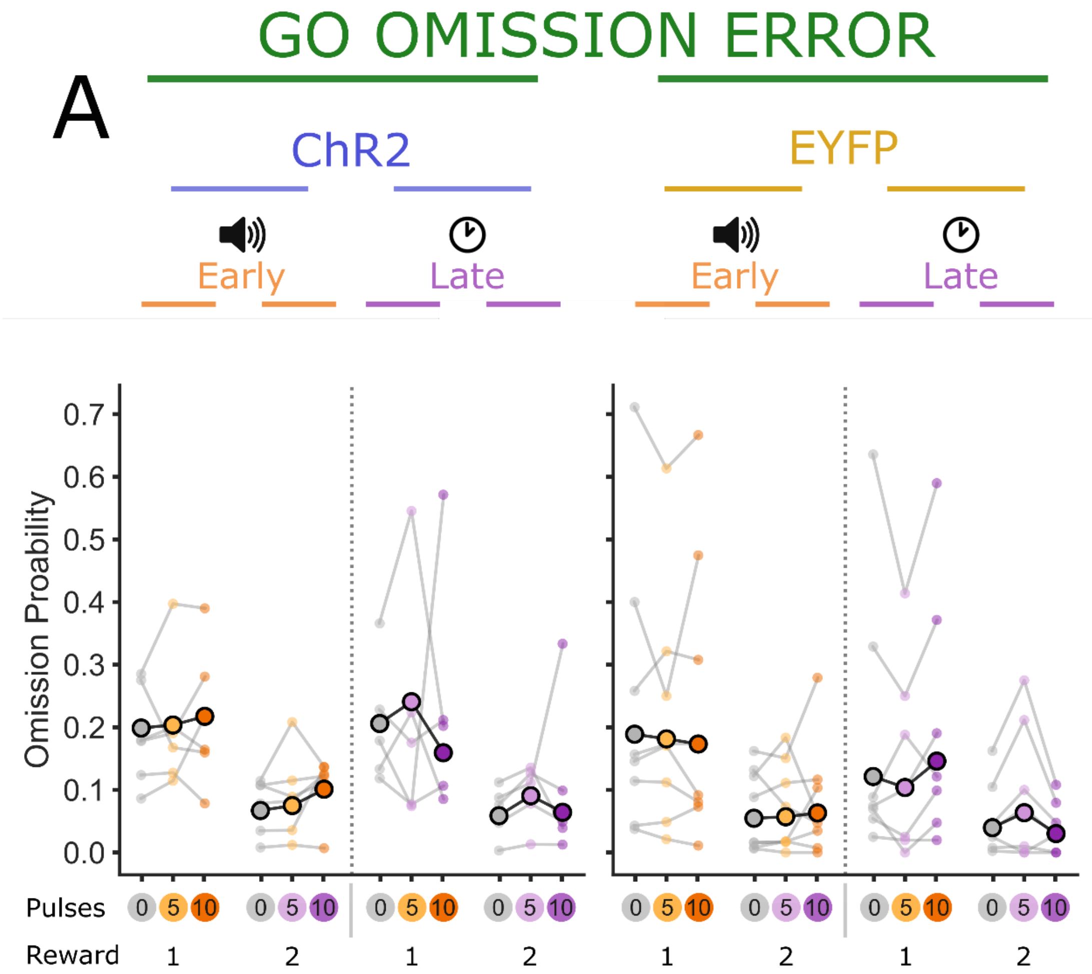


Figure 6. Stimulation does not significantly influence Go Response Omission errors.

A) Probability of making a response omission across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent omission probability over all trials. Faded points represent omission probability for individual animals, data from the same individuals joined by a faded line. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

both the EYFP and ChR2 groups, is likely attributable to a non-specific distracting effect of stimulation.

We next investigated the probability of selecting the wrong lever on a Go trial via logistic mixed-effect regression model fit using all the trials. This analysis demonstrated a timing * group interaction [p =0.003]. To examine this effect in more detail, models were seperately fit to data from the ChR2 group and data from the EYFP group. This showed that the ChR2

animals were more likely to select to wrong lever in the early stimulation block [Figure 7A, significant main effect of timing block p <0.001], and when a 10p stimulation was applied [significant main effect of stimulation magnitude p =0.001; pairwise comparisons: no stimulation vs 10 pulses p =0.004], and although this effect appeared most prominent in the early block, these effects did not significantly interact [stimulation magnitude * timing block interaction p = 0.143]. No significant effects related to stimulation were detected in the EFYP model, though EFYP animals did generally make fewer errors in the late stimulation block when the reward on offer was small [reward * timing block interaction, p = 0.023: pairwise comparisons when reward is small, early block vs late block p < 0.001].

VTA stimulation selectively slows travel latencies.

Prior experiments have found that activiating D1Rs could have complex effects on trial latencies, tending to speed how quickly animals reacted to a Go cue and exited the nose poke, but also, when administered systemically, slowing travel times(Grima et al., 2022). To examine whether optogenetic stimulation recapitualted these effects, linear mixed effect models were used to predict the effects of stimulation on action latencies.

Unlike following NAc D1R activation, animals reacted no faster to the Go cue after receiving stimulation of either magnitude [no main effect of stimulation magnitude or any interactions all ps > 0.104]. A larger reward being on offer speeded latencies [Main effect of reward, p = 0.037], consitent with vigour or motivation scaling with value of reward on offer, and in general the ChR2 group were faster to iniatiate action than the EYFP controls

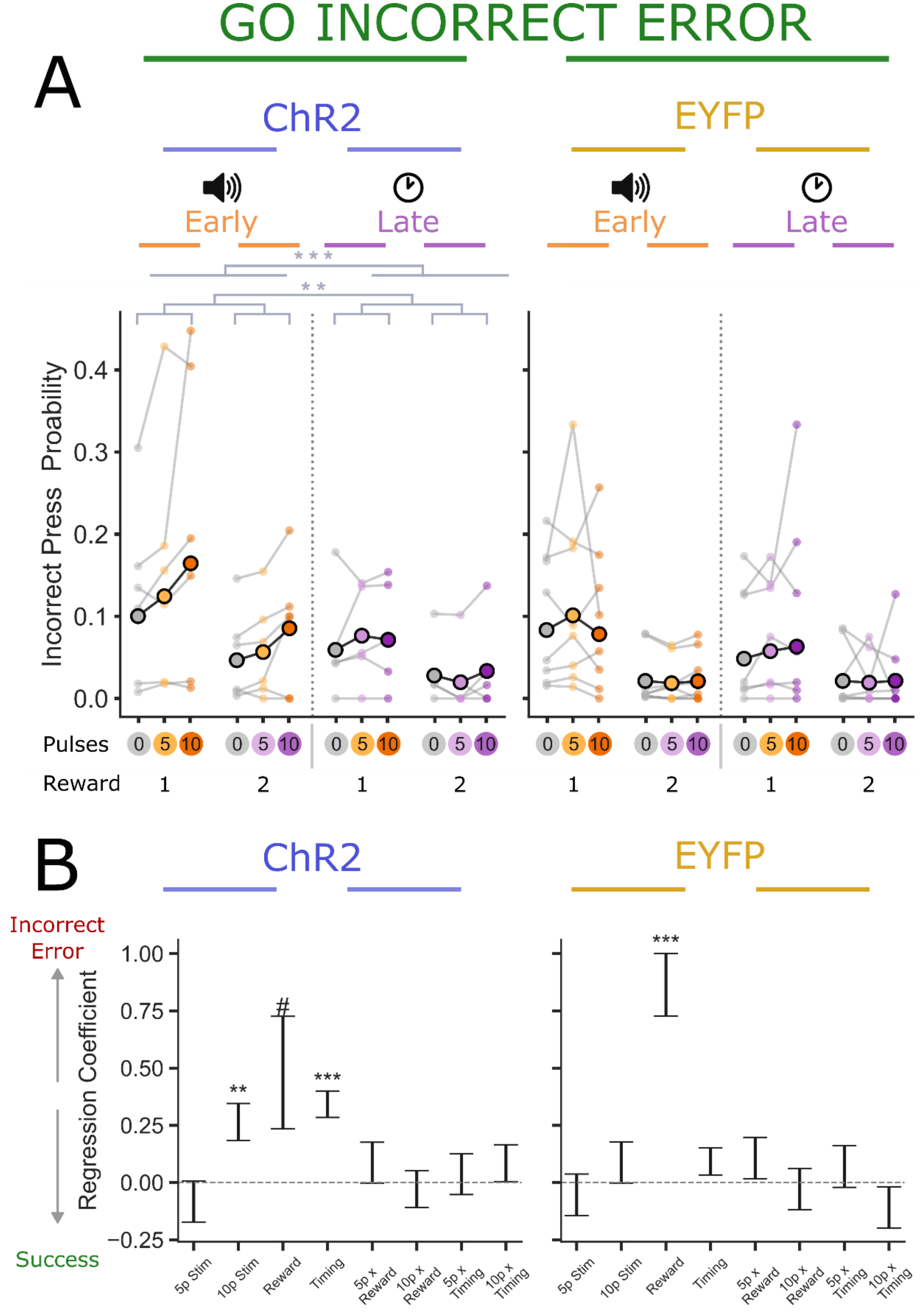


Figure 7. The higher magnitude stimulation and early timing block increase probability of an Incorrect Press error

A) Probability of incorrect press error across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent incorrect press probability over all trials. Faded points represent incorrect press probability for individual animals, data from the same individuals joined by a faded

line. B) Mixed-effects logistic regression on data from the ChR2 group (left) or the EYFP group predicting probability of incorrect press based upon stimulation, time of delivery, size of reward and their interactions. Three-way interactions coefficients not shown. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

[Main effect of group, $p = 0.041$]. There was also a trend for reaction times across all trials to be faster in the late stimulation block [Main effect of stimulation timing block $p = 0.07$].

After exiting the nosepoke, animals must then move to the lever on Go trials. The effects of stimulation on the latency to travel to the lever were found to be significantly modulated by all three other factors, and the 4-way interaction between all factors also trended towards significance [stimulation magnitude * timing block interaction $p < 0.001$; stimulation magnitude * reward interaction, $p = 0.04$, stimulation magnitude * group interaction $p = 0.011$; stimulation magnitude * timing block * reward * group interaction, $p = 0.095$]. To fully parse which effects were contributing to these interactions two models were fit to the dataset divided by group. In ChR2 animals, stimulation did not speed but *slowed* the latency at which animals moved from the nosepoke to the lever. This occurred in a manner that depended on the timepoint of stimulation delivery, with both the 5p and 10p stimulation only slowing latencies when applied early [Figure 9. Stimulation magnitude * timing block interaction; pairwise comparisons: early stimulation block, 0p vs 5p $p <0.001$, 0p vs 10p $p< 0.001$]. The effect of stimulation in ChR2 animals was not significantly modulated by the reward size on offer [all ps > 0.091] and in the model fit to just EYFP data no significant effects were detected [all ps > 0.094].

In summary, though stimulation did not affect the overall performance of ChR2 animals on Go trials, disruptions in efficient action execution were detected. These came in the form

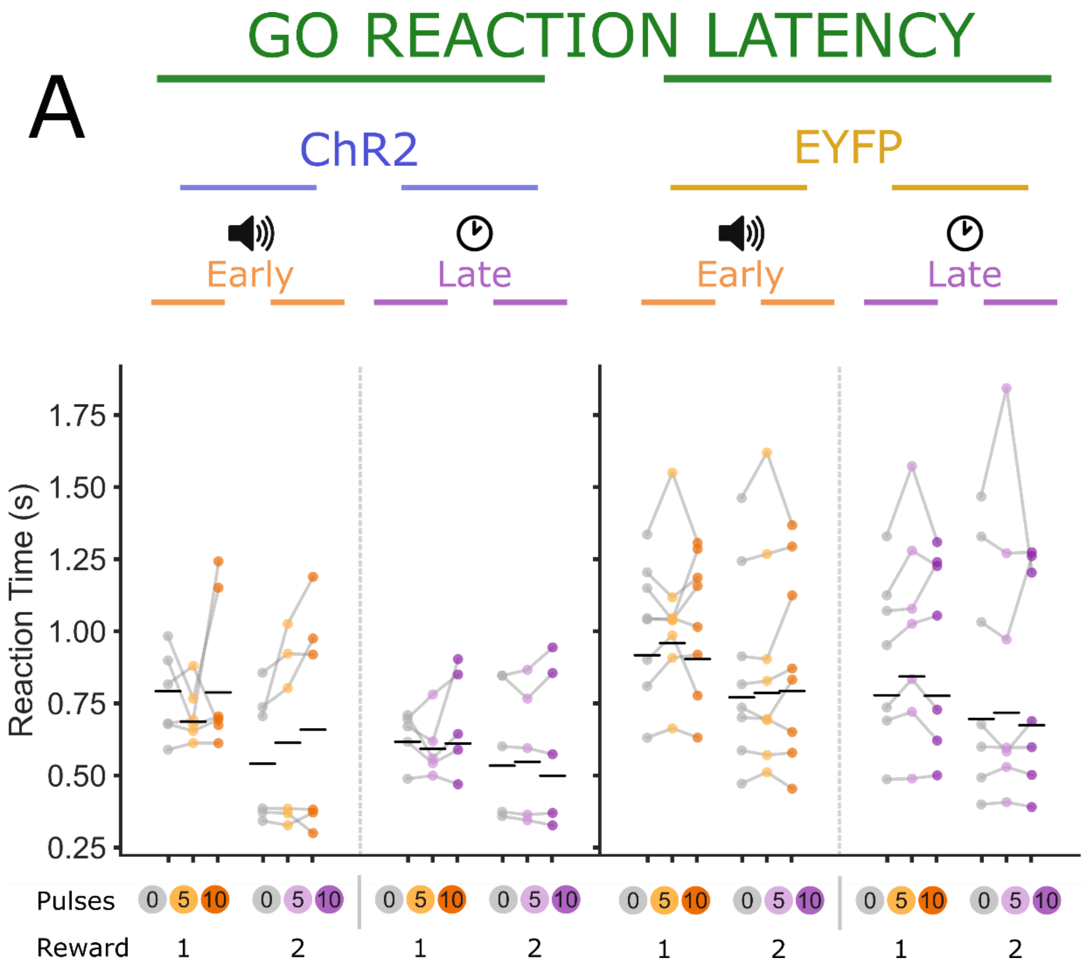


Figure 8. Stimulation does not significantly alter Reaction latency.

A) Reaction latency across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent mean latency over all trials. Faded points represent mean latency for individual animals, data from the same individuals joined by a faded line. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

---

of an increased likelihood to select the wrong lever, particularly when stimulation occurred early in the trial and when the 10 pulse stimulation was applied, and slower travel latencies to the lever following either stimulation early in the trial.

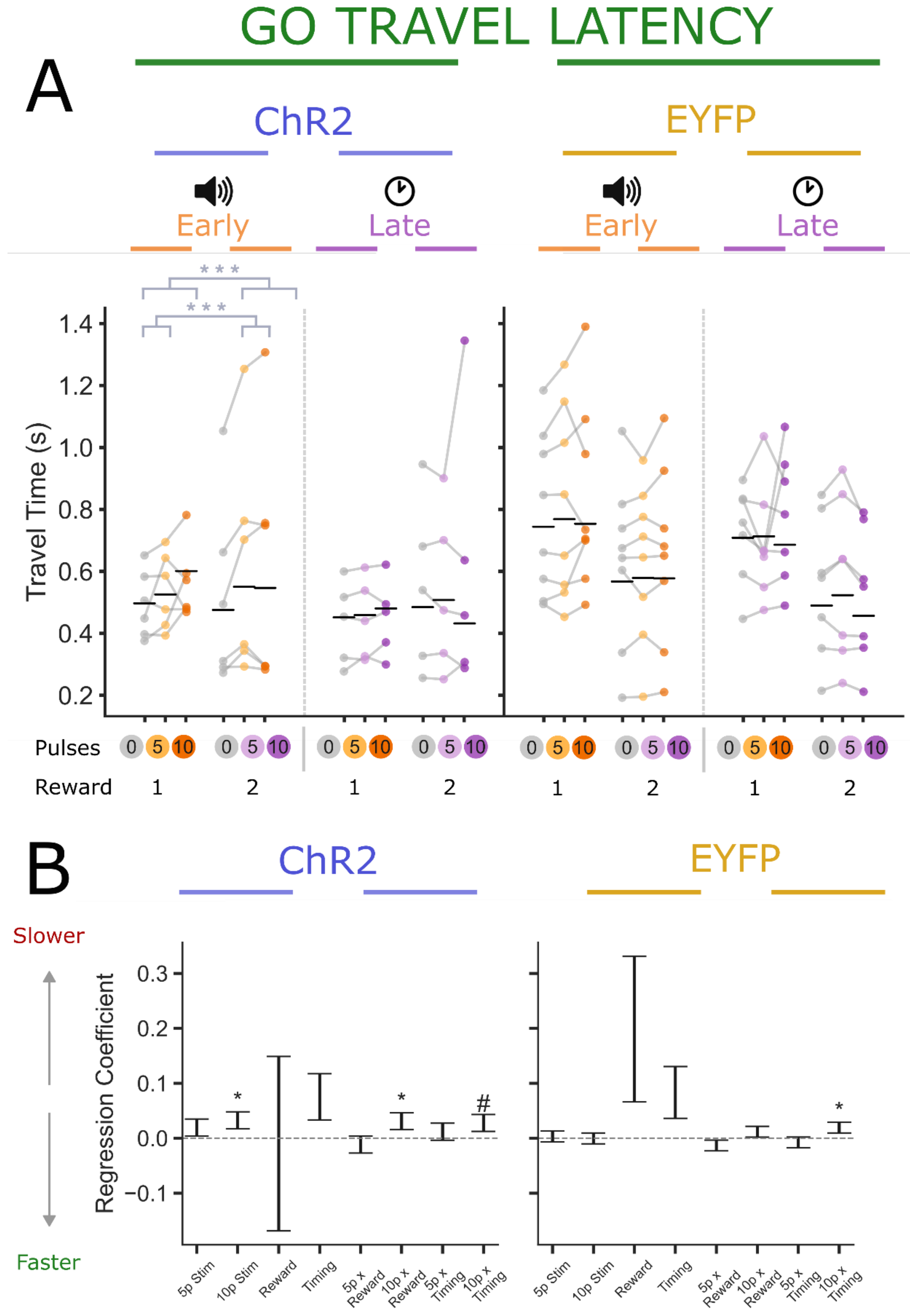


Figure 9. Stimulation slows Travel latency independent of magnitude.

A) Average Travel latency across conditions. Left panel, ChR2 animals; right panel EYFP animals. Left sub-panels with orange data from early blocks. Right sub-panels with purple data from late blocks. Filled points represent mean latency over all trials. Faded points represent mean latency for individual animals, data from the same individuals joined by a faded line. B) Mixed-effects logistic

regression on data from the ChR2 group (left) or the EYFP group predicting travel latency based upon stimulation, time of delivery, size of reward and their interactions. Three-way interactions coefficients not shown. # $p < 0.1$, * $p < 0.05$, ** $p<0.001$, *** $p<0.0001$.

3.4 - Discussion

Understanding the extent to which different facets of dopaminergic signalling may contribute to reward processing, movement initiation, or both, has proven to be a difficult challenge. As a paradigm that dissects action requirements and differing reward predictive cues the Go/No-Go task is a useful tool in progressing towards a comprehensive understanding of dopamine function in reward guided action.

Here we demonstrate that optogenetically evoked dopamine transients promote action over inaction, in a manner that depends both on the timing of these release events in relation to task relevant cues, and upon the effective dopamine concentration that is evoked downstream. These effects were felt most strongly on, but were not restricted to, trials receiving stimulation Optogenetic activation of VTA neurons proximal to cue presentation also affected Go trials. However, rather than acting to augment the effects of reward, which consistently improve and speed performance, stimulation actually disrupted action selection and vigour. Together, these results indicate that phasic dopamine transients do not act in a manner consistent with RPE theory, which would predict little effect on ongoing behaviour - but rather can causally promotion action initiation. These results align with previous studies demonstrating that NAc release is supressed following a cue instructing that action must be withheld (Syed et al., 2016).

Phasic dopamine transients causally promote transitions to action in a manner conditional on features of stimulation.

Chronic stimulation of dopamine transmission at D1Rs has been observed to promote impulsive action, both when testing cued action restraint (Grima et al., 2022) and in paradigms that measure action postponement such as the 5CSRTT (Pezze et al., 2007). However, the role that cue-evoked dopamine transients play in reward guided action has remained the subject of much debate. Here we demonstrate that in well trained animals, optogenetic activation of the VTA dopamine neurons can promote transitions to action even in the presence of a cue instructing restraint – but, crucially, this occurs in a conditional manner.

The shorter 5 pulse stimulation train used in these experiments is likely to evoke dopamine release closer to levels that might be encountered endogenously, with both the duration of stimulation and number of pulses being within the range of physiological burst firing patterns(Grace & Bunney, 1984; Hyland et al., 2002). Conversely, the longer 10 pulse stimulation is more likely to elicit supraphysiological release, and we know that in anesthetised animals the peak NAc dopamine release evoked by the 10 pulse stimulation is around double that evoked by the 5 pulse stimulation. However, when delivered shortly following No-Go cue presentation, there is no difference between the effect of the two stimulations – both promote inappropriate action initiation. This result maps to that of previous studies showing that NAc D1R activation mostly increases erroneous head exits following cue presentation, rather than later in the holding period(Grima et al., 2022).

Alternatively, when stimulation is delayed relative to cue presentation, and delivered when animals are already engaged in restraining action, the two stimulation trains now have

opposing effects. The 10-pulse stimulation again prompts inappropriate action, while the shorter 5-pulse stimulation now effects behaviour in the opposite direction – making successful restraint more likely.

Previous studies have found that activating SNc dopamine neurons in mice not engaged in a behavioural task produced movement only when a high magnitude stimulation was applied; stimulation titrated to within physiological ranges akin to unsignalled reward delivery had no effect (Coddington & Dudman, 2018). Taken together with our study, these results indicate that stimulation of dopamine neurons can be qualitatively different depending on magnitude. High magnitude stimulation appears to have a generally disruptive pro-movement effect. Conversely low magnitude stimulation appears to be sensitive to the state of the animal, in our study only promoting a switch to action initiation only when animals are primed to make a decision about whether or not act.

However, it is important to note that even supraphysiological stimulation of the VTA in mice in the Coddington and Dudman study was unable to promote movement, even at the highest magnitude. This highlights the relationship of VTA dopamine and *reward-associated* movement in particular. Indeed, other studies that have recorded movement effects following VTA activation have only observed them after animals have learnt to perform actions to obtain or consume rewards – and only at timepoints where initiating reward seeking action is a valid behavioural strategy. Hamid et al found that stimulation concurrent with a cue signalling trial availability decreased latencies to initiate reward seeking (Hamid et al., 2016). The authors propose a value-based account for this effect, suggesting that VTA stimulation evokes an increase in reward expectation that mediates an increased motivation to work. On the other hand, Yin et al posit a largely motor based

mechanism for movement effects resulting from VTA activation (Hughes R et al., 2020). In their task stimulation following (though not prior to) training on a self-initiated timing task results in the production of forward force, arguably a predictable effect given that the majority of VTA neurons they recorded had firing rates tightly correlated with forward force exertion.

Neither the data in the Hughes et al., nor Hamid et al., studies straightforwardly align with the effect of phasic VTA dopamine stimulation reported here. An increase in motivation would be expected to improve and speed, not impair performance. Equally, net forward force produced by VTA activation would be oppositional to the increase in backwards movement out of the nosepoke that we observe on failed trials. Instead, our results suggest that phasic dopamine stimulation may promote the initiation of reward seeking actions in a context dependent manner, which could enable a reinterpretation of these two experiments. The higher likelihood to begin reward seeking observed by Hamid et al could just as justifiably be mediated by an increase in the value of *acting*, rather than the value of the state itself. Equally, forward force is not the only behavioural effect recorded by Hughes et al, who also report an increase in anticipatory licking following stimulation; it is therefore possible that a more generalised reward approach action strategy involves both licking and *forward* movement as components, and is detected as forward force production in within the head fixed recording setup. Notably, the forward force producing effects of stimulation in this task are only observed *after* animals have experienced rewards within the context.

Effects of phasic dopamine stimulation on action restraint are not restricted to online behaviour.

Reports exist for activation of dopamine neurons either mediating effects upon behaviour immediately following stimulation or acting to change behaviour on future trials. Though clear effects are observed in this study for dopamine shaping online decisions about when to act, alterations in behaviour were also seen to extend to and accumulate on unstimulated trials.

Generally, without manipulation of dopamine neurons, error rates on No-Go trials are stable or even reduce slightly across a behavioural session (possibly as satiety acts to reduce the incentive drive to retrieve reward). By contrast, we observed in the ChR2 group that the number of inappropriate No-Go responses *increased* relative to the start of each session, as the number of stimulations an animal receives accumulates. As drugs that boost dopamine transmission have been shown to impair action restraint both in the current task (Grima et al., 2022), and related paradigms such as 5CSRTT (Pattij, Janssen, Vanderschuren, et al., 2007), this result could be tied to extracellular concentrations of dopamine increasing across the course of a session. However, impairment on unstimulated trials is only observed during the early block. Optogenetic activation should evoke similar release in both timing blocks – making a concentration-based explanation unlikely.

Many previous studies stimulating dopamine neurons observe reinforcement, with the option or action occurring concurrently with the stimulation more likely to be selected again (Hamid et al., 2016; Saddoris et al., 2015). Stimulation in this task could be reinforcing the behavioural strategy at the time of stimulation, in addition to changing online behaviour. In the early block, both the 5 pulse and 10 pulse stimulation have a pro-action

initiation effect – and therefore any reinforcing effects may be expected to also promote premature action. Though beyond the scope of this thesis, a more extensive analytic approach may involve employing logistic regressions to predict behaviour on the current trial according to features of preceding trials, such as action requirement, success and presence or absence of stimulation. This method would enable a more nuanced understanding of how previous experience that may influence successful restraint. For example, should the effect on unstimulated No-Go performance be due to reinforcement of premature action, then recently experiencing many stimulated No-Go trials that resulted in erroneous head exits, rather than successful restraint, may be expected to increase the probability of error on upcoming trials.

An accumulating impairment of unstimulated No-Go performance is not observed in the late block, when simulations are no longer delivered proximal to cue presentation. Previous studies have demonstrated that the potency or nature of the behavioural effects of dopamine neuron activation can differ when delivered at different timepoints within behavioural tasks. Lee et al demonstrated that while inhibiting dopamine neurons can disrupt anticipatory licking on future trials, this effect was timing dependent – being most efficacious when applied in a select window after reward delivery (2020). Hamid et al found that dopamine neuron stimulation would only reinforce future behaviour if delivered after a choice had been made, stimulation at the start of trial would instead promote action (2016).

This body of work suggests that phasic dopamine signals may be interpreted differently depending on timing, or proximity to different behavioural or environmental events. It has been posited that interneuron populations, in particular cholinergic interneurons, within

the striatum may define periods when dopamine can affect learning and periods when it can drive ongoing behaviour (Berke, 2018). It is possible for the differences observed between the early and late block could be attributed to differences in striatal receptivity to dopamine. Further studies would be needed to determine when patterns of cholinergic interneuron activity that have been proposed to regulate windows of plasticity (Morris et al., 2004; Yamanaka et al., 2018) occurred relative to task events and the timepoints of stimulation to confirm this hypothesis.

VTA stimulation disrupts aspects of correct action execution when delivered proximal to cue presentation.

Even in light of the impairment of No-Go performance following VTA dopamine activation - stimulation could still be hypothesised to boost performance on Go trials, where NAc dopamine release strongly reflects upcoming reward size(Syed et al., 2016). The larger reward offered on Go Large trials consistently results in better performance and faster movement latencies within this task (Grima et al., 2022). Should VTA activation act in a similar way to increasing reward size, we may expect a similar effect upon behaviour.

However, in no case did stimulating dopamine neurons result in improved performance on Go trials or more efficient reward pursuit in the form of faster action latencies. To the contrary, some small but significant disruptions to action execution were observed.

When the high magnitude stimulation was delivered, ChR2 animals were more likely to execute the wrong action and select the incorrect lever. This effect appears to be driven by timing of stimulation, with a numerically greater increase in these errors in the early

block. Note, however, that as these effects do not interact, and as animals encountered the early block first, it cannot be ruled out that incorrect action selection following a 10 pulse stimulation is compensated for as animals gain greater experience within the task rather than depending on being delivered concurrently with the cue.

A similar pattern of behavioural effects has been observed with systemic, but not intra-NAc D1 agonist delivery(Grima et al., 2022), indicating that these effects may arise from mesolimbic dopamine release in sites outside of the NAc. It has been demonstrated that driving dopamine neurons in the *SNc* can bias the selection of competing actions (Howard et al., 2017). Should stimulation of the VTA produce a similar effect this could explain an increase in incorrect errors, alongside the slowing of travel latencies that we also observe - as beginning to initiate alternative actions could comprise swift movement to the correct lever.

Summary

Here we have found that phasic activation of VTA dopamine neurons can promote action over inaction in a conditional manner. Applying a high magnitude stimulation promoted movement, while a low magnitude stimulation thought to be closer to physiological activity patterns promoted action when delivered after cue presentation, but strengthened restraint if delivered while animals were already engaged in withholding action. While these effects where felt most strongly on the trial receiving stimulation, we also found a concomitant effect for activation to bias toward premature action on future trials, and slight disruption to ongoing action performance.

4. Do mPFC subregions make separable contributions to reward guided action control?

The mPFC is implicated in executing and inhibiting action appropriately, and is functionally connected to the striatum, motor cortex and VTA - suggesting a key role in reward guided action control. However, subregions within the mPFC may make separable contributions to these functions. To determine the importance of the mPFC to reward guided action control, and to dissect regional differences in function, we inactivated the PL, IL and MO subregions while animals either made or withheld actions for reward. We found that while only the PL and IL are necessary for cued action restraint, all three subregions make a crucial contribution to action execution in a complex task space.

Contributions

For the experiments presented in this chapter, Mason Silveira and Mark Walton designed and planned the experiments, Mason Silveira conducted surgeries and collected behavioural data and George Jenkins analysed and presented the data with input from Mark Walton.

4.1 - Introduction

This thesis demonstrates that phasic dopamine transients can causally influence reward guided action. Prior work has shown that phasic transients are supressed when a cue instructs that action must be withheld (Syed et al., 2016), suggesting the involvement of regulatory mechanisms in shaping both mesolimbic dopamine and associated reward guided action.

One potential region of interest is medial frontal cortex (mPFC). Rodent mPFC has been implicated in a range of functions relevant to appropriate action control, including restraining impulsivity (Bari & Robbins, 2013; Kim & Lee, 2011) and using contextual cues to guide behaviour flexibly (Sharpe et al., 2019; Sharpe & Killcross, 2015). Rodent mPFC projects to both medial striatum, including NAc (Öngür & Price, 2000; Vertes, 2004), and to the VTA (Carr & Sesack, 2000; Geisler & Wise, 2008), and can alter both VTA dopamine neurone firing (Jo & Mizumori, 2016; Lodge, 2011; Tong et al., 1996) and downstream dopamine release (Jackson et al., 2001), making this region a potential source of regulatory afferent input to shape reward-guided action selection and inhibition.

The mPFC can be divided into subregions along a dorsoventral axis based on anatomical connectivity and cytoarchitecture, including the prelimbic cortex (PL), infralimbic cortex (IL) and medial orbital cortex (MO) (Öngür & Price, 2000), and disrupting the function of each has been shown to have distinct effects on behaviour (Capriles et al., 2003; Gourley & Taylor, 2016; Hardung et al., 2017b; Sierra-Mercado et al., 2010). A strong understanding of how the mPFC may regulate action control, and to what degree different subregions may exert separable influence on this process, remains unknown.

However, the mPFC is functionally connected to several other movement-related circuits, including dorsal striatum (Terra et al., 2020) and motor cortex (Narayanan & Laubach, 2006), giving this region multiple routes by which motor behaviour maybe shaped.

Appropriate action control involves using information, such as that provided by context or cues, to select the action strategy that best matches the current environmental state. Disrupting activity within the PL subregion in particular can impair the *activation* of appropriate behavioural responses, to both context and cue-associated information (Ishikawa et al., 2008b, 2008a; Sharpe & Killcross, 2015). Conversely, the more ventral IL subregion has been more strongly implicated in inhibiting inappropriate responding, as IL inactivation in discriminative stimuli tasks increases operant responding to cues that do *not* signal reward availability (Ishikawa et al., 2008a). These results map onto a more general proposition of the PL and IL promoting 'going' and 'stopping'(Gourley & Taylor, 2016), or the activation and inhibition of behaviour – placing the two subregions in opponent roles.

However, a number of studies specifically investigating the ability to withhold from action suggests there may be more nuance to the roles of mPFC subregions in action control. In the 5CSRTT task, which tests the capacity to postpone from acting while waiting for a stimulus, initial lesion studies found a role for ventral mPFC (mostly IL) in the inhibition of action while lesions that only encompassed more dorsal mPFC (including the PL) left inhibition intact (Chudasama & Muir, 2001; Passetti et al., 2002b). However later studies employed more refined reversible inactivations confirmed that the PL makes a contribution to action restraint (Paine et al., 2011; Pezze et al., 2014), and disruptions of PL in more simple reaction time tasks increase premature responding (Narayanan et al., 2006;

Risterucci et al., 2003). This body of work suggests that both the PL and IL are required for successful action postponement, consistent with connectivity of more dorsal regions of mPFC (including PL) to several other movement-related circuits, including dorsal striatum (Terra et al., 2020) and motor cortex (Narayanan & Laubach, 2006). Moreover,in a recent study optogenetic inhibition suggested that transient inhibition of IL effected a *reduction* in premature actions (Hardung et al., 2017). This study also selectively inhibited the MO, a subregion that has not been as intensely studied, and found no effect on premature responding but rather a slowing of reaction latencies.

Tasks that require animals to initiate *and* restrain action present a more complex environment where the appropriate action strategy must be matched to the current task state. Under these conditions, the mPFC has been shown to contribute to the restraint of action, independent of the subregion manipulated. Inactivating the mPFC in a subregion specific manner found that disrupting the PL, IL, or MO all result in a reduction in the ability of animals to restrain reward seeking action to avoid punishment (J. Verharen et al., 2019). A recent study using a task where cues indicated whether animals must either make, or refrain from making, a specific action to obtain reward found PL or IL inactivations to selectively disrupt restraint, while sparing reward seeking that required an operant response to be made (Capuzzo & Floresco, 2020).

While the aforementioned studies require appropriate action control, they only require one specific action, such as a lever press or magazine entry to be withheld. This allows compensatory mechanisms to take place (J. Verharen et al., 2019) and may not measure true restraint of *all* action. In addition, in light of the evidence implicating the mPFC in value based decision making (Orsini et al., 2018; Sul et al., 2010), many of the aforementioned

studies use paradigms with rewarding or aversive outcomes of a single magnitude. Therefore, understanding of how mPFC subregions contribute to regulation of action in a space where different action requirements must be balanced in the face of differing reward expectations is unknown.

Testing the effect of region-specific inactivations within the Go/No-Go task enables their contributions to cue guided action control to be examined. The paradigm provides an environment where multiple cues are used to guide behaviour, and action must be both restrained and initiated – functions that are associated with an intact mPFC. In addition, the mPFC is anatomically and functionally connected to the mesolimbic dopamine system and can shape activity and release in this pathway (Jackson et al., 2001; Jo et al., 2013; Jo & Mizumori, 2016; Lodge, 2011). Our comprehensive understanding of the effect of dopaminergic manipulations in this task will enable any effects of mPFC inactivation to be placed in the context of disrupting descending control of the mesolimbic dopamine system.

Aims

This experiment aims to determine the causal role of mPFC subregions by examining the effects of inactivating PL, IL and MO during reward guided action and restraint. In addition, any effects on performance will be compared to previous studies manipulating dopamine transmission within the task to ascertain the likelihood that mPFC regulation of action may involve interactions within the mesolimbic dopamine system. To address these questions animals well trained in the Go/No-Go task were implanted bilaterally with cannulae targeted to either the PL, IL or MO and task performance was recorded following infusions

of a mixture of the $GABA_a$ agonist baclofen and the $GABA_b$ agonist muscimol to reversibly inactivate the targeted region.

If the PL and IL are involved in the activation and inhibition of behaviour respectively, we would expect inactivation of PL to selectively impair Go performance while inactivation of IL would selectively impair No-Go performance. Alternatively, if we base predictions on recent tasks where action must be inhibited and executed within the same task, we may expect targeted inactivation of both regions to impair No-Go performance while sparing Go trials. While there is less evidence outlining how the MO may contribute to reward guided action initiation and restraint, based on approach-avoidance tasks where the MO is shown to contribute to restraint (J. Verharen et al., 2019), we would expect No-Go performance to be impaired following MO inactivation.

4.2 - Methods

All procedures were carried out in accordance with the UK Animals (Scientific Procedures) Act (1986). Animal (n=32) were trained in the Go/No-Go task and were maintained on a twelve-hour light/dark cycle. Animals housed in groups of three. All testing occurred in the light phase. During training and testing, access to food was restricted such that animal bodyweights were kept between 85%-90% of their free feeding weight. Water was available ad libitum. Animals were trained in two cohorts, both consisting of 16 animals. Prior to surgery for cannulae implantation for this study, animals in the first cohort had received systemic injections of the norepinephrine transporter (NET) inhibitor atomoxetine (saline, 0.1 mg/kg, 0.3 mg/kg, 1.0 mg/kg) in a counterbalanced Latin square design (data not presented here). Animals in the second cohort had received the MAGL inhibitor ABD-1970 by oral gavage (vehicle, 1mg/kg, 10mg/kg) (Data presented in Chapter 5).

Behavioural Task

Details of the behavioural task, training and apparatus are described in Chapter 2 (Methods). To summarise, animals were trained in the Go/No-Go task and tested when performance across trial types was stable. For these experiments each session lasted until 100 rewards had been received or 60 minutes had elapsed, whichever came first. Once animals had completed testing on the full task, they were retrained for one week on a simplified version of the task with only No-Go trials present (simplified Just-No-Go) and retested. Following testing in the simplified Just-No-Go task, animal underwent a week of retraining and testing in a version of the task with only Go trials present (simplified Just-Go).

Behavioural measures

Measures of performance and latency are outlined in Chapter 2 (Methods) but briefly reiterated in the table below.

Pharmacological manipulations

Details of surgical and infusion procedures are described in Chapter 2 (Methods), but to briefly reiterate animals were implanted with stainless steel guide cannula bilaterally into the medial prefrontal cortex to allow targeting of MO and IL (Cohort 1) and MO and PL (Cohort 2). Animals in cohort 1 were implanted with cannula that enabled the infusions in

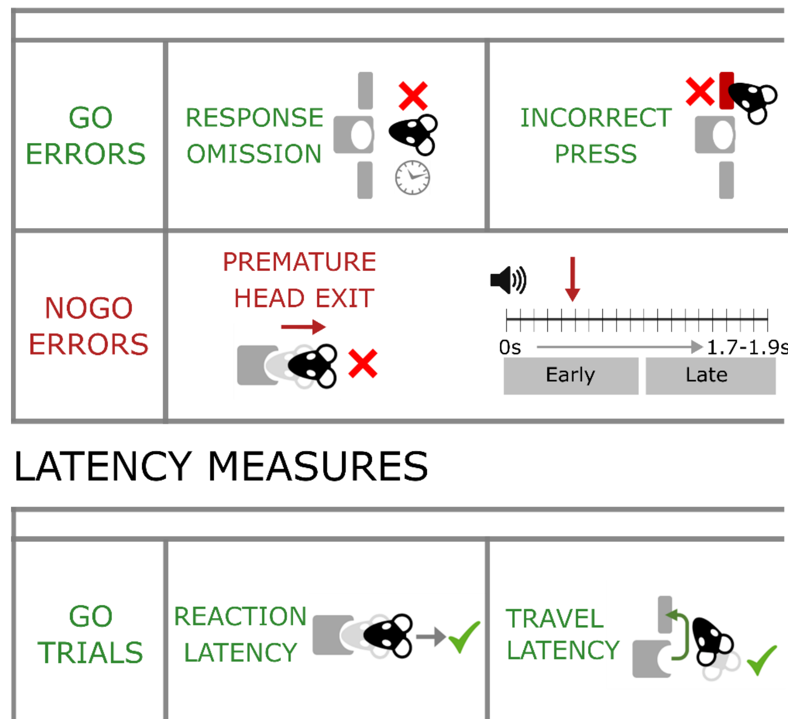


Table 1. Summary of performance and latency measures within the GoNoGo task

---

two sites, one more dorsal and one more ventral, with the aim to inactivate both the PL and the IL. However, cannulae were implanted consistently too ventral, so only data from the more dorsal infusions (IL/MO) and not the more ventral (dorsal peduncular cortex/unretrievable) were used in the final analysis. Animals in cohort 2 were implanted with cannulae that only targeted one site (PL/MO). Final cohort numbers were as follows, PL: n = 14, IL: n = 11, MO: n = 11. After achieving stable post-operative performance, animals received bilateral infusions with either saline or a mixture of the $GABA_B$ agonist baclofen and the $GABA_A$ agonist muscimol (Sigma–Aldrich), prepared separately at 0.5 µg/µL in 0.9% saline, and mixed together in equal volumes to form a 0.25 µg/µL solution. The baclofen/muscimol cocktail was injected bilaterally at a rate of 0.4 µL/min after which

injectors were left in place for a further minute. Testing then commenced 10 mins after completion. Infusions occurred on a 3-day cycle, with a baseline session, infusion day and washout day. Infusions occurred in a counterbalanced order and, to minimize carryover effects, at least one week separated every infusion round. In cohort 1, each animal received one infusion of vehicle and one infusion of baclofen/muscimol in the full task condition at each of the two cannulae sites - for a total of four infusions per animal. In cohort 2, animals receive one vehicle and one baclofen/muscimol infusion in the full task condition, and an additional vehicle infusion and baclofen/muscimol in both the simplified just-No-Go condition and the simplified Just-Go condition for a total of 6 infusions per animal. Further details of the infusion procedure, including histological recovery of projected locations of the injector tips can be found in Chapter 2 (Methods).

Statistical approaches

As outlined in Chapter 2 (Methods), a repeated measures ANOVA was used for analysis of data collected in these experiments. This included the effect of treatment as a two-level within-subjects factor (vehicle, inactivation) that of reward size as a two-level factor within-subjects (small reward, large reward) and that of region as a three-level between subjects' factor (PL, IL, MO). For analysis of the different experimental conditions, the effect of task type was also included as a two-level within subjects' factor (full task, simplified Just-Go or simplified Just-No-Go) For skewed latency distributions, a log-transformation was applied in order to meet the normality assumptions of statistical tests.

Video Tracking

The DeepLabCut toolbox (Nath et al., 2019) was used to track the position of the animal during each trial, as outlined in Chapter 2. Briefly, a network was trained to track body parts of the animal and features of the operant box. This enables trajectories of the animal's position on each trial to be obtained. On successful trials trajectories were tracked until 0.75s after entry into a region defined around the food magazine, or 5s after cue presentation – whichever came first. On failed Go trials, tracking lasted from cue onset until 5s after cue presentation and on failed No-Go trials tracking lasted from error onset until 5s after error onset. Files where trajectories could not be confidently recovered or where video and behavioural files could not be aligned were discarded. Only animals with video data passing inclusion criteria in every experimental condition were included in analysis (n = 31/32).

To analyse the order in which rats visited different zones of the operant chamber, the behavioural box was divided into a 3x3 grid, consisting of a left and right lever zone and a nosepoke zone on one side of the chamber, a food magazine zone on the other side of the chamber, and 5 neutral zones that did not contain any task relevant box features.

4.3 - Results

To determine whether there were pre-existing differences between animals with cannula implanted in different regions, we examined baseline performance. We looked at the last session of training, after animals had undergone surgery but before infusions had begun. There were no differences in Go performance between animals with cannula implanted in

different regions [No significant main effect of region or region * reward interaction all Fs < 1.1, all ps > 0.34]. When examining pre-infusion No-Go performance, we did find a reward size * region interaction [$F_{2,33} = 5.55$ $\eta_p^2 = 0.264$, $p = 0.009$; no significant main effect of region $F_{1,33} = 1.093$, $\eta_p^2 = 0.053$, $p = 0.348$]; however, there were no significant differences between groups with cannula implanted in different regions at either level of reward [pairwise comparisons: all comparisons at small or large reward $p > 0.091$]. Analysis revealed that the interaction was instead driven by the PL group in particular performing significantly better on No-Go trials when a large reward was on offer [pairwise comparisons: PL, small vs large $p < 0.001$].

First, to examine whether inactivation of mPFC subregions affected overall task engagement, we examined whether inactivation would alter the number of trials animals completed in a session. We found no effect of inactivation on the total number of trials completed [No significant main effect of treatment or treatment * region interaction, all Fs < 1.92, all ps > 0.174]. However, there were 2 animals from the PL group whose total number of trials after inactivation were > 3 standard deviations from the mean (4 and 5 trials), and so were labelled as outliers and excluded from subsequent analyses of performance and latency.

Distinct effects of inactivation of the mPFC subregions on action restraint.

Before investigating effects of inactivation on cued action restraint on No-Go trials, we first examined the frequency of aborted trials, where animals made a nosepoke to initiate a trial but exited the port before the pre-cue period elapsed. In none of the groups did

inactivation change the propensity of aborted trials [no significant main effect of treatment or treatment *region interaction all Fs < 0.49 all ps > 0.59], suggesting that inactivation of mPFC subregions did not induce a general increase in premature responding.

By contrast, inactivation impaired performance on No-Go trials in manner that depended on the region being targeted and reward size on offer [significant treatment * region * reward interaction: $F_{2,31} = 4.878$, $\eta_p^2 = 0.239$, $p = .0014$; no significant treatment * region interaction $F_{2,31} = 1.875$, $p = 0.17$, Figure 1]. Following PL infusions, No-Go performance was similarly impaired on both large and small reward trials [pairwise comparisons: PL group, both small reward trials and larger reward trials, vehicle vs inactivation $p < 0.001$]. When IL was inactivated, there was only a consistent impairment on small reward No-Go trials [pairwise comparisons: IL group, small reward trials, vehicle vs inactivation $p < 0.001$; large reward trials, vehicle vs inactivation $p = 0.091$]. In contrast, although there was a trend towards worse performance on large reward trials, overall No-Go performance was not significantly altered on either trial type when the MO was inactivated [pairwise comparisons: MO group, small reward trials, vehicle vs inactivation $p = 0.613$, large reward trials, vehicle vs inactivation $p = 0.071$].

Previous work has shown that there appear to be different processes involved in successful action restraint just after No-Go cue presentation or near the end of the No-Go period. Therefore, to determine how mPFC subregion inactivation altered *when* NoGo errors were occurring, the distribution of erroneous head exits in NoGo trials were plotted as a proportion of all NoGo head exits.

As can be observed in Figure 2, inactivation appeared to increase the likelihood of head exits at different timepoints depending on which region was perturbed. Specifically, PL

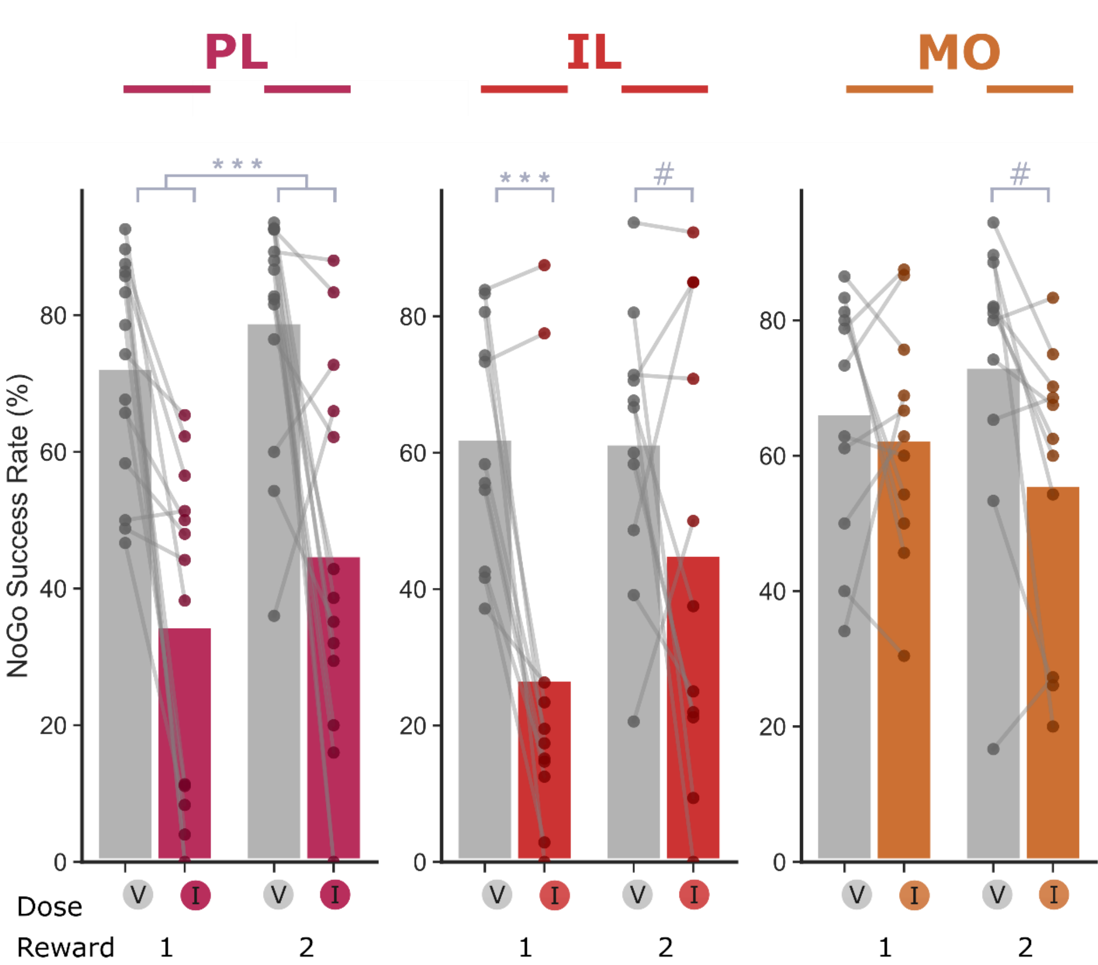


Figure 1. Inactivation impairs NoGo performance in a subregion specific manner.

Effect of inactivation upon NoGo performance. Vehicle sessions in grey. Inactivation sessions in pink (PL), red (IL) and orange (MO). Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

inactivation resulted in an increased likelihood to exit the nosepoke on both small and large reward trials, with a higher in the *latter* portion of the holding period for large trials in particular. While animals in the IL group displayed the greatest elevation in premature exits on small reward trials (Figure 2A). In agreement with examination of overall success rates,

nosepoke exits distributions from animals in the MO group were similar across inactivation and control sessions.

In order to investigate the apparent regional differences in the probability of head exit during NoGo trials, the holding period was divided into an 'early' (<800ms) portion and 'late' (>800ms) portion (Figure 2B). A four-way ANOVA revealed a significant four-way interaction [significant treatment * region * reward * error period interaction $F_{2,31} = 3.649$ $\eta_p^2 = 0.191$, $p = .0038$; three way interactions including treatment and region: all Fs > 3.78, all ps < 0.034].

To understand what was driving differences across the three regions, we broke down this effect by examining performance separately on small and large reward trials. When the reward on offer was small, inactivation increased the overall proportion of head exits in the PL and IL group, but not the MO group [significant treatment * region interaction: $F_{2,31} = 5.309$ $\eta_p^2 = 0.255$, $p = 0.01$, pairwise comparisons: PL, vehicle vs inactivation $p < 0.001$; IL, vehicle vs inactivation $p < 0.001$; MO, $p = 0.613$] consistent with a decrease in No-Go success rate. While head exits were more likely in the late period [significant main effect of error period: $F_{1,31} = 16.422$, $\eta_p^2 = 0.346$, $p < 0.001$] this was not altered by inactivation [no treatment * error period interaction: $F_{2,31} = 0.533$, $\eta_p^2 = 0.017$, $p = 0.471$]. Conversely, on large reward trials inactivation promoted head exits in the PL group only, and this increase was selective for the late period [significant treatment * region * error period interaction: $F_{2,31} = 5.209$ $\eta_p^2 = 0.252$, $p = 0.011$, pairwise comparisons: PL, early period, vehicle vs inactivation $p = 0.684$; late period, vehicle vs inactivation $p < 0.001$; IL and MO, all $p > 0.76$]. Together, this shows errors where increased differentially following

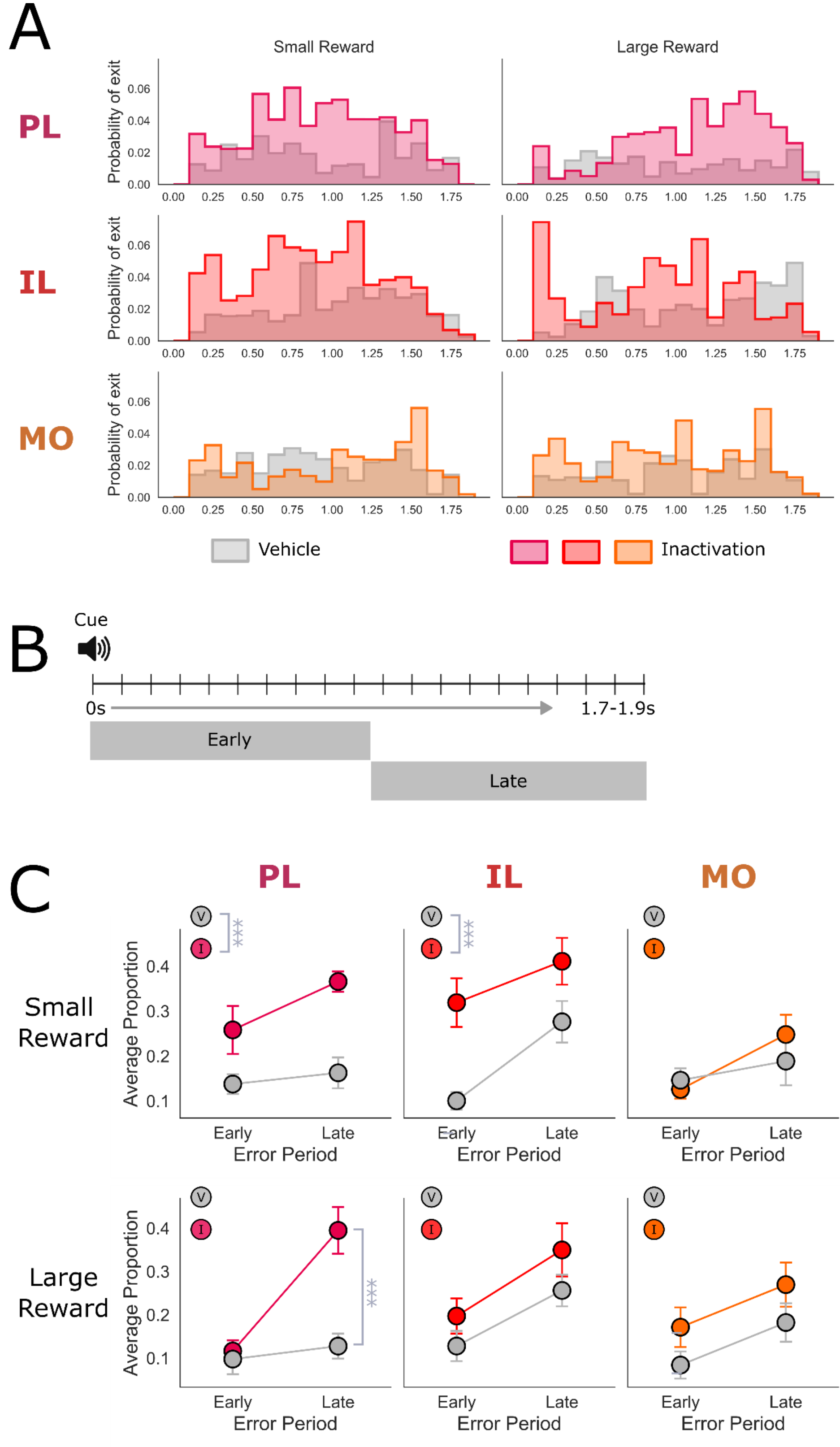


Figure 2. Inactivation differentially impairs erroneous head exit probability.

A) Average probability histograms of head exit times on Failed NoGo trials for Small Reward (left) or Large Reward (right) trials in vehicle sessions (grey), or inactivation sessions for PL (pink, top), IL (red, middle), MO (orange, bottom) calculated as probability over all trials, both successful and

failed. 0.1s bin. B) The NoGo holding time was divided into early exits (>800s) and late exits (<800ms). C) Mean proportions of nosepoke exits that were early or late dependent on treatment (grey: vehicle, pink (PL)/red, (IL) /orange (MO)): inactivation. # $p < 0.1$, * $p < 0.05$, **$p<0.001$,***$p<0.0001$.

---

inactivation of the three regions. PL inactivation increased errors in both periods on small reward trials, but only late errors on large reward trials. IL inactivation increased errors only on small reward trials, independent of error period, and MO inactivation had no effect on error proportions.

To examine whether inactivation altered behaviour after a No-Go error has been made, we next used video tracking to examine the first destination animals moved to after exiting the nosepoke early on NoGo trials. We focused on whether animals, in the 5s after prematurely exiting the nosepoke, would first move to the magazine or to one of the levers. Though generally animals more often moved first to a lever than the magazine on small reward trials, this effect did not interact with inactivation [significant reward * outcome interaction $F_{1,28} = 16.853$ $\eta_p^2 = 0.376$, $p < 0.001$: pairwise comparisons: on small reward trials, lever vs magazine $p < 0.001$], indicating that inactivation did not alter the pattern of zone visitation on erroneous NoGo trials [Figure 3B, no significant treatment * outcome or treatment * reward * outcome interaction $F_{1,28} < 0.56$ $p > 0.465$].

In summary, the effect of inactivation upon performance in NoGo trials depended on the region targeted, impairing action restraint when targeted to the PL or IL, but largely sparing performance when targeted to the MO. However, the nature of impairment also differed between PL and IL inactivations. Animals receiving IL inactivations were selectively

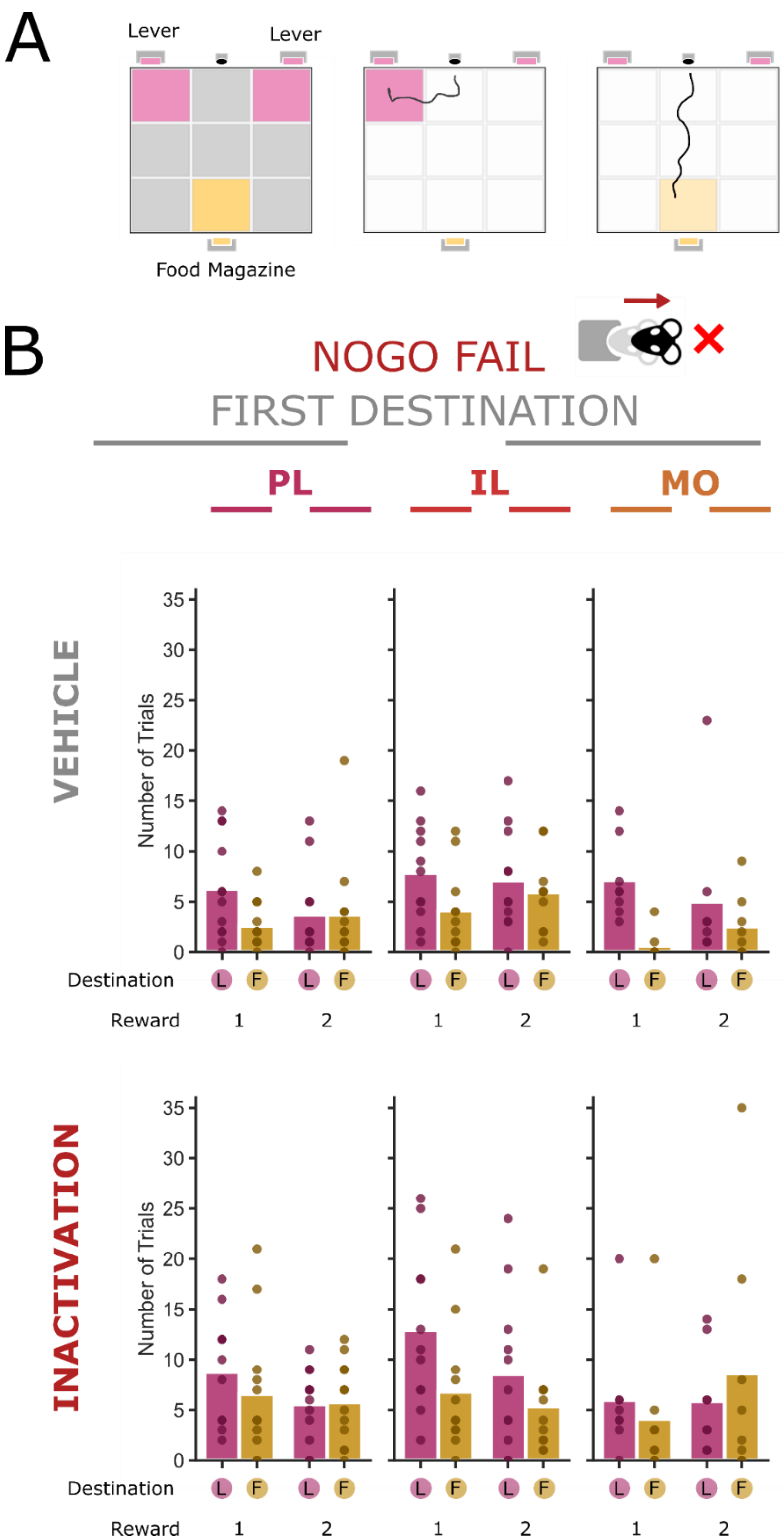


Figure 3. Inactivation does not alter visitation patterns on NoGo Fail trials.

A) Left, the operant box was divided into a 9x9 grid with one square for each lever (pink), one for the magazine (yellow) and 5 neutral zones. Middle example trajectory where animal moved a lever first after exiting the nosepoke. Right, example trajectory where animals moved to the magazine

first. B) Top, number of NoGo Fail trials where animals where animals first visited a lever (pink) or food magazine (yellow) in vehicle sessions. Data displayed from left to right for PL, IL MO groups. Bottom, same as Top but for inactivations sessions. Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

impaired on small reward trials, while those receiving PL inactivations were impaired on both small and large reward trials, with elevations in premature responses clustering particularly in the latter portion of the holding period when the reward on offer was large. However, infusions did not influence the first destination animals moved to after making an error, indicating that impairments in action restraint after inactivation could not be linked to a selective increase in drive to attempt reward collection or make an operant response.

Impairments of Go performance are uniform across regions

In contrast to the differential effects on NoGo performance, inactivation of all regions resulted in a consistent and substantial impairment of performance on Go trials [significant main effect of treatment: $F_{1,31} = 128.373$, $\eta_p^2 = 0.80$, $p < 0.001$; no treatment * region interaction: $F_{2,31} = 0.544$, $p = 0.586$, Figure 4]. To understand the nature of this impairment,

GO SUCCESS

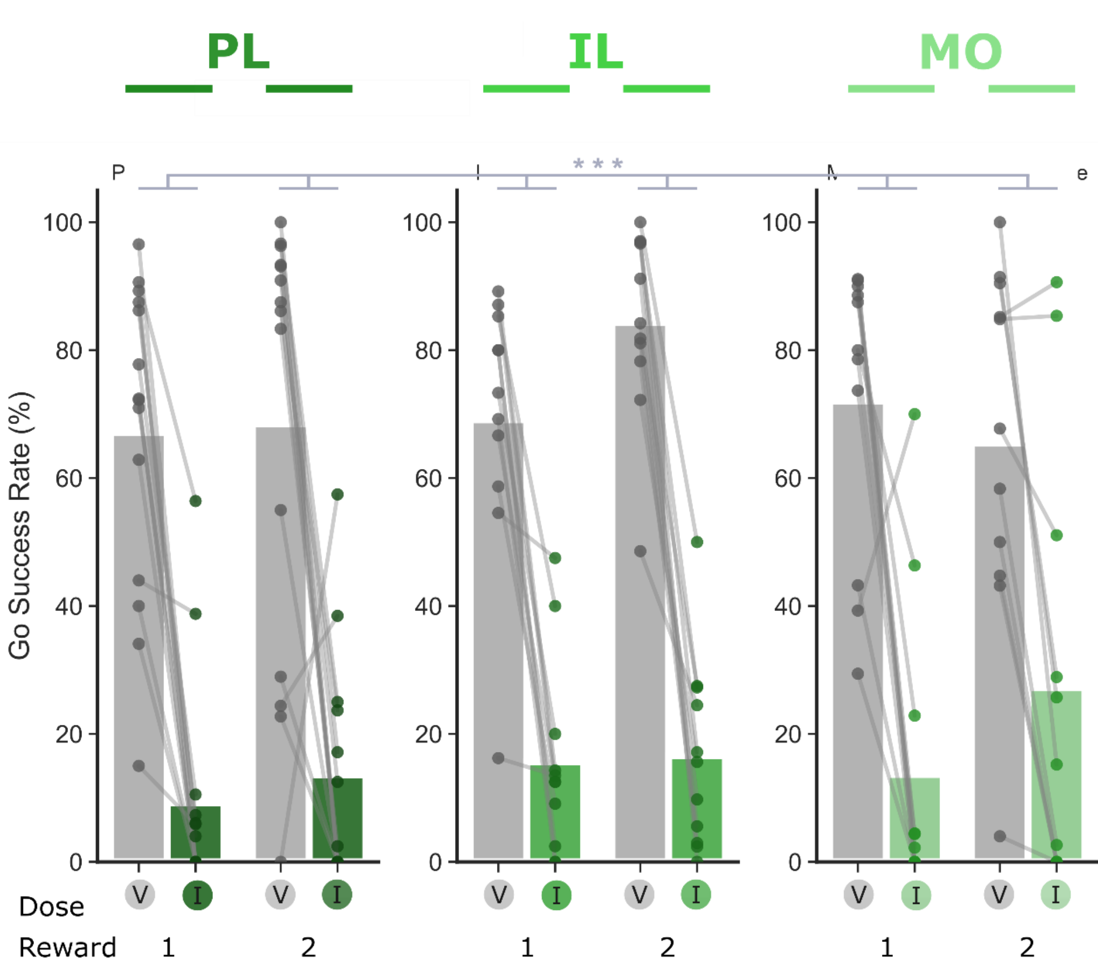


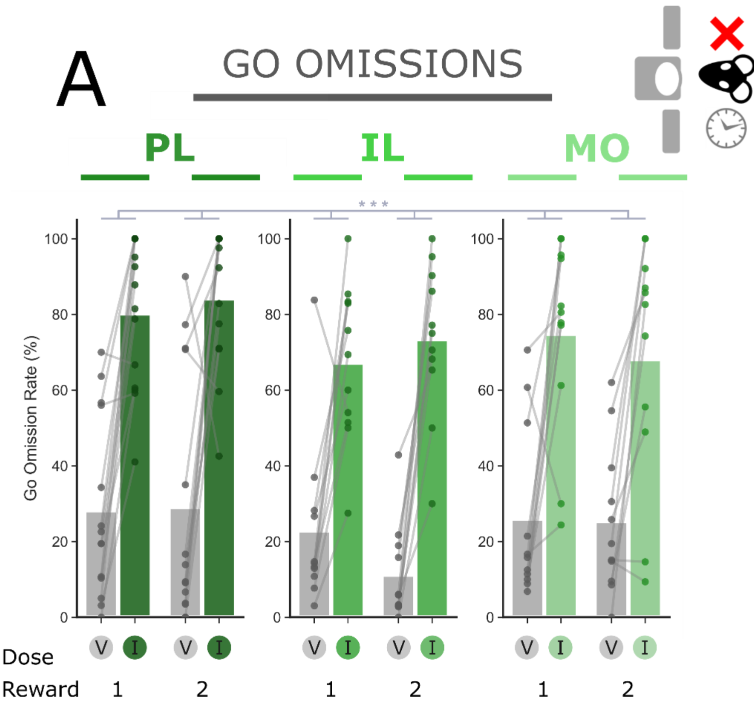
Figure 4. Inactivation impairs Go performance across subregions.

Effect of inactivation upon Go performance. Vehicle sessions in grey. Inactivation sessions in dark green (PL) mid green, (IL) and light green (MO). Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

---

the types of errors that animals made were analysed. When making a Go error, animals could either respond on the wrong lever (Go Incorrect Press) or fail to complete the required two lever presses on the correct lever within a 5s response window (Go Response Omission). Inactivation resulted in a large increase in the proportion of omitted trials, independent of region or reward on offer [Figure 5, Significant main effect of treatment:$F_{1,31} = 99.727$, $\eta_p^2 = 0.763$, $p < 0.001$; interactions between treatment * region:

GO ERRORS



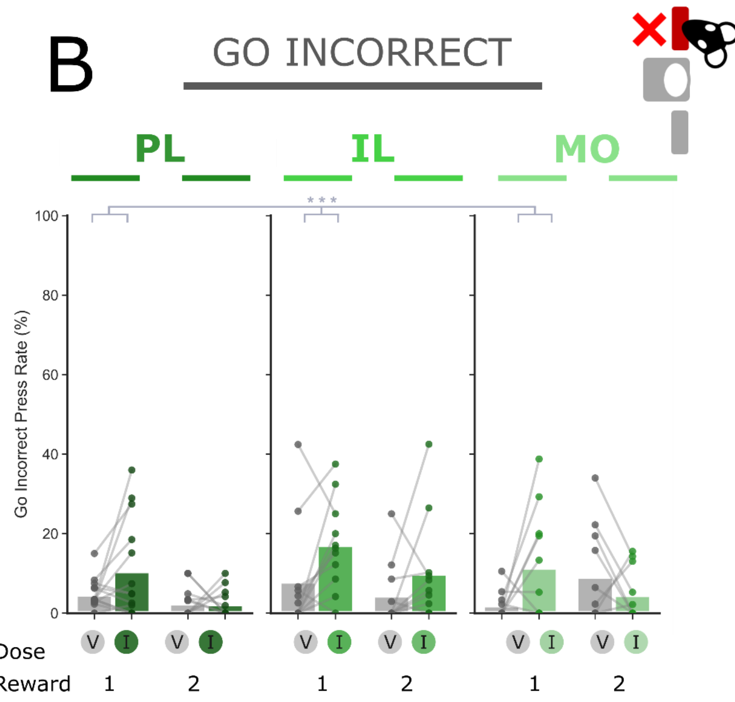


Figure 5. Inactivation increases Go errors across subregions.

A) Effect of inactivation upon Response Omission errors. Vehicle sessions in grey. Inactivation sessions in dark green (PL) mid green, (IL) and light green (MO). Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001. B) same as in A but for Incorrect press errors

---

all F < 1.523, p > 0.24]. Conversely, lever selection errors were only increased significantly on small reward trials [Figure 5, Significant treatment * reward interaction: $F_{2,31}$ = 9.024, η_p^2= 0.225, p = 0.005; significant pairwise comparisons on small reward trials: vehicle vs inactivation p < 0.001], but again this effect was equivalent across the subregions inactivated [treatment * region * reward interaction: $F_{2,31}$ = 1.102, p = 0.345].

To understand what might be driving the substantial increase in Omissions, we next examined response latencies as animals initiated and executed action on successful Go trials, to determine if behaviour was also impaired when animals completed trial requirements.

Inactivation may have affected the time it took animals to react to the Go cue and exit the nosepoke, with reaction times on large reward trials in particular displaying a trend to be slower in inactivation sessions [significant treatment * reward interaction: F1,16 = 4.829, η 2= 0.232, p = 0.043, pairwise comparisons: large reward, vehicle vs inactivation p = 0.002]. Again, this possible effect occurred across all groups regardless of the region targeted [no significant treatment * region *reward interaction, all F2,16 = 1.226, p = 0.32]. The effect upon latency to travel from the nosepoke to the lever was slowed in a manner that did depend on the region being perturbed [significant treatment *region interaction F1,16 = 6.011, η 2= 0.429, p = 0.011].~~The time it took animals to react to the Go cue and exit the nosepoke was slowed by inactivation in reward dependent manner [significant~~

~~treatment * reward interaction: $F_{1,16}$ = 4.829, η_p^2 = 0.232, p = 0.043, pairwise comparisons: large reward, vehicle vs inactivation p = 0.002]. However, again, this effect occurred across all groups, regardless of the region targeted [no significant treatment * region *reward interaction, all $F_{2,16}$ = 1.226, p = 0.32]. The effect upon latency to travel from the nosepoke to the lever *did* depend on the region being perturbed [significant treatment *region interaction $F_{1,16}$ = 6.011, η_p^2 = 0.429, p = 0.011].~~

Specifically, while inactivation of either PL or IL slowed travel times, inactivation of MO had no reliable effect [pairwise comparisons: IL, vehicle vs inactivation p < 0.001; PL, vehicle vs inactivation p = 0.019, MO, vehicle vs inactivation p =

REACTION LATENCY (S)		Vehicle	**Inactivation**	
PL	Small reward	0.82 ± 0.06	0.90 ± 0.10	
	Large reward	0.73 ± 0.09	1.35 ± 0.14	**
IL	Small reward	0.89 ± 0.07	1.15 ± 0.13	
	Large reward	0.77 ± 0.12	1.25 ± 0.19	**
MO	Small reward	0.93 ± 0.1	1.07 ± 0.13	
	Large reward	0.84 ± 0.1	1.31 ± 0.17	**

Table 2. Reaction latencies following inactivation. Presented as mean ± SEM.

TRAVEL LATENCY (S)		Vehicle	**Inactivation**	
PL	Small reward	1.36 ± 0.13	1.6 ± 0.23	*
	Large reward	1.1 ± 0.14	1.28 ± 0.1	*
IL	Small reward	1.04 ± 0.14	1.59 ± 0.14	***
	Large reward	0.77 ± 0.11	1.7 ± 0.28	***
	Small reward	1.05 ± 0.08	1.02 ± 0.17	

MO	Large reward	1.1 ± 0.17	1.11 ± 0.1

Table 3. Travel latencies following inactivation. Presented as mean ± SEM.

0.744]. Note, however, that latency measures could only be obtained from animals that completed Go trials of both reward sizes successfully in both vehicle and inactivation condition, meaning that the group sizes were reduced to; PL: n = 9, IL: n= 6, MO: n= 4.

Inactivation induces changes in underlying behaviour on Omission trials.

Inactivation induced a strong and significant impairment of Go trial performance across all subregions targeted that was largely mediated by an increase in Response Omission errors. While on successful Go trials or Incorrect Press trials animals must perform a stereotyped action sequence, behaviour on Response Omission trials can be more variable. To understand whether a change in behavioural strategy may underlie the increased Omission error rate, and whether changes may differ dependent on the subregion inactivated, video tracking again was employed.

We first examined the trajectories animals made on Go trials. The box was divided into a 9x9 grid, and for each trial individual squares were scored as 1 if the animal passed through that square during the trial and 0 if not. These were averaged to give proportional occupancy maps for each animal. A score of 0.5 for a square would indicate that square was occupied during 50% of trials. These proportional occupancy maps were then averaged across rats to give a single map for each trial type and condition.

Representing trajectories this way demonstrates that when animals successfully complete Go trials following inactivation, their stereotyped movements are similar to those made in

vehicle sessions (Figure 6). However, on Omission trials trajectories seem to change, with animals appearing to occupy a path through the box from the nosepoke to the lever less often in inactivation sessions (Figure 7).

~~We first examined the trajectories animals made on Go trials. Representing trajectories as proportional occupancy maps demonstrates that when animals successfully complete Go trials following inactivation, their stereotyped movements are similar to those made in vehicle sessions (Figure 6). However, on Omission trials trajectories seem to change, with animals appearing to occupy a path through the box from the nosepoke to the lever less often in inactivation sessions (Figure 7).~~

To determine if this was the case, we first examined the initial action made on Omission trials. An Omission trial could result from an animal attempting the right sequence of actions, moving first to the correct lever as they would on a successful trial, but failing to complete the action sequence correctly. Alternatively, animals could begin to execute a different action, or no action at all.

Patterns of initial zone visitation were altered following inactivation [Figure 8B, significant treatment* zone interaction $F_{2,56} = 10.092$, $\eta_p^2 = 0.265$, $p < 0.001$]. In the vehicle condition animals often initiated the correct action sequence, most frequently moving directly to the lever indicated by the cue rather than the wrong lever or the magazine [pairwise comparisons, vehicle, correct lever vs magazine $p < 0.001$; correct lever vs wrong lever $p < 0.001$]. Inactivation altered the strategies animals employed on Omission trials, as following infusion of baclofen/muscimol animals no longer overwhelming move more towards the correct lever, but now just as frequently move directly to the magazine [pairwise comparison: inactivation, correct lever vs magazine $p = 0.88$]. This effect occurred

independent of region targeted, though a three-way interaction trended toward significance [treatment *region * zone interaction: $F_{4,36} = 2.419$, $p = 0.077$]. Inactivation also increased the overall frequency of erroneous visits to each of the three zones [pairwise comparison: all outcomes, vehicle vs inactivation all ps < 0.005], consistent with an overall increase in Omission rates.

In addition to examining the first destination animals moved to, we also investigated the cumulative likelihood that animals would visit task-relevant zones in the 5s following cue presentation. In line with less frequent direct correct lever visits relative to other areas in the box following inactivation, the probability of entering the correct lever square within the 5s animals had to respond was also reduced [Figure 8, Significant main effect of treatment: $F_{1,28} = 20.317$, $\eta_p^2 = 0.42$, $p < 0.001$]. Again, this effect did not depend on the region targeted [no significant treatment * region interaction $F_{2,28} = 0.049$, $p = 0.952$]. The decrease in correct lever visitation was not driven by animals spending more time near the magazine, as animals were no more or less likely to visit the magazine zone within the same time period [no significant main effect of treatment or interactions, all Fs < 1.471, all ps > 0.23].

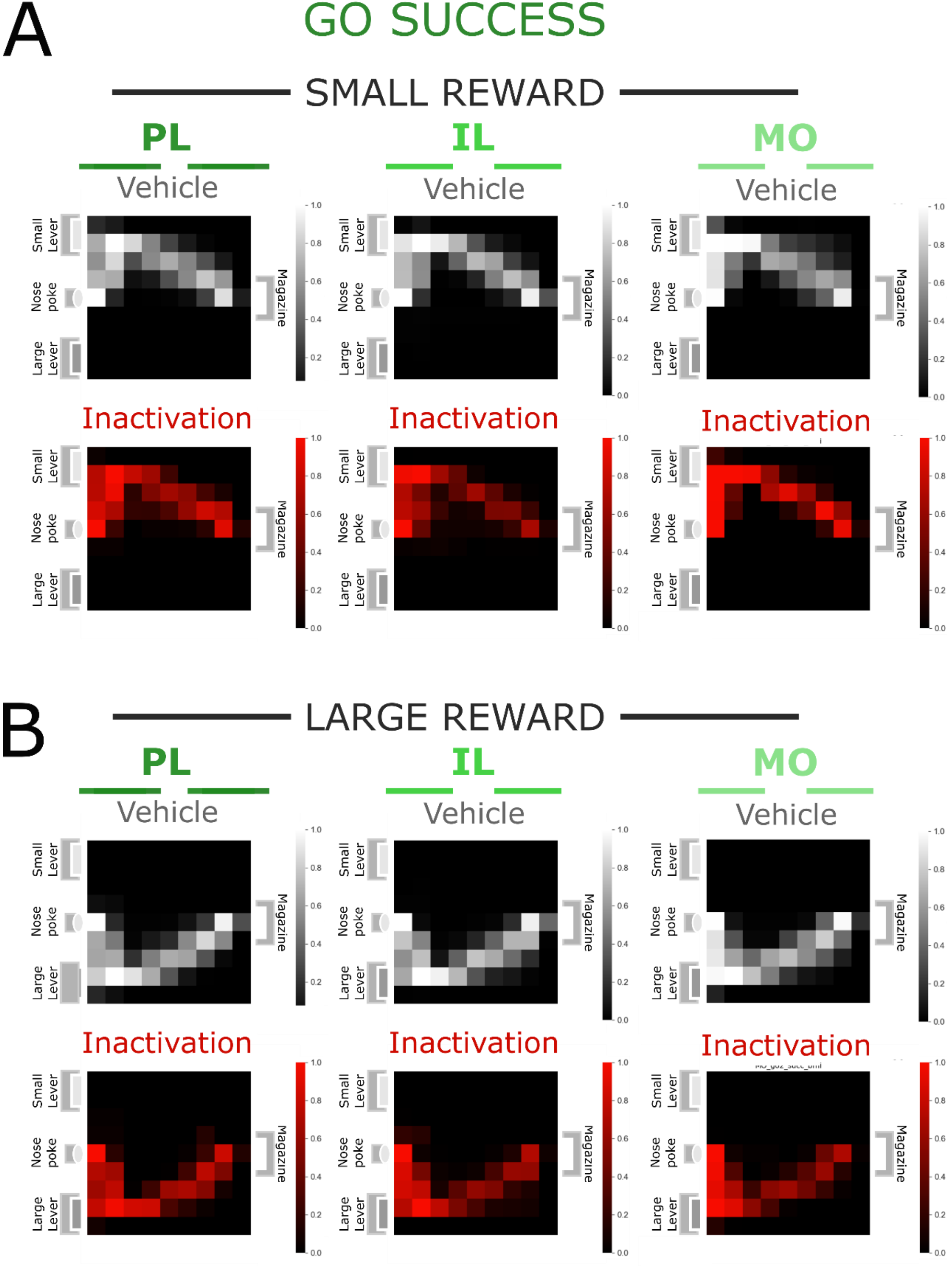


Figure 6. Proportional occupancy on Successful Go trials

Heat maps displaying the average proportion of trials in each condition where each square in a 9x9 grid of the box floor was occupied eg a value of 0.5 means animals moved through that square on half of trials. A) Small reward Successful Go trials in vehicle sessions (top) and inactivation session (bottom) for the PL (left), IL (middle) and MO (right) groups. B) same as in A but for Large reward Successful Go trials.

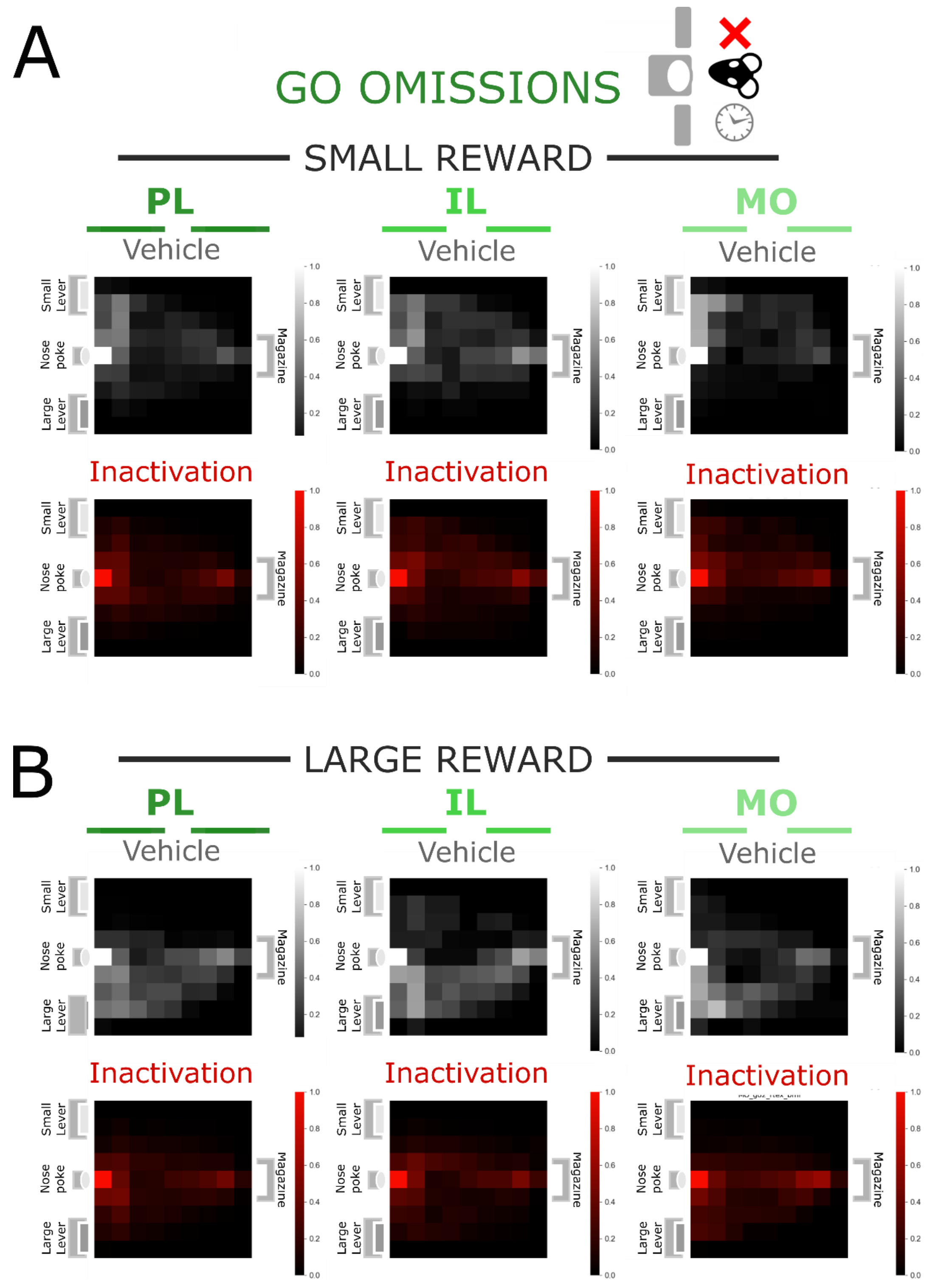


Figure 7. Proportional occupancy on Omission trials

Heat maps displaying the average proportion of trials in each condition where each square in a 9x9 grid of the box floor was occupied eg a value of 0.5 means animals moved through that square on half of trials. A) Small reward Omission trials in vehicle sessions (top) and inactivation sessions (bottom) for the PL (left), IL (middle) and MO (right) groups. B) same as in A but for Large reward Omission trials.

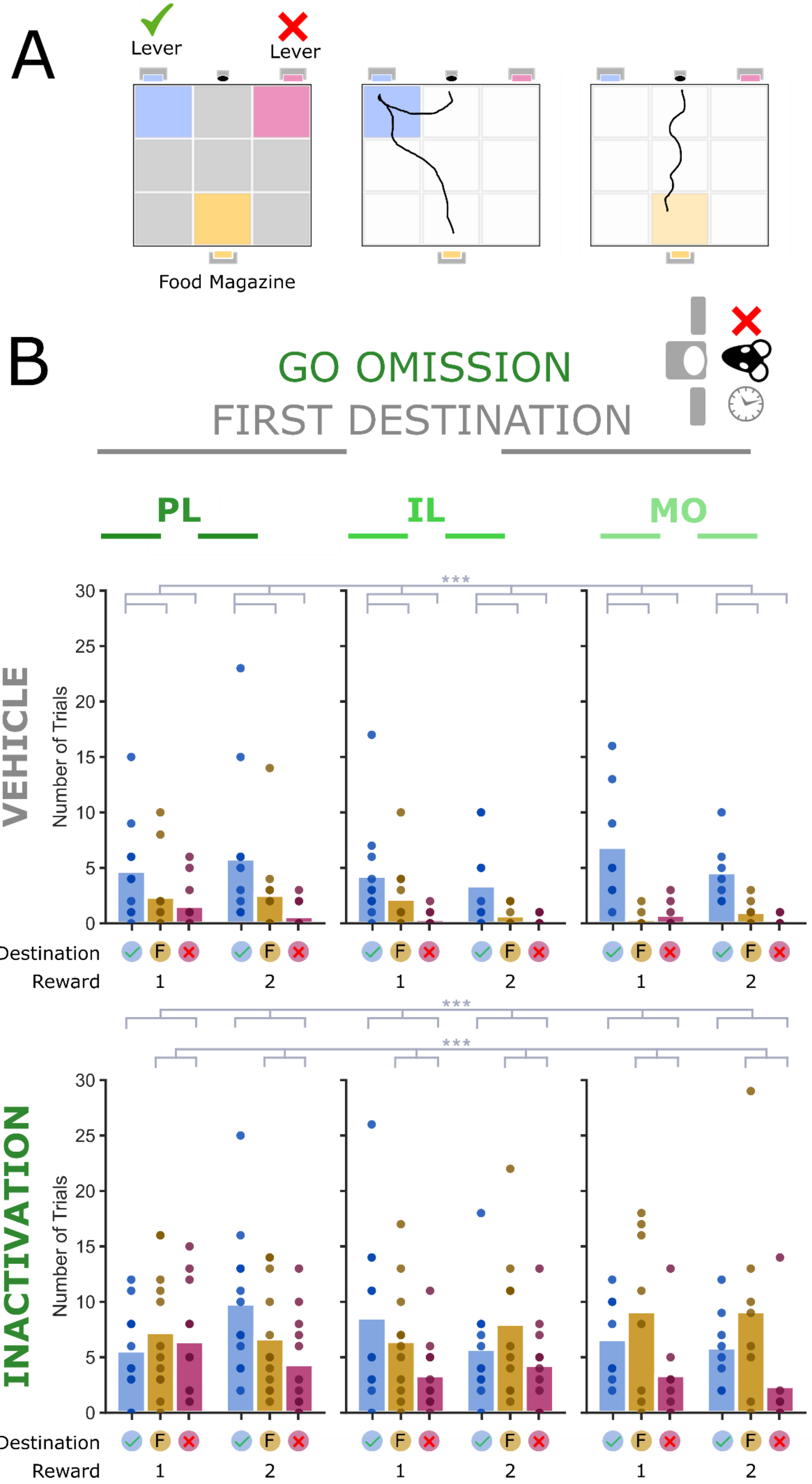

A
Lever
Lever
Food Magazine
B
GO OMISSION
FIRST DESTINATION
PL
IL
MO
VEHICLE
Number of Trials
Destination
Reward
INACTIVATION
Number of Trials
Destination
Reward

Figure 8. Inactivation alters visitation patterns on Response Omission trials

A) Left, the operant box was divided into a 9x9 grid with one square fir the nosepoke, one for each lever (correct: blue, wrong: pink), one for the magazine (yellow) and 5 neutral zones. Middle example trajectory where animal moved the correct lever first. Right, example trajectory where animals moved the magazine first. B) Top, number of Omission trials animals first visited the correct lever (blue) in correct lever (pink) or food magazine (yellow) in vehicle sessions. Data displayed from left to right for PL, IL MO groups. Bottom, same as Top but for inactivations sessions. Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

~~To understand how inactivation may alter general locomotion we also examined general immobility across the course of a session - with an immobile bout defined as animals remaining in the same zone for more than 5s. To account for the fact that sessions with an increased error rate would take longer to complete, and therefore present more opportunities for immobile bouts, the normalized rate of immobile events per min was examined. Animals were no more likely to enter into immobile bouts over the course of the inactivation sessions compared to control sessions [Figure 9, no significant main effect of treatment or treatment * region interaction all $Fs < 0.61$, all $ps > 0.5$]. However, there was a small but significant increase in the average duration of immobile periods [significant main effect of treatment $F_{2,28} = 7.334$, $p = 0.011$], which also not depend on region [no significant region * treatment interaction $F < 0.264$, $p > 0.26$].~~

To understand how inactivation may alter general locomotion we also examined general immobility across the course of a session - with an immobile bout defined as animals remaining in the same zone for more than 5s. To account for the fact that sessions with an increased error rate would take longer to complete, and therefore present more opportunities for immobile bouts, the normalized rate of immobile events per min was

examined. Animals were no more likely to enter into immobile bouts over the course of the inactivation sessions compared to control sessions [Figure 9, no significant main effect of treatment or treatment * region interaction all Fs < 0.61, all ps > 0.5]. However, there was a small but significant increase in the average duration of immobile periods [significant main effect of treatment $F_{2,28} = 7.334$, $p = 0.011$], which also not depend on region [no significant region * treatment interaction $F < 0.264$, $p > 0.26$].

In summary, inactivation targeted to either the PL, IL or MO results in an equivalent behavioural effect on Go trials. Animals display significant impairments in Go performance,

Figure 9. Immobility analysis

A) Effect of inactivation upon the rate of immobile bouts per minute. Vehicle sessions in grey. Inactivation sessions in dark green (PL) mid green, (IL) and light green (MO). Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$. B) same as in A but for average immobile bout duration.

mediated by an increase in small reward Incorrect press errors and a dramatic increase in Response Omission errors. Video tracking demonstrated that this increase in Omissions in all cases was mediated by a reduction in the likelihood of initiating the correct action and moving towards the correct lever following cue presentation.

Inactivation induced impairments are not rescued by reducing task complexity.

Inactivation of mPFC subregions results in profound impairments on Go trial performance, as well as subregion specific impairments in No-Go performance. One possible mechanism by which mPFC inactivation may impair performance is through disrupting animals' ability to differentiate or switch between Go and No-Go trials. Attempting to perform the incorrect action requirement would likely result in a failed trial in either case, as moving before the end of the holding period is a No-Go error, and waiting too long on Go trials increases the likelihood of timing out and omitting the trial. To examine this, following completion of Go/No-Go testing, animals in the second cohort (PL: n=10; MO n=7) were

retrained on two simplified versions of the task – first, a version with just NoGo trials (simplified Just-No-Go) and then a version with only Go trials (simplified Just-Go) – before being tested following infusions as before.

In an analysis including NoGo trials from both the standard task condition and the modified just NoGo condition, inactivation impaired NoGo performance [Figure 11, Significant main effect of treatment: $F_{1,15} = 12.66$, $\eta_p^2 = 0.458$, $p = 0.003$; trend treatment * reward * region interaction: $F_{1,1}5 = 3.385$ $\eta_p^2 = 0.18$, $p = 0.086$]. Importantly, no treatment by task version interaction was detected [all treatment * task version interactions, $Fs < 1.521$, $p > 0.236$], indicating that reducing the complexity of that task did not change the effect of inactivation. Similarly, although overall Go performance was improved by removing No-Go trials [Figure 12, Main effect of task version $F_{1,15} = 8.037$, $\eta_p^2 = 0.349$, $p = 0.013$], Go success rates were still dramatically compromised in the simplified "Just Go" condition following inactivation of either PL or MO [Main effect of treatment $F_{1,15} = 85.142$, $\eta_p^2 = 0.85$, $p < 0.001$; task * treatment interactions: all $Fs < 1.19$, all $ps > 0.291$].

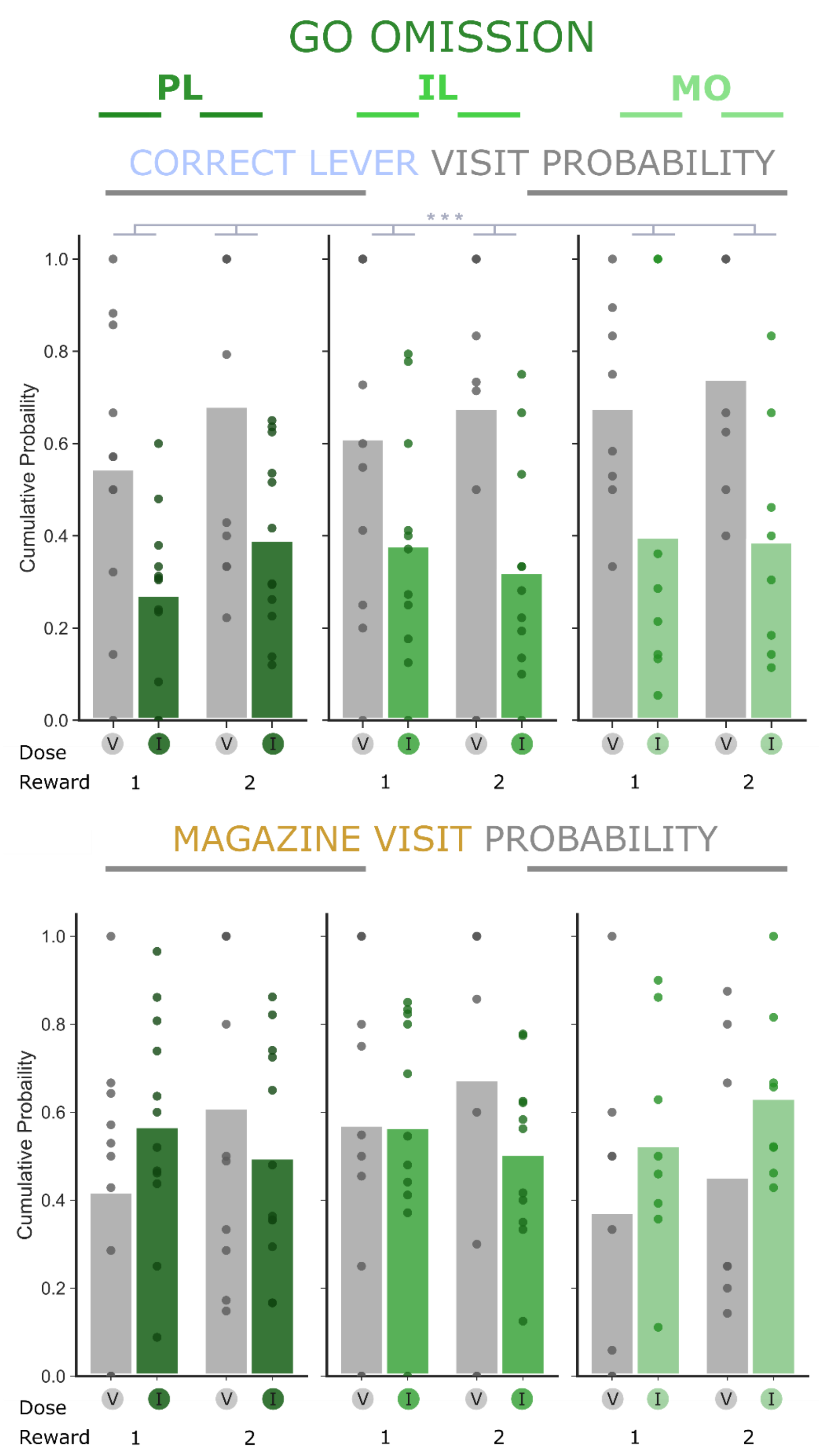


Figure 10. Inactivation reduces the probability to visit the correct lever.

Cumulative probability to visit the correct lever square in the 5s following Go cue presentation for vehicle sessions (grey) and inactivation session for the PL (dark green, left), IL (mid green, middle), MO (light green, right). Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$

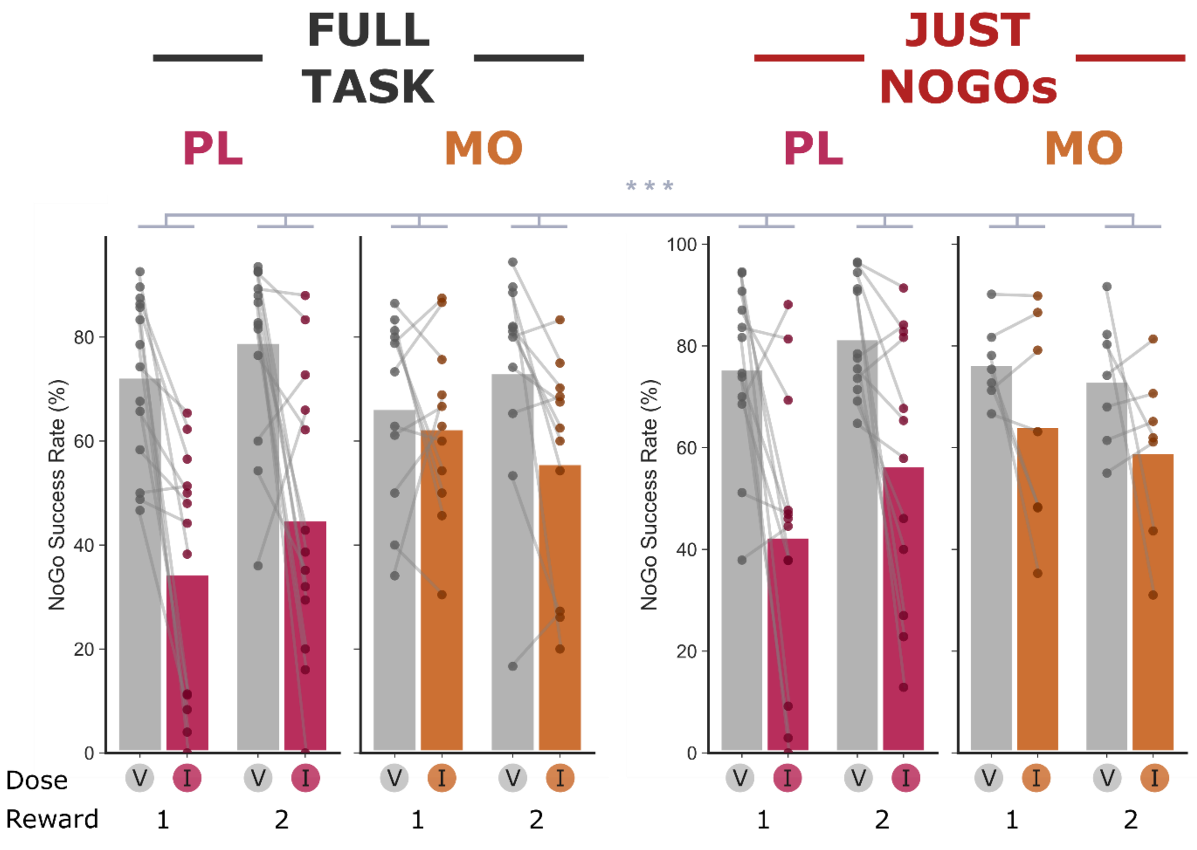


Figure 11. Reducing task complexity does not rescue effect of inactivation on NoGo performance

Left, Effect of inactivation upon NoGo performance in the full task condition, vehicle sessions in grey. Inactivation sessions in pink (PL) and orange (MO). Right, same as in top but in a reduced version of the task including only NoGo trials. Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

While inactivation impaired Go performance in both task versions, the patterns of Go errors underlying this impairment could differ between the full task and simplified versions. To determine if this was the case, we compared how inactivation changed Response Omission rates and Incorrect Press rates across both versions of the task. In line with an overall improvement in Go performance in the simplified Just-Go condition, animals made fewer Omission errors when only Go cues were present. However, inactivation of either PL or MO uniformly increased Response Omission errors [Figure 13,

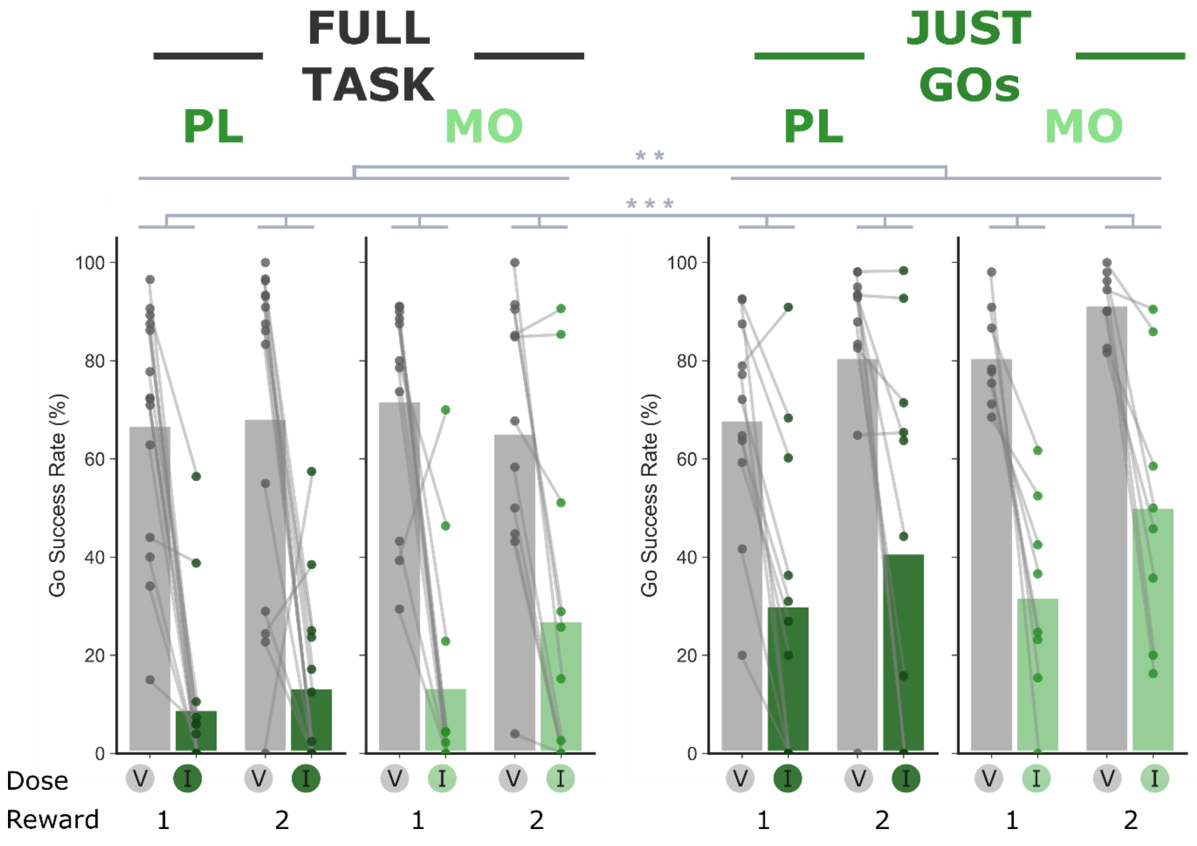


Figure 12. Reducing task complexity improves Go performance but does not rescue effect of inactivation.

Left, Effect of inactivation upon Go performance in the full task condition, vehicle sessions in grey. Inactivation sessions in dark green (PL) and light green (MO). Right, same as in top but in a reduced version of the task including only Go trials. Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

Main effect of treatment $F_{1,15}$ = 75.0.82, η_p^2= 0.833, p < 0.001], again in manner that did not interact with task version [no significant treatment * task version interactions all F < 1.372 p > 0.26]. Incorrect press error rates were equivalent across task versions [Figure 13, no significant main effect of task version $F_{1,15}$ =2.399, p = 0.14], and in line with the previous analysis incorrect press errors were only reliably increased on small reward trials [significant treatment * reward interaction: $F_{1,15}$ = 16.087, η_p^2= 0.517, p =0.001]. However, interactions between treatment and task version trended towards significance [trend

treatment * region * task version interaction: $F_{1,15} = 4.289$, $\eta_p^2 = 0.222$, $p = 0.056$; trend treatment * task version interaction: $F_{1,15} = 3.825$, $\eta_p^2 = 0.203$, $p = 0.06$], driven by MO animals in particular making more incorrect press errors in the simplified Just-Go condition [pairwise comparisons: MO, inactivation session, Just-Go vs full task $p = 0.007$; PL, inactivation session, Just-Go vs full task $p = 0.627$]

In summary, even on simplified versions of the task, inactivation still produced impairments in performance. This indicates that reduced performance following inactivation is not caused by an inability to switch between competing action requirements.

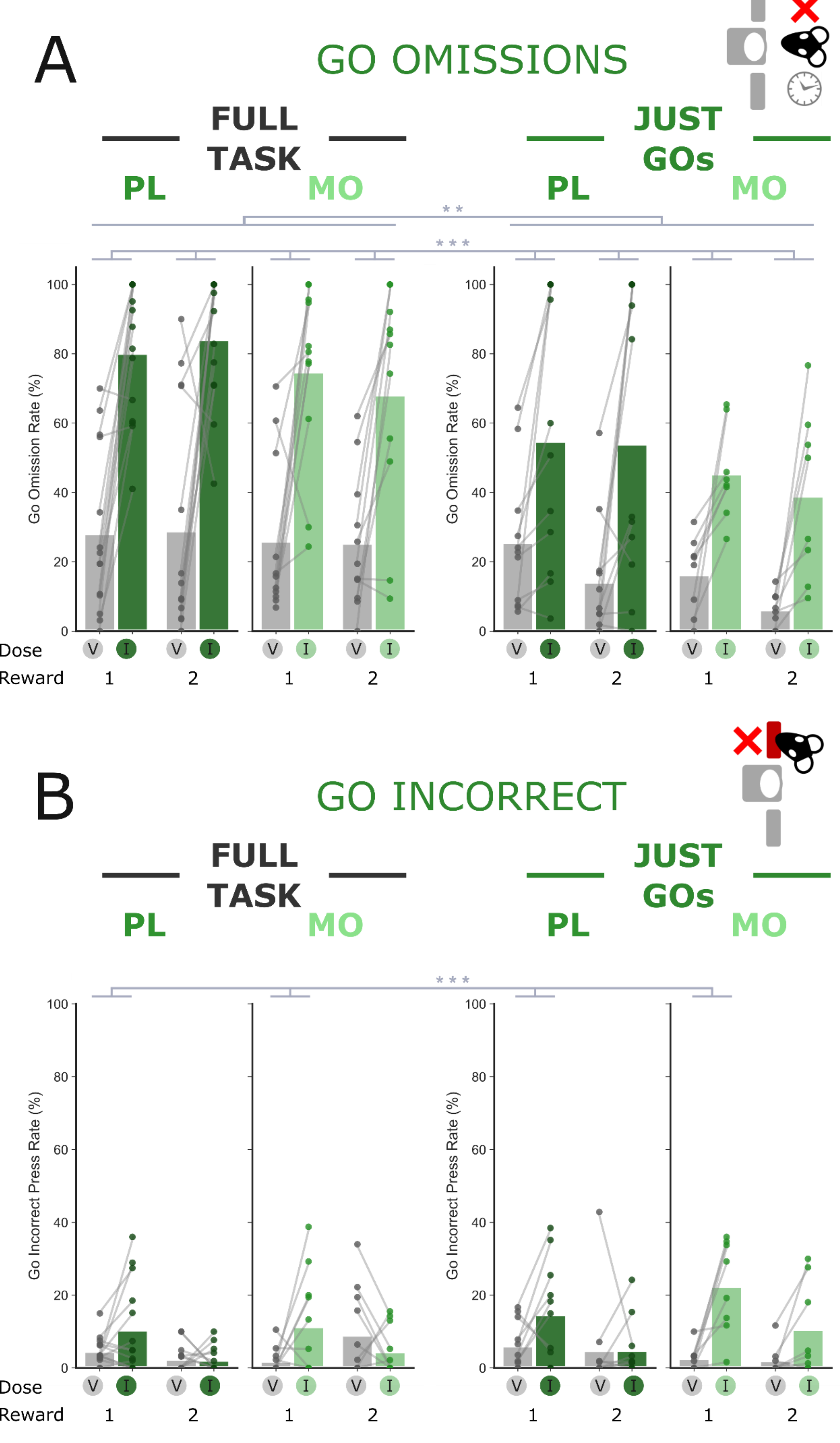

A
GO OMISSIONS
FULL TASK
JUST GOs
PL
MO
PL
MO
Go Omission Rate (%)
Dose
Reward
B
GO INCORRECT
FULL TASK
JUST GOs
PL
MO
PL
MO
Go Incorrect Press Rate (%)
Dose
Reward

Figure 13. Reducing task complexity reduces Go Response Omissions but does not rescue effect of inactivation on either type of error.

A) Left, Effect of inactivation upon Response omission performance in the full task condition, vehicle sessions in grey. Inactivation sessions in dark green (PL) and light green (MO). Right, same as in left but in a reduced version of the task including only Go trials B) same as in A but for Incorrect Press trials Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

4.4 - Discussion

The mPFC has been implicated in numerous components of reward guided action. To determine how subregions of the mPFC may contribute in a differential or unified manner, we selectively inactivated either PL, IL or MO subregions as they performed a Go/No-Go task that taxes animals' ability to make and withhold responses, and to switch between these different strategies. Inactivation impaired No-Go performance, but not homogeneously. Rather, impairment on No-Go trials was organised on a dorsoventral axis: PL inactivation disrupted action restraint on both small and large reward trials, with premature responses predominantly occurring later in the holding period, IL inactivation only affected on small reward trials, while MO inactivation did not produce any reliable impairment. On Go trials however, impairment occurred profoundly and indiscriminately of the region inactivated. These effects on behaviour will be discussed in more detail below.

Subregion specific inactivation of the mPFC produces a gradient of impairment on a dorsoventral axis.

While it has been argued that IL plays a particularly important role in the inhibition of behaviour (Chudasama & Muir, 2001; Passetti et al., 2002b), others have found the PL to also meaningfully contribute to the restraint or postponement of actions(Capuzzo & Floresco, 2020; Paine et al., 2011; Pezze et al., 2014; J. Verharen et al., 2019). Our results align better with the latter group of studies, as inactivation of both PL and IL caused a significant impairment on No-Go trials. However, while disruptions in both the PL and IL affected success rates in the same direction, there are some clear differences between the two regions in how that impairment on behaviour is expressed.

Infusions targeted to IL only reliably disrupted action restraint on trials where a small reward is on offer. This may seem surprising given that IL is classically prominently associated with the inhibition of actions, across both aversive and appetitive domains (Capuzzo & Floresco, 2020; Moscarello & LeDoux, 2013). Our results demonstrates that the contribution that IL makes to suppressing behaviour needs to be refined to consider the context in which behaviour is to be inhibited. On No-Go Large trials there is a greater reward for successful restraint and a greater cost of premature action. A component of action inhibition that scales with the current gains/losses for acting inappropriately must be mediated by circuits external to the IL, as IL inactivated animals are more likely to withhold action successfully with the promise of a large reward.

Previous experiments in this thesis have confirmed that phasic dopamine transients can promote action over inaction, and prior studies have shown that mPFC inactivation can increase VTA burst firing evoked by reward predictive cues (Jo et al., 2013; Jo & Mizumori, 2016), indicating that the mPFC may act endogenously to regulate dopamine release. Therefore, one mechanism by which mPFC inactivation may impair NoGo performance

would be dysregulation of mesolimbic dopamine. While IL inactivation did not selectively increase errors within the first 800ms of the holding period, examining the distributions themselves indicates that IL errors are more prominent closer to cue presentation. As dopaminergic manipulations within the task are restricted to, or most strongly promote, earlier errors proximal to cue presentation (Grima et al., 2022), the effects of IL disruption on No-Go trials could conceivably in part be mediated by dopamine dysregulation.

In line with previous studies where action initiation and restraint must be balanced (Capuzzo & Floresco, 2020; J. Verharen et al., 2019), inactivation of the PL in the Go/No-Go task impaired the ability of animals to withhold action. Unlike IL infusions, this effect occurred independent of the reward size on offer; both No-Go Small and No-Go Large trials were affected. Moreover, PL inactivation produced a different distribution of No-Go errors, with erroneous head exits clustered toward the end of the holding period. An increase in errors towards the end of a No-Go trial indicates that there is likely an extra-dopaminergic component to PL contributions to restraint; NAc D1R infusions mostly increased error probability in the early half of the holding period. As reversible inactivations of the PL has been demonstrated to impair processes such as attention (Pezze et al., 2014), and interval timing (J. Kim et al., 2013), it is possible that impairment may result from inability to remain in a focused state and maintain restraint until the end of the holding period.

Unlike the PL and IL, the MO has not been the focus of such intense study in the domain of reward guided action. Some previous studies position the MO as contributing to successful action restraint in the same manner as the IL and PL(J. Verharen et al., 2019). However, within the Go/No-Go task we do not observe significant impairment of No-Go performance following MO inactivation. This effect may not be that surprising, given that disruption of

the MO has been shown to affect vigour in instrumental tasks (Münster & Hauber, 2018) but not following the presentation of reward predictive cues (Jenni et al., 2023), and it is the latter that appears to interfere with successful restraint in No-Go trials in our task design. Therefore, our results may align better with findings indicating that the MO makes a separable contribution to action restraint when compared to more dorsal mPFC subregions (Hardung et al., 2017a).

Effects of mPFC inactivation on action execution are uniform across regions.

Unlike No-Go trials, the effect of sub-region specific inactivations on Go trials resulted in a comparable impairment in animals' ability to execute the correct action on Go trials. Though disrupting function caused animals to be more likely to press the large reward associated lever on Go small trials, the most prominent impairment was driven by a dramatic increase in Omissions.

Further analysis demonstrated that this increase in Omissions was mediated by a shift in underlying behaviour. When animals made an Omission error in vehicle sessions, they often still initiated the appropriate action and moved to the correct lever. However, inactivation of any mPFC subregion decreased the likelihood of correct action initiation – animals were less likely to even begin the correct action sequence, not just less likely to complete it. This effect contrasts with other manipulations such as NAc D1R blockade(Grima et al., 2022), which increased the frequency of Omission errors without altering behaviour on Omission trials, indicating the prefrontal inactivation may not just render animals in an 'unfocused state' but rather interfere with the process of correct action selection.

The profound disruption on Go trials contrasts with previous findings, where inactivation of either the PL or IL selectively impaired restraint but left the ability of animals to execute lever pressing for reward intact (Capuzzo & Floresco, 2020). One possible explanation for these differences is the more complex task space that the Go/No-Go paradigm presents, requiring greater prefrontal engagement to successfully navigate action selection and execution. Though previous reports have found that mPFC inactivation does not disrupt the ability to discriminate between cues (van Holstein & Floresco, 2019), these experiments occurred in simpler settings where there was no requirement to initiate a trial by sustaining a nose poke response and only two cues were used to guide conditional discrimination between 2 levers. Within the Go/No-Go task there are three possible action strategies that lead to reward (go to the left lever, go to the right lever, remain in nosepoke), and four different cues guide appropriate action. The greater number of cues used could make selecting the appropriate action more difficult, while the more complex action space could contribute to the general disruption of action processes, such that animals are more likely to fail to complete the correct action in the allotted time.

To test the possibility that cue complexity contributed to the disruption of performance in the current experiment, we retrained and tested animals from the second cohort (PL and MO only) in a reduced versions of the task. This consisted of simplified Just-Go and simplified Just-No-Go variants, where only two cues were present throughout the behavioural session. However, in these reduced versions of the task, the same patterns of impairment were still present, suggesting that impairments did not arise from difficulty mapping the appropriate action to the presented cue. However, one caveat to this conclusion is that the animals had a relatively short period on the reduced task versions (3

days) relative to the extensive training they had on the full version of the task. Animals do display an improvement in performance on Go trials in the simplified version of the task, indicating that their behaviour is sensitive to simplified task requirements, but further experiments where animals only ever experience a Just-Go version of the task would be needed in order to exclude cue complexity as a factor in the impairment of Go performance. Additional experiments may also benefit from modifying the task such that four cues are present, but both Small and Large Go trials require presses to made on the same lever. In this variant only one action would ever be associated with reward, while maintaining the conflict between action initiation and restraint. Testing effects of regional mPFC inactivation in this version of the task would resolve whether the effects observed in the present study dependent on the presence of multiple rewarded actions.

When examining general immobility, inactivation did not make animals more likely to enter into immobile bouts when compared to vehicle session. This observation is in line with the fact that only the probability of lever approach, and not magazine approach, is decreased on Go Omission trials following inactivation – suggesting that the impairment on Go trials is not a result of generalized impairment of movement. Animals do show an increase in the average length of immobile bouts following inactivation, but the greatly increased error rates on inactivation sessions make it difficult to ascribe this effect specifically to disruption of prefrontal activity. The generally lower reward rate that animals experience on inactivation sessions could result in extended pauses in between trials as other 'immobile' behaviors such as grooming or resting become more valuable to engage in.

Summary

By selectively inactivating the mPFC in a targeted manner we found that inactivation of the PL and IL, but not the MO, resulted in separable impairments in action restraint. Conversely, on Go trials, inactivation severely impaired performance irrespective of the region targeted. This effect contrasts with previous studies that found action restraint to be impaired, but initiation spared, following PL or IL inactivation (Capuzzo & Floresco, 2020). The relative increase in complexity in the Go/No-Go task may play a role in the more profound impairment we observed, though simplifying the number of cues and action requirements did not rescue effects of inactivation.

5. How does potentiated endocannabinoid signalling effect action initiation and restraint?

Endocannabinoids endogenously regulate mesolimbic dopamine in an activity dependent manner, and compounds that potentiate endocannabinoid signalling are of interest from a therapeutic perspective. To determine how endocannabinoid signalling may affect action initiation and restraint, we dosed animals with a novel compound that inhibits the breakdown of the endocannabinoid 2-AG. We found that boosting endocannabinoid signalling selectively impaired action execution by increasing the likelihood of being in an unfocused state.

Contributions

For the experiments presented in this chapter, George Jenkins, Mason Silveira and Mark Walton designed and planned the experiments with collaborators at A/S Lundbeck (Francois Gastembide, Benjamin Hall, Gunnar Sorensen), George Jenkins and Mason Silveira collected behavioural data, and George Jenkins analysed and presented the data with input from Mark Walton.

5.1 - Introduction

Endocannabinoid signalling acts on a local scale to shape afferent input, and thereby regulate endogenous activity, in a multitude of circuits across the brain. The endocannabinoid system has been demonstrated to make important contributions to regulation of mesolimbic dopamine release under natural conditions (Cheer et al., 2007; Oleson et al., 2012). During periods of sustained activity, endocannabinoids are synthesised and released from dopamine neurons and are thought to inhibit afferent GABAergic input to the VTA, increasing the firing rate of VTA dopamine neurons and thus potentiating dopamine release (Mátyás et al., 2008; M. Melis et al., 2004; Ohno-Shosaku & Kano, 2014).

This endocannabinoid-mediated boost to VTA activity is important for the expression of natural reward seeking behaviour. Disruption of endocannabinoid signalling, both systemically and local to the VTA, can reduce both evoked dopamine release and the latency and vigour with which reward seeking behaviour is executed (Alvarez-Jaimes et al., 2009; Cheer et al., 2007; Feja et al., 2020; Oleson et al., 2012). The endocannabinoid system thereby forms a key component of the regulatory systems that ensure mesolimbic dopamine release, and associated dopamine dependent behaviours, are situationally appropriate (Oleson et al., 2021; Peters et al., 2020).

Pharmacological intervention targeting the release or breakdown of endocannabinoids has the potential to amplify signalling when dopamine neurons are endogenously active, while making little difference at times when dopamine neurons are quiescent. These compounds have generated increasing therapeutic interest in recent years (Bedse et al., 2018; Gil-Ordóñez et al., 2018; Petrosino & Di Marzo, 2010) as modulations of endocannabinoid

availability avoid the of side effects produced by direct cannabinoid receptor agonists (Tai & Fantegrossi, 2014; Wiley et al., n.d.).

Endogenous endocannabinoids are active only transiently before being broken down by enzymes. The endocannabinoid 2-arachidonylglycerol (2-AG), which is thought to mediate activity-induced disinhibition of the VTA (Covey et al., 2017), is metabolised by the enzyme monoacylglycerol lipase (MAGL) (Parsons & Hurd, 2015). Compounds that inhibit this enzyme extend the duration of 2-AG signalling, and have been demonstrated both to increase phasic NAc dopamine release evoked by reward-predictive cues and reduce the latency with which reward seeking actions are initiated (Covey et al., 2018; Oleson et al., 2012; Wenzel et al., 2018). In line with the literature that situates dopaminergic transmission as motivating reward-directed behaviour (Berridge, 2007; Bromberg-Martin et al., 2010; Salamone & Correa, 2012), MAGL inhibition has promise in the treatment of disorders characterised by an amotivational state. Indeed, a recent study demonstrated that MAGL inhibition rescued deficits in dopaminergic signalling and motivated behaviour presented in a mouse model of Huntington's disease (Covey et al., 2018).

However, a thorough understanding of how the endocannabinoid system may work in concert with dopaminergic circuits to regulate reward guided action is lacking. The majority of experiments investigating the effects of endocannabinoid manipulations occur in simple instrumental (Alvarez-Jaimes et al., 2009; Covey et al., 2018; Simonnet et al., 2013; Solinas & Goldberg, 2005) or cued reward seeking tasks (Feja et al., 2020; Oleson et al., 2012). Dopaminergic contributions to these behavioural paradigms are well understood, making it easier to relate effects to the dopaminergic system, but the complexity of the behavioural effects that can be examined are limited.

Testing the effect of endocannabinoid manipulations in the Go/No-Go task enables effects upon a wide range of action related behavioural variables to be tracked and to be examined. In addition, we now possess a strong understanding of both dopamine signalling and the effects of dopaminergic manipulations within the task, enabling effects of endocannabinoids to be compared and related the dopamine system.

Aims

This experiment aims to determine the effect of boosting endocannabinoid signalling upon reward guided action and restraint and determine whether changes in behaviour are consistent with endocannabinoids acting via the mesolimbic dopamine system, and acting in an activity dependent manner. To answer these questions well trained animals were dosed with ABD-1970, an MAGL inhibitor and tested in the Go/No-Go task.

Should ABD-1970 result in a 2-AG mediated boost to dopaminergic release that is activity dependent, we may expect elevated cue-evoked dopamine – resulting in amplified vigour and hastened latencies on Go trials. In this case we would predict these effects on Go trials to occur with little impact on No-Go trials – where dopamine release following No-Go cue presentation is suppressed.

Conversely, should the inhibition of MAGL via ABD-1970 result in an augmented dopamine release that is not constrained to period of elevated dopamine neurone activity, we may expect to see changes to vigour on Go trials accompanied by impaired performance on No-Go trials.

Finally, if treatment with ABD-1970 significantly alters endocannabinoid signalling in brain regions other than the VTA – we may record a different pattern of effects that does not align with dopaminergic manipulations.

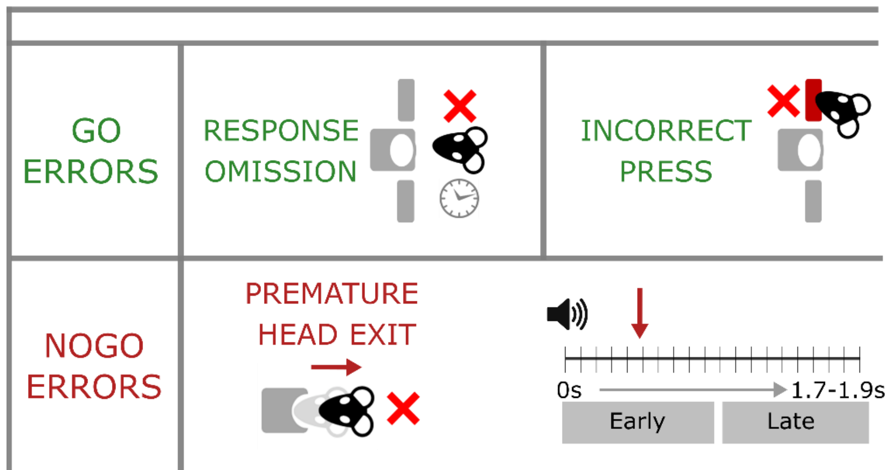
5.2 - Methods

All procedures were performed in line with the UK Animal (Scientific Procedure) Act 1986 and in accordance with the University of Oxford animal use guidelines and were approved by the local ethical review panel at the Department of Experimental Psychology, University of Oxford. Male adult Sprague Dawley rats (n=15) approximately 8 weeks old at the beginning of training, were trained in the Go/No-Go task and were maintained on a twelve-hour light/dark cycle. All testing occurred in the light phase. Animals were housed in groups of three. Food was restricted such that animal bodyweights were kept between 85%-90% of their free feeding weight. Water was available ad libitum. Details of the behavioural task, training and apparatus are described in Chapter 2 (Methods).

Behavioural measures

Measures of performance and latency are outlined in Chapter 2 (Methods) but briefly reiterated in the table below.

PERFORMANCE



LATENCY MEASURES

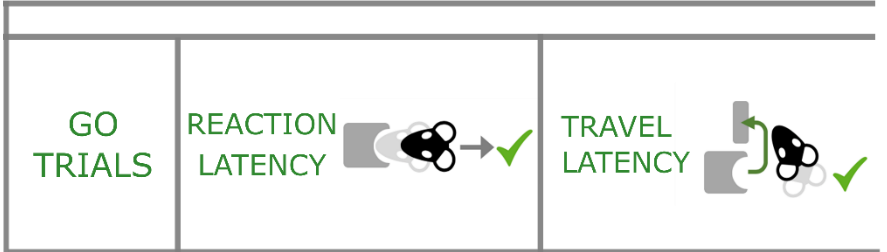


Table 1. Summary of performance and latency measures within the GoNoGo task

---

Pharmacological compounds and preparation

ABD-1970 (Lundbeck) was dissolved in 0.5% methylcellulose solution (Sigma-Aldrich). Animals underwent two treatment sessions, one with a low dose (1mg/kg) and another with a high dose (10mg/kg), in addition to a vehicle session. The order of treatment was counterbalanced across animals, and the experimenter was blinded to the treatment identity on each dosing day. All solutions were delivered orally via gavage and 30 minutes was allowed after dosing before animals were transferred to the experimental context and subsequently tested.

Statistical approaches

As outlined in Chapter 2 (Methods), a repeated measures ANOVA was used for analysis of data collected in these experiments. This included the effect of dose as a three-level factor (vehicle, low dose, high dose) and that of reward size as a two-level factor (small reward, large reward). For skewed latency distributions, a log-transformation was applied in order to meet the normality assumptions of statistical tests.

Video Tracking

The DeepLabCut toolbox (Nath et al., 2019) was used to track the position of the animal during each trial, as outlined in Chapter 2. Briefly, a network was trained to track body parts of the animal and features of the operant box. This enables trajectories of the animal's position on each trial to be obtained. Trajectories were analysed from cue onset until either: 0.75s after entry into the food magazine on successful trials or 5s after cue presentation on failed trials. To analyse the order in which rats visited different zones of the operant chamber, the behavioural box was divided into a 3x3 grid, consisting of a left and right lever zone and a nosepoke zone on one side of the chamber, a food magazine zone on the other side of the chamber, and 5 neutral zones that did not contain any task relevant box features.

5.3 - Results

We wanted to examine how extending the duration of endogenous endocannabinoid signalling may affect reward guided action restraint and initiation within the Go/No-Go task. Specifically, we wanted to observe how a low and high dose of the MAGL inhibitor ABD-1970 may affect performance on Go and No-Go trials as animals make or withhold action to obtain reward.

MAGL inhibition does not affect No-Go performance.

Treatment with the MAGL inhibitor ABD-1970 did not alter performance on No-Go trials relative to the vehicle condition [No significant main effect of drug or drug * reward interaction, all Fs < 0.82, all ps > 0.45, Figure 1].

While neither dose of ABD-1970 resulted in a significant change in the *frequency* that No-Go errors were made, we also examined No-Go trials in more detail to determine if any effect upon *when* errors were occurring was evident with MAGL inhibition. In order to do so, the distribution of erroneous head exits in No-Go trials were plotted as a proportion of all No-Go head exits (Figure 2A). However, again neither the low nor the high dose effected a noticeable change in these distributions. To quantify this, the No-Go holding period was divided into an 'early' (<800ms) portion and 'late' (>800ms) portion (Figure 2B). Animals were more likely to make an error in the late period, but this was not effected by drug [Significant main effect of error period: $F_{1,14} = 10.494$, $\eta_p^2 = 0.428$, $p = 0.006$; no significant main effect of dose or interactions: all Fs < 0.82, all ps > 0.45, Figure 2C]

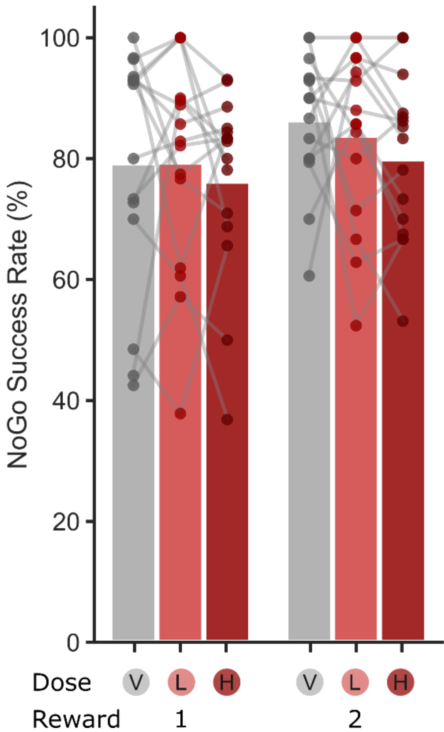


Figure 1. ABD-1970 does not affect No-Go Performance.

Effect of ABD-1970 upon No-Go performance. Vehicle sessions in grey, low and high dose drug sessions in light and dark red respectively. Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

MAGL inhibition selectively impairs Go performance at the highest dose.

While ABD-1970 did not alter performance on No-Go trials, the same was not true for Go trials, with the highest dose producing a significant impairment of Go performance relative to the low dose or vehicle condition. [Figure 3, significant main effect of drug: $F_{2,14}$ = 4.72, η_p^2= 0.252, p = .0017; pairwise comparisons: vehicle vs high dose p = 0.011, low dose vs high dose p = 0.048]. Unlike No-Go trials, performance on Go trials was also demonstrated

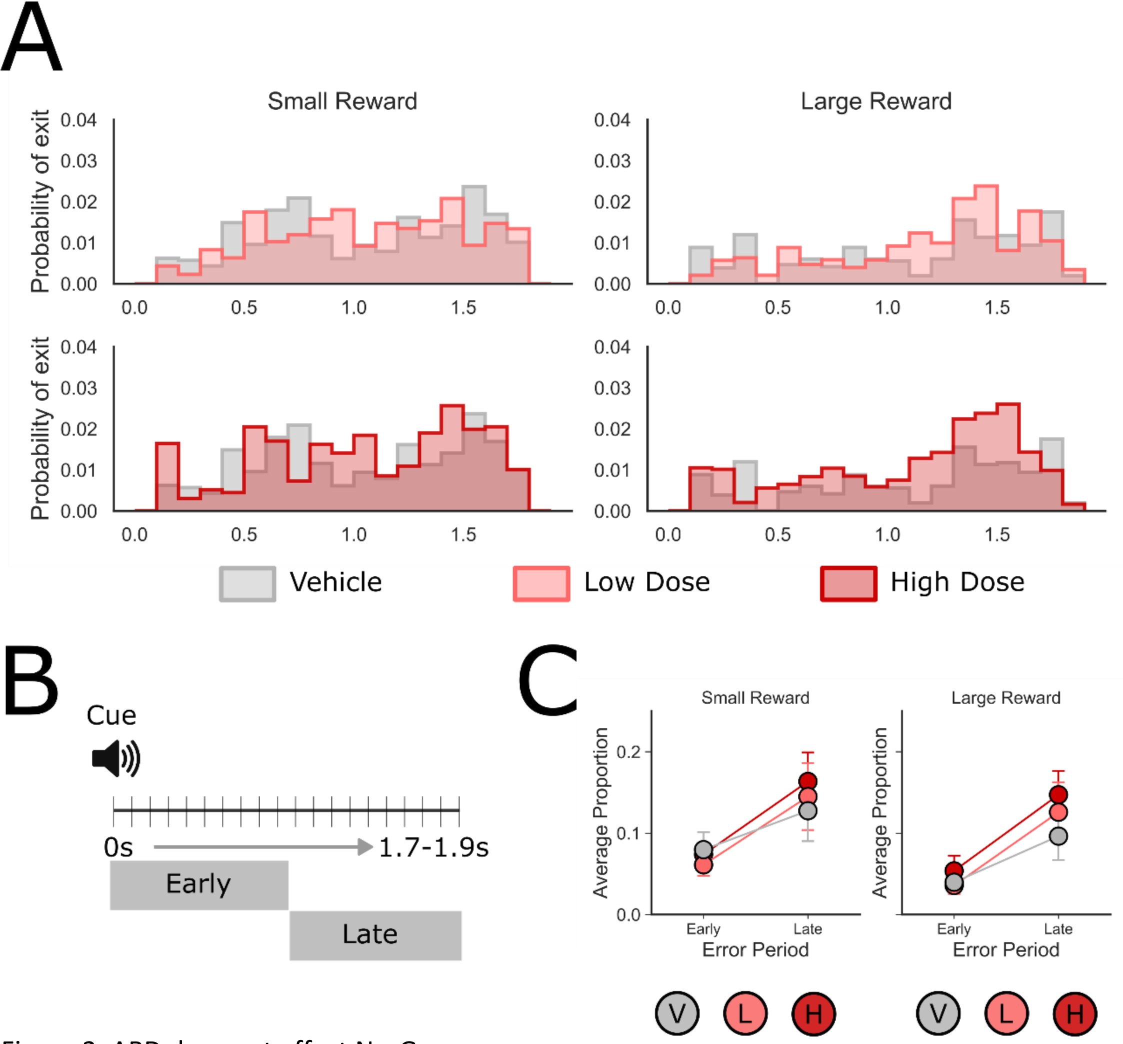


Figure 2. ABD does not affect No-Go errors.

A) Average probability histograms of head exit times on Failed No-Go trials for Small Reward (left) or Large Reward (right) trials in sessions receiving vehicle (grey, top and bottom), low dose of ABD-1970 (light red, top) or high dose of ABD (dark red, bottom) calculated as probability over all trials, both successful and failed. 0.1s bins. B) The No-Go holding time was divided into early exits (>800s) and late exits (<800ms). C) Mean proportions of nosepoke exits that were early or late dependent on dose of drug (grey: vehicle, light red: low dos, dark red: high dose)

---

to be sensitive to the quantity of reward on offer, though this effect was not modulated by drug treatment [Main effect of reward: $F_{1,14} = 18.827$, $\eta_p^2 = 0.574$, $p < 0.001$; No significant drug * reward interaction, $F_{1,14} = 0.003$, $p = 0.996$].

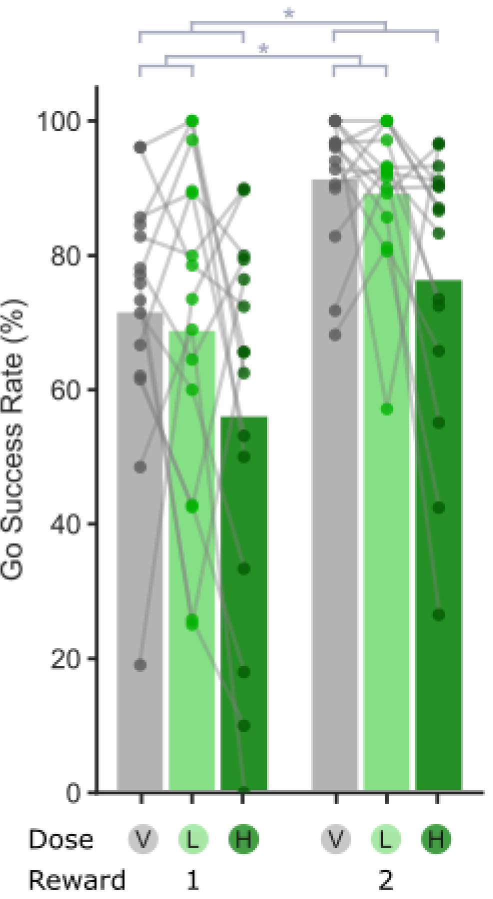


Figure 3. ABD selectively impairs Go performance at the highest dose only.

Effect of ABD upon Go performance. Vehicle sessions in grey, low and high dose drug sessions in light and dark green respectively. Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

In order to examine the effects of the dose-dependent impairment of Go performance in more detail, the proportions of Omission and Incorrect press errors were examined. Examining Omission errors demonstrated that animals had significantly higher omission rates after receiving the high, but not the low dose of ABD-1970 [Significant main effect of drug: $F_{2,28} = 3.65$, $\eta_p^2 = 0.207$, $p = 0.038$; significant pairwise comparisons: Vehicle vs high dose $p = 0.022$, Figure 4]. In line with the effect upon Go performance overall, animals were

more likely to fail to respond on small reward Go trials relative to large reward Go large trials, but this effect did not interact with the drug [Significant main effect of reward: $F_{1,14} = 17.84$, $\eta_p^2 = 0.56$, $p < 0.001$; drug * reward interaction $p = 0.96$, $F = 0.035$]. While there was a numeric increase in the rate of Incorrect press errors following treatment with the high dose of ABD-1970, this did not reach significance [No significant main effect of drug or drug *reward interaction, all Fs < 2.67, ps > 0.087, Figure 4].

MAGL inhibition does not speed latencies.

~~While~~ While we have previously showed that the high dose impaired Go performance, we examined the latencies at which animals initiated and executed action to determine if MAGL inhibition may still promote fast and vigorous action. There was no reliable change in latencies to exit the nosepoke following a Go cue after ABD-1970 treatment [No main effect of drug: F2,26 = 1.92, p > 0.16; drug * reward interaction F2,26 = 2.12, p = 0.093, Figure 5]. The latency to move from the nosepoke to the lever was also not hastened, if anything there was a small trend suggesting that travel latencies might be slower at the highest dose of the drug [Significant main effect of drug: F2,26 = 3.77, η 2= 0.225, p = .0036; pairwise comparisons: Vehicle vs high dose p = 0.043, Figure 5]. Travel latencies, were on average shorter when a larger reward was on offer [Significant main effect of reward: F2,26 = 5.718, η 2= 0.305, p = .0033] but again this reward sensitivity did not interact with drug [No significant dose * reward interaction: F = 0.4, p = 0.674]~~we have previously showed that the high dose impaired Go performance, we examined the latencies at which animals initiated and executed action to determine if MAGL inhibition may still promote fast and vigorous action. There was no reliable change in latencies to exit the nosepoke following a~~

~~Go cue after ABD-1970 treatment [No main effect of drug: $F_{2,26}$ = 1.92, p > 0.16; drug * reward interaction $F_{2,26}$ = 2.12, p = 0.093, Figure 5]. When examining the latency to move from the nosepoke to the lever, there was a small but significant *slowing* at the highest dose of the drug [Significant main effect of drug: $F_{2,26}$ = 3.77, η_p^2 = 0.225, p = .0036; pairwise comparisons: Vehicle vs high dose p = 0.043, Figure 5]. Travel latencies, were on average shorter when a larger reward was on offer [Significant main effect of reward: $F_{2,26}$ = 5.718, η_p^2 = 0.305, p = .0033] but again this reward sensitivity did not interact with drug [No significant dose * reward interaction: F = 0.4, p = 0.674]~~

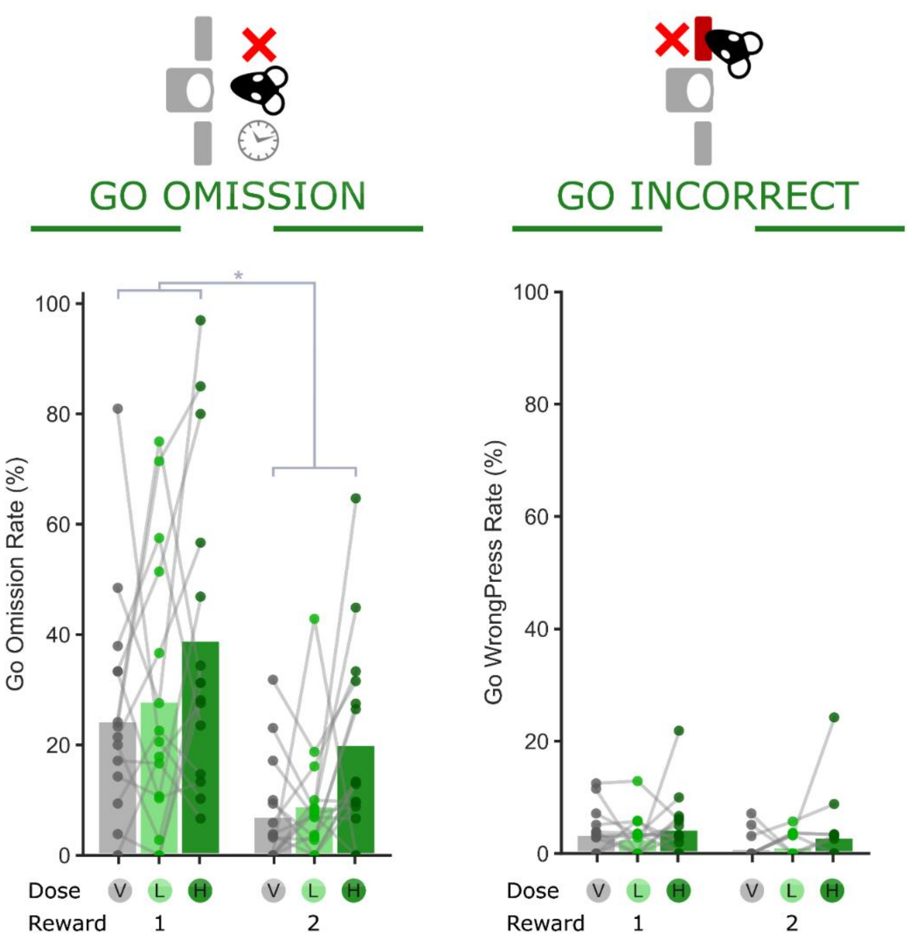


Figure 4. ABD selectively increases Response Omission errors at the highest dose only.

Left Effect of ABD upon Go Response Omission Error Rate. Vehicle sessions in grey, low and high dose drug sessions in light and dark red respectively. Bars show mean, points show data from individual animals. Right, same as in Left but for Incorrect Press Error. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$.

MAGL inhibition does not change behavioural strategy on omission trials.

Our initial analysis demonstrated that the highest dose of ABD-1970 impaired action execution on Go trials, increasing the likelihood that animals would fail to complete therequired action within the allotted time. To understand this effect, we went on to examine behaviour on Go trials in more detail by applying video tracking. On Successful Go trials animals made largely similar trajectories, independent of receiving either vehicle or ABD-1970, first visiting the lever and then the magazine (Figure 6A). While on Successful Go trials animals most often perform a stereotyped action sequence in order to complete lever

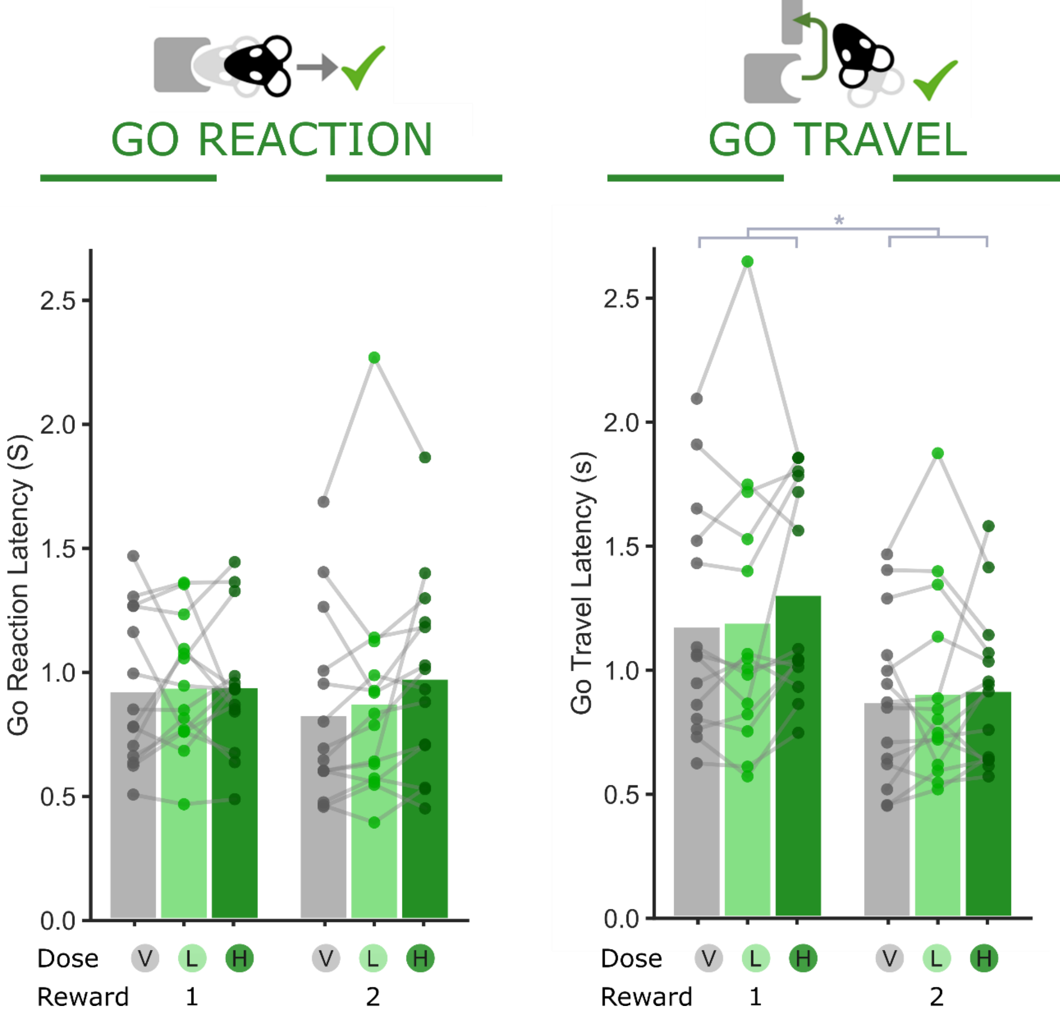
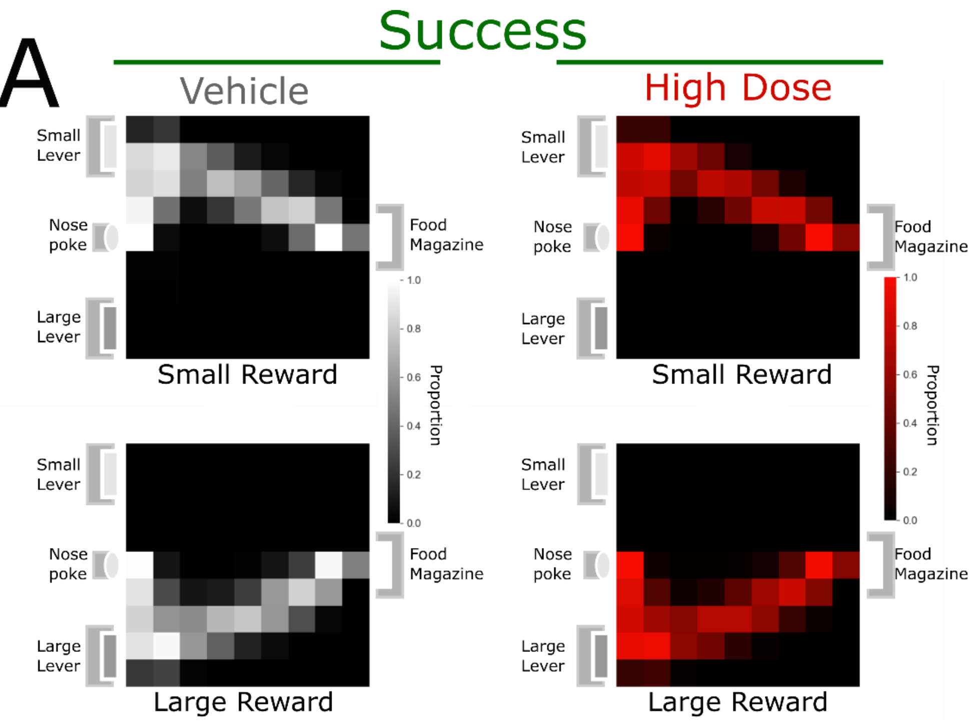


Figure 5. ABD slows travel latencies.

Left Effect of ABD upon Go Response Omission Error Rate. Vehicle sessions in grey, low and high dose drug sessions in light and dark red respectively. Bars show mean, points show data from individual animals. Right, same as in Left but for Incorrect Press Error. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001

press requirements within the allotted time, behaviour on Response Omission trials can be more variable. However, despite the high dose of ABD-1970 increasing the frequency of Response Omission errors, the trajectories animals made on Omission trials after receiving ABD-1970 are also largely similar to those in vehicle sessions (Figure 6B).



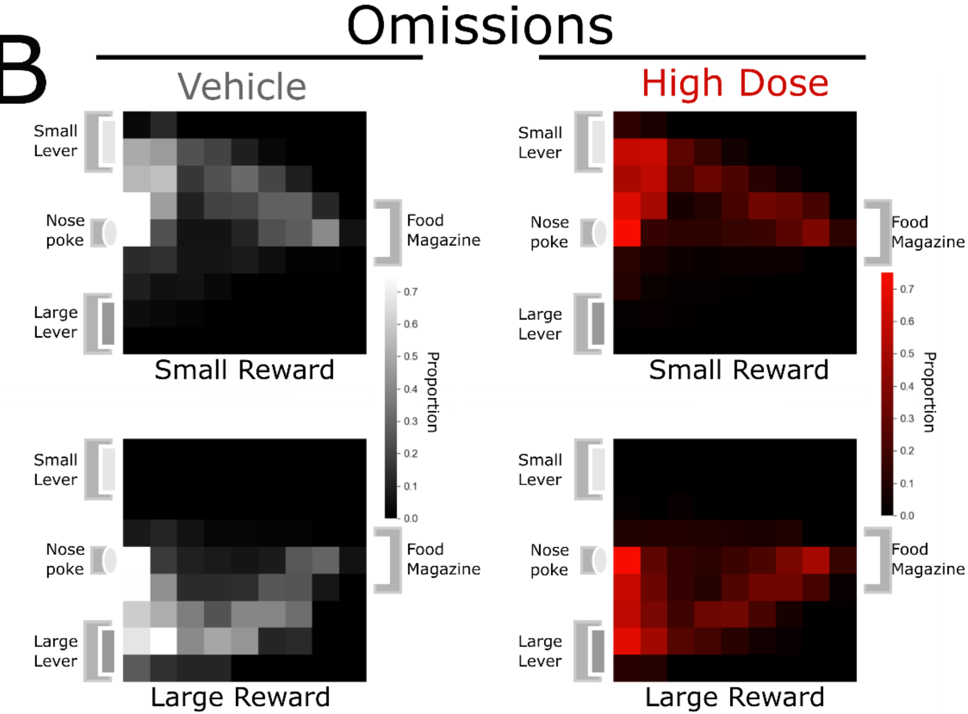


Figure 6. Proportional occupancy on Successful Go trials and Omission trials

Heat maps displaying the average proportion of trials in each condition where each square in a 9x9 grid of the box floor was occupied eg a value of 0.5 means animals moved through that square on half of trials. A) Successful Go trials in vehicle sessions (left) on small (top) and large(bottom) reward right. Right same as in left for abd-1970 high dose sessions. B) same as in A but for Omission trials.

To quantify the degree to which behaviour on Omission trials may or may not differ after treatment with ABD-1970 we first examined the initial action made on Omission trials. In order to examine the frequency with which different initial actions were occurring, we recorded the numbers of Omission trials where animals moved first to one of three task relevant box zones: the correct lever zone, the wrong zone or the food magazine zone (Figure, 7A).

The high dose of ABD-1970 did not alter initial visitation patterns on omission trials, as there was no selective increase in visits to a particular zone [No significant dose * zone interaction: $F_{2,14} = 2.334$, $p = 0.135$]. Independent of receiving vehicle or ABD, animals were most likely to initiate the correct action and move towards the lever indicated by the cue, rather than the wrong lever or the magazine [Figure 7B, significant main effect of zone $F_{1,14} = 12.13$, $\eta_p^2 = 0.464$, p = <0.001, pairwise comparisons, correct lever vs wrong lever $p < 0.001$; correct lever vs food magazine $p = 0.002$]. However, ABD did increase the total number of omission trials where animals moved to one of the three task-relevant zones [Figure 7B. Significant main effect of dose: $F_{1,14} = 6.225$, $\eta_p^2 = 0.308$, $p = .0026$], consistent with an overall increase in omission rate.

While ABD-1970 did not affect the initial action animals made, it may have disrupted the sequence of selected actions. Therefore, we also examined visitation patterns in the box over the 5s following Go cue presentation, to determine if patterns of movement during Omission trials were altered. However, there was no change in the cumulative probability of visiting the food magazine after receiving the high dose of ABD-1970 [No significant main effect of dose, or dose * reward interaction, all Fs < 0.48, all ps > 0.51, Figure 8], nor were animals

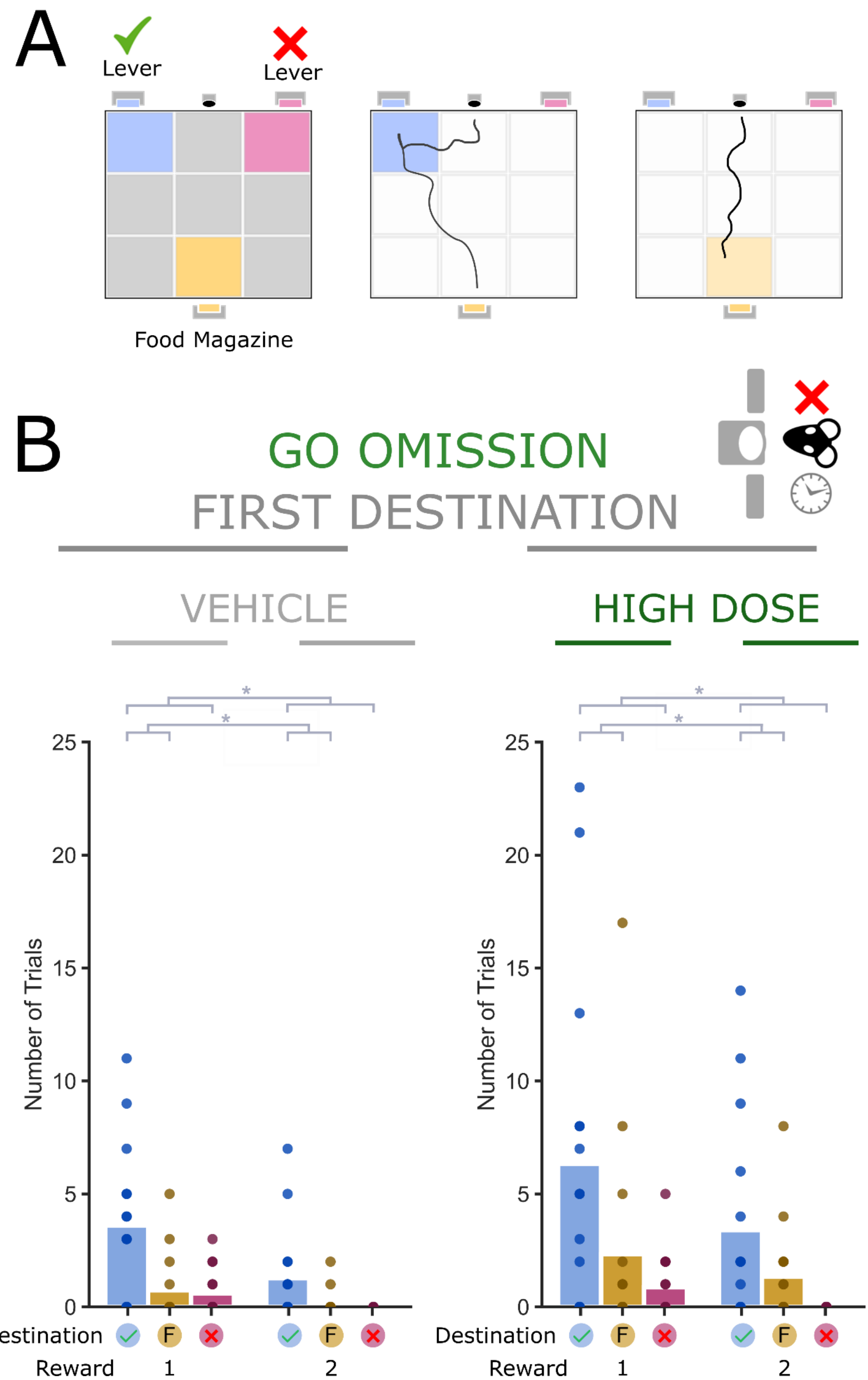


Figure 7. ABD-1970 does not alter visitation patterns on omission trials.

A) Left, the operant box was divided into a 9x9 grid with one square for each lever (correct: blue, wrong: pink), one for the magazine (yellow) and 5 neutral zones. Middle example trajectory where animal moved the correct lever first. Right, example trajectory where animals moved the magazine first. B) Left number of omission trials where animals first visited the correct lever (blue) in correct lever (pink) or food magazine (yellow) in vehicle sessions. Right, same as Left but for high dose of ABD-1970 sessions. Bars show mean, points show data from individual animals. # p < 0.1, * p < 0.05, **p<0.001, ***p<0.0001.

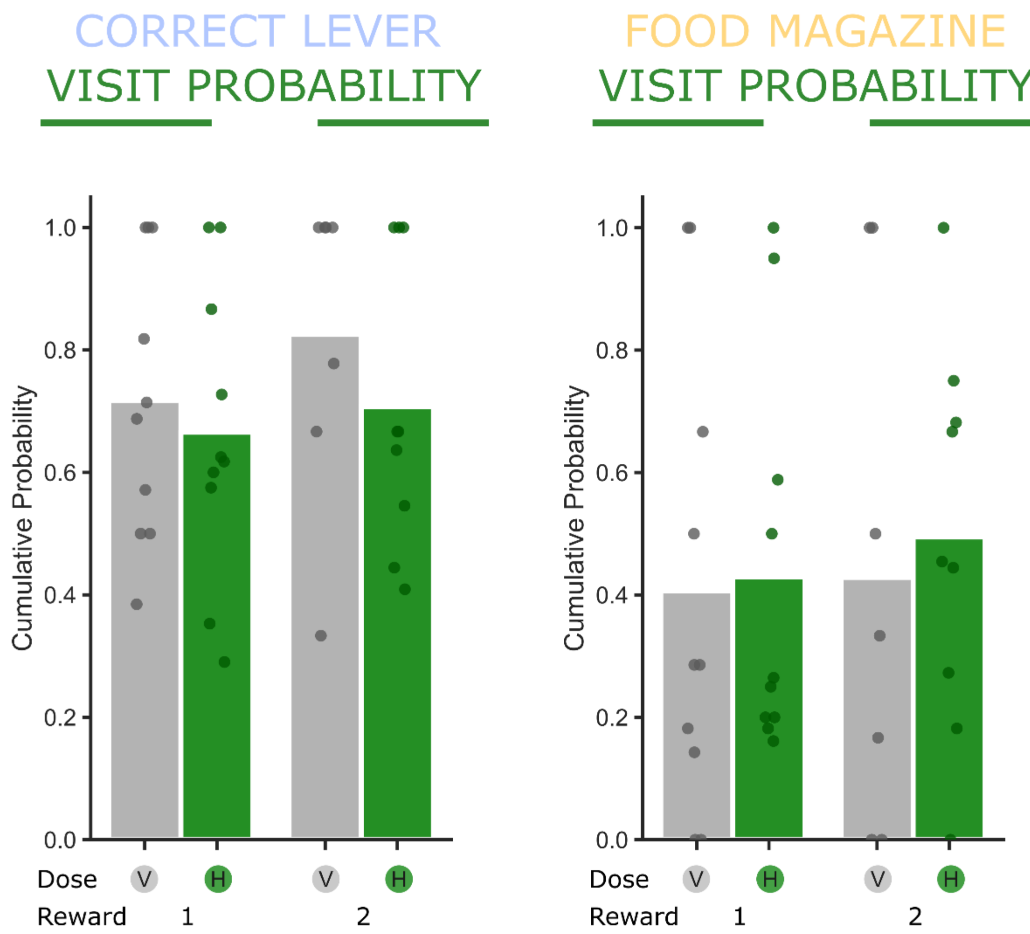


Figure 8. MAGL inhibition does not alter correct lever or food magazine visit probability

Cumulative probability to visit the correct lever square in the 5s following Go cue presentation for vehicle sessions (grey) and high dose sessions (dark green. Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, **$p<0.001$, ***$p<0.0001$

more or less likely to visit the correct lever zone across the duration of Omission trials [No significant main effect of dose, or dose * reward interaction, all Fs < 0.58, all ps > 0.47]

Finally, while the high dose of ABD-1970 did not alter the likelihood of initiating the correct action on Go trials, we also examined general immobility to assess whether the drug might alter locomotion outside of Go trials. Analysis found no difference in either the normalized

rate of immobile epochs [Figure 9, No significant main effect of dose, Fs = 0.201, p > 0.66], nor the average duration of an immobile bout No significant main effect of dose, Fs = 1.387, p > 0.27] between sessions receiving vehicle or the high dose of ABD-1970.

In summary, by tracking animals' movements across a trial demonstrated that while the high dose of ABD-1970 increased the frequency of Omission trials, animals' behaviour when making an omission was comparable between vehicle and high-dose sessions. Animals did not alter their initial actions and were equally likely to visit key areas within the box such as the correct lever and the food magazine. This finding aligns with the visual similarities between the average trajectories on Omission trials after receiving either vehicle or ABD-1970.

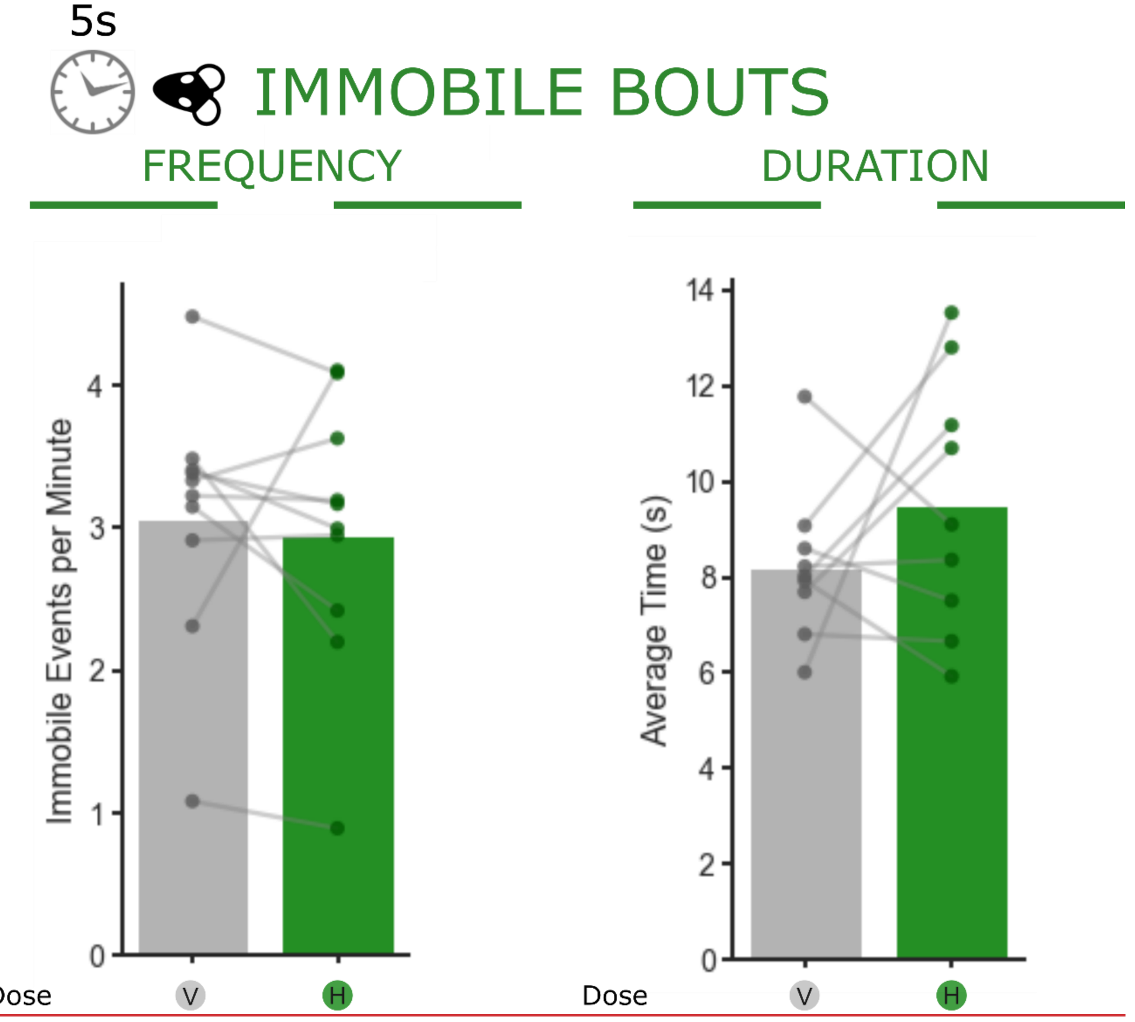


Figure 9. Immobility analysis

A) Effect of ABD upon the rate of immobile bouts per minute. Vehicle sessions in grey. High dose drug sessions in dark green. Bars show mean, points show data from individual animals. # $p < 0.1$, * $p < 0.05$, ** $p<0.001$, *** $p<0.0001$. B) same as in A but for average immobile bout duration.

5.4 - Discussion

The endocannabinoid system provides crucial regulatory tone that sculpts dopamine release and shapes appropriate behaviour under physiological conditions. We wanted to determine whether boosting signalling of the endocannabinoid 2-AG by inhibiting MAGL could selectively improve motivation, without impacting impulse control, potentially via actions within the mesolimbic dopamine system. To investigate this, we tested the effect of the novel MAGL inhibitor ABD-1970 in the Go/No-Go task, where we possess an extensive understanding of dopaminergic manipulations. In line with predictions that extended endocannabinoid signalling should not affect action restraint if any modulation of dopamine release is activity dependent, we saw no difference in No-Go performance following treatment with either dose of the MAGL inhibitor ABD-1970. However, instead of boosting motivation and improving Go trial performance, we instead observed a selective impairment on Go trials, driven by an increase in omission errors. These effects on behaviour will be discussed in more detail below.

Potentiated endocannabinoid signalling does not impair cued action restraint.

Unlike classical neurotransmitters, endocannabinoids are synthesised and released in an activity-dependent manner (Alger & Kim, 2011). Therefore, potentiating endocannabinoid signalling should only effect dopamine release in situations where VTA neurons are highly

active – and prior work has shown MAGL inhibitors boost NAc dopamine to phasic events such as reward predictive cues(Covey et al., 2018).

Our previous studies indicate that suppressed mesolimbic dopamine levels following cue presentation on No-Go trials are critical for action restraint (Grima et al., 2022; Syed et al., 2016). As shown in the previous chapter, optogenetically stimulating the VTA at this timepoint consistently promotes inappropriate action, and NAc D1R agonism produces an increased probability of impulsive action in the time window following the cue.

Elevating levels of 2-AG with the MAGL inhibitor did not affect No-Go performance within the Go/No-Go task – indicating that the drug does not likely boost cue-driven dopamine release. This result could be due to ABD-1970 not mediating an increase in mesolimbic dopamine at any timepoint. Alternatively, this could also be consistent with a model whereby the endocannabinoid system modulates dopamine release within reward guided action in an activity dependent manner. As there are no period of elevated activity in dopamine neurons whilst animals withhold action in the Go\No-Go task, there is no elevated dopamine release for endocannabinoid signalling to potentiate.

Previous studies have investigated the contributions of endocannabinoid signalling to action restraint using pharmacological manipulations, though the experiments have largely been confined to the 5-choice serial reaction time task (5CSRTT). Within the 5CSRTT, activating the endocannabinoid system has no effect on premature action, even when using CB1R agonists (Pattij, Janssen, Schepers, et al., 2007). Unlike MAGL inhibition, this manipulation should induce sustained activation of CB1R receptors independent from the activity of neuronal populations such as VTA dopamine neurons. Conversely CB1R antagonism *reduces* premature action (de Bruin et al., 2011; Pattij, Janssen, Schepers, et

al., 2007). These findings have been suggested to be mediated by endocannabinoid actions on the mesolimbic dopamine system. The pro-restraint effects of CB1R blockade have been speculated to occur via the disinhibition of CB1R expressing GABAergic inputs to the VTA – which would increase inhibition onto the VTA and prevent activity that maintains baseline levels of premature responding (Wiskerke & Pattij, 2015). The null effect of CB1R agonism have been posited to occur from near-maximal endogenous endocannabinoid tone within the 5CSRTT, though further experiments measuring endocannabinoid release will be needed to determine this.

Comparing results across these two behavioural paradigms is a difficult task. The 5CSRTT has some fundamental differences to the Go/No-Go task presented in this thesis, in that animals must refrain making a response while they wait for a stimulus to be presented (which, in key blocks, is often delayed), rather than being instructed to withhold an action by a cue signalling a future reward. In addition - unlike the Go/No-Go task, an understanding of how mesolimbic dopamine release fluctuates within the 5CSRTT is lacking. The effects of phasic dopamine manipulation indicates that key differences exist between the two paradigms, as optogenetically activating the VTA in the Go/No-Go task, but not the 5CSRTT, promotes premature action (Flores-Dourojeanni et al., 2021). This aligns with suggestions that the psychological constructs underlying postponing action until a stimulus arrives and cued action restraint may differ (Bari & Robbins, 2013; Härmson et al., 2022), along with the role mesolimbic dopamine plays in both processes. In theory, activating endocannabinoid signalling could mediate changes in dopamine, and subsequent changes in behaviour, in one task and not the other. Testing the effects of CB1R agonism within the Go/No-Go task would determine whether the lack of impact on No-Go

trials in this study is owed to the either a lack of any boost to dopamine release whatsoever, the activity-dependence of the manipulation, or a ceiling effect of endogenous cannabinoid signalling in cued action restraint, as is proposed for action postponement within the 5CSRTT.

MAGL inhibition increases the propensity to be in an unfocused state on Go trials

Previous studies have related elevated 2-AG signalling following MAGL inhibition to increased motivation and a faster latency to initiate reward seeking following presentation of a cue signalling reward availability (Covey et al., 2018; Feja et al., 2020; Oleson et al., 2012). Translating these behavioural effects into the Go/No-Go task, we may expect a speeding of the latencies at which animal initiate action or travel from the lever to the nosepoke. However, the results we observe are not consistent with a boost in motivation, as following the highest dose animals are no faster to react to a Go cue nor to travel to the lever - with travel times in particular possibly being slower at the high dose. These effects are combined with an increase in omissions, where animals fail to complete the required action sequence, an effect also contrary to a more motivated reward pursuit state.

Hastened reaction times, as predicted following MAGL inhibition, is an effect observed in prior studies activating NAc core D1Rs or boosting NAc core dopamine tone with amphetamine in the Go/No-Go task (Grima et al., 2022; Härmson et al., 2022). However, endocannabinoid signalling is activity dependent, and thereby might be thought to align more closely with manipulations of VTA activity at the timepoint of cue presentation when

the dopamine release is endogenously elevated, rather than chronic but anatomically specific pharmacological manipulations. When optically activating the VTA proximal to cue presentation, we did also observe a slowing of travel times. However, stimulation was not seen to affect omissions, and did not result in an overall impairment of Go performance, even at the highest magnitude. As such the effects of MAGL inhibition do not straightforwardly align with a boost in dopamine neurone activity at cue presentation.Previous studies have related elevated 2-AG signalling following MAGL inhibition to increased motivation and a faster latency to initiate reward seeking following presentation of a cue signalling reward availability (Covey et al., 2018; Feja et al., 2020; Oleson et al., 2012). Translating these behavioural effects into the Go/No-Go task, we may expect a speeding of the latencies at which animal initiate action or travel from the lever to the nosepoke. However, the results we observe are not consistent with a boost in motivation, as following the highest dose animals are no faster to react to a Go cue and are in fact slightly *slower* to travel to the lever. These effects are combined with an increase in omissions, where animals fail to complete the required action sequence, an effect also contrary to a more motivated reward pursuit state.

Hastened reaction times, as predicted following MAGL inhibition, is an effect observed in prior studies activating NAc core D1Rs or boosting NAc core dopamine tone with amphetamine in the Go/No-Go task(Grima et al., 2022; Härmson et al., 2022). However, endocannabinoid signalling is activity dependent, and thereby might be thought to align more closely with manipulations of VTA activity at the timepoint of cue presentation when the dopamine release is endogenously elevated, rather than chronic but anatomically

~~specific pharmacological manipulations. When optically activating the VTA proximal to cue presentation, we did also observe a slowing of travel times. However, stimulation was not seen to affect omissions, and did not result in an overall impairment of Go performance, even at the highest magnitude. As such the effects of MAGL inhibition do not straightforwardly align with a boost in dopamine neurone activity at cue presentation.~~

Conversely, the effects of systemic D1R activation demonstrate more overlap with the effects of ABD-1970, as a systemic D1R agonist slows travel times and impairs Go performance(Grima et al., 2022). However, in the case of systemic D1 manipulation the increase in Go errors is largely driven by animal selecting the incorrect lever. This is a fundamentally different impairment to the elevated omissions observed with ABD-1970, and unlike ABD-1970 the systemic D1R agonist also impairs No-Go performance. The difference between these patterns of effect indicates that MAGL inhibition is also unlikely to mediate a prolonged and distributed increase in dopamine release.

An increased likelihood to make an Omission error could be due to number of different causes - such as decreased motivation to act following a Go cue, or disruption to the process of selecting the correct action according to the presented cue. In both these cases behaviour on omission trials would be expected to change – with animals moving less often to the correct lever. However, when ABD-1970 treated animals omit responding, the pattern of behaviour is strikingly similar to when omission errors occur in vehicle sessions. Animals still move directly towards the correct lever most frequently under MAGL inhibition, and trajectories are similar both on and off the drug. The highest dose of ABD-1970 appears to increase the likelihood animals are in an unfocused state, and attempt the

correct action but fail to complete it, rather than producing a new behavioural drive that interferes with trial completion.

This effect shares some features with that of NAc D1R *blockade*, where the propensity of animals to omit trials was increased without changing the underlying behaviour on erroneous trials (Grima et al., 2022). MAGL inhibition has been demonstrated to boost VTA dopamine release, making the correlation of these effects appear initially unlikely. However, activation of CB1R within the NAc can also *reduce* dopamine release via suppression of excitatory afferents to the NAc (Covey et al., 2017; Mateo et al., 2017). Prior studies have suggested that the observed increase in both dopamine and motivational drive following MAGL inhibition results from preferential enhancement of disinhibitory 2-AG signalling within VTA (Covey et al., 2018). Therefore, one possible explanation for the opposing effects observed in this study is that the differential task requirements in the Go/No-Go task result in differential endogenous endocannabinoid signalling patterns within the mesolimbic dopamine pathway, such that boosting 2-AG transmission favours inhibition within the NAc over disinhibition of the VTA.

In line with this theory, the aforementioned studies that observe pro-motivational effects of MAGL inhibition occur in relatively simple tasks such as cued-reward seeking and progressive ratio (Covey et al., 2018; Feja et al., 2020; Oleson et al., 2012). In these paradigms only one action has positive valence. The Go/No-Go task presents a more complex action space, where a single action such a pressing a particular lever is only advantageous on a subset of trials. The more complex environment of the Go/No-Go task may recruit a more distributed network of brain structures, including circuits outside of the

mesolimbic dopamine pathway that may also be influenced by endocannabinoid signalling in a manner sufficient to disrupt behaviour.

Understanding how endocannabinoid signalling may vary on a moment-by-moment basis within behavioural tasks will be advantageous in any future studies aiming to relate endocannabinoids, mesolimbic dopamine and reward guided action. The recent advent of the fluorescent endocannabinoid sensor $GRAB_{eCB2.0}$ (Dong et al., 2022) will enable endocannabinoid signalling dynamics to be related to dopamine release with a sharper temporal resolution than methods such as microdialysis, and allow changes in endocannabinoid signalling in response to different behavioural requirements of pharmacological interventions to be tracked in vivo.

6. General Discussion

Here, the findings of this thesis are summarised. Technical limitations of the techniques utilising the preceding chapters are considered, and suitable future directions for research are suggested.

6.1 - Introduction

Selecting the appropriate behavioural strategy for a given situation is key to obtaining rewarding outcomes, and therefore for survival. Reward guided action control is a complex, multistep process requiring computations about reward value, the optimal strategy to choose and when best to execute it (Rangel et al., 2008). A crucial component of this process is deciding when best to initiate reward seeking action, and when instead action should be withheld. The midbrain dopamine system stands at the intersection of two rich fields of study, being implicated in both movement and reward and is therefore well positioned to contribute to the control of reward guided action. This thesis aimed to understand the role that mesolimbic dopamine in particular may play in reward guided action initiation and restraint, alongside complementary circuits and systems that may act in concert with dopamine to shape behavioural control.

Throughout this thesis a common behavioural paradigm has been used to investigate effects of manipulations across three systems: the mesolimbic dopamine pathway, the endogenous endocannabinoid systems, and the medial prefrontal cortex. The Go/No-Go task employed in this thesis enables the impact of said manipulations on performance under both differing conditions of action requirements and differing reward expectations to be explored and compared against previous findings (Grima et al., 2022; Harmson et al. 2022).

Utilising optogenetics allowed us to determine the capacity for phasic activation of VTA dopamine neurons to promote action over inaction, and the dependence of this effect upon concurrent cue presentation at low stimulation magnitudes. These experiments also found a concomitant effect for activation to bias toward premature action on *future* trials, as well as a small disruption to action selection and invigoration. Through targeted pharmacological inactivation of parts of mPFC in a subregion specific manner, it has been discerned that the PL and IL subregions, but not the MO, make distinct contributions to successful action restraint when a cue signals future reward. This result contrasts with the other finding from this experiment, that all three regions are critically necessary for successful action execution in a complex task space. Finally, investigations into the contributions of endocannabinoid regulation have determined that inhibiting the breakdown of the endocannabinoid 2-AG, which has been shown to augment reward-cue evoked dopamine release in previous studies (Covey et al., 2018), did not boost motivation or change impulse control as has been previously described in simpler behavioural tasks. Rather, potentiated endocannabinoid signalling increased the likelihood of animals being in an unfocused state and failing to execute action successfully.

Together, this work elucidates the causal role for mesolimbic dopamine in reward guided action, and highlights which contributing circuits may act in conjunction to shape reward guided action more generally.

6.2 - Summary and implications of key findings

Chapter 3 – Phasic dopamine transients conditionally promote action over inaction

Historically, dopamine has long been linked to reward and reinforcement (Olds & Olds, 1958). More recent studies lead to the formulation and definition of the reward prediction error hypothesis, proposing that phasic activity in the midbrain dopamine system signals the difference between expected and actual rewarding outcomes, and drives learning to shape future behaviour (Schultz et al., 1997). However, recent studies have demonstrated modulations of dopamine neurons in the absence of explicit rewards (Coddington & Dudman, 2018; Dodson et al., 2016), and others have aligned phasic elevations in mesolimbic dopamine release with the initiation of reward seeking actions (Roitman et al., 2004; Syed et al., 2016). Therefore, a key open question has been how to link together these separate conceptions of dopamine as a *reinforcement* signal and dopamine as an *action / activation* signal.

To determine whether phasic dopamine stimulation could causally influence the likelihood of action initiation, In Chapter 3 we employed optogenetics to activate VTA dopamine on fast timescales akin to those observed endogenously upon presentation of reward-predicting cues. We found that optogenetically evoked dopamine transients can promote

action over inaction, but in alignment with previous studies this pro-movement effect was only observed conditionally (Da Silva et al., 2018; Hamid et al., 2016; Hughes et al., 2020). Within the current study, this dependence was linked to both the magnitude of the stimulation applied and its timing relative to instructive cues. Applying a high magnitude stimulation either early or late during a No-Go cue while animals were refraining from moving generally promoted movement to be initiated, while a lower magnitude stimulation, arguably closer to physiological activity patterns, promoted action when delivered after cue presentation, but *strengthened restraint* if delivered while animals were already engaged in withholding action.

Though these effects were most prominent on the trial on which optical stimulation was delivered, we also found stimulation induced a bias towards premature action. This effect accumulated *within* a session, such that animals expressing ChR2 became increasingly likely to act prematurely compared to EYFP controls, but did not persist *across* testing days.

We also observed a slight disruption to efficient action execution following stimulation, resulting in ChR2 animals being more likely to press the incorrect lever and slower to move from the nosepoke to the lever following stimulation. Similar results have been observed with systemic, but not intra-NAc D1 agonist delivery (Grima et al., 2022), suggesting that these effects may arise from mesolimbic dopamine release in sites outside of the NAc.

Contrary to a purely RPE perspective, this study does not suggest that phasic dopamine transients do not induce an improvement in ongoing or future performance consistent with boosting value or motivation. Rather these results indicate that phasic dopamine transients can causally influence the likelihood of action on a moment by moment basis, and that the effects of VTA activation are modulated according to the time window when

activation occurs, and the effective window of downstream dopamine concentration that follows activation. In addition to effects on online performance, we also observe extended changes in behaviour that accumulate across an hour long session - but do not persist to the next day. That we observe effects on future behaviour, but not enduringly - or as strongly as those on moment-by-moment behaviour, may indicate that the presence and potency of alterations to behavioural strategy following dopamine neuron stimulation are not universal. Rather they may depend on the timing of stimulation relative to internal and external states (Hamid et al., 2016; K. Lee et al., 2020) – and potentially how dopamine release aligns with period of differential striatal receptivity (Berke, 2018).

With a role confirmed for elevated mesolimbic dopamine in promoting action, we next moved to investigate the contributions of other circuits within the brain that may act to shape to appropriate action control, and could possibly mediate a restraint induced suppression of dopamine release.

Chapter 4 – inactivation of mPFC subregions differentially impacts action restraint but produces similar impairments in action execution.

The medial prefrontal cortex is implicated in ensuring appropriate behavioural responses to cues (Ishikawa et al., 2008a, 2008b) and the inhibition of action responding (Capuzzo & Floresco, 2020; Paine et al., 2011; Pezze et al., 2014), suggesting that this structure contributes to action control, though different subregions of the mPFC may make separable contributions to this function (Hardung et al., 2017b). The mPFC is functionally connected to a range of motor circuits throughout the brain, in addition to mesolimbic

dopamine system, and mPFC activity can modulate release of dopamine in downstream regions such as the NAc (Jackson et al., 2001). To dissect the contributions of mPFC subregions to reward guided action initiation and restraint, the subregions of the mPFC were independently inactivated with a mixture of the GABA agonists muscimol and baclofen.

Effects on action restraint showed subregion specificity. Targeting the most rostral region, MO, had no effect on No-Go performance. After inactivation of IL, performance on small reward trials was affected, while PL inactivation disrupted performance on all No-Go trials, mediated by a selective increase in head exits occurring towards the end of the holding period. These differences in behavioural effect may indicate that the two regions perform different functions in situations requiring action restraint. The IL may provide a general inhibition, separable from the drive to that makes restraint more important when greater costs/benefits are at stake – while the PL may contribute to restraint by enabling focus to be maintained for the duration of the holding period. The involvement of both the PL and IL, but not the MO, is largely consistent with studies showing that the two more dorsal regions are required for cued action restraint (Capuzzo & Floresco, 2020; J. Verharen et al., 2019), and indicates that the MO may perform a separate functional role.

Conversely it was found that inactivations sharply impaired performance on Go trials independent of the region targeted. This drop in performance was mediated in an increased likelihood to press the preferred lever, and a dramatic increase in the proportion of trials where animals failed to complete the correct response. This was accompanied by a reduction in the probability of animals to initiate and sustain the correct sequence of actions. This dramatic impairment contrasts with the preservation of responding on

simpler tasks (Capuzzo & Floresco, 2020), and ability of animals to track stimulus-reward associations following mPFC inactivation when only two cues are used to guide behaviour (van Holstein & Floresco, 2019). Impairments were not rescued by reducing trial complexity, indicating that disruption to performance cannot simply due to an inability to switch in the same session between differing action requirements. Nor did animals display a reduction in motivation, as similar number of trials were completed overall following infusions of baclofen/muscimol or vehicle. Further experiments will be need to determine the nature of behavioural disruption following inactivation of mPFC sub regions.

These results indicate the critical involvement of the mPFC in reward guided behaviour in a complex task space, and highlight the subregion specific role of the PL and IL in cued action restraint.

Inactivation of mPFC subregions has effects that could include, but extend beyond a role purely in modulating dopamine release in situations where action must be controlled appropriately, and further experiments will be required to closely understand how medial prefrontal regions may cooperate with mesolimbic dopamine to shape behaviour. In order to expand knowledge of how other regulatory systems, known to sculpt endogenous dopamine release, may contribute to successful action control in Chapter 5 we examined the effect of manipulating endocannabinoid signalling the Go/No-Go task.

Chapter 5 – MAGL inhibition promotes an unfocused states where omission errors are more likely.

Endocannabinoids form a key component of the systems that endogenously regulate activity in the mesolimbic dopamine pathway, elevating VTA firing in an activity dependent manner. Regulation by the endocannabinoid system has been demonstrated to shape both evoked dopamine release and associated reward seeking behaviour (Covey et al., 2018; Oleson et al., 2012). Compounds that potentiate endocannabinoid signalling by inhibiting their breakdown have been demonstrated to boost motivation and hasten reward seeking action in simple paradigms, and are of increasing therapeutic interest.

In Chapter 5 we employed a novel compound, the MAGL inhibitor ABD-1970, within the Go/No-Go task. We found that extending endocannabinoid signalling did not affect No-Go performance. However, unlike previous reports in simpler tasks, ABD-1970 did not produce effects consistent with an increase in motivation, but rather impaired performance at the highest dose. This impairment was not mediated by a change in underlying behaviour, but rather an increased likelihood to be in an unfocused state and fail to complete the trials action requirements. This finding aligned better with the previously reported effects of NAC D1R *blockade* (Grima et al., 2022), and therefore align more closely to a parallel strand of previous work indicating that endocannabinoid action can also *suppress* NAc dopamine release (Covey et al., 2017; Mateo et al., 2017), in addition to elevating VTA dopamine neuron activity. These results indicate that the endocannabinoid signalling may differ between simple and complex tasks, and call for more close examination of the relationship between dynamic changes in endocannabinoid and dopamine release.

6.3 - Limitations, technical considerations, and future directions

Further consideration of action restraint tasks

A number of tasks, including the Go\No-Go tasks utilised in this thesis, have been employed to assesses the capacity of animals to inhibit behaviour. Behavioural inhibition paradigms involve different aspects of behavioural control, including action cancellation, action restraint and action postponement (Bari & Robbins, 2013). The inhibitory component of the Go\No-Go task contains aspects of both action restraint and action postponement, and thus yields similarities to classical go/nogo tasks and the 5CSRTT. In these paradigms animals must withhold from responding until the presentation or offset of a cue – with any operant response made during the waiting time considered impulsive.

This process is similar to the restraint animals must exercise on No-Go trials, but with some key differences. In the Go\No-Go task the maximum nosepoke hold, including both the precue and No-Go holding periods, is 2.4s. This is relatively brief when compared to alternative paradigms, in 5CSRTT the usual period of restraint is often 5-9s, while in other inhibitory tasks action can be required to be withheld for up to 15s (Capuzzo & Floresco, 2020; Verharen et al., 2019). Thus, we are likely unable to determine the effect of experimental manipulations upon mechanisms contributing to extended restraint when using the Go\No-Go task. However, via the effect of a shorter optogenetic VTA stimulation, we are still able to differentiate between the earlier and later halves of the holding period. This finding dovetails with previous studies demonstrating the selective impact of dopaminergic or serotonergic manipulations on early or late errors respectively .(Grima et al., 2022; Härmson et al., 2022).Together these results suggest the ~2.4s holding period on

No-Go trials is still able to capture separable mechanisms underlying restraint in the face of early cue driven and later anticipatory drivers to move.

In the Go\No-Go task, animals must first make a nosepoke hold for the pre-cue period in order to initiate a trial, then on No-Go trials the hold must be maintained for further 1.7-1.9s. This is in contrast to other paradigms such as the 5CSRTT, where animals may need to make a sustained nosepoke in order to begin a trial, but no further hold is required. This difference in behavioural state provides an additional dimension along which Go and No-Go trials could be said to be oppositional. On No-Go trials animals must remain in the same state, while on Go trials animals must switch.

As a result of having to make a sustained hold for the duration of a No-Go trial, the nature of action restraint itself in the Go\No-Go task differs from more classical paradigms. In classical go/nogo tasks or the 5CSRTT animals are restricted from making one specific movement – often interacting with a nosepoke, while all other actions are unpunished. This allows the expression of other behaviours, such general locomotion or approach to a reward-associated food magazine, that could be said to aid animals in withholding from nose poking. In one approach-avoidance task, where entry to a food magazine had to be withheld while a cue was presented, animals demonstrated ‘attract repel’ behaviour – effectively initiating and curtailing reward approach several times over the course of the waiting period (Verharen et al., 2019). In the Go\No-Go task animals must remain in the poke for the duration of the trial, preventing the ambulation or reward approach that animals can perform in other paradigms. While No-Go trials restrict movement to a greater degree that other paradigms, animals are not required to cease all action completely. Rotational or forward movement in the nosepoke is not punished, and animals can even

briefly exit and re-enter during the holding period. As such, findings from this thesis cannot be generalised to the true restraint of all action or movement. Despite this, the extended nosepoke hold required by the Go\No-Go task ensure that the pavlovian drive to approach reward must be inhibited in order to succeed on No-Go trials, animals are still able to make movements, just not movements that bring them closer to reward.

Specificity of optogenetic approach.

The optogenetic approach employed in Chapter 3 of this thesis allowed for fine temporal control of VTA dopamine neurons, enabling the time dependence of stimulation-induced effects on action to be determined. While other methods such as electrical stimulation offer a similar time course of activation, optogenetics has the added advantage of restricting manipulations to genetically identified population of neurons, in this case VTA dopamine neurons.

However, dopaminergic afferents from the VTA innervate a range of cortico-limbic structures. The approach employed in the current study would therefore evoke dopamine release in a range of downstream structures, and prevents the localisation of effects to any one region, such the NAc core. In addition, we are unable to tightly link the effects of this study to VTA *dopamine*, as VTA neurons are known to co-release glutamate (Stuber et al., 2010). As well as co-release, light delivery itself has the potential to be the source of non-specific effects. Illumination can heat surrounding tissue and activate neurons indiscriminately (Owen et al., 2019). However, this effect is a much greater concern when applying continuous illumination, with the pulsed stimulation trains used in these experiments far less likely to generate heating related effects.

Finally, we have an understanding of dopamine release dynamics within the task from previous studies, and an understanding of effects of the optogenetic stimulation on dopamine release in anesthetised animals. However, we were not consistently able to combine stimulation, recording and behaviour in same behaving animals (the successful recordings unfortunately coincided with the less successful targeting / transfection approaches: see Chapter 2 Methods). This approach would have allowed us to make confident statements about how dopamine evoked by our high and low stimulation compared to natural events, such as unexpected rewards and reward-predictive cues.

Considerations towards satiety induced changes

Unlike the pharmacological manipulations, the optogenetic experiments detailed in Chapter 3 allowed individual animals to be tested in repeated sessions. This made the collection of greater number of trials possible, which in turn enabled us to leverage more powerful analytical approaches including mixed effects modelling – which were crucial in effectively analysing the complex dataset.

To facilitate the collection of a large volume of data we increased the maximum number of rewards that could be obtained per session from 100 to 300. Though increasing the yield of trials per session, one drawback of this approach is the near-certain inducement of satiety, which is likely to result in within session changes in behaviour as motivation decreases. Animals move faster and are less likely to omit on Go trials when a large reward is on offer, thus as reward value decreases with satiety we might expect to see slower movement latencies and greater omission rates towards the end of behavioural session.

Conversely, the decreased incentive drive to pursue reward as animal satiate may be advantageous when restraining action on No-Go trials.

In addition to within session changes in baseline behaviour, it is also possible that the effects of optogenetic VTA stimulation may change with satiety. Dopamine release in the NAc has been shown to differ depending on satiated state(Ahn & Phillips, 1999; Papageorgiou et al., 2016) and though beyond the scope of this thesis it is possible that effects of VTA stimulation upon action may differ in magnitude or effectiveness in a sated versus hungry state. Even in the face of these caveats, the effective collection of a large dataset and the analytical approaches enabled remain valuable.

Considerations regarding pharmacological manipulations

Unlike the precise temporal control afforded by optogenetic stimulation, the muscimol/baclofen infusions employed in Chapter 4 and the MAGL inhibitors used in Chapter 5 induce chronic effects that extend across the entirety of a behavioural session, and both are accompanied by specific additional considerations that will be discussed below.

Unlike lesions, inactivations using GABA agonists are reversible, and avoid the development of compensatory activity between surgery and behavioural testing that can be present following lesions. However, we aimed to specifically inactivate subregions of the mPFC, a task complicated by the distribution of PL, IL and MO on a dorsoventral axis. Cannulae implantation into more ventral regions would necessitate some damage to subregions located more dorsally. Though the dramatic difference between performance

in inactivation and vehicle sessions indicates that effects are largely mediated by the baclofen/muscimol infusions, the contributions that implantation damage may make should be taken into account when interpreting results from this chapter.

In addition, in these experiments it is not possible to determine the extent of the spread of baclofen/muscimol from the infusion site. Though different patterns of behavioural effects indicate that infusions did not result in inactivation of a large homogenous area, regardless of targeting, it is important to bear in mind that we can only state where infusions were made, and not where in the mPFC was inactivated.

In Chapter 5, we employed a MAGL inhibitor, that extends signalling of the endocannabinoid 2-AG, via oral gavage. Both the mechanism or drug action and the method of delivery align with growing interest in these compounds from a therapeutic perspective. However, endocannabinoid signalling occurs in a local manner within numerous circuits in the brain, and the systemic nature of our approach forgoes localising any effect of our manipulation. This is of particular importance when considering that within the context of the mesolimbic dopamine system, endocannabinoid actions at VTA cell bodies and within the NAc are thought to have opposing effects on dopamine release (Covey et al., 2017, 2018; Mateo et al., 2017; Oleson et al., 2012). In addition, cannabinoid receptors are widely distributed throughout the central nervous system, with systemic manipulations likely to affect activity in other regions important for behavioural control such as the hippocampus or prefrontal cortex (Busquets-Garcia et al., 2015; Tsou et al., 1998). Though less advantageous in terms of discerning potential therapeutic viability, local infusions of drugs that act via the endocannabinoid system would enable the

contributions of local endocannabinoid regulation to dopamine release in situations of action control to be determined more thoroughly.

Finally, an additional consideration regards animal reuse, as the cohort receiving the MAGL inhibitor ABD-1970 in Chapter 5 went on to receive prefrontal inactivations in Chapter 4. CB1R are expressed widely across the brain (Tsou et al., 1998), including within the mPFC, so it is likely that increased levels of 2-AG following MAGL inhibition would have induced some change in activity within mPFC circuits. However, any effects on levels of the MAGL enzyme itself are likely to be transient, as MAGL function is near completely restored 72 hours following ABD-1970 treatment (Clapper et al., 2018). As the experiments in Chapters 4 and 5 were separated by a span of over five weeks it is highly unlikely for 2-AG breakdown to remain impaired by the time animals received mPFC inactivations. On the other hand, despite the MAGL-treated cohorts baseline behaviour seeming unaltered, endocannabinoid signalling can induce plasticity (Pan et al., 2008) – meaning that it is possible for there to be some lasting effects of ABD-1970 dosage. Despite this potential drawback, animal reuse does have the advantage of reducing the total number of experimental animals used – in line with the 'reduction' principle of animal research.

6.4 - Future directions

Defining effects of mesolimbic dopamine on action with pathway specific precision

As previously discussed, despite prior research specifically implicating NAc dopamine transmission in reward guided action, the experiments in Chapter 3 involve activation of

the VTA as a whole. A clear first step would be to definitively localise effects recorded in Chapter 3 to the VTA-NAc projection. This could be achieved, either by retrogradely transfecting this pathway, via targeting optical fibres to the NAC, or by repeating the experiment in Chapter 3 in conjunction with intra-NAc dopamine antagonist infusion. The latter option has the advantage of tying any behavioural effect occluded by antagonists specifically to NAc *dopamine*, rather than glutamate transmission.

Prefrontal -mesolimbic interactions in reward guided restraint

The experiments recounted in Chapter 3 of this thesis provide evidence for the causal relationship between phasic mesolimbic dopamine and the initiation of reward guided action. In Chapter 4 we demonstrated the particular contributions of the PL and IL subregions to successful action restraint. Yet these experiments occurred in separate animals, and the question of to what extent prefrontal circuits are acting in concert with, or external to, the mesolimbic dopamine pathway remains unanswered.

Selective transfection of projections between the mPFC subregions and afferent targets, such as the NAc or amygdala, with chemogenetic constructs would enable facets of the disruption recorded in Chapter 4 to be defined with greater anatomical precision. Following the identification of specific pathways that contribute functionally to action control, the use of inhibitory opsins such as GtACR (Mohammad et al., 2017) would enable prefrontal projections to be disrupted on a trial-by-trial basis. Combining this approach with a method of recording such as FCV or fibre photometry would determine whether silencing this component of the circuit would occlude the pattern of dopamine suppression

observed on No-Go trials, in addition to any effects upon action restraint. These experiments would allow for a more nuanced understanding of how projection specific circuits within the mPFC may shape dopamine release in situations of action control, and determine to what degree the impairments of action restraint observed in Chapter 4 may be owed to dysregulated dopamine. Following the identification of the functional components in these circuits, employing recording techniques such as electrophysiology would enable understanding of how information may flow through the circuit.

Concluding Remarks

I have demonstrated the capacity for phasic dopamine signals to causally promote action over inaction, and determined the dependence of this effect upon proximity to external cues and magnitude of the phasic signal. I have also demonstrated crucial involvement of mPFC in reward guided action control, and distinct involvement of PL and IL in action restraint. I report that potentiating endocannabinoid signalling can impair focus, rather than boost motivation, in complex behavioural task. I outline the importance of mesolimbic dopamine to reward guided action and highlight how external circuits and systems may interact with mesolimbic dopamine to guide action appropriately.

References

Ahn, S., & Phillips, A. G. (1999). Dopaminergic Correlates of Sensory-Specific Satiety in the Medial Prefrontal Cortex and Nucleus Accumbens of the Rat. *Journal of Neuroscience*, *19*(19), RC29–RC29.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, *12*(10), 366–375.

Alex, K. D., & Pehek, E. A. (2007). Pharmacologic mechanisms of serotonergic regulation of dopamine neurotransmission. *Pharmacology & Therapeutics*, *113*(2), 296–320.

Alger, B. E., & Kim, J. (2011). Supply and demand for endocannabinoids. *Trends in Neurosciences*, *34*(6), 304–315.

Alvarez-Jaimes, L., Polis, I., & Parsons, L. H. (2009). Regional Influence of Cannabinoid CB 1 Receptors in the Regulation of Ethanol Self-Administration by Wistar Rats. *Open Neuropsychopharmacol J*, *2*, 77–85.

Atwood, B. K., & MacKie, K. (2010). CB2: a cannabinoid receptor with an identity crisis. *British Journal of Pharmacology*, *160*(3), 467–479.

Bamford, N. S., Robinson, S., Palmiter, R. D., Joyce, J. A., Moore, C., & Meshul, C. K. (2004). Dopamine Modulates Release from Corticostriatal Terminals. *Journal of Neuroscience*, *24*(43), 9541–9552.

Bamford, N. S., Wightman, R. M., & Sulzer, D. (2018). Dopamine's Effects on Corticostriatal Synapses during Reward-Based Behaviors. *Neuron*, *97*(3), 494–510.

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44–79.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious Mixed Models. *ArXiv:1506.04967*.

Beaulieu, J. M., & Gainetdinov, R. R. (2011). The Physiology, Signaling, and Pharmacology of Dopamine Receptors. *Pharmacological Reviews*, *63*(1), 182–217.

Bedse, G., Bluett, R. J., Patrick, T. A., Romness, N. K., Gaulden, A. D., Kingsley, P. J., Plath, N., Marnett, L. J., & Patel, S. (2018). Therapeutic endocannabinoid augmentation for mood and anxiety disorders: comparative profiling of FAAH, MAGL and dual inhibitors. *Translational Psychiatry 2018 8:1*, *8*(1), 1–14.

Berke, J. D. (2018). What does dopamine mean? In *Nature Neuroscience* (Vol. 21, Issue 6, pp. 787–793). Nature Publishing Group.

Berndt, A., Schoenenberger, P., Mattis, J., Tye, K. M., Deisseroth, K., Hegemann, P., & Oertner, T. G. (2011). High-efficiency channelrhodopsins for fast neuronal stimulation at low light levels. *Proc Natl Acad Sci U S A .*, *108*(18), 7595–7600.

Berridge, K. C. (2007). The debate over dopamine’s role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391–431.

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: “liking”, “wanting”, and learning. *Current Opinion in Pharmacology*, *9*(1), 65–73.

Berridge, K. C., Venier, I. L., & Robinson, T. E. (1989). Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: implications for arousal and anhedonia hypotheses of dopamine function. *Behavioral Neuroscience*, *103*(1), 36–45.

Blanco-Pozo, M., Akam, T., & Walton, M. E. (2023). Dopamine-independent state inference mediates expert reward guided decision making. *BioRxiv*, 2021.06.25.449995.

Boekhoudt, L., Voets, E. S., Flores-Dourojeanni, J. P., Luijendijk, M. C., JMJ Vanderschuren, L., & Adan, R. A. (2016). Chemogenetic Activation of Midbrain Dopamine Neurons Affects Attention, but not Impulsivity, in the Five-Choice Serial Reaction Time Task in Rats. *Nature Publishing Group*, *42*, 1315–1325.

Boekhoudt, L., Wijbrans, E. C., Man, J. H. K., Luijendijk, M. C. M., de Jong, J. W., van der Plasse, G., Vanderschuren, L. J. M. J., & Adan, R. A. H. (2018). Enhancing excitability of dopamine neurons promotes motivational behaviour through increased action initiation. *European Neuropsychopharmacology*, *28*(1), 171–184.

Bolam, J. P., Hanely, J. J., Booth, P. A. C., & Bevan, M. D. (2000). Synaptic organisation of the basal ganglia. *The Journal of Anatomy*, *196*(4), 527–542.

Brog, J. S., Salyaf, A., Deutch, A. Y., Zahm, D. S., & and, D. (1993). The patterns of afferent innervation of the core and shell in the “accumbens” part of the rat ventral striatum: immunohistochemical detection of retrogradely transported. *Wiley Online Library*, *338*(2), 255–278.

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834.

Burns, H. D., Van Laere, K., Sanabria-Bohórquez, S., Hamill, T. G., Bormans, G., Eng, W. S., Gibson, R., Ryan, C., Connolly, B., Patel, S., Krause, S., Vanko, A., Van Hecken, A., Dupont, P., De Lepeleire, I., Rothenberg, P., Stoch, S. A., Cote, J., Hagmann, W. K., …

Hargreaves, R. J. (2007). [18F]MK-9470, a positron emission tomography (PET) tracer for in vivo human PET brain imaging of the cannabinoid-1 receptor. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(23), 9800–9805.

Busquets-Garcia, A., Desprez, T., Metna-Laurent, M., Bellocchio, L., Marsicano, G., & Soria-Gomez, E. (2015). Dissecting the cannabinergic control of behavior: The where matters. *BioEssays*, *37*(11), 1215–1225.

Cachope, R., & Cheer, J. F. (2014). Local control of striatal dopamine release. *Frontiers in Behavioral Neuroscience*, *8*(MAY), 188.

Capriles, N., Rodaros, D., Sorge, R. E., & Stewart, J. (2003). A role for the prefrontal cortex in stress- and cocaine-induced reinstatement of cocaine seeking in rats. *Psychopharmacology*, *168*(1–2), 66–74.

Capuzzo, G., & Floresco, S. B. (2020). Prelimbic and Infralimbic Prefrontal Regulation of Active and Inhibitory Avoidance and Reward-Seeking. *Journal of Neuroscience*, *40*(24), 4773–4787.

Carr, D. B., & Sesack, S. R. (2000). Projections from the Rat Prefrontal Cortex to the Ventral Tegmental Area: Target Specificity in the Synaptic Associations with Mesoaccumbens and Mesocortical Neurons. *Journal of Neuroscience*, *20*(10), 3864–3873.

Cazorla, M., deCarvalho, F. D., Chohan, M. O., Shegda, M., Chuhma, N., Rayport, S., Ahmari, S. E., Moore, H., & Kellendonk, C. (2014). Dopamine D2 receptors regulate the anatomical and functional balance of basal ganglia circuitry. *Neuron*, *81*(1), 153–164.

Chang, C. Y., Esber, G. R., Marrero-Garcia, Y., Yau, H. J., Bonci, A., & Schoenbaum, G. (2015). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nature Neuroscience 2016 19:1*, *19*(1), 111–116.

Cheer, J. F., Marsden, C. A., Kendall, D. A., & Mason, R. (2000). Lack of response suppression follows repeated ventral tegmental cannabinoid administration: an in vitro electrophysiological study. *Neuroscience*, *99*(4), 661–667.

Cheer, J. F., Wassum, K. M., Heien, M. L. A. V., Phillips, P. E. M., & Wightman, R. M. (2004). Cannabinoids Enhance Subsecond Dopamine Release in the Nucleus Accumbens of Awake Rats. *Journal of Neuroscience*, *24*(18), 4393–4400.

Cheer, J. F., Wassum, K. M., Sombers, L. A., Heien, M. L. A. V, Ariansen, J. L., Aragona, B. J., Phillips, P. E. M., & Wightman, R. M. (2007). *Brief Communications*.

Chudasama, Y., Doobay, V. M., & Liu, Y. (2012). Hippocampal-Prefrontal Cortical Circuit Mediates Inhibitory Response Control in the Rat. *Journal of Neuroscience*, *32*(32), 10915–10924.

Chudasama, Y., & Muir, J. L. (2001). Visual attention in the rat: A role for the prelimbic cortex and thalamic nuclei? *Behavioral Neuroscience*, *115*(2), 417–428.

Clapper, J. R., Henry, C. L., Niphakis, M. J., Knize, A. M., Coppola, A. R., Simon, G. M., Ngo, N., Herbst, R. A., Herbst, D. M., Reed, A. W., Cisar, J. S., Weber, O. D., Viader, A., Alexander, J. P., Cunningham, M. L., Jones, T. K., Fraser, I. P., Grice, C. A., Ezekowitz, R. A. B., … Blankman, J. L. (2018). Monoacylglycerol lipase inhibition in human and

rodent systems supports clinical evaluation of endocannabinoid modulatorss. *Journal of Pharmacology and Experimental Therapeutics*, *367*(3), 494–508

Coddington, L. T., & Dudman, J. T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nature Neuroscience*, *21*(11), 1563–1573.

Coddington, L. T., & Dudman, J. T. (2019). Learning from Action: Reconsidering Movement Signaling in Midbrain Dopamine Neuron Activity. In *Neuron* (Vol. 104, Issue 1, pp. 63–77). Cell Press.

Cogliati Dezza, I., Cleeremens, A., & Alexander, W. (2022). Independent and Interacting Value Systems for Reward and Information in the Human Brain. *ELife*, *11*.

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, *482*(7383), 85–88.

Collins, A. L., Aitken, T. J., Greenfield, V. Y., Ostlund, S. B., & Wassum, K. M. (2016). Nucleus Accumbens Acetylcholine Receptors Modulate Dopamine and Motivation. *Neuropsychopharmacology 2016 41:12*, *41*(12), 2830–2838.

Collins, A. L., Greenfield, V. Y., Bye, J. K., Linker, K. E., Wang, A. S., & Wassum, K. M. (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Scientific Reports*.

Colwill, R. M., & Rescorla, R. A. (1988). Associations Between the Discriminative Stimulus and the Reinforcer in Instrumental Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*(2), 155–164.

Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable electrode mapping study. *Brain Research*, *185*(1), 1–15.

Covey, D. P., Dantrassy, H. M., Yohn, S. E., Castro, A., Conn, P. J., Mateo, Y., & Cheer, J. F. (2018). Inhibition of endocannabinoid degradation rectifies motivational and dopaminergic deficits in the Q175 mouse model of Huntington's disease. *Neuropsychopharmacology 2018 43:10*, *43*(10), 2056–2063.

Covey, D. P., Mateo, Y., Sulzer, D., Cheer, J. F., & Lovinger, D. M. (2017). Endocannabinoid modulation of dopamine neurotransmission. *Neuropharmacology*, *124*, 52–61.

Cui, G., Jun, S. B., Jin, X., Pham, M. D., Vogel, S. S., Lovinger, D. M., & Costa, R. M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*, *494*(7436), 238–242.

Dahlstrom, A. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system. I. Demonstration of monoamines in the cell bodies of brain stem neurons. *Acta Physiol. Scand.*, *62*(232), 1–55.

Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience and Biobehavioral Reviews*, *28*(7), 771–784.

Dalley, J. W., & Roiser, J. P. (2012). Dopamine, serotonin and impulsivity. *Neuroscience*, *215*, 42–58.

Da Silva, J. A., Tecuapetla, F., Paixão, V., & Costa, R. M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature*, *554*(7691), 244–248.

Dayan, P., Niv, Y., Seymour, B., & D. Daw, N. (2006). The misbehavior of value and the discipline of the will. *Neural Networks*, *19*(8), 1153–1160.

Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature Neuroscience*, *10*(8), 1020–1028.

de Bruin, N. M. W. J., Lange, J. H. M., Kruse, C. G., Herremans, A. H., Schoffelmeer, A. N. M., van Drimmelen, M., & De Vries, T. J. (2011). SLV330, a cannabinoid CB1 receptor antagonist, attenuates ethanol and nicotine seeking and improves inhibitory response control in rats. *Behavioural Brain Research*, *217*(2), 408–415.

de Haan, R., Lim, J., van der Burg, S. A., Pieneman, A. W., Nigade, V., Mansvelder, H. D., & de Kock, C. P. J. (2018). Neural Representation of Motor Output, Context and Behavioral Adaptation in Rat Medial Prefrontal Cortex During Learned Behavior. *Frontiers in Neural Circuits*, *12*, 75.

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neurosciences*, *13*(7), 281–285.

Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge University Press.

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

Doig, N. M., Moss, J., & Bolam, J. P. (2010). Cortical and Thalamic Innervation of Direct and Indirect Pathway Medium-Sized Spiny Neurons in Mouse Striatum. *Journal of Neuroscience*, *30*(44), 14610–14618.

Dong, A., He, K., Dudok, B., Farrell, J. S., Guan, W., Liput, D. J., Puhl, H. L., Cai, R., Wang, H., Duan, J., Albarran, E., Ding, J., Lovinger, D. M., Li, B., Soltesz, I., & Li, Y. (2022). A fluorescent sensor for spatiotemporally resolved imaging of endocannabinoid dynamics in vivo. *Nature Biotechnology*, *40*(5), 787.

du Hoffmann, J., & Nicola, S. M. (2014). Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(43), 14349–14364.

Eagle, D. M., Bari, A., & Robbins, T. W. (2008). The neuropsychopharmacology of action inhibition: Cross-species translation of the stop-signal and go/no-go tasks. *Psychopharmacology*, *199*(3), 439–456.

Eagle, D. M., Wong, J. C. K., Allan, M. E., Mar, A. C., Theobald, D. E., & Robbins, T. W. (2011). Contrasting Roles for Dopamine D1 and D2 Receptor Subtypes in the Dorsomedial Striatum but Not the Nucleus Accumbens Core during Behavioral Inhibition in the Stop-Signal Task in Rats. *Journal of Neuroscience*, *31*(20), 7349–7356.

Egerton, A., Allison, C., Brett, R. R., & Pratt, J. A. (2006). Cannabinoids and prefrontal cortical function: insights from preclinical studies. *Neuroscience and Biobehavioral Reviews*, *30*(5), 680–695.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience 2017 18:9*, *18*(9), 547–558.

Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H. J., Ornelas, S., Koay, S. A., Thiberge, S. Y., Daw, N. D., Tank, D. W., & Witten, I. B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*, *570*(7762), 509–513.

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

Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The Role of Medial Prefrontal Cortex in Memory and Decision Making. *Neuron*, *76*(6), 1057–1070.

Fattore, L., Melis, M., Fadda, P., Pistis, M., & Fratta, W. (2010). The endocannabinoid system and nondrug rewarding behaviours. *Experimental Neurology*, *224*(1), 23–36.

Feja, M., Leigh, M. P. K., Baindur, A. N., McGraw, J. J., Wakabayashi, K. T., Cravatt, B. F., & Bass, C. E. (2020). The novel MAGL inhibitor MJN110 enhances responding to reward-predictive incentive cues by activation of CB1 receptors. *Neuropharmacology*, *162*, 107814.

Ferenczi, E. A., Zalocusky, K. A., Liston, C., Grosenick, L., Warden, M. R., Amatya, D., Katovich, K., Mehta, H., Patenaude, B., Ramakrishnan, C., Kalanithi, P., Etkin, A.,

Knutson, B., Glover, G. H., & Deisseroth, K. (2016). Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science (New York, N.Y.)*, *351*(6268), aac9698.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., & Akil, H. (2010). A selective role for dopamine in stimulus–reward learning. *Nature 2010 469:7328*, *469*(7328), 53–57.

Floresco, S. B., Yang, C. R., Phillips, A. G., & Blaha, C. D. (1998). Association Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *European Journal of Neuroscience*, *10*(4), 1241–1251.

Flores-Dourojeanni, J. P., Van Rijt, C., Van Den Munkhof, M. H., Boekhoudt, L., Luijendijk, M. C. M., Vanderschuren, L. J. M. J., & Adan, R. A. H. (2021). Temporally Specific Roles of Ventral Tegmental Area Projections to the Nucleus Accumbens and Prefrontal Cortex in Attention and Impulse Control. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *41*(19), 4293–4304.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolotti, G. (2005). Neuroscience: Parietal lobe: From action organization to intention understanding. *Science*, *308*(5722), 662–667

Fouriezos, G., Hansson, P., & Wise, R. A. (1978). Neuroleptic-induced attenuation of brain stimulation reward in rats. *Journal of Comparative and Physiological Psychology*, *92*(4), 661–671.

Fouriezos, G., & Wise, R. A. (1976). Pimozide-induced extinction of intracranial self-stimulation: response patterns rule out motor or performance deficits. *Brain Research*, *103*(2), 377–380.

Gabbott, P. L. A., Warner, T. A., Jays, P. R. L., Salway, P., & Busby, S. J. (2005). Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. *Journal of Comparative Neurology*, *492*(2), 145–177.

Gan, J. O., Walton, M. E., & Phillips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience*, *13*(1), 25–27.

Geisler, S., & Wise, R. A. (2008). Functional implications of glutamatergic projections to the ventral tegmental area. *Reviews in the Neurosciences*, *19*(4–5), 227–244.

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 Dopamine Receptor-regulated Gene Expression of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429–1432.

Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of Striatal Projection Systems by Dopamine. *Annu Rev Neurosci*

German, D. C., & Manaye, K. F. (1993). Midbrain dopaminergic neurons (nuclei A8, A9, and A10): Three-dimensional reconstruction in the rat. *Journal of Comparative Neurology*, *331*(3), 297–309.

Gershman, S. J., Guitart-Masip, M., & Cavanagh, J. F. (2021). Neural signatures of arbitration between Pavlovian and instrumental action selection. *PLOS Computational Biology*, *17*(2), e1008553.

Gessa, G., Melis, M., Muntoni, A., & Diana, M. (1998). Cannabinoids activate mesolimbic dopamine neurons by an action on cannabinoid CB1 receptors. *European Journal of Pharmacology*, *341*(1), 39–44.

Gil-Ordóñez, A., Martín-Fontecha, M., Ortega-Gutiérrez, S., & López-Rodríguez, M. L. (2018). Monoacylglycerol lipase (MAGL) as a promising therapeutic target. *Biochemical Pharmacology*, *157*, 18–32.

Glass, M., Dragunow, M., & Faull, R. L. M. (1997). Cannabinoid receptors in the human brain: a detailed anatomical and quantitative autoradiographic study in the fetal, neonatal and adult human brain. *Neuroscience*, *77*(2), 299–318.

Gourley, S. L., & Taylor, J. R. (2016). Going and stopping: dichotomies in behavioral control by the prefrontal cortex. *Nature Neuroscience 2016 19:5*, *19*(5), 656–664.

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1), 1–24.

Grace, A. A., & Bunney, B. S. (1984). The control of firing pattern in nigral dopamine neurons: burst firing. *Journal of Neuroscience*, *4*(11), 2877–2890.

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences*, *30*(5), 220–227.

Grima, L. L., Panayi, M. C., Härmson, O., Syed, E. C. J., Manohar, S. G., Husain, M., & Walton, M. E. (2022). Nucleus accumbens D1-receptors regulate and focus transitions to reward-seeking action. *Neuropsychopharmacology 2022*, 1–11.

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202.

Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action Dominates Valence in Anticipatory Representations in the Human Striatum and Dopaminergic Midbrain*. Journal of Neuroscience*

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal Pathways in Primates Form an Ascending Spiral from the Shell to the Dorsolateral Striatum. *Journal of Neuroscience*, *20*(6), 2369–2382.

Halladay, L. R., & Blair, H. T. (2015). Distinct ensembles of medial prefrontal cortex neurons are activated by threatening stimuli that elicit excitation vs. inhibition of movement. *Journal of Neurophysiology*, *114*(2), 793–807.

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

Hardung, S., Epple, R., Jäckel, Z., Eriksson, D., Uran, C., Senn, V., Gibor, L., Yizhar, O., & Diester, I. (2017a). A Functional Gradient in the Rodent Prefrontal Cortex Supports Behavioral Inhibition. *Current Biology*, *27*(4), 549–555.

Hardung, S., Epple, R., Jäckel, Z., Eriksson, D., Uran, C., Senn, V., Gibor, L., Yizhar, O., & Diester, I. (2017b). A Functional Gradient in the Rodent Prefrontal Cortex Supports Behavioral Inhibition. *Current Biology*, *27*(4), 549–555.

Härmson, O., Grima, L. L., Panayi, M. C., Husain, M., & Walton, M. E. (2022). 5-HT2C receptor perturbation has bidirectional influence over instrumental vigour and restraint. *Psychopharmacology*, *239*(1), 123–140.

Harrison, A. A., Everitt, B. J., & Robbins, T. W. (1997). Central 5-HT depletion enhances impulsive responding without affecting the accuracy of attentional performance: Interactions with dopaminergic mechanisms. *Psychopharmacology*, *133*(4), 329–342.

Harrison, A. A., Everitt, B. J., & Robbins, T. W. (1999). Central serotonin depletion impairs both the acquisition and performance of a symmetrically reinforced go/no-go conditional visual discrimination. *Behavioural Brain Research*, *100*(1–2), 99–112.

Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Accumbens Symmetrically Encodes a Reward Prediction Error Term. *Journal of Neuroscience*, *34*(3), 698–704.

Horst, N. K., & Laubach, M. (2013). Reward-related activity in the medial prefrontal cortex is driven by consumption. *Frontiers in Neuroscience*, *7*(7 APR), 56.

Howard, C. D., Li, H., Geddes, C. E., & Jin, X. (2017). Dynamic Nigrostriatal Dopamine Biases Action Selection. *Neuron*, *93*(6), 1436-1450.e8.

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

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature 2013 500:7464*, *500*(7464), 575–579.

Howlett, A. C., & Mukhopadhyay, S. (2000). Cellular signal transduction by anandamide and 2-arachidonoylglycerol. *Chemistry and Physics of Lipids*, *108*(1–2), 53–70.

Hughes R, Bakhurin K, Petter E, Watson G, Kim N, Freidman A, & Yin H. (2020). Ventral tegmental dopamine neurons control the impulse vector during motivated behavior. *Current Biology*, *30*(14), 2681–2694.

Humphries, M. D., & Prescott, T. J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Progress in Neurobiology*, *90*(4), 385–417.

Hyland, B. I., Reynolds, J. N. J., Hay, J., Perk, C. G., & Miller, R. (2002). Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience*, *114*(2), 475–492.

Ibáñez-Sandoval, O., Tecuapetla, F., Unal, B., Shah, F., Koós, T., & Tepper, J. M. (2010). Electrophysiological and morphological characteristics and synaptic connectivity of tyrosine hydroxylase-expressing neurons in adult mouse striatum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(20), 6999–7016.

Ikemoto, S. (2010). Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory. *Neuroscience and Biobehavioral Reviews*, *35*(2), 129.

Ikemoto, S., & Bonci, A. (2014). Neurocircuitry of drug reward. *Neuropharmacology*, *76*(PART B), 329–341.

Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008a). Contributions of the amygdala and medial prefrontal cortex to incentive cue responding. *Neuroscience*, *155*(3), 573–584.

Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008b). Dorsomedial Prefrontal Cortex Contribution to Behavioral and Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, *28*(19), 5088–5098.

Jackson, M. E., Frost, A. S., & Moghaddam, B. (2001). Stimulation of prefrontal cortex at physiologically relevant frequencies inhibits dopamine release in the nucleus accumbens. *Journal of Neurochemistry*, *78*(4), 920–923.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry*, *79*, 368–376.

Jenkins, G., & Walton, M. (2020). Dopamine: Don't Underestimate the Force. *Current Biology*, *30*(14), R824–R826.

Jenni, N. L., Symonds, N., & Floresco, S. B. (2023). Medial orbitofrontal cortical regulation of different aspects of Pavlovian and instrumental reward seeking. *Psychopharmacology*, *240*(3), 441–459.

Jones, J. L., Day, J. J., Aragona, B. J., Wheeler, R. A., Wightman, R. M., & Carelli, R. M. (2010). Basolateral Amygdala Modulates Terminal Dopamine Release in the Nucleus Accumbens and Conditioned Responding. *Biological Psychiatry*, *67*(8), 737–744.

Jo, Y. S., Lee, J., & Mizumori, S. J. Y. (2013). Effects of Prefrontal Cortical Inactivation on Neural Activity in the Ventral Tegmental Area. *Journal of Neuroscience*, *33*(19), 8159–8171.

Jo, Y. S., & Mizumori, S. J. Y. (2016). Prefrontal Regulation of Neuronal Activity in the Ventral Tegmental Area. *Cerebral Cortex*, *26*(10), 4057–4068.

Kamigaki, T., & Dan, Y. (2017). Delay activity of specific prefrontal interneuron subtypes modulates memory-guided behavior. *Nature Neuroscience 2017 20:6*, *20*(6), 854–863.

Kamin, L. (1968). Attention-like processes in classical conditioning. *In Symposium on the Prediction of Behaviour*.

Keiflin, R., Pribut, H. J., Shah, N. B., & Janak, P. H. (2019). Ventral Tegmental Dopamine Neurons Participate in Reward Identity Predictions. *Current Biology*, *29*(1), 93-103.e3.

Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, *6*(OCTOBER 2012), 30039.

Kim, C. K., Ye, L., Jennings, J. H., Pichamoorthy, N., Tang, D. D., Yoo, A. C. W., Ramakrishnan, C., & Deisseroth, K. (2017). Molecular and Circuit-Dynamical Identification of Top-Down Neural Mechanisms for Restraint of Reward Seeking. *Cell*, *170*(5), 1013-1027.e14.

Kim, H., Malik, A. N., Mikhael, J. G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S. J., & Uchida, N. (2020). A Unified Framework for Dopamine Signals across Timescales. *Cell*, *183*(6), 1600-1616.e25.

Kim, J., Ghim, J. W., Lee, J. H., & Jung, M. W. (2013). Neural Correlates of Interval Timing in Rodent Prefrontal Cortex. *Journal of Neuroscience*, *33*(34), 13834–13847.

Kim, S., & Lee, D. (2011). Prefrontal cortex and impulsive decision making. *Biological Psychiatry*, *69*(12), 1140–1146.

Koch, M., Schmid, A., & Schnitzler, H. U. (2000). Role of nucleus accumbens dopamine D1 and D2 receptors in instrumental and pavlovian paradigms of conditioned reward. *Psychopharmacology*, *152*(1), 67–73.

Ko, D., & Wanat, M. J. (2016). Phasic Dopamine Transmission Reflects Initiation Vigor and Exerted Effort in an Action- and Region-Specific Manner. *Journal of Neuroscience*, *36*(7), 2202–2211.

Kravitz, A. V., Freeze, B. S., Parker, P. R. L., Kay, K., Thwin, M. T., Deisseroth, K., & Kreitzer, A. C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature 2010 466:7306*, *466*(7306), 622–626.

Kupchik, Y. M., Brown, R. M., Heinsbroek, J. A., Lobo, M. K., Schwartz, D. J., & Kalivas, P. W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nature Neuroscience 2015 18:9*, *18*(9), 1230–1232.

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2343–2348.

Lee, K., Claar, L. D., Hachisuka, A., Bakhurin, K. I., Nguyen, J., Trott, J. M., Gill, J. L., & Masmanidis, S. C. (2020). Temporally restricted dopaminergic control of reward-conditioned movements. *Nature Neuroscience 2020 23:2*, *23*(2), 209–216.

Lee, S.-H., Ledri, M., Tóth, B., Marchionni, I., Henstridge, C. M., Dudok, X., Kenesei, K., Barna, L., Szabó, S. I., Renkecz, T., Oberoi, X., Watanabe, M., Limoli, C. L., Horvai, G., Soltesz, I., & Katona, I. (2015). Multiple forms of endocannabinoid and endovanilloid signaling regulate the tonic control of GABA release. *Soc Neuroscience*, *35*(27), 10039–10057.

Legault, M., Rompré, P.-P., & Wise, R. A. (2000). Chemical Stimulation of the Ventral Hippocampus Elevates Nucleus Accumbens Dopamine by Activating Dopaminergic Neurons of the Ventral Tegmental Area. *Journal of Neuroscience*, *20*(4), 1635–1642.

Lex, A., & Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn Mem*, *15*(7), 283–491.

Lodge, D. J. (2011). The Medial Prefrontal and Orbitofrontal Cortices Differentially Regulate Dopamine System Function. *Neuropsychopharmacology 2011 36:6*, *36*(6), 1227–1236.

Lovibond, P. F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 225–247.

Luo, A. H., Tahsili-Fahadan, P., Wise, R. A., Lupica, C. R., & Aston-Jones, G. (2007). Luo et al (Wise) 2011- Linking context with reward: a functional circuit from hippocampal CA3 to ventral tegmental area. *Genome Project Data Processing Subgroup*, *3*, 4233.

Mackintosh, N. J. (1983). *Conditioning and associative learning.* Oxford University Press.

Mannella, F., Gurney, K., & Baldassarre, G. (2013). The nucleus accumbens as a nexus between values and goals in goal-directed behavior: A review and a new hypothesis. *Frontiers in Behavioral Neuroscience*, *7*(OCT), 51817.

Margolis, E. B., Lock, H., Chefer, V. D., Shippenberg, T. S., Hjelmstad, G. O., & Fields, H. L. (2006). κ opioids selectivity control dopaminergic neurons projecting to the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(8), 2938–2942.

Mateo, Y., Johnson, K. A., Covey, D. P., Atwood, B. K., Wang, H.-L., Zhang, S., Gildish, I., Cachope, R., Bellocchio, L., Guzmán, M., Morales, M., Cheer, J. F., & Lovinger, D. M. (2017). Endocannabinoid Actions on Cortical Terminals Orchestrate Local Modulation of Dopamine Release in the Nucleus Accumbens. *Neuron*, *96*(5), 1112-1126.e5.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

Mátyás, F., Urbán, G. M., Watanabe, M., Mackie, K., Zimmer, A., Freund, T. F., & Katona, I. (2008). Identification of the sites of 2-arachidonoylglycerol synthesis and action imply retrograde endocannabinoid signaling at both GABAergic and glutamatergic synapses in the ventral tegmental area. *Neuropharmacology*, *54*(1), 95–107.

McClure, S. M., Daw, N. D., & Read Montague, P. (2003). A computational substrate for incentive salience. *Trends in Neurosciences*, *26*(8), 423–428.

McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., & Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity encoding in the nucleus accumbens. *Neuron*, *78*(5), 910–922.

Melis, M., Pistis, M., Perra, S., Muntoni, A. L., Pillolla, G., & Gessa, G. L. (2004). Endocannabinoids Mediate Presynaptic Inhibition of Glutamatergic Transmission in Rat Ventral Tegmental Area Dopamine Neurons through Activation of CB1 Receptors. *Journal of Neuroscience*, *24*(1), 53–62.

Melis, T., Succu, S., Sanna, F., Boi, A., Argiolas, A., & Melis, M. R. (2007). The cannabinoid antagonist SR 141716A (Rimonabant) reduces the increase of extra-cellular dopamine release in the rat nucleus accumbens induced by a novel high palatable food. *Neuroscience Letters*, *419*(3), 231–235.

Miller, E. K., & Cohen, J. D. (2003). An Integrative Theory of Prefrontal Cortex Function. *Annu Rev Neurosci*, *24*, 167–202.

Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, *72*(2), 1024–1027.

Mogenson, G. J., Jones, D. L., & Yim, C. Y. (1980). From motivation to action: Functional interface between the limbic system and the motor system. *Progress in Neurobiology*, *14*(2–3), 69–97

Mohammad, F., Stewart, J. C., Ott, S., Chlebikova, K., Chua, J. Y., Koh, T. W., Ho, J., & Claridge-Chang, A. (2017). Optogenetic inhibition of behavior with anion channelrhodopsins. *Nature Methods 2017 14:3*, *14*(3), 271–274.

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J.-M. T., Vinson, L. T., Patriarchi, T., Tian, L., Kennedy, R. T., & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, 1.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(5), 1936–1947.

Moreira, F. A., Jupp, B., Belin, D., & Dalley, J. W. (2015). Endocannabinoids and striatal function: implications for addiction-related behaviours. *Behavioural Pharmacology*, *26*, 59–72.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, *43*(1), 133–143.

Moscarello, J. M., & LeDoux, J. E. (2013). Active Avoidance Learning Requires Prefrontal Suppression of Amygdala-Mediated Defensive Reactions. *Journal of Neuroscience*, *33*(9), 3815–3823.

Münster, A., & Hauber, W. (2018). Medial Orbitofrontal Cortex Mediates Effort-related Responding in Rats. *Cerebral Cortex*, *28*(12), 4379–4389.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience 2008 9:11*, *9*(11), 856–869.

Narayanan, N. S., Horst, N. K., & Laubach, M. (2006). Reversible inactivations of rat medial prefrontal cortex impair the ability to wait for a stimulus. *Neuroscience*, *139*(3), 865–876.

Narayanan, N. S., & Laubach, M. (2006). Top-Down Control of Motor Cortex Ensembles by Dorsomedial Prefrontal Cortex. *Neuron*, *52*(5), 921–931.

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols 2019 14:7*, *14*(7), 2152–2176.

Nicola, S. M., Yun, I. A., Wakabayashi, K. T., & Fields, H. L. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance during a Discriminative Stimulus Task. *Journal of Neurophysiology*, *91*(4), 1840–1865.

Nicola, S., Taha, S., Kim, S., & Fields, H. (2005). Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. *Neuroscience*.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520.

Ohno-Shosaku, T., & Kano, M. (2014). Endocannabinoid-mediated retrograde modulation of synaptic transmission. *Current Opinion in Neurobiology*, *29*, 1–8.

Olds, J., & Milner, P. (1954). POSITIVE REINFORCEMENT PRODUCED BY ELECTRICAL STIMULATION OF SEPTAL AREA AND OTHER REGIONS OF RAT BRAIN. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–427.

Olds, J., & Olds, M. (1958). Positive Reinforcement Produced by Stimulating Hypothalamus with Iproniazid and Other Compounds. *Science*.

Oleson, E. B., Beckert, M. V., Morra, J. T., Lansink, C. S., Cachope, R., Abdullah, R. A., Loriaux, A. L., Schetters, D., Pattij, T., Roitman, M. F., Lichtman, A. H., & Cheer, J. F. (2012). Endocannabinoids Shape Accumbal Encoding of Cue-Motivated Behavior via CB1 Receptor Activation in the Ventral Tegmentum. *Neuron*, *73*(2), 360–373.

Oleson, E. B., Hamilton, L. R., & Gomez, D. M. (2021). Cannabinoid Modulation of Dopamine Release During Motivation, Periodic Reinforcement, Exploratory Behavior, Habit Formation, and Attention. *Frontiers in Synaptic Neuroscience*, *13*, 29.

Omelchenko, N., Synapse, S. S.-, & 2009, undefined. (2009). Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Wiley Online Library*, *63*(10), 895–906.

Öngür, D., & Price, J. L. (2000). The Organization of Networks within the Orbital and Medial Prefrontal Cortex of Rats, Monkeys and Humans. *Cerebral Cortex*, *10*(3), 206–219.

Ookschot, D. E. (1996). Total Number of Neurons in the Neostriatal, Pallidal, Subthalamic, and Substantia Nigral Nuclei of the Rat Basal Ganglia: A Stereological Study Using the Cavalieri and Optical Disector Methods. *THE ,JOURNAL OF COMPAKATIVE NEUROLOGY*, *366*, 580–599.

Orsini, C. A., Hernandez, C. M., Bizon, J. L., & Setlow, B. (2018). Deconstructing value-based decision making via temporally selective manipulation of neural activity: Insights from

rodent models. *Cognitive, Affective, & Behavioral Neuroscience 2018 19:3*, *19*(3), 459–476.

Ostlund, S. B., Wassum, K. M., Murphy, N. P., Balleine, B. W., & Maidment, N. T. (2011). Extracellular Dopamine Levels in Striatal Subregions Track Shifts in Motivation and Response Cost during Instrumental Conditioning. *Journal of Neuroscience*, *31*(1), 200–207.

Overton, P. G., Tong, Z. Y., & Clark, D. (1996). A pharmacological analysis of the burst events induced in midbrain dopaminergic neurons by electrical stimulation of the prefrontal cortex in the rat. *Journal of Neural Transmission*, *103*(5), 523–540.

Owen, S. F., Liu, M. H., & Kreitzer, A. C. (2019). Thermal constraints on in vivo optogenetic manipulations. *Nature Neuroscience 2019 22:7*, *22*(7), 1061–1065.

Owesson-White, C. A., Roitman, M. F., Sombers, L. A., Belle, A. M., Keithley, R. B., Peele, J. L., Carelli, R. M., & Wightman, R. M. (2012). Sources contributing to the average extracellular concentration of dopamine in the nucleus accumbens. *Journal of Neurochemistry*, *121*(2), 252–262.

Paine, T. A., Slipp, L. E., & Carlezon, W. A. (2011). Schizophrenia-Like Attentional Deficits Following Blockade of Prefrontal Cortex GABAA Receptors. *Neuropsychopharmacology 2011 36:8*, *36*(8), 1703–1713.

Pan, B., Hillard, C. J., & Liu, Q. S. (2008). Endocannabinoid Signaling Mediates Cocaine-Induced Inhibitory Synaptic Plasticity in Midbrain Dopamine Neurons. *Journal of Neuroscience*, *28*(6), 1385–1397.

Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., Mensh, B. D., Karpova, A. Y., & Dudman, J. T. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell*, *162*(6), 1418–1430.

Papageorgiou, G. K., Baudonnat, M., Cucca, F., & Walton, M. E. (2016). Mesolimbic Dopamine Encodes Prediction Errors in a State-Dependent Manner. *Cell Reports*, *15*(2), 221–228.

Parsons, L. H., & Hurd, Y. L. (2015). Endocannabinoid signalling in reward and addiction. *Nature Reviews Neuroscience 2015 16:10*, *16*(10), 579–594.

Passetti, F., Chudasama, Y., & Robbins, T. W. (2002a). The Frontal Cortex of the Rat and Visual Attentional Performance: Dissociable Functions of Distinct Medial Prefrontal Subregions. *Cerebral Cortex*, *12*(12), 1254–1268.

Passetti, F., Chudasama, Y., & Robbins, T. W. (2002b). The Frontal Cortex of the Rat and Visual Attentional Performance: Dissociable Functions of Distinct Medial Prefrontal Subregions. *Cerebral Cortex*, *12*(12), 1254–1268.

Pattij, T., Janssen, M. C. W., Schepers, I., González-Cuevas, G., De Vries, T. J., & Schoffelmeer, A. N. M. (2007). Effects of the cannabinoid CB1 receptor antagonist rimonabant on distinct measures of impulsive behavior in rats. *Psychopharmacology*, *193*(1), 85–96.

Pattij, T., Janssen, M. C. W., Vanderschuren, L. J. M. J., Schoffelmeer, A. N. M., & van Gaalen, M. M. (2007). Involvement of dopamine D1 and D2 receptors in the nucleus

accumbens core and shell in inhibitory response control. *Psychopharmacology*, *191*(3), 587–598.

Peters, K. Z., Oleson, E. B., & Cheer, J. F. (2020). A Brain on Cannabinoids: The Role of Dopamine Release in Reward Seeking and Addiction. *Cold Spring Harbor Perspectives in Medicine*, a039305.

Petrosino, S., & Di Marzo, V. (2010). FAAH and MAGL inhibitors: therapeutic opportunities from regulating endocannabinoid levels. *Current Opinion in Investigational Drugs (London, England : 2000)*, *11*(1), 51–62.

Pezze, M., Dalley, J., & Robbins, T. (2007). Differential roles of dopamine D1 and D2 receptors in the nucleus accumbens in attentional performance on the five-choice serial reaction time task. *Neuropsychopharmacology*, *32*(2), 273–283.

Pezze, M., McGarrity, S., Mason, R., Fone, K. C., & Bast, T. (2014). Too Little and Too Much: Hypoactivation and Disinhibition of Medial Prefrontal Cortex Cause Attentional Deficits. *Journal of Neuroscience*, *34*(23), 7931–7946.

Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V., Wightman, R. M., & Carelli, R. M. (2003). Erratum: Subsecond dopamine release promotes cocaine seeking. *Nature*, *422*(6932), 614–618.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience 2008 9:7*, *9*(7), 545–556.

Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. *N Black, A. & Prokasy, W. (Eds), Classical Conditioning II: Current Research and Theory. Appleton-Century-Crofts, New York, NY*, 64–99.

Reynolds, J. N. J., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature 2001 413:6851*, *413*(6851), 67–70.

Rhodes, S. E. V., & Killcross, A. S. (2007). Lesions of rat infralimbic cortex enhance renewal of extinguished appetitive Pavlovian responding. *The European Journal of Neuroscience*, *25*(8), 2498–2503.

Risterucci, C., Terramorsi, D., Nieoullon, A., & Amalric, M. (2003). Excitotoxic lesions of the prelimbic-infralimbic areas of the rodent prefrontal cortex disrupt motor preparatory processes. *European Journal of Neuroscience*, *17*(7), 1498–1508.

Robbins, T. W., & Everitt, B. J. (2007). A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). *Psychopharmacology*, *191*(3), 433–437.

Roberts, A. (2020). Prefrontal Regulation of Threat-Elicited Behaviors: A Pathway to Translation. *Https://Doi.Org/10.1146/Annurev-Psych-010419-050905*, *71*, 357–387.

Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*, *10*(12), 1615–1624.

Roitman, M. F., Stuber, G. D., Phillips, P. E. M., Wightman, R. M., & Carelli, R. M. (2004). Dopamine Operates as a Subsecond Modulator of Food Seeking. *J Neurosci*, *24*(6), 1265–1271.

Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol*, *63*(3), 592–606.

Saddoris, M. P., Sugam, J. A., Stuber, G. D., Witten, I. B., Deisseroth, K., & Carelli, R. M. (2015). Mesolimbic Dopamine Dynamically Tracks, and Is Causally Linked to, Discrete Aspects of Value-Based Decision Making. *Biological Psychiatry*, *77*(10), 903–911.

Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, *76*(3), 470–485.

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461–482.

Salamone, J. D., Cousins, M. S., & Snyder, B. J. (1997). Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. *Neuroscience and Biobehavioral Reviews*, *21*(3), 341–359.

Salamone, J. D., Wisniecki, A., Carlson, B. B., & Correa, M. (2001). Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience*, *105*(4), 863–870.

Satoh, T., Nakai, S., Sato, T., & Kimura, M. (2003). Correlated Coding of Motivation and Outcome of Decision by Dopamine Neurons. *Journal of Neuroscience*, *23*(30), 9913–9923.

Saunders, A., Oldenburg, I. A., Berezovskii, V. K., Johnson, C. A., Kingery, N. D., Elliott, H. L., Xie, T., Gerfen, C. R., & Sabatini, B. L. (2015). A direct GABAergic output from the basal ganglia to frontal cortex. *Nature 2015 521:7550*, *521*(7550), 85–89.

Saunders, B. T., Richard, J. M., Margolis, E. B., & Janak, P. H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nature Neuroscience*, *21*(8), 1072–1083.

Schlicker, E., & Kathmann, M. (2001). Modulation of transmitter release via presynaptic cannabinoid receptors. *Trends in Pharmacological Sciences*, *22*(11), 565–572.

Scholl, J., Kolling, N., Nelissen, N., Wittmann, M. K., Harmer, C. J., & Rushworth, M. F. S. (2015). The Good, the Bad, and the Irrelevant: Neural Mechanisms of Learning Real and Hypothetical Rewards and Effort. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *35*(32), 11233–11251.

Schultz, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *Journal of Neurophysiology*, *56*(5), 1439–1461.

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schultz, W. (2019a). Recent advances in understanding the role of phasic dopamine activity. In *F1000Research* (Vol. 8). F1000 Research Ltd.

Schultz, W. (2019b). Recent advances in understanding the role of phasic dopamine activity [version 1; peer review: 3 approved]. In *F1000Research* (Vol. 8). F1000 Research Ltd.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–1599.

Schultz, W., & Romo, R. (1990a). Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *Journal of Neurophysiology*, *63*(3), 607–624.

Schultz, W., & Romo, R. (1990b). Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *Journal of Neurophysiology*, *63*(3), 607–624.

Schultz, W., Ruffieux, A., & Aebischer, P. (1983). E ,xi i_’mental Bran Research The Activity of Pars Compacta Neurons of the Monkey Substantia Nigra in Relation to Motor Activation*. In *Exp Brain Res* (Vol. 51).

Sesack, S., Deutch, A., … R. R.-J. of C., & 1989, undefined. (1989). Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: An anterograde tract-tracing study with Phaseolus vulgaris. *Wiley Online Library*, *290*(2), 213–242.

Sharpe, M. J., & Killcross, S. (2015). The prelimbic cortex uses higher-order cues to modulate both the acquisition and expression of conditioned fear. *Frontiers in Systems Neuroscience*, *8*(JAN), 235.

Sharpe, M. J., Stalnaker, T., Schuck, N. W., Killcross, S., Schoenbaum, G., & Niv, Y. (2019). An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. *Annual Review of Psychology*, *70*, 53–76.

Sierra-Mercado, D., Padilla-Coreano, N., & Quirk, G. J. (2010). Dissociable Roles of Prelimbic and Infralimbic Cortices, Ventral Hippocampus, and Basolateral Amygdala in the Expression and Extinction of Conditioned Fear. *Neuropsychopharmacology 2011 36:2*, *36*(2), 529–538.

Simonnet, A., Cador, M., & Caille, S. (2013). Nicotine reinforcement is reduced by cannabinoid CB1 receptor blockade in the ventral tegmental area. *Addiction Biology*, *18*(6), 930–936.

Singmann, H., & Kellen, D. (2019). An Introduction to Mixed Models for Experimental Psychology. In *New Methods in Cognitive Psychology* (pp. 4–31).

Smith, Y., Raju, D. V., Pare, J. F., & Sidibe, M. (2004). The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends in Neurosciences*, *27*(9), 520–527.

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., Sousa, N., & Rodrigues, A. J. (2016). *Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation*.

Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neuroscience & Biobehavioral Reviews*, *68*, 370–386.

Solinas, M., & Goldberg, S. R. (2005). Motivational Effects of Cannabinoids and Opioids on Food Reinforcement Depend on Simultaneous Activation of Cannabinoid and Opioid Systems. *Neuropsychopharmacology 2005 30:11*, *30*(11), 2035–2045.

Sombers, L. A., Beyene, M., Carelli, R. M., & Wightman, R. M. (2009). Synaptic Overflow of Dopamine in the Nucleus Accumbens Arises from Neuronal Activity in the Ventral Tegmental Area. *Journal of Neuroscience*, *29*(6), 1735–1742.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Publishing Group*.

Stuber, G. D., Hnasko, T. S., Britt, J. P., Edwards, R. H., & Bonci, A. (2010). Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(24), 8229–8233.

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron*, *66*(3), 449–460.

Sulzer, D., Cragg, S. J., & Rice, M. E. (2016). Striatal dopamine neurotransmission: regulation of release and uptake. *Basal Ganglia*, *6*(3), 123–148.

Surmeier, D. J., Song, W. J., & Yan, Z. (1996). Coordinated Expression of Dopamine Receptors in Neostriatal Medium Spiny Neurons. *Journal of Neuroscience*, *16*(20), 6579–6591.

Syed, E. C. J., Grima, L. L., Magill, P. J., Bogacz, R., Brown, P., & Walton, M. E. (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature Neuroscience*, *19*(1), 34–36.

Szabo, B., Siemes, S., & Wallmichrath, I. (2002). SHORT COMMUNICATION Inhibition of GABAergic neurotransmission in the ventral tegmental area by cannabinoids. *European Journal of Neuroscience*, *15*(12), 2057–2061.

Tai, S., & Fantegrossi, W. E. (2014). Synthetic Cannabinoids: Pharmacology, Behavioral Effects, and Abuse Potential. *Current Addiction Reports*, *1*(2), 129–136.

Terra, H., Bruinsma, B., de Kloet, S. F., van der Roest, M., Pattij, T., & Mansvelder, H. D. (2020). Prefrontal Cortical Projection Neurons Targeting Dorsomedial Striatum Control Behavioral Inhibition. *Current Biology*, *30*(21), 4188-4200.e5.

Threlfell, S., Lalic, T., Platt, N. J., Jennings, K. A., Deisseroth, K., & Cragg, S. J. (2012). Striatal Dopamine Release Is Triggered by Synchronized Activity in Cholinergic Interneurons. *Neuron*, *75*(1), 58–64.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science (New York, N.Y.)*, *307*(5715), 1642–1645.

Tong, Z.-Y., Overton, P. G., & Clark, D. (1996). *Stimulation of the Prefkontal Cortex in the Rat Induces Patterns of Activity in Midbrain Dopaminergic Neurons Which Resemble Natural Burst Events*. *22*, 195–208.

Tsou, K., Brown, S., Sañudo-Peña, M. C., Mackie, K., & Walker, J. M. (1998). Immunohistochemical distribution of cannabinoid CB1 receptors in the rat central nervous system. *Neuroscience*, *83*(2), 393–411.

Tsutsui-Kimura, I., Matsumoto, H., Akiti, K., Yamada, M. M., Uchida, N., & Watabe-Uchida, M. (2020). Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *ELife*, *9*, 1–39.

van Elzelingen, W., Warnaar, P., Matos, J., Bastet, W., Jonkman, R., Smulders, D., Goedhoop, J., Denys, D., Arbab, T., & Willuhn, I. (2022). Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. *Current Biology*, *32*(5), 1163-1174.e6.

Van Gaalen, M. M., Brueggeman, R. J., Bronius, P. F. C., Schoffelmeer, A. N. M., & Vanderschuren, L. J. M. J. (2006). Behavioral disinhibition requires dopamine receptor activation. *Psychopharmacology*, *187*(1), 73–85.

van Heukelum, S., Mars, R. B., Guthrie, M., Buitelaar, J. K., Beckmann, C. F., Tiesinga, P. H. E., Vogt, B. A., Glennon, J. C., & Havenith, M. N. (2020). Where is Cingulate Cortex? A Cross-Species View. *Trends in Neurosciences*, *43*(5), 285–299.

van Holstein, M., & Floresco, S. B. (2019). Dissociable roles for the ventral and dorsal medial prefrontal cortex in cue-guided risk/reward decision making. *Neuropsychopharmacology 2019 45:4*, *45*(4), 683–693.

Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology*.

Verharen, J., Van Den Heuvel, M. ., Luijendijk, M., Vanderschuren, & Adan, R. (2019). Corticolimbic Mechanisms of Behavioral Inhibition under Threat of Punishment. *Journal of Neuroscience*, *39*(22), 4353–4364.

Vertes, R. P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse*, *51*(1), 32–58.

Vertes, R. P. (2006). Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neuroscience*, *142*(1), 1–20.

Vicente, A. M., Galvão-Ferreira, P., Tecuapetla, F., & Costa, R. M. (2016). Direct and indirect dorsolateral striatum pathways reinforce different action strategies. In *Current Biology* (Vol. 26, Issue 7, pp. R267–R269). Cell Press.

Voorn, P., Vanderschuren, L. J. M. J., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. A. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, *27*(8), 468–474.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature 2001 412:6842*, *412*(6842), 43–48.

Wanat, M. J., Kuhnen, C. M., & Phillips, P. E. M. (2010). Delays Conferred by Escalating Costs Modulate Dopamine Release to Rewards But Not Their Predictors. *Journal of Neuroscience*, *30*(36), 12020–12027.

Wang, H., & Pickel, V. M. (2002). Dopamine D2 receptors are present in prefrontal cortical afferents and their targets in patches of the rat caudate-putamen nucleus. *The Journal of Comparative Neurology*, *442*(4), 392–404.

Ward, S. J., Walker, E. A., & Dykstra, L. A. (2007). Effect of Cannabinoid CB1 Receptor Antagonist SR141714A and CB1 Receptor Knockout on Cue-Induced Reinstatement of Ensure® and Corn-Oil Seeking in Mice. *Neuropsychopharmacology 2007 32:12*, *32*(12), 2592–2600.

Wassum, K. M., Ostlund, S. B., & Maidment, N. T. (2012). Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. *Biological Psychiatry*, *71*(10), 846–854.

Watabe-Uchida, M., & Uchida, N. (2018). Multiple Dopamine Systems: Weal and Woe of Dopamine. *Cold Spring Harbor Symposia on Quantitative Biology*, *83*, 83–95.

Wenzel, J. M., Oleson, E. B., Gove, W. N., Cole, A. B., Gyawali, U., Dantrassy, H. M., Bluett, R. J., Dryanovski, D. I., Stuber, G. D., Deisseroth, K., Mathur, B. N., Patel, S., Lupica, C. R., & Cheer, J. F. (2018). Phasic Dopamine Signals in the Nucleus Accumbens that Cause Active Avoidance Require Endocannabinoid Mobilization in the Midbrain. *Current Biology*, *28*(9), 1392-1404.e5.

Wiley, J., pharmacology, B. M.-E. journal of, & 2003, undefined. (n.d.). Cannabinoid pharmacological properties common to other centrally acting drugs. *Elsevier*. Retrieved May 15, 2023, from

Wilson, R. I., & Nicoll, R. A. (2001). Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature 2001 410:6828*, *410*(6828), 588–592.

Winstanley, C. A., Zeeb, F. D., Bedard, A., Fu, K., Lai, B., Steele, C., & Wong, A. C. (2010). Dopaminergic modulation of the orbitofrontal cortex affects attention, motivation and impulsive responding in rats performing the five-choice serial reaction time task. *Behavioural Brain Research*, *210*(2), 263–272.

Wise, R. A. (1978). Neuroleptic attenuation of intracranial self-stimulation: Reward or performance deficits? *Life Sciences*, *22*(7), 535–542.

Wise, R. A. (1982). Neuroleptics and operant behavior: The anhedonia hypothesis. *Behavioral and Brain Sciences*, *5*(1), 39–53.

Wise, R. A. (2008). Dopamine and reward: The anhedonia hypothesis 30 years on. *Neurotoxicity Research 2008 14:2*, *14*(2), 169–183.

Wise, R. A., & Schwartz, H. V. (1981). Pimozide attenuates acquisition of lever-pressing for food in rats. *Pharmacology Biochemistry and Behavior*, *15*(4), 655–656.

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends in Neurosciences*, *31*(12), 599–608.

Wiskerke, J., & Pattij, T. (2015). The cannabinoid system and impulsive behavior. *Cannabinoids in Neurologic and Mental Disease*, 343–364.

Wiskerke, J., Stoop, N., Schetters, D., Schoffelmeer, A. N. M., & Pattij, T. (2011). Cannabinoid CB1 Receptor Activation Mediates the Opposing Effects of Amphetamine on Impulsive Action and Impulsive Choice. *PLOS ONE*, *6*(10), e25856.

Wu, S., Qian, C., Shen, X., Zhang, X., Huang, Y., Chen, S., & Wang, Y. (2022). Spike prediction on primary motor cortex from medial prefrontal cortex during task learning. *Journal of Neural Engineering*, *19*(4), 046025

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G. C. R., Urakubo, H., Ishii, S., & Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science*, *345*(6204), 1616–1620.

Yamanaka, K., Hori, Y., Minamimoto, T., Yamada, H., Matsumoto, N., Enomoto, K., Aosaki, T., Graybiel, A. M., & Kimura, M. (2018). Roles of centromedian parafascicular nuclei of thalamus and cholinergic interneurons in the dorsal striatum in associative learning of environmental events. *Journal of Neural Transmission*, *125*(3), 501–513.

Yan, Z., Song, W. J., & Surmeier, D. J. (1997). D2 dopamine receptors reduce N-type Ca2+ currents in rat neostriatal cholinergic interneurons through a membrane-delimited, protein-kinase-C- insensitive pathway. *Journal of Neurophysiology*, *77*(2), 1003–1015.

Yapo, C., Nair, A. G., Clement, L., Castro, L. R., Hellgren Kotaleski, J., & Vincent, P. (2017). Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. *The Journal of Physiology*, *595*(24), 7451–7475.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience 2006 7:6*, *7*(6), 464–476.

Yokel, R. A., & Wise, R. A. (1976). Attenuation of intravenous amphetamine reinforcement by central dopamine blockade in rats. *Psychopharmacology*, *48*(3), 311–318.

Yun, I. A., Nicola, S. M., & Fields, H. L. (2004). Contrasting effects of dopamine and glutamate receptor antagonist injection in the nucleus accumbens suggest a neural mechanism underlying cue-evoked goal-directed behavior. *European Journal of Neuroscience*, *20*(1), 249–263.

Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, *24*(12), 2923–2933.

Zelikowsky, M., Bissiere, S., Hast, T. A., Bennett, R. Z., Abdipranoto, A., Vissel, B., & Fanselow, M. S. (2013). Prefrontal microcircuit underlies contextual learning after hippocampal loss. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(24), 9938–9943.

Zhou, F.-M., Wilson, C., & Dani, J. (2003). Muscarinic and Nicotinic Cholinergic Mechanisms in the Mesostriatal Dopamine Systems. *Neuroscientist*, *9*, 23–36.