Investigating the neurochemical basis of action initiation, selection, and inhibition

Laura L. Grima

St. John's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2018

Acknowledgements

The image of a lone scientist toiling away at the bench in isolation might appear in many people's minds at first mention of the scientific vocation, but I am incredibly fortunate that this depiction could not be further from my own experiences as a Ph.D. student. In fact, if anything, these last four years have been as much shaped by the many wonderful individuals with whom I have had the privilege to work than the science itself, and for that I am incredibly thankful.

Two people who have done much of this shaping are my supervisors, Mark Walton and Masud Husain. Masud, thank you for your calm guidance and words of wisdom ("a page every day!"), and for your kindness, warmth, and support. I have learnt much from your eye for neat experimental design and uncanny ability to see the woods for the trees in a data set, not to mention your effectiveness at communicating your science to others. And Mark, you have gone above and beyond in your role to act not just as a supervisor, but as a mentor. Thank you for being so generous with your time and energy, for letting me shamelessly appropriate many of your ideas, and for ensuring that every piece of work I did – be it a poster, conference abstract, or indeed this very thesis – was as good as it could be. You have taught me the importance of a thorough and rigorous approach to neuroscience, behaviour, and psychology. Thank you for allowing me to make mistakes, for never making me feel stupid for asking a question, and for always being so encouraging.

Being situated in not one, but two labs full to the brim with smart and passionate scientists has enabled me to benefit both from the wisdom and from the friendship of many very talented individuals. From the Walton lab, I must begin by thanking Emilie Syed for taking on a student who had never so much as picked up a rat and for training me in the dark art of voltammetry, the results of which ended up spawning this whole thesis. Importantly, you showed me that neuroscience could both be an exciting challenge but also a lot of fun, so thank you.

I must also thank Marios Panagi on multiple counts: firstly, for being so integral to the collection of the local infusion data presented in Chapter 5, which in many ways was a turning point of this thesis, and secondly for being the 'highly specific old school stats man' that I could consult at any point for advice and guidance – on statistics, science more generally, and everything else in between. Similarly, Oliver Harmson was a key part of the experiments and data collection for several of the chapters presented here. Oliver, thank you for your enthusiasm and positivity, for being a real driving force behind these studies, and for always keeping me on my toes.

There are many others from the Walton lab who have contributed to other aspects of this thesis outside of experimental work directly. I have to thank Lauren Burgeno for feeding me delicious food, leaving post-it notes of profanity-laden reassurance on my computer, and being an all-round splendid friend during the writing up period. Tom JP, thank you for coming to my aid whenever I was in a MATLAB hole, for sending cat videos when they were much needed, and for ensuring that a good night out never came to a premature end. Emilie 'tooth breaker' Werlen has also been a voice of

reason throughout writing up; thank you for your encouragement and for those fun, candid conversations we've had after a few cocktails. Similarly, Anna Huber has been a calming and lucid presence throughout the Ph.D., so thank you. Thomas 'Metacam' Akam, thank you for introducing me to the world of Python and mouse training – I look forward to getting back to rummaging soon. And Clio Korn, thank you for being a wonderful conference companion, for being an inspiration both scientifically and as a friend, and for always reminding me to have "courage!"

My time in Oxford began with me joining the Husain lab as a Master's student before continuing onto a Ph.D., and I can't think of a more mentally stimulating group of individuals to have been surrounded by as I delved into my first substantial research projects. In particular, Edwin Dalmaijer has been my Husain lab sibling and an excellent source of scientific conversation, coding counsel, and lewd jokes. Thank you for celebrating the good times with me, but also for being there on the inevitable occasion that things weren't going quite to plan. Sean Fallon, thank you for your generous friendship and candid advice. I have always admired your ability to design sharp, clever studies and to ask equally sharp questions. I must also thank Matt Apps, who, instead of being perturbed by me following him from Royal Holloway to Oxford, was willing to take me on as a student that first year, and who has remained a good friend and scientific guide ever since. Kathrin Giehl, you're one of my favourite people. Thank you for your wonderful companionship and for all of the fun times we've had. Similarly, Annie Sillence, Campbell Le Heron, Yuen-Siang Ang, Kinan Muhammed, Nahid Zokaei, and Sanjay Manohar, thank you all for making my time in the Husain lab so memorable and full of great discussions, scientific and non-scientific alike.

There are other individuals outside of these labs without whom this work could not have been completed. Katie Hewitt, there is no doubt that the BNU would fall apart without you. Thank you for enduring my pestering of you with patience and kindness. David Bannerman, as PPL holder you are integral to all of the work done here. Thank you for taking on that responsibility whilst somehow maintaining your exuberance, and also for being a real source of positivity throughout the Ph.D. I must thank Greg Daubney, both for histology and for letting me use his secret stash of syringes whenever I forgot to bring my own. Liz Tunbridge and Rafal Bogacz, thank you for your invaluable advice and feedback at each important milestone of this thesis. Thank you, Nick Yeung, for being on the ball as DGS, and thank you Kate Nation for your support as college advisor.

In the second summer of the Ph.D., I was fortunate to be able to spend four months at the Champalimaud Centre for the Unknown. Joe Paton warmly welcomed me into his lab in sunny Lisbon, and the experience further cemented my desire to continue working in science. Joe, thank you for your time and hospitality. Your entire lab was both scientifically rigorous but also wonderfully congenial. I must thank Sofia Soares and Bruno Cruz in particular for their time spent showing me the fibre photometry ropes, not to mention introducing me to all that the beautiful city of Lisbon had to offer.

My home these past four years has been St. John's College – both in the literal and in the figurative sense. Whilst living here I have been fortunate to have had a family of sorts, a group of college mates who have become close friends. Thanks especially to Mikayla Hunter, Sally Le Page, Ellie Milnes-Smith, Chris Arran, Anthony Payne, Kieran Campbell, Olivier Lennon, Marina Lambrakis, Eden Tanner, and everyone else in the Fab House for your camaraderie – and for summer evenings spent in the college bar, Bananagrams in the pub, movie nights in the MCR, and for ensuring that I was never alone in taking advantage of all the free cheese and wine college had to offer.

Last, but most certainly not least, I must thank my parents. Thank you for your unconditional love and support, and for always being there for me. This thesis is for you.

Investigating the neurochemical basis of action initiation, selection, and inhibition

Laura L. Grima
St. John's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2018

Abstract

Two central monoamine systems, dopamine and serotonin, are thought to be integral to the promotion or inhibition of goal-directed action respectively, but how this information is conveyed in tandem with reward-related signals is not yet fully understood. In particular, whilst reward prediction error encoding by midbrain dopamine activity and release is well replicated, this contrasts with a wealth of pharmacological evidence emphasising dopamine's role in *activation of behaviour* for reward. In Chapter 3, we establish a novel cue-instructed behavioural paradigm where animals must either enact or entirely restrain movement with the promise of receiving a small or large reward: the Go/No-Go task. We show that the phasic dopaminergic signal is shaped both by reward size on offer and by movement required to garner reward. To establish whether this phasic signal is causally *driving* behaviour, in Chapter 4 we systemically manipulate D1-like receptors – thought to be most sensitive to phasic dopamine release – to show that global modulation of this receptor mediates action initiation, inhibition, and execution, and biases action selection towards highly rewarding responses. In Chapter 5 the ability to initiate or inhibit cue-driven action exclusively is localised to D1-like receptors in the core of the nucleus accumbens. D2-like receptors have been considered important for the inhibition of action, but more recently it has been hypothesised that they may in fact work synergistically with D1-like receptors to promote reward-guided action. In Chapter 6 we show that these receptors indeed gate goal-directed behaviour, but only when movement is required. Finally, the serotonergic 5-HT_{2C} receptor has been demonstrated as key to the ability to wait for reward but whether this is dependent on movement has not yet been explored. Therefore, in Chapter 7, we systemically and locally apply a 5-HT_{2C} receptor ligand to show that disruption of this receptor's normal function reduces behavioural inhibition whilst simultaneously improving both speed and ability to enact goal-directed behaviour, but that the locus of these effects is extra-accumbens. In Chapter 8, we summarise these findings and discuss their implications for our understanding of how two of the brain's principal neurotransmitter systems, dopamine and serotonin, regulate motivated behaviour in the context of action and inaction.

This research was funded by a 50/50 Studentship from the Economic and Social Research Council and St. John's College, Oxford.

Table of Contents

1 Introduction

Chapter abstract:
Understanding how the brain translates motivation into action is an important question in neuroscience, and two monoamines – dopamine and serotonin – play key roles in shaping goal-directed action for reward. Here, the anatomy and physiology of each is described, and current hypotheses of their contributions each to reward processing and motivated action are evaluated. We then outline the thesis aims: to use a novel paradigm to *bridge* these perspectives and understand, using methods at both fast and slow timescales, how these neurotransmitters uniquely shape behaviour for reward.

1. The brain: from motivation to action

An unappealing lump of grey flesh it may be, but the mammalian brain is in many ways the finest example of evolutionary adaptation gone right. Whilst some may consider complex cognition and, in the case of humans, consciousness as the brain's *raison d'etre*, a perhaps more ecologically rational view is that these characteristics have arisen as a corollary of the brain's fundamental 'purpose': to enact behaviour on the basis of sensory input, for the benefit of the organism.

This may seem like a straightforward enough goal to achieve, and the brain an abundantly complex organ for such a simple ambition. Yet the intricacy of the task at hand becomes apparent when one considers the range of behaviours possible, even when limiting these behaviours to bodily movements – and the decisions, both explicit and implicit, that need to be made about those movements: to act or not to act? If to act, when? And if to act, which movements to enact? And how – which specific sequences of movement?

However, these questions are immaterial if one does not have the *motivation* to act in the first place. Motivation can be operationalised as the energizing of behaviour and has been considered in terms of directional, goal-directed action, as well changes in general levels of arousal. In both cases, motivation is the latent drive that compels organisms to act – the impetus behind all behaviour – and thus allows organisms to regulate their external and internal environments (Robbins & Everitt, 1996; Salamone & Correa, 2012).

An additional consideration is that the very environment that organisms are operating in is not static, but dynamic; it constantly and unexpectedly changes such that targets of directed behaviour may alter, goals may move or become unavailable, and outcomes may not occur when or in the form anticipated. The features of the world are probabilistic. Because of this, a recent zeitgeist for understanding the brain revolves using reward-based or other pertinent information to develop *predictions* that can then guide behaviour (Friston, 2010).

Two monoamines, dopamine and serotonin, are thought to be integral to the execution and inhibition of motivated behaviour respectively, but are also thought to play important roles in processing of reward-related information. Whilst the scope of this thesis is not such that it can address these broad and interconnected ideas in their entirety, the importance of motivated behaviour and prediction underlie its aim: to attempt to understand the role of two monoamines, dopamine and serotonin, in shaping goal-directed action for reward.

2. *Anatomy and physiology of the mammalian dopaminergic system*

Over 60 years ago, 3-hydroxytyramine, more commonly known as dopamine, was first identified in the human brain (Montagu, 1957). A short time afterwards, Carlsson and colleagues (1958) demonstrated that dopamine, originally thought of as simply a precursor of noradrenaline, was in fact a neurotransmitter in its own right, capable of carrying messages between neurons and inducing a postsynaptic response. Since

then, dopamine has become arguably one of the most intensely-studied of the neurotransmitters, and research into understanding its role in both normal and abnormal cognition and behaviour has influenced the applied fields of psychiatry, psychopharmacology, and neurology amongst others (Iversen & Iversen, 2007).

2.1. Dopaminergic nuclei and major projections

Dopamine is a monoamine of the subtype catecholamine, an organic compound comprised of a catechol and a side-chain amine. Catecholamines are synthesised from the amino acid L-tyrosine, though the exact resultant neurotransmitter depends on the presence of specific enzymes within the neuron (Brady, Siegel, Albers, & Price, 2012). A neuron that contains tyrosine hydroxylase and dopa decarboxylase can produce dopamine and is thus a dopaminergic neuron.

Whilst dopamine itself can be found across the brain, the cell bodies of dopamine-producing neurons are concentrated in a relatively small number of regions; there are nine dopamine-producing nuclei in the mammalian brain, A8 to A16 (Figure 1; Björklund & Dunnett, 2007), three of which – A8, A9, and A10 – are located in the midbrain. The substantia nigra pars compacta (SNc; A9) and the ventral tegmental area (VTA; A10) in particular are perhaps the best studied of these nuclei due to their contributions to the pathology of Parkinson's Disease, and reward processing and reinforcement, respectively (Tritsch & Sabatini, 2012). Projections from the SNc to the striatum, and from the VTA to the limbic and cortical areas, lead to dopamine release in these regions and make up three of the four major dopaminergic projections in the mammalian brain: the nigrostriatal, mesolimbic, and mesocortical

pathways (Anden et al., 1964; also see Figure 1). The fourth major pathway, the tuberoinfundibular projection (not illustrated here), is quite unlike the previous three in terms of its function and instead influences the release of hormones such as prolactin (Weiner & Ganong, 1978). Because of this it is not discussed further.

However, it should be noted that the aforementioned traditional depiction of these pathways is oversimplified; neurons in the SNc also innervate cortical and limbic areas (Bentivoglio & Morelli, 2005). Similarly, whilst there is substantial evidence that these three pathways are distinct both anatomically and functionally, their cells of origin in the midbrain are intermingled within the SN-VTA complex (Björklund & Dunnett, 2007). This is also reflective of the fact that the inputs to these SNc and VTA dopamine neurons have been shown to arise from a diverse array of other regions, including motor and somatosensory areas (Watabe-Uchida, Zhu, Ogawa, Vamanrao, & Uchida, 2012).

Within the basal ganglia, it is not just the striatum that is innervated by midbrain dopaminergic neurons; parts of the globus pallidus, ventral pallidum, and subthalamic nucleus also receive input (Hassani, Francois, Yelnik, & Feger, 1997; Lindvall & Björklund, 1979). Therefore, the downstream effects within the basal ganglia of the activation of these neurons extend further than the caudate-putamen, allowing dopamine to have a more comprehensive influence on the modulation of basal ganglia circuitry.

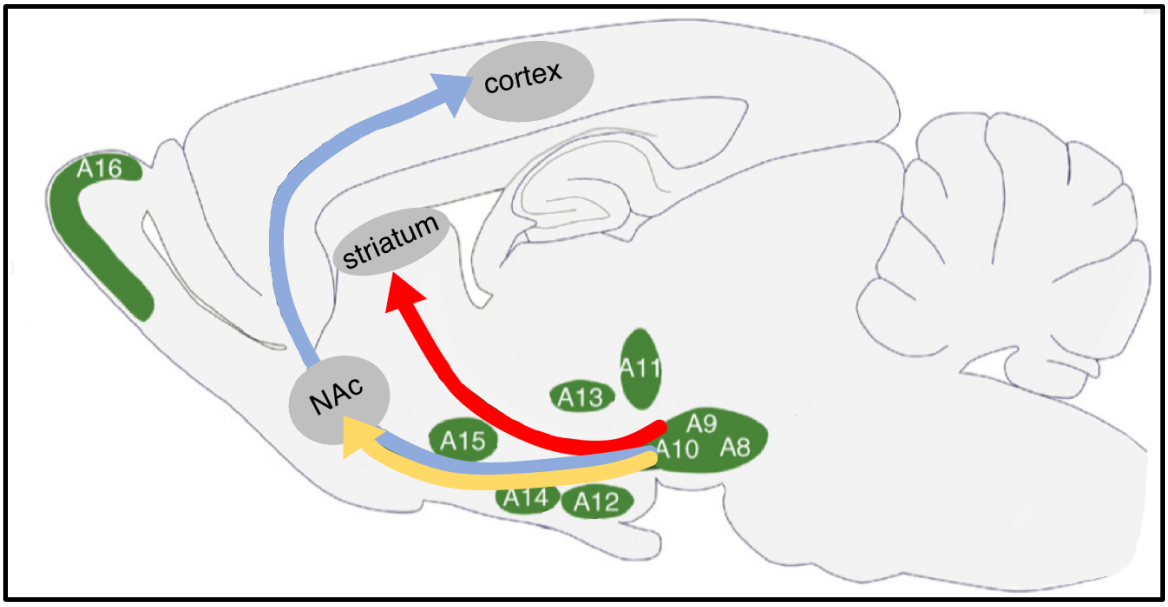


Fig. 1. Key dopaminergic nuclei and projections of the rodent brain in a sagittal view (modified from Björklund & Dunnett, 2007). There are nine major cell groups (green). Whilst not a comprehensive illustration of all dopaminergic efferents, three important projections are also highlighted: nigrostriatal (red), mesolimbic (yellow), and mesocortical (blue).

2.2. The striatum and nucleus accumbens

The striatum is a major subcortical structure of the basal ganglia, acting as its primary input pathway – and, of particular bearing to this thesis, it has been heavily implicated in both movement and reward processing (Kravitz & Kreitzer, 2012). Reflecting both structural and functional differences, the striatum is customarily divided into dorsal and ventral components (Figure 2). Dorsal striatum is comprised of the caudate and putamen, and can itself be further divided into dorsolateral and dorsomedial parts, whilst ventral striatum is comprised of the nucleus accumbens and olfactory tubercle (Gerfen & Surmeier, 2011). Aside from its aforementioned midbrain dopaminergic afferents, the striatum also receives inputs from much of the cortex with a ventromedial-to-dorsolateral topography (such that some argue the sharp dorsal/ventral dichotomy is not quite accurate; Voorn, Vanderschuren, Groenewegen, Robbins, & Pennartz, 2004), as well as the thalamus (Bolam, Hanley,

Booth, & Bevan, 2000) and limbic structures. It integrates these inputs to become the source of the direct and indirect pathways of the basal ganglia: two parallel circuits thought to be of particular importance for movement initiation and selection (Albin, Young, & Penney, 1989), although it should be noted that this direct/indirect dichotomy might only hold for dorsal, not ventral striatum, as discussed in more detail in section 4.4.

Striatal neurons fall into two main classes: medium spiny neurons (MSNs) – which constitute the vast majority of striatal neurons (Kemp & Powell, 1971) – and aspiny interneurons, which can be either cholinergic or GABAergic (Marin, Anderson, & Rubenstein, 2000). MSNs themselves are GABAergic and project either to the substantia nigra pars reticulata (SNr) or the globus pallidus (GP), constituting the direct and indirect pathways respectively. Importantly, they are strongly innervated by midbrain dopaminergic neuron axons that synapse onto MSN dendrites and spine necks (Smith, Bennett, Bolam, Parent, & Sadikot, 1994).

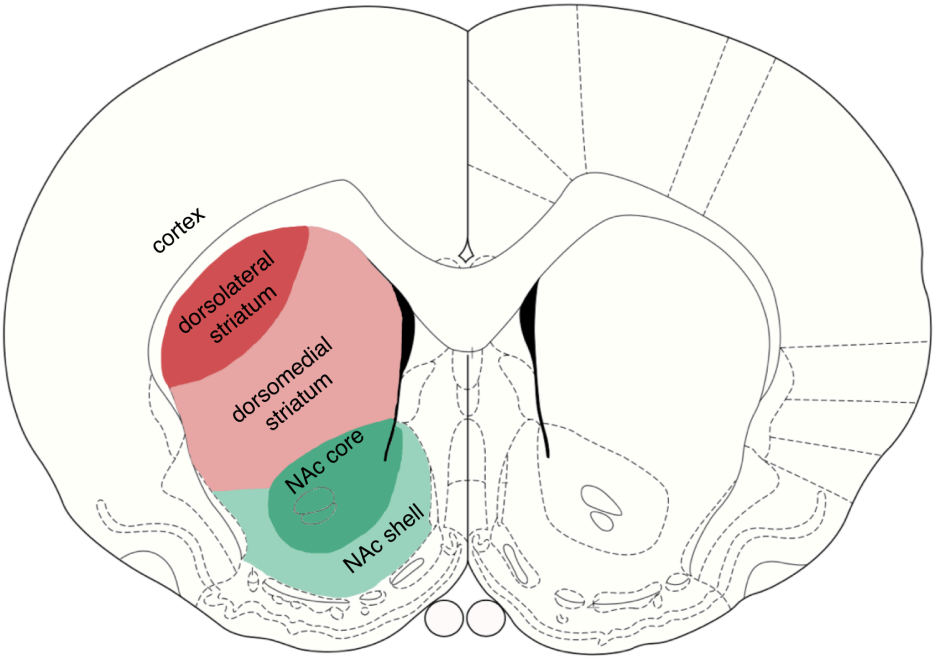


Fig. 2. Coronal schematic illustrating the striatal subdivisions (modified from Paxinos & Watson, 2009).

Within the striatum, the nucleus accumbens (NAc), also often considered as the whole of the ventral striatum (although originally the olfactory turbercle was included, as this region is also heavily innervated by dopaminergic neurons; Gerfen & Surmeier, 2011), is an area of independent interest for several reasons; whilst the neurochemistry of the cells here is similar to that of its dorsal counterpart, it is innervated more by limbic than cortical or thalamic structures (O'Donnell & Grace, 1995; Swanson & Cowan, 1975) and, importantly, its dopaminergic innervation principally originates from the VTA rather than the SNc portion of the midbrain (Phillipson & Griffiths, 1985). This has led to it being termed the 'limbic-motor interface', where information about reward and motivational drive are integrated to guide behavioural output (Mogenson, Jones, & Yim, 1980), in contrast to the dorsal striatal's role in sensorimotor integration (Robbins & Everitt, 1992). This

dopaminergic innervation, however, has a mediolateral topography reflective of the main subdivision of the accumbens: into core and shell regions (Ikemoto, 2007; Figure 2). Therefore, dopamine neurons in the posteromedial VTA generally project to the ventromedial striatum, including the medial accumbens shell, whilst anteromedial VTA largely projects to the ventrolateral striatum, including the accumbens core as well as the lateral shell.

Beyond this disparity in innervation, the NAc core and shell each have some unique features that further allow their differentiation; in rats, the core has a higher cell density (Meredith, Agolia, Arts, Groenewegen, & Zahm, 1992) and greater dopamine and serotonin metabolism (Deutch & Cameron, 1992) than the shell. This is in addition to differences in both their afferents and efferents (Salgado & Kaplitt, 2015). A further anatomical subdivision is that between rostral and caudal NAc, with some considering the rostral pole to be its third major component anatomically and functionally (Zahm & Brog, 1992).

2.2. Two modes of activity: tonic and phasic

Canonically, dopamine neurons have two modes of firing activity: tonic and phasic (Grace, 1991). 'Tonic' refers to slow changes in the steady baseline levels of activity and dopamine release in downstream structures; a sort of spontaneous 'ticking away' of dopamine neurons (Schultz, 2007b). 'Phasic' dopamine, in contrast, refers to a sharp increase or decrease of firing rate – bursting activity – leading to large changes in dopamine concentration downstream that can last several seconds (Grace, 1991). Whilst the origins of these signals is separable physiologically, these two modes of

activity can also interact such that they facilitate each other (Lodge & Grace, 2006). Whether they may have separable contributions to behaviour is discussed subsequently.

3. *Mesolimbic dopamine function: from reward to action and back again*

As the previous section illustrated, dopamine can be found across the brain, but its behavioural function is not homogenous across these regions, and the mesolimbic pathway in particular is a stalwart in the reward and motivation literatures. However, our understanding of the exact nature and function of the signal relayed by mesolimbic dopamine has evolved over the years; as Salamone and Correa (2012) put it, we humans are excellent story tellers, but over time the coherence of the scientific narrative we have established for dopamine has weakened until a new, seemingly more cohesive narrative takes its place – which itself begins to fall apart, and the cycle continues. So here is a short review of dopamine's transformation, from a neurotransmitter that *responds* to reward, to one that *activates behaviour* for reward, to one that *learns* about reward – and, finally, to the proposal of a hypothesis that attempts to integrate these aspects using our now more comprehensive understanding of the physiology of this neural system.

3.1. The pleasure neurotransmitter: 'liking' reward

Initial evidence linking mesolimbic dopamine to reward evaluation came from electrical self-stimulation studies (Olds & Milner, 1954), supported further by

findings using neurophysiological and pharmacological approaches (Yokel & Wise, 1975). Mesolimbic dopamine was shown to be elicited by both primary natural rewards such as food, water, and sex (as reviewed in Schultz, Dayan, & Montague, 1997) as well as many different drugs of abuse (Chiara & Imperato, 1988; Pontieri, Tanda, Orzi, & Di Chiara, 1996). This led to the highly influential dopamine hypotheses of reward (Wise, 1978), as well as to the anhedonia hypothesis (Wise, 1982): that dopamine in the mammalian brain plays an important role in the subjective pleasure experienced by the receipt of reward – or, as Wise (1980) put it, 'yumminess'.

However, it was soon realised that there were some caveats to this hypothesis, and that the narrative of dopamine as the 'pleasure neurotransmitter' was a misinterpretation of its function in the brain. In particular, lesioning of the ascending dopaminergic projections does not reduce positive reactions to sweet taste (Berridge & Robinson, 1998; Berridge, Venier, & Robinson, 1989), and has little effect on consumption of food or reward (Koob, Riley, Smith, & Robbins, 1978). Additionally, studies using approaches with a high temporal resolution also seemed to suggest that levels of dopamine would rise *before* the consumption or acquisition of reward (Phillips, Atkinson, Blackburn, & Blaha, 1993). Additionally, mesolimbic dopamine release to natural reward was found to be contingent on the state of the animal, requiring food deprivation to be elicited (Bassareo & Di Chiara, 1999). Thus the *motivation* for reward seemed to be important to this signal. This and other evidence suggested that distinguishing between appetitive and consummatory (or 'end of sequence') phases of behaviour – between engaging in motivated behaviour that leads

to reward, and the completion of this behaviour in ingesting or attaining reward – is important in attempting to understand the function of mesolimbic dopamine (Robbins & Everitt, 2007; Robbins & Everitt, 1992, 1996).

3.2. The motivational neurotransmitter: driving behaviour for reward

Incentive motivation can be thought of as a type of priming, whereby a previously neutral stimulus acquires motivational importance through association with a primary reward and has therefore assimilated drive-like effects (Berridge & Robinson, 1998). One formulation of incentive motivation is Berridge's concept of incentive salience, also more colloquially termed by him as 'wanting', against the hedonic pleasure of consuming reward, or 'liking' (Berridge, 1996). There is evidence that dopamine may play a key role in the former but – as discussed previously – not the latter. More specifically, incentive salience involves the transformation of a reward-related neural representation, such as a cue for reward (conditioned stimulus, or CS) or the reward itself (unconditioned stimulus, or UCS; Berridge, 2007) into something that incentivises behaviour. The degree of incentive salience determines how attractive and 'wanted' those stimuli are (Berridge, 2007). This is distinguishable from more cognitive forms of 'wanting' that involve explicit or declarative goals or forms of planning (Berridge, Robinson, & Aldridge, 2009). Importantly, here incentive salience is not merely a passive construct but actively promotes behaviour, specifically the approach of and consumption of rewards (Berridge & Robinson, 1998).

Whilst discriminating between the subjective hedonic pleasure associated with experiencing reward and the subjective desire to achieve reward is an important one,

it is also somewhat psychologically ambiguous and therefore difficult to parse mechanistically, as pointed out by Robbins & Everitt (2007). Indeed, Salamone & Correa (2002) also raise the caveat that 'wanting' seems to be a multidimensional construct: it has both a directional aspect (appetite to consume food) and an activational aspect (behavioural activation for the initiation and sustainment of instrumental actions to retrieve that food). They cite the fact that low doses of dopamine antagonists do not impair appetite for food, but do impair *instrumental* expressions of motivation, as evidence that this is an important distinction (Salamone et al., 1991).

The suggestion that the dopaminergic system could provide the energetic impetus behind activation of behaviour has previously been proposed by Robbins and Everitt (1992). This activational aspect of motivated behaviour – as opposed to general arousal – is an important component of an organism's ability to overcome costs of time, effort, distance, and so on in order to achieve reward (or indeed to avoid punishment) – and therefore to survive. Such activation can be elicited simply by presentation of motivational stimuli, such as sucrose pellets, but also by food deprivation, novelty, or conditioned stimuli, and can lead to increased locomotion, feeding, and other behaviours (Salamone, 1988; Robbins & Everitt, 1992). In the case of cues paired with appetitive reinforcers, change in mesolimbic dopamine levels and therefore in activation consequently enhances motivated behaviour by increasing general responsivity to those cues, facilitating the impact of stimulus-reward associations on goal-directed behaviour (Robbins & Everitt, 1992).

3.3. Activation-sensorimotor hypotheses: exerting effort for reward

Whilst there is much support for the hypothesis that nucleus accumbens dopamine is important for the activation of behaviour (B. Lex & Hauber, 2010), as well as energy expenditure (Beeler, Frazier, Zhuang, & Salamone, 2012), Salamone's conception of mesolimbic dopamine has an additional element beyond its importance in activating behaviour: he also advocates that said dopaminergic activation of behaviour is a resource that can be allocated dependent on effort requirements. The archetypal paradigm he has used to demonstrate this offers rodents a choice between lever pressing to obtain a preferred food or consumption of freely-available lab chow (Salamone et al., 1991). Dopamine receptor antagonism alters response allocation, shifting behaviour from a predominance of lever pressing to increased lab chow consumption (Salamone, Correa, Farrar, & Mingote, 2007). This melds with more recent proposals that mesolimbic dopamine may calculate the utility of an action, allowing mammals to overcome response costs to garner reward (Phillips, Walton, & Jhou, 2007). However, whilst these effects are clearly mediated by accumbens dopamine, they may also be dependent on top-down influences, with areas such as the medial frontal cortex also having been shown to regulate effort-based decision-making (Walton, Bannerman, & Rushworth, 2002).

A more recent activation-sensorimotor theory – the flexible approach hypothesis – has sought to combine the clear invigorative effects of accumbens dopamine on behaviour, particularly that which is effortful, with the incentive motivation conception and the capacity for dopamine to increase responding to reward-predictive cues (Nicola, 2010). In noting that lesioning of the NAc does not always

influence decision-making to exert greater effort for larger reward (Walton et al., 2009) but that animals are less able to seek reward following reward-predicting cues with similar accumbens manipulation (Nicola, 2007), Nicola (2010) states that accumbens dopamine is required for reward-seeking behaviour 'only when the specific actions required to obtain reward are variable across instances of reward availability'. In practice this means that if the behaviour required to garner reward is the same, repeated action, accumbens dopamine is not required – but it is required if different actions are required to reach a goal. This proposal is also able to account for differences in effects of dopaminergic manipulations to effort-based decision-making; perhaps Walton et al. (2009) found little effect because animals in this paradigm had to initiate each trial in a central magazine, meaning that the left or right actions when making a choice were very stereotyped. In a similar but subtly different paradigm, Ghods-Sharifi & Floresco (2010) found effects on choices when cues indicating the long, 40s choice period were presented without a trial initiation action, meaning that animals might need to flexibly approach a lever in order to select it.

3.4. The teaching neurotransmitter: learning about reward

One commonality amongst incentive salience, behavioural activation, and flexible approach is that they all presuppose a role for mesolimbic dopamine not just in terms of influencing behavioural responses to reward itself, but also to *cues* that predict upcoming reward – for example, by claiming that these cues become adorned with value, or that they drive behaviour. However, is dopamine implicated in the process

of transference of value or behavioural activation to conditioned stimuli? And is it necessary for this learning?

There is much evidence that cues that have been learnt to reliably predict reward cause phasic elevations in accumbens dopamine (Day, Jones, Wightman, & Carelli, 2010; Roitman, Stuber, Phillips, Wightman, & Carelli, 2004). Initially simplistic hypotheses linking dopamine to learning suggest that dopamine might 'stamp' in stimulus-response or stimulus-stimulus associations if they occur prior to reward – in other words, dopamine is necessary for reinforcement (Wise, 2004). However, it is evidence from methods that allow precise temporal measurement of the dopaminergic system that has led to the development of a more sophisticated hypothesis of dopamine's role in learning; a hypothesis that, since its conception, has arguably influenced the field to such an extent that it has become the new definition of 'what dopamine does': that phasic midbrain dopamine, and its associated mesolimbic dopamine release, encodes a reward prediction error.

The data to support this hypothesis was initially generated by Schultz (Schultz, 1998; Schultz, Apicella, & Ljungberg, 1993; Schultz, Dayan, & Montague, 1997), who, using electrophysiological recordings in the macaque midbrain, neatly demonstrated that putative dopamine neurons increase their firing rate in response to an unpredicted reward (figure 3a), but that if, over many trials, the animal learned to associate a conditioned stimulus as a reliable predictor of upcoming reward, this response would shift from the (now fully predicted) reward to the cue predictive of reward (figure 3b). Additionally, if the conditioned stimulus was given but the expected reward

subsequently omitted the neuronal activity *decreased* at the time of expected reward (figure 3c). Therefore, dopamine plays a role in learning specifically when expectations about reward have been violated.

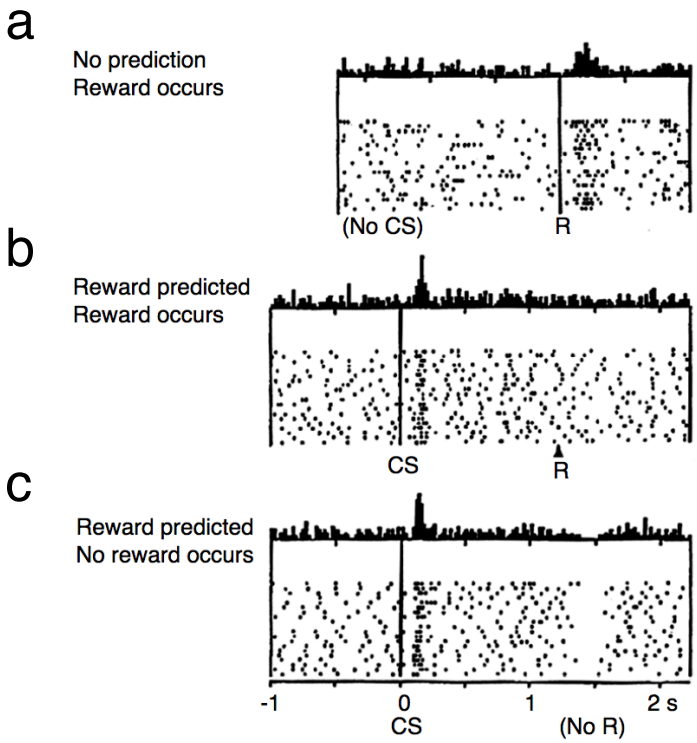


Fig. 3. Electrophysiological recordings from putative midbrain dopaminergic neurons in the macaque in response to (a) unpredicted reward, (b) a conditioned stimulus (CS) that has been associated with upcoming reward, and (c) the omission of expected reward. Adapted from Schultz, Dayan, & Montague (1997).

The appeal of this hypothesis comes not just from its intuitive explanation of dopaminergic responses in Pavlovian conditioning, but also from the fact that these prediction error responses map beautifully onto a class of reinforcement learning model, temporal difference (TD) learning (Sutton & Barto, 1998). TD models propose that the reward prediction error term acts as a teaching signal, conveying the difference between a weighted average of future expected reward and present

information – for example a cue, or reward itself – that leads to a revision of expectations which is then used to change downstream to change behaviour (Glimcher, 2011). The expected future value part of the equation is calculated by summing a weighted average of recently experienced reward to the degree of violation of that expectation, multiplied by a learning rate.

Implicit in its name, an important feature of such models is use of a continuous representation of time, and it is this that allows updating of the value function not just for the moment in time when reward is given, but to previous moments in time, allowing eventual updating of value to time periods when stimuli predictive of reward occur and for those stimuli themselves to be imbued with value (Sutton & Barto, 1998). Thus the critical link between the midbrain dopamine responses as seen in Figure 3 and the TD class of learning models was made (Montague, Dayan, & Sejnowski, 1996).

Since the proposal of midbrain dopamine neurons acting as an analogue for reward prediction error was first put forward, it has been confirmed using behavioural paradigms such as blocking (Waelti, Dickinson, & Schultz, 2001) and conditioned inhibition (Tobler, Dickinson, & Schultz, 2003), across rewards with different properties or dimensions (Lak, Stauffer, & Schultz, 2014), and across species (Cohen, Haesler, Vong, Lowell, & Uchida, 2012; Takahashi et al., 2016; O'Doherty, Dolan, & Schultz, 2006). Similarly, necessary suppositions for TD learning to take place in the brain have also been tested. For example, for a reward prediction error computation to occur, the neurons involved need to be able to perform subtraction to produce the

necessary error term; this has been demonstrated experimentally in VTA dopamine neurons (Eshel, Bukwich, Hemmelder, Tian, & Uchida, 2015).

Causal evidence has added weight to the case. Timed inhibition of dopamine neurons disrupts learning (Hamid et al., 2016), and the inducement of an 'artificial' prediction error by optogenetic activation of dopamine neurons at the appropriate time is sufficient to increase cue-driven reward-seeking behaviour (Steinberg et al., 2013). Pharmacological manipulation in tandem with measurement of dopamine release in the accumbens has also shown that the role of dopamine in the behavioural expression of such stimulus-reward learning is specific to the attribution of incentive salience to reward cues, not to all stimulus-reward associations (Flagel et al., 2011)

Additionally, reward prediction error signals have also been identified in accumbens dopamine release (Day, Roitman, Wightman, & Carelli, 2007). However, prediction errors detected in the accumbens have some different properties to those that arise in the midbrain; for example, due to the inherently low baseline firing rate of dopamine neurons, in the VTA the magnitude of change in response to a negative reward prediction error is less than that of a positive reward prediction error, even if the absolute difference of both prediction errors is the same. Yet dopamine release in the nucleus accumbens *symmetrically* encodes positive and negative reward prediction errors (Hart, Rutledge, Glimcher, & Phillips, 2014).

Thus the evidence in support of reward prediction error encoding by midbrain dopamine neurons and downstream striatal dopamine release is substantial.

However, several of the aforementioned studies have also provided hints that there is more to the data, and in recent years there has been a slow but sure unravelling of the previously seemingly unassailable reward prediction error hypothesis in its original form. For example, there is much evidence that dopamine neurons might also be responsive to non-rewarding stimuli, particularly salient or novel stimuli, which is not predicted by the reward prediction error model (Horvitz, 2000). This could however be accounted for by a theory which posts that all such stimuli may be of some motivational relevance, and this salience signal has been recognised by Schultz himself as a separable but important component of the dopamine signal (Schultz, 2016). Indeed, this novelty response has recently been localised to a subpopulation of neurons that project to the tail of the striatum, suggesting that they may be independent from reward prediction error signaling neurons (Menegas, Babayan, Uchida, & Watabe-uchida, 2017).

Yet this is only one example of the necessary broadening of the reward prediction error hypothesis. The theory has now had to evolve to account not only for errors in reward prediction, but also for errors in the prediction of sensory features (Takahashi et al., 2017), processing of aversive stimuli (Matsumoto & Hikosaka, 2009b; although this is controversial - see Fiorillo, Song, & Yun, 2013), hidden states or contexts (Starkweather, Babayan, Uchida, & Gershman, 2017), cue-cue learning outside of value association (Sadacca, Jones, & Schoenbaum, 2016), and judgement of time beyond that necessary for TD learning (Soares, Atallah, & Paton, 2016), and more. Whilst it is possible to suggest that many of these functions are a logical extension of the basic premise initially put forward by Schultz, a modern prediction error theory

of dopamine is unavoidably more complex than its original conception. Counter to this, however, some have argued that this variability seen in the response properties of midbrain dopamine neurons may not necessarily reflect a host of functions, but that they may be integrated for the ultimate purpose of prediction error computation (Lau, Monteiro, & Paton, 2017).

It first be recognised that discrepancies between these theories could be explained by the fact that the prediction error aims to account for *phasic*, and not tonic fluctuations in dopamine activity and release. Much of the evidence cited in support for ideas such as incentive motivation or the flexible approach use broad manipulations of dopamine such as pharmacological approaches, often disrupting dopamine transmission both at slow and fast timescales and certainly having a profound influence on tonic dopamine levels. Similarly, fine anatomical localisation is not always possible with certain approaches such as electrophysiology; whilst it is possible to know the location of the neurons from which recordings are being taken, where these neurons are projecting is often more ambiguous in comparison to either pharmacological or optogenetic manipulations. This is important as, although not discussed in detail here, encoding of prediction error responses may be different when comparing the midbrain populations of the VTA and SNc, and responses downstream in the dorsal and ventral striatum, as well as in the core and shell of the nucleus accumbens, are also dissociable (Brown, Mccutcheon, Cone, Ragozzino, & Roitman, 2011). But regardless of these caveats, there has been building evidence that even phasic dopamine responses within the midbrain and striatum may play an important role in reward-guided movement, as discussed subsequently.

3.5. The neurotransmitter for reward-driven action: all of the above?

The fact that dopamine in the brain is, at some level, required for movement – particularly movement initiation – has long been established; Parkinson's Disease is identified as a neurodegenerative disorder in which a principal feature is that dopamine neurons in the substantia nigra are lost, causing a striatal dopamine deficiency and a behavioural phenotype that includes a host of motor symptoms such as bradykinesia (Poewe et al., 2017; though many non-motor symptoms have also been identified: Sinha, Manohar, & Husain, 2013; Chaudhuri & Schapira, 2009). However, this body of literature was not immediately amalgamated with basic neuroscience studies, particularly those attempting to understand the phasic dopamine response, and even relatively recently some have maintained that only very modest dopamine concentration fluctuations over long timescales can be seen in response to movement (Schultz, 2007a, although see: Schultz, Ruffieux, & Aebischer, 1983 for early evidence of movement encoding in non-human primate SNc neurons).

Yet there is certainly evidence that gross manipulations of levels of dopamine can enhance the neural encoding of rewarding *action* by neurons in the midbrain and striatum, separate from any effect on the representation of reward value itself (Guitart-Masip, Chowdhury, Sharot, Dayan, Duzel, & Dolan, 2012). Others have shown that the activity of midbrain dopamine neurons is sensitive to goal-directed movement parameters, but that the specific characteristics of this may be different in the SNc as compared to the VTA. For example, optogenetic activation of these separate populations of neurons can either invigorate intense locomotion, in the case of the SNc, or increase the value of reward-related cues, if stimulating the VTA

(Saunders, Richard, Margolis, & Janak, 2017). Indeed, SNc activity has been increasingly associated with the encoding of specific movement kinematics (Barter et al., 2015), as well as the initiation (Dodson et al., 2016; Howe & Dombeck, 2016; Jin & Costa, 2010), likelihood, and vigour (Alves, Tecuapetla, Paixão, & Costa, 2018) of movement, even if they are not reward-driven but spontaneous.

Whilst the SNc may be sensitive to the specific features of movements, studies recording from or manipulating the VTA have often reinforced its role in reward processing, reflecting the reward prediction error model (Howe & Dombeck, 2016). However, there is evidence that the SNc:action, VTA:reward dichotomy may not always hold. The VTA may well encode movement – *if* it is purposeful movement that drives the organism to achieve a goal. Recordings of VTA neurons have revealed a population of neurons that multiplex movement and motivation (Puryear, Kim, & Mizumori, 2010). Neuronal ensemble analyses have also shown that these same neurons are also modulated by the sustainment of action necessary to garner reward (Wood, Simon, Koerner, Kass, & Moghaddam, 2017).

Downstream of the VTA, dopamine release in the nucleus accumbens also does not abide strictly by the reward prediction error and seems also to be influenced by action for reward. In particular, levels of accumbens dopamine has been shown to ramp up before animals initiate drug-seeking behaviours (Phillips, Stuber, Heien, Wightman, & Carelli, 2003), during motivated approach to a lever for sucrose – with peak dopamine release correlating to time of lever press (Roitman et al., 2004), and during completion of a sequence of actions to garner reward (Wassum, Ostlund, &

Maidment, 2012). Intriguingly, the latter study was also able to relate the amplitude of dopamine release preceding initiation of an action sequence to the speed with which that sequence was completed, implicating striatal dopamine levels in mediating the vigour of reward-oriented actions. A similar effect of vigour has been localised to the nucleus accumbens core, but not shell, and again only for actions leading to reward but not for unreinforced behaviour (Ko & Wanat, 2016).

Indeed, Lloyd & Dayan (2015) have posited that this ramping seen across many studies may reflect a discounted vigour parameter, allowing the animals to weigh up the cost of increased vigour with the benefit of obtaining reward sooner. This is in contrast to an earlier model of the dopaminergic contribution to response vigour, which suggested that it is tonic, and not phasic, dopamine that contributes to the calculation of a long-run average rate of reward and thus the opportunity cost of actions over time, thereby mediating the vigour of responding (Niv, Daw, Joel, & Dayan, 2007). The authors here do admit evidence that relates vigour to phasic dopamine release, however (Watanabe et al., 2001) and suggest that the effect of phasic signaling on overall dopaminergic tone might explain these effects (Phillips & Wightman, 2004).

Therefore, whilst phasic dopamine release in the accumbens certainly has features of a reward prediction error (Day et al., 2007; Flagel et al., 2011) it also seems to be somehow also related to the movements or actions required to garner reward. Some have suggested that this invigoration of movement by reward-related cues, as per incentive salience, could be distinct from a role in this signal of reporting a prediction

error (McClure, Daw, & Montague, 2003). But others have suggested that these two functions – as well as our understanding of the role of the different timescales of dopamine – could be united, bringing together literatures on motivation and incentive salience with the role of dopamine in learning and prediction. Hamid et al., (2016) conducted a study that tied together many of these features; by using a combination of microdialysis, fast-scan cyclic voltammetry, and optogenetics in a task where animals had to perform a series of actions for reward, they were able to disambiguate the learning and motivation aspects of the accumbens dopamine signal. They solidified many of the conclusions already hinted at by previous studies – that minute-to-minute changes in dopamine track reward rate, but that this can be extracted from changes to the baseline of the phasic dopamine signal; that phasic dopamine release in the accumbens core reflected temporally discounted future reward; and that inhibition of this signal affected the moment-to-moment probability, or hazard rate, of engaging in work. Additionally, optogenetic manipulation of these neurons had differential effects on behaviour dependent on the timing of stimulation; activating midbrain dopamine neurons before the execution of a response increased its vigour, whilst this same stimulation but at the time of feedback reinforced the previous action. Thus the same dopaminergic signal can convey different information – movement and reward – dependent on the context in which the signal is elicited.

However, truly dissociating between motivation and motor behaviour is a tricky task (Salamone & Correa, 2002; Wise, 2004) and whilst there is clearly much evidence for a role of nucleus accumbens core dopamine in motivating action for reward at

multiple timescales, no study to date has parametrically dissected what degree of an influence goal-directed action has on this signal.

4. Dopamine receptor pharmacology

4.1. Discovery of the dopamine receptor subtypes

Dopamine receptors, as with almost all other neurotransmitter receptors, come in different forms – although this had not been appreciated until the late 1970s, when the two initial dopamine receptor subtypes were discovered: D1 and D2 (Kebabian & Calne, 1979). The authors found that these receptors could be differentiated in multiple ways: by applying pharmacological agents, by studying their physiology, and by the fact that they couple to and activate different G protein complexes. In the following decade, the existence of further dopamine receptor subtypes was proposed but not formally recognised (Andersen et al., 1990). Indeed, some argued that what was being observed in these experiments was the result of heightened or lessened affinity states of the D1 or D2 receptors, rather than evidence of as yet unidentified receptors (Leff & Creese, 1985).

It took the advent of recombinant DNA technology and molecular cloning techniques to finally reveal 3 additional receptor subtypes beyond D1 and D2: D3 (Sokoloff, Giros, Martres, Bouthenet, & Schwartz, 1990), D4 (Van Tol et al., 1991), and D5 (Sunahara et al., 1991) receptors. However, several properties of each of these receptors allowed them to be grouped into D1-like (encompassing D1 and D5 receptors) and D2-like (encompassing D2, D3, and D4 receptors). These properties include the genomic

organisation of the receptors which also reflects the D1-like and D2-like subgrouping (Hiranita & Collins, 2015), and the fact that, although all dopamine receptors are G-protein coupled, D1-like and D2-like receptors are activated by different G-proteins in separable cascades of processes – arguably the predominant feature by which they are now distinguished.

4.2. Molecular properties of the dopamine receptor subtypes

D1-like and D2-like receptors (from here on referred to as D1R and D2R) couple to and activate different G protein complexes to either increase or decrease adenylyl cyclase (AC) activity (and are therefore all metabotropic). D1Rs activate the $G\alpha_{s/olf}$ family of G proteins, stimulating cyclic adenosine monophosphate (cAMP) production by AC and activating protein kinase A (PKA), whilst D2Rs instead couple to the $G\alpha_{i/o}$ family of G proteins to inhibit AC and thus limit PKA activation (Beaulieu & Gainetdinov, 2011). Dopamine receptors, particularly $G\alpha_{i/o}$ coupled receptors – D2Rs – can also modulate intracellular calcium levels outside of the cAMP/PKA cascade to regulate ligand- and voltage-gated ion channels, for example by stimulating inhibitory G protein-activated inwardly rectifying potassium channels (GIRK) (Beckstead, Grandy, Wickman, & Williams, 2004).

The receptor subtypes also differ in terms of their baseline affinity states. The affinity of a receptor can be defined as the strength of the intermolecular force between the receptor and its ligand – in this case, dopamine. If a receptor is high-affinity, there are strong binding forces between it and its ligand, meaning that the ligand will take longer to dissociate from the receptor (Strange, 2008). Conversely, ligands will

dissociate from low-affinity receptors more quickly. This has an intuitive effect on the concentration of the ligand required for maximal occupation of a receptor: a low concentration is required for occupation of a high-affinity receptor, and a high concentration required of a low-affinity receptor.

As discussed in section 2.3, dopamine release can be thought of as having two modes: tonic and phasic. Whilst tonic dopamine release is characterised by slow changes at lower volumes of the neurotransmitter, phasic dopamine release implicates a larger volume being emitted at a shorter timescale. D2Rs are reported to have an affinity for dopamine that is 10- to 100-fold greater than that of D1Rs (Beaulieu & Gainetdinov, 2011). Accordingly it is thought that D2Rs are more sensitive to the lower levels of tonic dopamine release, whilst D1R are instead more readily activated by phasic dopamine events (Richfield et al., 1989; though the high-affinity property of D2Rs has recently been questioned: Yapo et al., 2017). This distinction has prompted models of the respective behavioural contributions of these receptor subtypes, implicating D1R more heavily in reward learning and D2R in overall levels of motivation (Missale, Nash, Robinson, Jaber, & Caron, 1998). However, recent accounts are more nuanced, suggesting that D2Rs are also sensitive to pauses in dopamine neuron firing (Dreyer, Herrik, Berg, & Hounsgaard, 2010). In fact, D2Rs may even be sensitive to phasic *increases* of dopamine release; in one study, artificial stimulation of dopamine terminals was able to evoke D2R inhibitory postsynaptic currents in under a second, indicating that these receptors would be able to encode this pulsed stimulation as individual temporal events, even despite elevated basal levels of dopamine (Marcott, Mamaligas, & Ford, 2014).

Once activated, be it by phasic or tonic dopamine release, it is widely accepted that D1Rs and D2Rs respectively have excitatory and inhibitory effects on their expressing neurons (Surmeier, Ding, Day, Wang, & Shen, 2007). These effects can occur due to changes in dendritic excitability and glutamatergic signalling; in the case of D1R, the aforementioned increase in cAMP levels leads to the activation of protein kinase A (PKA), setting in motion a cascade of phosphorylation of various targets within the cell and trafficking of AMPA and NMDA receptors that effect the excitability of the cell (Hallett, Spoelgen, Hyman, Standaert, & Dunah, 2006), allowing it to be more responsive to synaptic release of glutamate (Surmeier et al., 2007). Similarly, D2R signalling is also able to alter the glutamate receptor function of neurons that possess them, although this instead involves downregulation of AMPA receptor currents, therefore increasing inhibition of the cell (Cepeda, Buchwald, & Levine, 1993). This distinction between the excitatory and inhibitory effects of D1Rs and D2Rs respectively, as well as the anatomy of their downstream projections, has been a key feature of a prominent theory of their function in the striatum: the direct/indirect pathway model (Bolam et al., 2000), as discussed in section 4.4.

D2Rs themselves can be subdivided further molecularly based on two distinct isoforms that arise due to alternative gene splicing: D2S (short) and D2L (long) variants (Giros et al., 1989). This is in contrast to the genes that encode D1Rs which do not generate splice variants (Tritsch & Sabatini, 2012). When originally discovered the two D2R isoforms were thought to have the same function, but further research showed that their localisation and thus function differed, with D2L acting

predominantly at postsynaptic sites and D2S serving as presynaptic autoreceptors (Usiello et al., 2000, athough more recent research questions this distinction: Gantz et al., 2015). The properties and functions of D2 autoreceptors are explored in greater detail subsequently.

4.3. Dopamine receptor localisation and interactions in the striatum

D1Rs can be found exclusively on non-dopaminergic neurons, and within the striatum these consist predominantly of GABAergic medium spiny neurons (MSNs) which make up approximately 95% of the striatal population of neurons (Yager, Garcia, Wunsch, & Ferguson, 2015). D2Rs, on the other hand, whilst also being found predominantly on non-dopaminergic neurons, can also be located on dopaminergic neurons as presynaptic receptors, or autoreceptors, allowing for feedback inhibition of dopamine release (Ford, 2014). In both cases, these receptors are predominantly found extrasynaptically (Caille, Dumartin, & Bloch, 1996) – evidence that dopamine can act via volume transmission, being released directly into the extracellular fluid (Fuxe et al., 2010).

The presence of D2 autoreceptors is specific to the striatum; dopamine neurons projecting to the prefrontal cortex have few or no autoreceptors and are therefore insensitive to autoreceptor-feedback mediated inhibition (Cubeddu, Hoffman, & Talmaciu, 1990; Lammel et al., 2008). In particular, striatal autoreceptors can be found at the axon terminals in projection areas such as the NAc (Missale et al., 1998). They allow for the modulation of the activity of dopamine neurons directly, via activation of a potassium current (Ford, 2014) and for the modulation of dopamine

release in the striatum through several mechanisms. D2 autoreceptors on nerve terminals can inhibit the probability of vesicular dopamine release (Suaud-Chagny, Ponec, & Gonon, 1991), and they can also adjust dopamine synthesis by decreasing tyrosine hydroxylase – a much slower and long-lasting mechanism separable from the subsecond feedback inhibition that occurs subsequent to burst firing of dopamine neurons (Wolf & Roth, 1990). Alternatively, autoreceptors can alter the reuptake of dopamine by the dopamine reuptake transporter (DAT) (Truong, Newman, Hanson, & Fleckenstein, 2004). In all cases this leads to a decrease of the release of dopamine (Ford, 2014). It should be noted that D3 receptors are also present on dopamine neurons (Rivet, Audinot, Gobert, Peglion, & Millan, 1994), although it is likely that D3 autoreceptors play a minor functional role in comparison to the D2 autoreceptors (Ford, 2014). Additionally, there is evidence that postsynaptic D2Rs (as well as postsynaptic D1Rs; Stamford, Kruk, & Millar, 1991; Zetterström, Sharp, & Ungerstedt, 1986) may also regulate dopamine release via feedback mechanisms, but this again is to a lesser extent than D2 autoreceptors and is dependent on striatal region, occurring more in dorsal striatum than the nucleus accumbens (Anzalone et al., 2012).

In addition to their presence on MSNs, D2Rs are also expressed by interneurons in the striatum. They can be found on GABAergic (Delle Donne, Sesack, & Pickel, 1997), cholinergic (Wang et al., 2006) and glutamatergic interneurons (Surmeier et al., 2007). This is in contrast to D1Rs which have low or no expression on most striatal interneurons (Tritsch & Sabatini, 2012). In total these interneuron populations make up only 5% of striatal neurons, but nevertheless they may make important

contributions to the dynamics of dopamine transmission; for example, the selective activation of cholinergic interneurons has been shown to augment phasic dopamine release in the NAc (Cachope et al., 2012) and hence they can exert a strong influence on goal-directed behaviour (Witten et al., 2010).

Whilst D1Rs and D2Rs therefore have clear differences both in terms of their localisation and their molecular properties, it is also worth noting that up to 15% of dorsal striatal MSNs may in fact co-express both receptor subtypes (Bertran-Gonzalez et al., 2008; Perreault, Hasbi, O'Dowd, & George, 2011). In addition, in the NAc, it has been suggested that the majority of these D1R/D2R co-expressing neurons consist of D1R-D2R heterodimers, meaning that their properties may be distinct from their component receptors (Perreault, Hasbi, O'dowd, & George, 2014). Some have regarded this as evidence that there may even be a third neuronal pathway for signalling in the basal ganglia (Perreault et al., 2011). However, their influence on behaviour has not yet been established.

Even when not regarding co-expression per se, there are notable interactions between D1Rs and D2Rs that should be recognised, particularly when discussing pharmacological methods that purport to exclusively manipulate one or the other receptor subtype. For example, whilst D1Rs are implicated in feedback control of midbrain dopaminergic neurons, their ability to do so also requires activation of D2Rs (Shi, Smith, Pun, Millet, & Bunney, 1997, Rahman & Mcbride, 2001). Likewise, the behavioural and neural effects of *postsynaptic* D2R stimulation are dependent on D1R activation (Walters, Bergstrom, Carlson, Chase, & Braun, 1987). In short, the receptor

subtypes act synergistically, augmenting – or, in some cases, suppressing (Uslello et al. 2000) – their downstream effects. Therefore, whilst much of this thesis reports on differential contributions of D1R and D2R to aspects of motivated behaviour, it is also recognised that the systems do not act in isolation and that their effects may work cooperatively. This theme of synergy amongst D1Rs and D2Rs has also emerged in research on the different projection pathways of D1R and D2R-expressing neurons.

4.4. Dopamine receptors and striatal projections

The classical model of basal ganglia circuitry proposes two parallel but opposing circuits from the input nucleus of the striatum to the various output nuclei for the purpose of motor control (Albin et al., 1989). With regards to the MSNs of the dorsal striatum specifically, they have canonically been thought to project to their downstream targets via two main pathways: striatonigral MSNs *directly* project to the output nuclei of the basal ganglia to stimulate behavioural activity – the globus pallidus internal, substantia nigra, and VTA (Figure 4a) – whilst striatopallidal MSNs *indirectly* project to the output nuclei via the globus pallidus externa and subthalamic nuclei, resulting in an inhibition of movement (Figure 4b; Soares-cunha, Coimbra, Sousa, & Rodrigues, 2016). Importantly, these two populations express different dopamine receptor subtypes; whilst direct MSNs express D1Rs as well as substance P and dynorphin, indirect MSNs express D2Rs in addition to adenosine receptor 2a and enkephalin (Gerfen, 1992; Gerfen & Surmeier, 2011). These MSNs are also to some extent topographically distributed, such that those that are more dorsal include both D1Rs and D2Rs, whilst more caudal striatal MSNs are almost entirely D1R-expressing (Gangarossa et al., 2013).

However, it is worth noting that this strict division of MSN dopamine receptor subtypes dependent on pathway has been questioned more recently, and may not involve a complete segregation (Cazorla et al., 2014; Saunders et al., 2015). Additionally, concurrent activation of both pathways during movement observed in some studies also suggests that the inhibition and activation dichotomy attributed to the direct and indirect pathways may not fully explain the differential effects of these pathways on behaviour (Calabresi, Picconi, Tozzi, Ghiglieri, & Di Filippo, 2014; Vicente, Galvao-Ferreira, Tecuapetla, & Costa, 2016).

The division in *ventral* striatum is also not quite as clear cut. The NAc projects to the SNc and VTA (together, the ventral mesencephalon) to constitute the striatomesencephalic direct pathway (Figure 4c), whilst it projects to the ventral pallidum and subthalamic nucleus before reaching the ventral mesencephalon as part of the striatopallidal indirect pathway (Figure 4d). As with dorsal striatum, there is a dorsolateral to ventromedial gradient to these projections encompassing the NAc core and shell (Zahm & Heimer, 1990), which also reflects differences in distributions of the receptor subtypes: while both can be found throughout the core, D2-MSNs are preferentially expressed in the medial and ventral shell (Gangarossa et al., 2013). And as with dorsal striatum, this disparity in distribution of receptor subtype again reflects the parallel pathways that make up the circuitry (Hikida, Kimura, Wada, Funabiki, & Nakanishi Shigetada, 2010; Stefanik, Kupchik, Brown, & Kalivas, 2013).

Yet here, whilst the ventral striatal direct pathway is indeed mediated by D1-MSNs, the indirect pathway is modulated by *both* types of MSNs (Kupchik et al., 2015). This has considerable implications for the interpretation of basal ganglia signalling as the ventral pallidum is an output nucleus, meaning that in the ventral striatum, D2-MSNs are projecting to an output nucleus directly and therefore both D1 and D2-MSNs can activate or suppress thalamic activity. Thus the prevalent model of the pathways opposing each other to permit and suppress movement respectively may not be accurate. The behavioural models associated with the basal ganglia pathways, particularly in terms of ventral striatal function, are discussed in more detail in section 5.1.

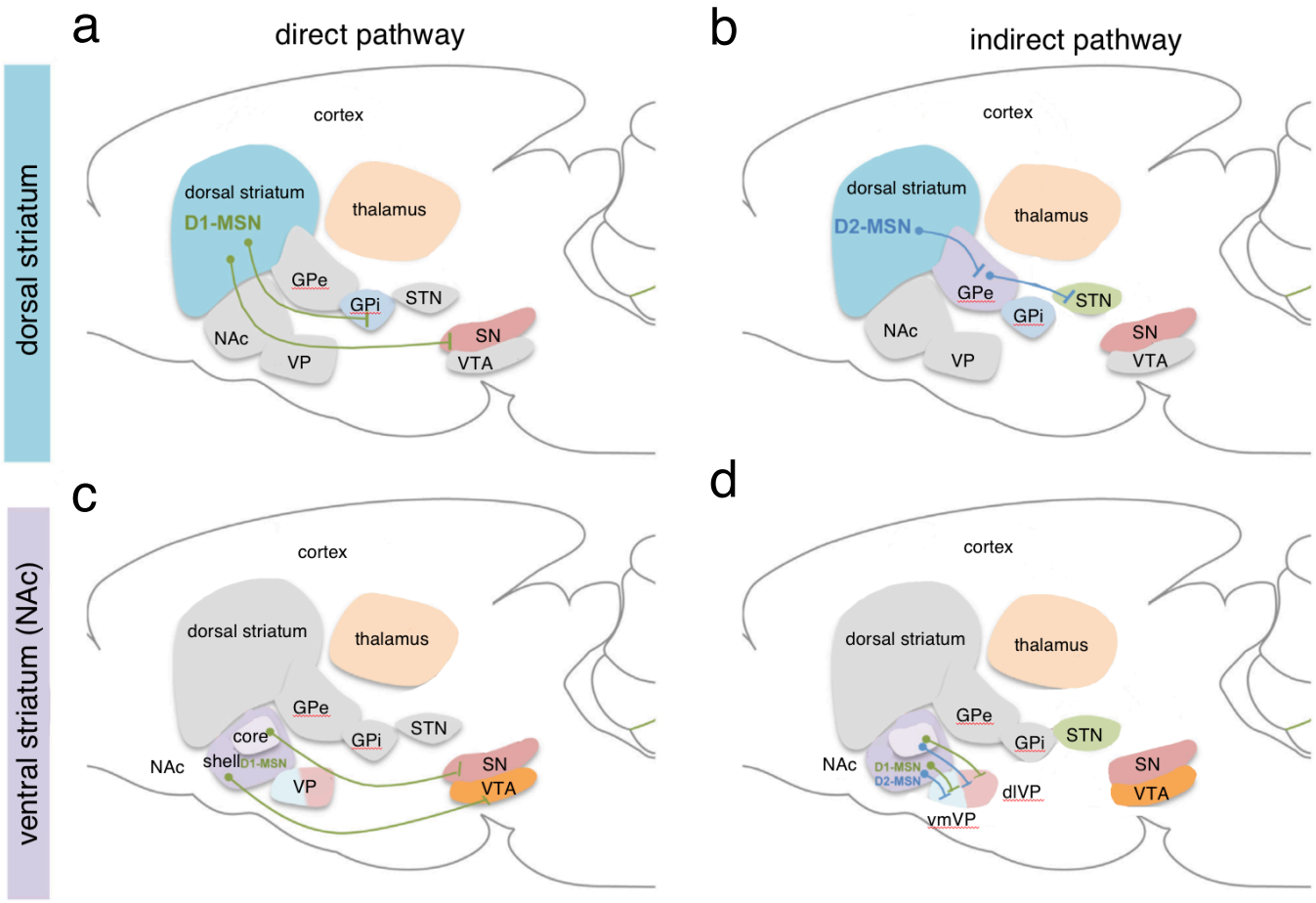


Fig. 4. Projections of D1R and D2R-expressing MSNs as part of the dorsal and ventral striatal direct and indirect pathways of the rodent brain in a sagittal view (modified from Soares-Cunha et al., 2016). (a) Direct pathway dorsal striatal MSNs express D1Rs and project directly to the basal ganglia output nuclei. (b) Indirect pathway dorsal striatal MSNs express D2Rs and project indirectly to the SN via the GPe and STN. (c) Direct pathway ventral striatal (NAc) MSNs from both the core and shell express D1Rs and project to the SN and VTA. (d) Indirect pathway ventral striatal MSNs, constituting both D1- and D2-MSNs, project to the VP and STN before reaching the output nuclei. The core projects predominantly to the dlVP, whils the shell innervates the vmVP, though both projections to the VP involve D1R and D2R-expressing MSNs. *SN: substantia nigra; GPi: globus pallidus pars interna; GPe: globus pallidus pars externa; STN: subthalamic nucleus; NAc: nucleus accumbens; VP: ventral pallidum; dlVP: dorsolateral pallidum; vmVP: ventromedial pallidum.*

4.5. Dopamine receptor expression in prefrontal cortex

Although the focus thus far has been on dopaminergic projections and receptor localisation to and in the striatum, dopamine is released and detected broadly throughout the central nervous system (Callier et al., 2003). D1Rs and D2Rs are indeed most notably found in the striatum, but subpopulations of pyramidal neurons, interneurons, and glial cells in cortical areas also express these receptor subtypes (Tritsch & Sabatini, 2012). However, the exact distribution and density of these

receptors and their afferent dopaminergic fibres varies both between species and across cortical areas, and is therefore less well-characterised than their striatal counterparts (Puig, Rose, Schmidt, & Freund, 2014). The prefrontal cortex (PFC) has been an area of particular focus due to the mesocortical pathway that principally projects there from the midbrain.

Unlike the striatum, where the concentration of D2Rs is particularly elevated (Missale et al., 1998), D1Rs are the most widely expressed receptor subtype of the pyramidal cells and GABAergic interneurons of the PFC (Tritsch & Sabatini, 2012). Nevertheless, D2Rs have also been detected on both pyramidal and GABAergic cells but to a lesser extent. Importantly, however, D2 autoreceptor-mediated feedback inhibition is principally a striatal feature; VTA dopamine neurons that project to the PFC have reduced D2Rs and GIRK channels and are therefore less sensitive to this modulation (Lammel et al., 2008).

With regards to expression on pyramidal neurons specifically, higher levels of the mRNA of both D1Rs and D2Rs has been found in deeper cortical layers, with only sparse expression on layer 2/3 pyramidal neurons (Santana, Mengod, & Artigas, 2009). This suggests that a considerable proportion of pyramidal cells may not be directly modulated by dopamine. Additionally, the subset of pyramidal neurons that do express dopamine receptors have heterogeneous downstream projections such that they are difficult to group functionally – these cells included corticostriatal, corticothalamic, and corticocortical neurons but in each case, only a proportion of these neurons expressed dopamine receptors (Gaspar, Bloch, & Le Moine, 1995). D1R

expression on GABAergic interneurons in the PFC is somewhat more homologous. D1R mRNA is present in up to 60% of all GABAergic cortical interneurons across the cortical layers in the rat, in comparison to the 17% of interneurons that express D2Rs (Santana et al., 2009) – although despite this lower expression, these D2Rs may play an important role in regulating the dopamine system (Tomasella et al., 2018).

The relative heterogeneity of dopamine receptor expression in the prefrontal cortex as compared to the striatum has not impeded efforts to understand the roles of D1Rs and D2Rs in cognition and behaviour using a range of methods, particularly pharmacological intervention. Key findings are reviewed in section 5.6, but to understand both these and related findings in the striatum, an explanation of the mechanisms of pharmacological intervention is necessary.

4.6. Mechanisms of pharmacological intervention

The mechanisms that allow the dopamine receptor subtypes to be differentially activated endogenously can also be recruited when applying pharmacological agents in order to understand the molecular and behavioural functions of that receptor. Drugs can act in multiple ways to alter neurotransmission – by mimicking the neurotransmitter and bonding to its receptor, by affecting neurotransmitter reuptake or clearance, or by enhancing or dampening neuronal responses to neurotransmitter activation of the receptor (Kandel, Schwartz, & Jessell, 2000). Generally, the effects of drugs that bind to receptors can act in one of two directions: either producing an effect within the cell as an agonist, or blocking the ability of a natural agonist to bind to the receptor as an antagonist (Pleuvry, 2004).

However, the exact effects of a drug depends on its properties. These properties are its affinity, in other words its propensity to form a reversible complex with the target receptor, and also its efficacy – its ability to produce a functional response (Ariens & Simonis, 1964). When a drug produces a maximal response it is considered a full agonist, whilst one that has no functional effect is a full antagonist. However, the effects of most drugs is somewhere between these two extremes, meaning that they cannot fully activate the receptors regardless of the concentration of the drug available. These are partial agonists (Pleuvry, 2004). Additionally, a drug that otherwise would act as a partial agonist can in practice function as a competitive antagonist if a full agonist is also present, leading to a net decrease in receptor activation (Bolonna & Kerwin, 2005). Finally, there are also inverse agonists, which act on receptors that are spontaneously active even in the absence of ligand binding. As the name suggests, inverse agonists inhibit this spontaneous activation but are unable to activate the receptor themselves, thus acting as functional antagonists (Nutt et al., 2017).

As described in sections 4.1 and 4.2, there are 5 dopamine receptors that can be grouped into two subtypes dependent on their G protein coupling: D1-like and D2-like (Beaulieu & Gainetdinov, 2011). Whilst there are many drugs available that are able to target one of these two subtypes, there are fewer that can target specific receptors, and so a D1 agonist often binds to both D1 and D5 receptors, for example. The D2 autoreceptor further complicates the picture as identifying the effects of its manipulation as being either pre- or postsynaptic can be difficult, though doseage

can have a role to play; broadly, higher doses of D2 drugs tend to lead to effects that are postsynaptically mediated whilst autoreceptor effects can be found at lower doses (Tanaka, Vincent, Nomikos, & Fibiger, 1992).

Finally, in some cases the effects of a drug may be counterintuitive when considering the wider circuitry of the dopaminergic system. For example, there is evidence that dopamine receptor agonists applied to striatal MSNs can lead to a decrease in dopamine release in the striatum over a period of tens of minutes as measured by microdialysis (Zetterström et al., 1986). Similarly, antagonists can in fact increase dopamine via a similar feedback mechanism (See, Sorg, Chapman, & Kalivas, 1991). This can have important implications for the interpretation of drug effects and is explored further in the General Discussion of this thesis, which shows further evidence of this phenomenon.

5. Separable contributions of the dopamine receptor subtypes to behaviour

5.1. Behaviour and the direct and indirect pathways across striatal subregions

The anatomical direct and indirect pathways of the basal ganglia, as described in section 4.4, have traditionally been thought of functionally as a motor control mechanism. To reiterate, according to this model, activation of the striatonigral direct pathway promotes movement, whilst activation of the indirect striatopallidal

pathway inhibits movement (Albin et al., 1989). Dopamine is able to regulate behavioural control via dopaminergic receptors in the striatum, with direct pathway D1 receptors carrying out a cortically-determined movement signal, and indirect pathway D2 receptors inhibiting any conflicting actions (Frank & O'Reilly, 2006; Frank, Santamaria, O'Reilly, & Willcutt, 2007) – although, as discussed in 4.4, there are some important caveats to this simplistic interpretation, particularly when comparing dorsal and ventral striatum.

On the whole, the Frank et al. (2006) model of behavioural activation and inhibition via D1 and D2 receptors is well-supported in the dorsal striatum; use of more modern methods such as optogenetics, which allow precise localisation and temporal control of neurons expressing these receptors, have confirmed a Parkinsonian behavioural phenotype when activating the indirect pathway – including freezing responses, bradykinesia, and decreased spontaneous action initiation – and increased locomotion with activation of the direct pathway (Kravitz et al., 2010). Similarly, another study using pharmacological manipulations in the dorsomedial striatum in a stop-signal reaction time task found that antagonism of D1 receptors speeded reaction times to stop, whilst antagonism of D2 receptors had the opposite effect, again illustrating the opposing roles of these receptors (Eagle et al., 2011). However, calcium recordings of these pathways in the dorsal striatum has indicated that, as animals initiated actions, both the direct *and* indirect pathways were concurrently activated (Cui et al., 2013).

Similarly, our understanding of the activation and inhibition of behaviour via the direct and indirect pathways in the ventral striatum is also ambiguous. For example, the Eagle et al. (2011) study found little effect on behavioural inhibition when D1 and D2 antagonists were infused into the core of the nucleus accumbens, whilst other studies *have* found a role for the ventral striatum in inhibition of behaviour (e.g. Christakou, Robbins, & Everitt, 2004). While this may at first appear inconsistent, what it in fact demonstrates is the importance of paying close attention to the psychological demands of different behavioural paradigms.

5.2. Components of behavioural inhibition

The ability to withhold movement or action when necessary is an important facet of adaptive behaviour. For example, in extreme circumstances such inhibition can sometimes allow for the avoidance of a threat or danger, but more generally it can also allow for more calculated, accurate, and appropriate actions to take place. In converse, behavioural *disinhibition* – or impulsivity – is usually defined on the basis of inappropriate, ill-conceived, or risky action, with reduced consideration of the consequences of such action (Durana & Barnes, 1993). However, as others including Dalley et al. (2011) have pointed out, this is very much a multi-faceted definition that can be parsed out into several components. Indeed, studies in both humans (Solanto et al., 2001) and animals (Winstanley, Dalley, Theobald, & Robbins, 2004) that have investigated the degree of correlation between behaviour on tasks measuring these components have often found a lack of relationship between them, suggesting that impulsivity is not a single, unitary construct. This is also supported by studies of the neural substrates of impulsivity (Dalley et al., 2011; Jupp & Dalley, 2014). One possible

taxonomy of inhibitory control has been proposed by Bari & Robbins (2013), which distinguishes between the postponement, restraint, and cancellation of action as forms of impulsive action, separate from impulsive choice (Figure 5). These components, their associated behavioural paradigms, and the contribution of the striatal dopaminergic regions and receptor subtypes to them are subsequently discussed in the order in which they may naturally arise in behaviour.

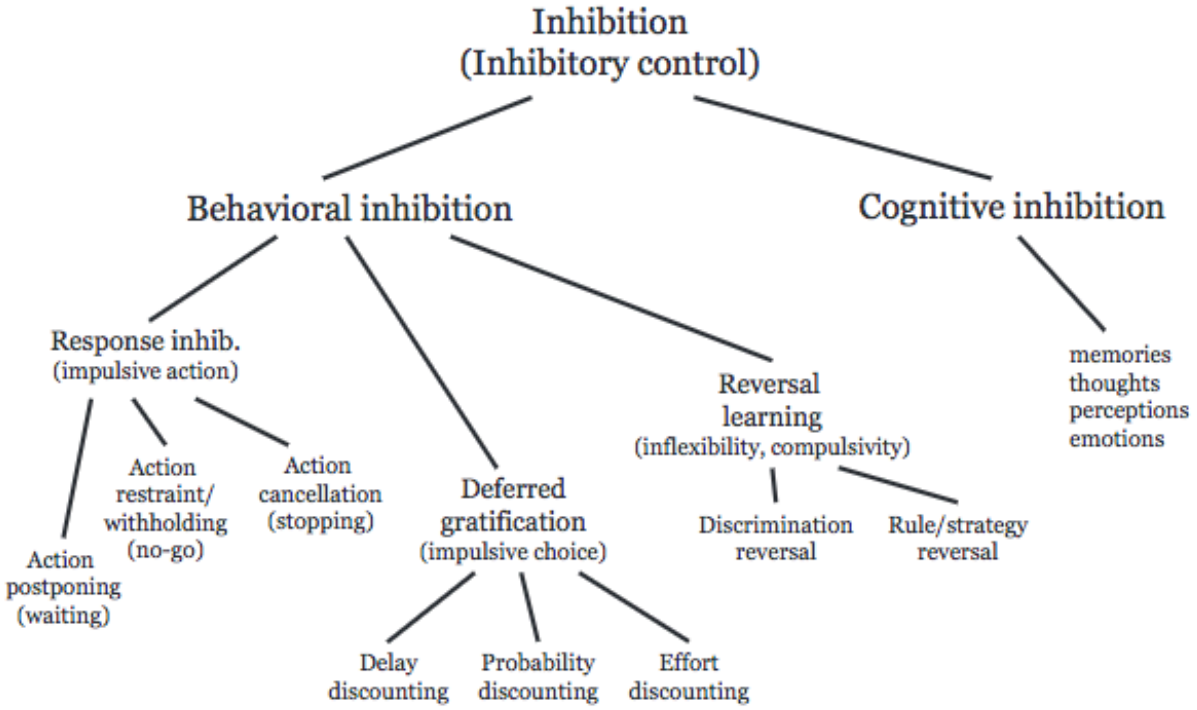


Fig. 5. A proposed taxonomy of response inhibition. Here the focus will be on behavioural inhibition. (Adapted from Bari & Robbins, 2013).

Impulsive choice

Before carrying out behaviour, one usually has to choose (either implicitly or explicitly) one course of action over another. In paradigms of impulsive choice, the decision is often between a sooner, smaller reward or a larger, later reward, with an increased propensity for the former indicating greater impulsivity. A role for

dopamine in such forms of decision-making has been proposed due to its mediation of the incentive value of rewards, although it has been debated whether higher levels of the neurotransmitter should bias individuals' preference towards more immediate outcomes – due to increasing incentive value of the sooner option – or towards the delayed outcome, allowing animals to overcome the cost of waiting for reward.

Indeed, lesion studies have shown that ablation of the core (Cardinal, Pennicott, Sugathapala, Robbins, & Everitt, 2001) but not shell (Pothuizen, Jongen-Rêlo, Feldon, & Yee, 2005) subregion of the NAc shifts choice to sooner, smaller rewards, supporting the latter hypothesis. Although a systemic study, this finding is further supported and extended by van Gaalen, van Koten, Schoffelmeer, and Vanderschuren (2006) who showed that D1R antagonism also increases impulsive choice. However, whilst another study also showed this same effect of D1R antagonism, systemic stimulation of D1Rs did not *reduce* impulsive choice (Koffarnus, Newman, Grundt, Rice, & Woods, 2011), though these effects may be asymmetrical due to the nature of pharmacological manipulation as a method.

Application of a D2R antagonist had no effect in this task, suggesting that impulsive choice may be exclusively mediated by D1Rs (van Gaalen et al., 2006). One important caveat is that this effect is sensitive to the parameters used in the task; when the value of the smaller, immediate reward is adjusted in order to identify an indifference point – when the rat is willing to choose either option with equal frequency, indicating that they are of approximately equal subjective value – a systemic D2R antagonist can in fact decrease the indifference point via a reduction in the value of delayed rewards,

suggesting an increase in impulsivity (Wade, De Wit, & Richards, 2000). In this same study, a D1R antagonist had no effect on impulsive decision-making.

Action postponement: waiting impulsivity

Being unable to withhold action until it is appropriate to make a movement is a form of behavioural disinhibition, or impulsivity, that is related to impulsive choice but lacks the selection between rewarded alternatives (Robinson et al., 2009). The ability to restrain action whilst waiting for a reward – in other words, to postpone the action – has been studied using several behavioural paradigms, the most well-known and oft-used being the 5-choice serial reaction time task (5-choice) (Robbins, 2002). Whilst the task was originally developed with the aim of understanding attentional processes, it also allows for metrics of both impulsivity and compulsivity. In the 5-choice, a brief visual stimulus appears in one of five nose pokes, and animals have to indicate their detection of said stimulus by responding in the appropriate poke as quickly as possible in order to achieve reward. However, prior to the presentation of the stimulus animals must withhold responding; any anticipatory responses into a port during inter-trial interval period (usually lasting between 5 and 9 seconds) are deemed premature and thus impulsive, and are usually punished by a timeout period.

Both systemic and local infusion of d-amphetamine into the NAc core in this task have been shown to increase impulsive responding (Pattij, Janssen, Vanderschuren, Schoffelmeer, & van Gaalen, 2007; Van Gaalen, Brueggeman, Bronius, Schoffelmeer, & Vanderschuren, 2006). Evidence as to whether similar effects are also found in the

shell subregion of the NAc is mixed; for example, using a simplified 'forced-choice' version of the 5-choice task, where only one port is available instead of 5, Murphy, Robinson, Theobald, Dalley, and Robbins (2008) found that lesions of the core enhanced impulsivity produced by systemic amphetamine, whilst lesions of the shell produced the converse effect. However, neither lesions of the core nor shell affected action disinhibition at baseline, suggesting that providing multiple options for the expression of operant responding may be important when attempting to elicit impulsive behaviour. Indeed, it has also been shown that rats that are categorised as 'high impulsive' on the 5-choice task also have steeper discounting functions in impulsive choice, suggesting that there might be overlap between these constructs (Robinson et al., 2009).

With regards to the separable contributions of the dopamine receptor subtypes to waiting impulsivity, both systemic (Harrison, Everitt, & Robbins, 1997; Van Gaalen et al., 2006) and local infusions into the NAc core or shell (Pattij et al., 2007) of a D1R antagonist have been shown to reduce impulsivity. However, as with the studies investigating impulsive choice, a D1R agonist applied systemically does not always increase impulsivity (Winstanley et al., 2010) although local infusion of a high dose of the same drug into the NAc did in fact reduce animals' ability to wait for reward (Pezze, Dalley, & Robbins, 2007). This effect does not appear to be entirely specific to D1Rs, however, as infusion of a D2R agonist also increased impulsivity in this same study, whilst a D2R antagonist in contrast had little effect (Pattij et al., 2007; Pezze et al., 2007). In short, both dopamine receptor families may play a role in withholding responding during a time delay for reward, and this effect can be localised to both

the NAc core and shell, with activation of dopamine receptors necessary for behavioural disinhibition.

As mentioned before, there is some overlap between the behavioural constructs measured in paradigms of delay discounting of reward, and those that study waiting impulsivity. However, there are some inconsistencies when comparing the effects of receptor specific manipulation in the two task types; in particular, D1R antagonism seems to *increase* impulsive choice whilst also *decreasing* waiting impulsivity. This may reflect fundamental differences between these forms of impulsivity, particularly in terms of the necessity of purposeful action in one but not the other. However, as mentioned previously, the lack of focused accumbens manipulations in delay discounting paradigms also makes meaningful direct comparison difficult.

Action cancellation: stopping impulsivity

The aspects of impulsivity discussed thus far have considered processes that occur up to the point of action initiation, but as discussed in Dalley et al. (2011), another form of behavioural disinhibition is stopping impulsivity: the ability (or lack thereof) to cancel an action that has already begun to be executed. In humans and rodents, a popular paradigm used to study this phenomenon is the stop-signal reaction time task (Eagle, Bari, & Robbins, 2008; Logan, Cowan, & Davis, 1984). Subjects are trained to respond to targets as quickly as possible in this task, which makes up a predominance of the trials, but on a proportion of trials a 'stop' signal is sounded, at which point the subject has to cancel their behavioural response on that trial. One

benefit of this experimental design is that the outcome measure is not simply whether a participant has been able to cancel the action, but also the speed at which they can do so – the stop-signal reaction time (Eagle et al., 2008).

It has been consistently shown that mesolimbic dopamine may only play a limited role cancelling actions that have already been initiated. Both a D1R and D2R antagonist administered systemically had no influence on the stop-signal reaction time (Bari & Robbins, 2013), and infusion of the same substances into the core of the NAc similarly had little effect (Eagle et al., 2011). This is in contrast to manipulations in the *dorsal* striatum in the same study, which either decreased (in the case of the D1R antagonist) or increased (with a D2R antagonist) impulsivity. Thus whilst there is a clear involvement of the NAc core (and possibly shell) in both choice and waiting impulsivity, and whilst D1Rs and D2Rs may contribute to the ability to inhibit action in these scenarios, the role they play in simple motor inhibition is negligible (Dalley et al., 2011).

5.3. Biasing action selection

Neurons in the dorsal striatum have been implicated in the selection of specific motoric actions, regardless of their goal-directed nature (Howard, Li, Geddes, & Jin, 2017). Manipulations of the dopaminergic receptor families in the dorsal striatum have also confirmed a role for these neurons in action selection, with unilateral optogenetic activation of D1R-expressing MSNs increasing choices of the contralateral port, whilst unilateral activation of D2R-expressing MSNs increased

choices of the ipsilateral port (Tai, Lee, Benavidez, Bonci, & Wilbrecht, 2012). However, there is less evidence that neurons in the NAc are also involved in such specific action selection. Indeed, one study that directly compared reward and choice encoding in dorsal and ventral striatal terminals found that whilst dopaminergic dorsomedial terminals were most influenced by a contralateral choice, those in the ventral striatum responded most to the consumption of reward and presentation of reward-predicting cues (Parker et al., 2016). A complementary proposal is that, in contrast to the dorsal striatal action selection strategy, which encodes associations between temporally predictable stimuli and actions, neurons in the accumbens may encode associations between temporally *unpredictable* stimuli and an appropriate action to maximise reward (Nicola, 2007).

Additionally, there is a proposal that D2 receptors might bias action selection when choosing between behavioural strategies which complements studies showing that their stimulation leads to behavioural perseveration and inflexibility (Haluk & Floresco, 2009). In their 'prepare and select' model of the striatal direct and indirect pathways, Keeler, Pretsell, & Robbins (2014) propose a mechanism for this by which MSNs expressing D2 receptors encode reward association strength via internalisation of those receptors (Bartlett et al., 2005), and so application of a D2 agonist, for example, would lead to an inhibition of pathways with a weak reward association strength and thus selection of the same, previously high-value, action. Another theory suggests that MSNs expressing D2Rs may instead be exclusively encoding the negative consequences, or costs, of an action, and thus a D2R agonist for example would thereby increase the costliness of an action such that an alternative, less costly

action would be preferred, without changing preference in a scenario where only payoff discrepant between the options (Bogacz, 2017). Thus whilst D1Rs or D2Rs may not play a role in selection of the very specific properties of an action, they may bias action selection when weighing up the costs and benefits of rewarding actions, even if overall levels of dopamine in the accumbens often do not reflect the decision ultimately made (Gan, Walton, & Phillips, 2010).

5.4. Action vigour and motivational drive

The traditional conception of the 'go' and 'no-go' striatal pathways implicates them as controllers of the ability, or decision, to inhibit or express behaviour (Frank & O'Reilly, 2006). However, the findings from Eagle et al., (2011) suggest that, in dorsomedial striatum at least, dopamine also plays a role at *speed* at which behavioural inhibition proceeds once the decision to withhold responding has taken place. There appears to be a similar story in the ventral striatum; for example, D1 antagonism in the nucleus accumbens core during the 5-choice task slows latencies to respond even in correct trials (Pezze et al., 2007). Therefore, it is not just the likelihood of responding but potentially the degree of *drive*, or incentive salience, of reward-directed behaviour that is modulated by accumbens D1 receptors.

This also complements findings from studies of effort-based decision-making, where D1 antagonism reduces choice of the high effort option (Yohn et al., 2015). Indeed, Salamone's work has demonstrated the necessity of accumbens dopamine, including the activation of NAc core D1Rs, in overcoming costs in order to achieve reward that is separated from the animal either by effort in terms of lever-pressing (Nowend,

Arizzi, Carlson, & Salamone, 2001) but also more general vigorous activity and – as discussed previously – delay when having to wait for reward; in other words, allowing organisms to "transcend the psychological distance that separates them from motivationally relevant stimuli" (Salamone & Correa, 2012). This is also consistent with a framework where the effects of these receptors are dependent on a calculation of utility – whether it is worth being motivated or initiating an action for reward – and relates to ideas discussed in section 4.5 connecting mesolimbic dopamine to the value of work.

The role that D2 receptors play in this framework of action utility is perhaps more ambiguous. However, a recent study showed that they too can contribute to motivational drive-like behaviours. Soares-Cunha et al. (2016) carried out a series of experiments using both Pavlovian-to-instrumental transfer and progressive ratio paradigms (the former providing a behavioural measure of degree of incentive salience, and the latter providing a measure of willingness to work for reward) and found that brief optogenetic activation of D2-expressing MSNs in the nucleus accumbens during reward-predictive cues augmented motivation. The same manipulation of D1 receptors led to a similar behavioural phenotype. The authors cite evidence showing that D2 receptors can also be rapidly activated by phasic dopamine release (Marcott, Mamaligas, & Ford, 2014) as a mechanism for these changes, and it builds upon work by Floresco (2007) that has also indicated a role for D2 receptors in motivational processes, as well as other work showing inhibition of the indirect pathway can lead to reduced likelihood of completing a sequence of actions for

reward (Tecuapetla et al., 2016). Therefore, both D1-like and D2-like receptors may contribute to the execution and vigour of actions for reward.

5.5. Behavioural responding to reward-predictive cues

Beyond the reward prediction error hypothesis of mesolimbic dopamine function, there is evidence that the D1 and D2 receptor subtypes may have unique roles in learning about reward and reinforcement in response to reward-predictive cues. However, this will not be discussed here as all behaviours studied in this thesis are well-learnt. Of more relevance is the contribution of these receptors to responding to value-laden cues. A series of studies by the Nicola lab have used a discriminative stimulus task to investigate this phenomenon. This task presents animals with rewarded (discriminative stimulus; DS) and non-rewarded (non-rewarded stimulus; NS) cues. For animals to garner reward during the DS, they must perform a correct action by nosepoking in the active port. Any responses made during the NS into either the active or inactive ports will not elicit a reward. In some variants of the task, a third cue is also presented – a probabilistically predictive stimulus (PS) which elicits rewards on a fraction of trials.

At baseline, animals will respond to nearly all DS cues, but Nicola, Taha, Kim, & Fields (2005) found that infusion of a dopamine re-uptake inhibitor into the nucleus accumbens core increased responses to the PS but not the NS. This supports the hypothesis of dopamine motivating reward-oriented behaviour and 'shifting the balance' in terms of performing action for reward is worthwhile, separable from an influence on general motoric activity. However, the infusion of both D1 and D2

antagonists infused into the nucleus accumbens leads to a decrease in responding to the DS, suggesting that both receptor subtypes are necessary for the expression of this behaviour (Yun, Wakabayashi, Fields, & Nicola, 2004). And in a more recent study, they utilised microelectrode arrays that allowed for the simultaneous recording of neuronal firing and the infusion of pharmacological agents locally to show that application of antagonists of either receptor subtype into the accumbens core selectively attenuates the excitation, but not inhibition, evoked by reward-predictive cues of MSNs in the region (du Hoffmann & Nicola, 2014).

This is in agreement with the aforementioned study by Soares et al. (2016) that showed that brief activation of either D1 or D2Rs could augment motivation when having to work for reward, particularly in terms of the potential synergestic relationship between these receptors in driving behaviour. Anatomically, this may be related to the comparative lack of segmentation between the D1R- and D2R-pathways in the NAc (Kupchik et al., 2015).

5.6. Prefrontal cortex and cognition

Dopamine in the prefrontal cortex has some quite distinct roles from the reward, motivation, and action-related functions that characterise dopamine transmission in the striatum. Arguably its best-studied aspect is its contribution to working memory; the ability to hold and manipulate items in memory (Baddeley, 2012). In animals, this function is typically measured using delayed response tasks (Seamans & Yang, 2004) and generally it has been found that local infusions of D1, but not D2, antagonists disrupt working memory performance (Goldman-Rakic, 1998). Similarly, application

of a D1 agonist systemically has been shown to facilitate information retrieval after a long delay, suggesting that these receptors may also play a role in long-term memory (Hotte, Naudon, & Jay, 2005).

Prefrontal dopamine also plays an important role in attentional processes, which is of particular relevance considering that many of the previously described inhibitory control tasks (and many operant tasks more generally) require attention for success. Stimulating VTA dopamine neurons that project to the medial prefrontal cortex has been shown to aid learning in a discrimination task (Popescu, Zhou, & Poo, 2016), and the D1 receptor specifically has been shown to mediate attentional performance in the 5-choice task (Granon et al., 2000). This is in contrast to findings in mice lacking D2 receptors; although their ability to acquire the task was reduced, these mice showed no deficits in attending to perceptual stimuli (Glickstein, DeSteno, Hof, & Schmauss, 2005). Therefore, the role of D2 receptors in attentional processes is less well established.

6. Anatomy, physiology, and receptor pharmacology of the mammalian serotonergic system

Often considered an 'opponent' of the dopamine system both pharmacologically and behaviourally (Daw, Kakade, & Dayan, 2002), serotonin, or 5-hydroxytrptamine (5-HT), is another member of the monoamine family and a neurotransmitter which exerts its effects throughout the central and peripheral nervous systems (Berger, Gray, & Roth, 2009). Serotonin has been implicated in a multitude of behaviours and

cognitions including emotional processing, motor behaviours, autonomic processes, arousal, sleep cycles, and appetite (Frazer & Hensler, 1999). However, the focus in this thesis will be its contributions to goal-directed behaviour and movement initiation and inhibition, with an emphasis on how serotonin may interact with the dopaminergic system to produce and influence these processes.

6.1. Serotonergic nuclei and major projections

In the brain, serotonin can be found in nine groups of cell bodies located in the brainstem – specifically the pons – and the midbrain (Dahlstrom & Fuxe, 1964), although the raphe nuclei in the midline of the brainstem produce the vast majority of this neurotransmitter. The efferent projections that arise from the raphe nuclei are extensive, with the caudal raphe innervating the spinal cord, whilst the rostral raphe nuclei, including the dorsal and medial raphe, releasing serotonin into multiple cortical and subcortical structures (Berger et al., 2009).

One such projection is that of the dorsal raphe to the striatum (Waselus, Galvez, Valentino, & Van Bockstaele, 2006). Similarly, it has been shown that 5-HT neurons directly synapse onto both dopaminergic and non-dopaminergic cells in the VTA (Van Bockstaele, Cestari, & Pickel, 1994). Therefore, the dorsal raphe has been heavily implicated in reward-seeking behaviours, both due to its innervation of brain areas related to reward processing, and its direct interaction with the dopaminergic system (Nakamura, 2013).

Similarly, the dorsal raphe nucleus receives input from areas such as the VTA and SNc (Peyron et al., 1995) amongst other innervations that convey state-based information, such as the hypothalamus (Celada, Casanovas, Paez, & Artigas, 2002), and negative motivational values, such as the lateral habenula (Matsumoto & Hikosaka, 2007). Thus whilst 5-HT has an influence on all of the major dopaminergic pathways, the projection from the dorsal raphe nucleus to the striatum may be a key player in behaviour for reward.

6.2. Characterisation and distribution of the serotonin receptors

The serotonin receptor family is large – in fact, there are more serotonin receptor subtypes than any other family of G-protein coupled (GPCR) neurotransmitter receptors, with 13 different GPCR receptors and one ligand-gated ion channel. These receptors have been divided into seven groups, 5-HT_1 to 5-HT_7, but can also be classified based on whether they couple to $G_{q/11}$, G_s, or $G_{i/o}$ proteins, with the one exception being the ionotropic 5-HT_3 receptor (Nichols & Nichols, 2008). This complexity extends to the anatomical distribution of receptors throughout the brain; although corresponding to the extensive nuclei projections, 5-HT receptors of various subtypes can be found throughout the substantia nigra, hippocampus, hypothalamus, amygdala, striatum, and frontal cortex (Charnay & Leger, 2010).

Of the serotonin receptors, the 5-HT_{2C} receptor has been strongly implicated in reward-related behaviours (Martin, Ballard, & Higgins, 2002). This receptor, perhaps because of its inherently high level of constitutive activity (Berg, Harvey, Spampinato,

& Clarke, 2005) is also able to regulate dopaminergic tone across all three major dopaminergic pathways, including modulation of mesolimbic dopaminergic function (Matteo, Giovanni, Mascio, & Esposito, 1999a). The 5-HT_{2C} receptor is an excitatory $G_{q/11}$-coupled receptor that can be found in the caudate nucleus, nucleus accumbens, olfactory tubercle and pyriform cortex as well as across the prefrontal cortex (Pazos & Palacios, 1985; Santana & Artigas, 2017), but along with the 1B and 2A subtypes, the 5-HT_{2C} receptor can be found at particularly high levels in the striatum (Wright, Seroogy, Lundgren, Davis, & Jennes, 1995). Until relatively recently, it was difficult to distinguish the 5-HT_{2C} receptor from its 2A and 2B counterparts, but selective agents for each of these subtypes have now been developed (Hoyer, Hannon, & Martin, 2002). Therefore, it is now possible to investigate its conjoint influence on behaviour and the mesolimbic dopamine system.

6.3. Serotonergic regulation of striatal MSNs

There is much evidence from both studies that have stimulated the dorsal raphe and local infusions of serotonin that 5-HT can affect striatal function, although there are data showing both inhibition (Davies & Tongroach, 1978; Olpe & Koella, 1977) and excitation of its MSNs (Navailles & De Deurwaerdère, 2011).

The irregularities between these studies may be due to differences in the consequences of activating the serotonergic receptor subtypes. Indeed, of the many 5-HT receptor subtypes, at least 9 have been identified as being expressed in the striatum (Hoyer et al., 2002). The 5-HT_{2C} receptor specifically contributes to striatal inhibition via indirect tonic inhibitory control over MSN excitability (Rueter, Tecott,

& Blier, 2000). There are two potential mechanisms for this control: both enhancement of the inhibitory action of cholinergic interneurons on striatal MSNs, and an increase in excitability of fast-spiking interneurons, thereby increasing GABAergic postsynaptic inhibition to also result in decreased striatal MSN activity (Miguelez, Morera-Herreras, Torrecilla, Ruiz-Ortega, & Ugedo, 2014).

6.4. Serotonergic regulation of mesolimbic dopamine release

Whilst serotonin receptors expressed in the striatum can modulate the activity of striatal MSNs, there is also an abundance of causal evidence that modulating the central 5-HT system can have knock-on effects on mesolimbic dopamine levels; specifically, lesioning of the raphe nuclei increases the production of dopamine in the nucleus accumbens, suggesting that 5-HT also exerts inhibitory control over the dopamine system (Herve et al., 1979).

Whilst the activation of some receptor subtypes has a stimulatory effect on dopamine release (Alex & Pehek, 2007), the inhibition of the dopamine system, as with the inhibition of striatal MSNs, is predominantly mediated by the 5-HT_{2C} receptor; in the midbrain, the baseline activity of VTA dopamine neurons is inhibited by the administration of 5-HT agonists, and application of a selective 5-HT_{2C} antagonist can reduce this effect (Di Giovanni, Di Matteo, Di Mascio, & Esposito, 2000). This is supported by studies using 5-HT_{2C} agonists (Gobert et al., 2000). Similarly, application of a 5-HT_{2C} antagonist alone leads to a dose-dependent increase in the basal firing rate of these same neurons as well as an increase in bursting activity (Matteo, Giovanni, Mascio, & Esposito, 1999b). Put simply, activation of these

receptors inhibits the spontaneous firing of dopamine neurons, whilst their blockade leads to the converse. However, this effect may be localised to the VTA portion of the midbrain specifically; application of a 5-HT reuptake inhibitor was found to have very little effect on dopamine neurons in the adjacent SNc (Prisco & Esposito, 1995).

The exact mechanism of this regulation of tonic dopamine activity has not yet been entirely elucidated, but a portion of 5-HT_{2C} receptors are localised on GABAergic interneurons which are able to inhibit dopamine neurons in the VTA (and SNc). Therefore, GABA-mediation may have a large part to play (Nakamura, 2013). However, to complicate matters, there is recent evidence that 5-HT_{2C} receptors are also localised on *dopamine* neurons in the VTA, though their contribution to the activation or suppression of these neurons – and whether their effects are differentiable behaviourally – is still unclear (Bubar, Stutz, & Cunningham, 2011). Feedback loops across brain areas may also be important; 5-HT receptors are also located postsynaptically in dopamine projection areas such as the striatum, and feedback from, for example, the nucleus accumbens to the VTA via GABAergic neurons could indirectly alter the excitability of dopamine neurons in these areas (Di Giovanni, Esposito, & Di Matteo, 2010).

Often serotonergic mediation of the mesolimbic dopamine system is discussed in terms of its effects on the basal, or tonic, activity and release of dopamine. But there is evidence that 5-HT_{2C} receptors may also be able to modulate phasic mesolimbic activity; applications of 5-HT_{2C} ligands at a concentration that does not affect tonic levels of dopamine in the nucleus accumbens are regardless able to affect stimulated

release due to cocaine (Alex & Pehek, 2007; Navailles, De Deurwaerdère, Porras, & Spampinato, 2004). However, these and other similar studies have been conducted in anaesthetised animals, and in behaving animals the results are more mixed, with recent evidence suggesting that antagonising 5-HT_{2C} receptors may not effect reward-related phasic dopamine release in the nucleus accumbens (Bailey et al., 2018).

6.5. Prefrontal serotonergic 2C receptors

Whilst there is much evidence that 5-HT_{2C} receptors play a role in regulating striatal dopamine and MSNs, there is less of an indication that this is also the case in the prefrontal cortex. This may be due to overall reduced levels of expression in this area; whilst 5-HT_{2C} receptor mRNA has been found to be expressed in the frontal cortex, it is at relatively weak levels, suggesting that these receptors are less concentrated in the region (Pompeiano, Palacios, & Mengod, 1994). However, they can be found at highest concentration in layers V and VI, particularly on GABAergic interneurons within this region (Liu, Bubar, Lanfranco, Hillman, & Cunningham, 2007). This suggests potential for an inhibitory feedback loop, whereby the activation of these 5-HT_{2C} receptors increases GABAergic activity, thereby inhibiting excitatory glutamatergic cells that project to the mesolimbic pathway (Sesack, Carr, Omelchenko, & Pinto, 2003).

Systemic administration of a 5-HT_{2C} agonist decreases dopamine release in the prefrontal cortex (Gobert et al., 2000), and conversely applying an antagonist increases dopamine levels, reflective of findings in the striatum (Millan, Dekeyne, & Gobert, 1998). There are several caveats, however. Firstly, this effect may be state-

dependent, (Pozzi, Acconcia, Ceglia, Invernizzi, & Samanin, 2002). Secondly, and more importantly, when such drugs are infused directly into the prefrontal cortex, there is little to no change in cortical basal dopamine levels (Alex, Yavanian, McFarlane, Pluto, & Pehek, 2005). This suggests that there is a lack of terminal-region 5-HT_{2C} receptor modulation of the mesocortical pathway, differentiating it from its nigrostriatal and mesolimbic counterparts (Alex & Pehek, 2007). However, it remains the case that 5-HT_{2C} receptors elsewhere in the brain can modulate dopamine in the prefrontal cortex, an important consideration when using systemic pharmacological approaches.

Beyond the modulation of the dopaminergic system, there are several sources of direct serotonergic influence over the excitability of neurons in the prefrontal cortex, for example by modulating synaptic plasticity (Anastasio et al., 2010; Carr, Cooper, Ulrich, Spruston, & Surmeier, 2002). However, it is the inhibitory 5-HT_{1A} and excitatory 5-HT_{2A} receptors that are most densely expressed on pyramidal neurons in the prefrontal cortex, and similarly these plus the 5-HT_{3A} receptor are also widely expressed on inhibitory interneurons in the cortex, with the 5-HT_{2C} receptor less abundant in the region (Puig & Gulledge, 2011).

7. Theories of serotonin function

Unlike the dopaminergic system, there is not yet a unified framework for the understanding of serotonergic function in the mammalian brain, but despite this, there have been attempts to develop a 'theory of serotonin' that encompasses its seemingly ubiquitous contribution to behaviour and cognition. These attempts have

classically fallen into two camps: serotonergic processing of aversive stimuli or events (Deakin & Graeff, 1991a), and the contribution of serotonin to behavioural inhibition (Soubrie, 1986). More recently, these camps have been coalesced by proposals that serotonin lies on the other side of the coin to dopamine; where dopamine signals the positives and drives behaviour forward, serotonin signals the negatives, and inhibits that drive (Cools, Nakamura, & Daw, 2011; Daw, Kakade, & Dayan, 2002; Tops, Russo, Boksem, & Tucker, 2009 - although see Boureau & Dayan, 2011). Here, both these classical and more modern interpretations will be reviewed, with a particular focus on the role of serotonin in behavioural inhibition.

7.1. Inhibition in affective contexts

The 5-HT system has long been associated with punishment (Tye, Everitt, & Iversen, 1977) and aversive processing (Deakin & Graeff, 1991). Of particular relevance here is the fact that aversion and *inhibition* are tightly coupled concepts; for example, mice may freeze in response to a noxious or threatening stimulus (Yilmaz & Meister, 2013). There is evidence that serotonin influences behavioural inhibition in an affective context, such as punishment-induced inhibition (Crockett, Clark, & Robbins, 2009) or inaction following uncontrollable punishment (Maier & Watkins, 2005), and neurons in the raphe nuclei have been shown to respond to inescapable shock during which animals inhibit movement (Takase et al., 2004). Indeed, in the Crockett, Clark, and Robbins (2009) study, there was a direct comparison of punishment-related behaviour and overall motor inhibition, and the authors found increases in the

former but not the latter during acute tryptophan depletion. This has led to proposals that serotonin uniquely tunes affective behaviours (Dayan & Huys, 2009).

However, whilst it seems intuitive to relate a response to negative stimuli to the inhibition of behaviour, it is sometimes preferable in situations of threat or aversion to perform an *active* defensive response (Blanchard, Blanchard, & Briebel, 2005). And whether a stimulus can be considered positive or negative in the first place depends on the context in which it is placed – the loss of an expected reward, or the obtainment of a smaller reward than expected could both credibly be conceived as forms of punishment. Thus Boureau and Dayan (2011, figure 5) have proposed a more refined framework that takes these complexities into account by the integration of two dimensions: affect (reward and punishment) against effect (invigoration and inhibition) (Ikemoto & Panksepp, 1999). Within this framework, there can be a role for serotonin in behavioural inhibition outside affect, whilst also taking account its interactions with punishment processing.

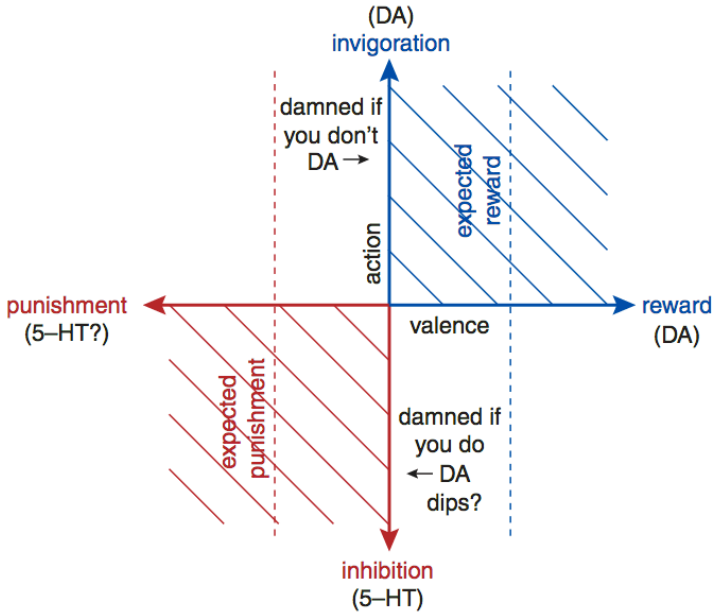


Fig. 5. The effect-affect plot of dopamine and serotonin function (adapted from Boureau & Dayan, 2011), demonstrating interactions between the 'effect' axis of invigoration and inhibition against the 'affect' axis of punishment and reward.

7.2. Inhibition in outside of affect: waiting for reward

The forms of inhibition in affective contexts mentioned in the previous section often involve immobility in the face of danger or punishment. But inhibition – or the lack of it, impulsivity – comes in various other forms that have been investigated formally using a range of behavioural paradigms, as described in section 5.1, though due to their importance for this thesis, they are briefly reiterated here in the context of serotonergic manipulations.

Impulsive choice

Whilst some have found no effect of forebrain serotonergic depletion on the discounting of delayed reward (Winstanley, Dalley, et al., 2004), others have often found a hypersensitivity to delay with a pharmacological reduction of serotonin (Denk et al., 2005; Schweighofer et al., 2008; Cardinal, 2006), as well as the converse with increased levels of serotonin (Bizot, Le Bihan, Puech, Hamon, & Thiébot, 1999). It has been suggested that this is due to an increased rate of discounting in such tasks, making the delayed reward the subjectively less valuable option (Mobini, Chiang, Ho, Bradshaw, & Szabadi, 2000). There is also evidence that this effect on impulsive choice, but not waiting impulsivity as measured by the 5-choice, may be mediated by the 5-HT_{2C} receptor (Paterson, Wetzler, Hackett, & Hanania, 2012).

Action disinhibition: waiting impulsivity

In tandem with the aforementioned link between reduced serotonin and increased impulsive choice, action is generally disinhibited with ablation of forebrain serotonin, resulting in an increase in the number of premature responses in the 5-choice (Harrison, Everitt, & Robbins, 1997) but also an increase in disinhibition as indicated by the speed and number of responses made in an autoshaping protocol (Winstanley, Dalley, et al., 2004), suggesting that this effect is not dependent on the availability of different options. Interestingly, Harrison et al. (1997) found that the expression of impulsivity was attenuated by administration of a systemic D1R antagonist, suggesting that this effect may be mediated by the dopaminergic system.

The comparatively simpler Go/No-Go task also allows for measurement of waiting impulsivity. In one form of this paradigm used by Fletcher (1993), animals are trained to respond primarily during periods of reinforcement, which are signalled by the presence or absence of a light cue. Drug infusions to suppress the activity of 5-HT neurons in either the dorsal or medial raphe impaired success rate in this task, though in the case of the dorsal raphe, it was due to *reduced* responding during the reward cue, whilst in the medial raphe animals *increased* responding during periods of non-reward. Thus there may be homogeneity across anatomically differentiable populations of serotonergic neurons in their contribution to withholding action for reward. A more structured form of the Go/No-Go task employed by Harrison, Everitt, and Robbins, (1999), in which animals were punished for erroneously responding during No-Go trials, replicated an inability to withhold responding when appropriate.

However, the degree of this effect, particularly when altering serotonin in the NAc, may be dependent on the length of the delay. Unlike several previous studies, Fletcher, Chambers, Rizos, and Chintoh (2009) found that selective 5-HT depletion from the NAc did not affect action disinhibition in the 5-choice, but did increase responding on using a differential-reinforcement of low-rate (DRL) schedule, during which animals are reinforced after set periodic delays (in this case, every 20s). The authors suggest that this may have been because of the difference in waiting times, which is two-to-four fold greater in the DRL task than the 5-choice and is thereby more difficult for animals to estimate. An alternative possibility is that serotonin has more of an effect when waiting is not explicitly cued, such as in the DRL, but is instead

internally-mediated, in contrast to tasks such as the 5-choice which are strongly cue-driven.

Studies using more modern methods have extended the hypothesis that serotonergic neurons, particularly in the dorsal raphe, promote waiting for reward. For example, one study showed that optogenetic stimulation of this population of neurons increased the ability of mice to withhold premature responding when waiting for a randomly delayed tone, although this same stimulation did not have an effect on reward-based place preference or decision-making paradigms, suggesting that it is not reinforcing *per se* (Fonseca, Murakami, & Mainen, 2015).

In terms of the serotonergic receptor subtypes, antagonism of the 5-HT_{2C} receptor increases premature responding, acting in the same direction as gross serotonin manipulations – and in converse to the 2A receptor, which is in fact able to decrease premature responding (Winstanley, Theobald, Dalley, Glennon, & Robbins, 2004). Notably, these effects are mediated by receptors in the nucleus accumbens, rather than prelimbic or infralimbic cortex, as determined by local infusions (Robinson et al., 2008). This localisation of the effects of global 5-HT depletion on impulsivity to subcortical sites has been reaffirmed by studies showing positive correlations between tonic levels of 5-HT in the medial prefrontal cortex and premature responses – in other words, in the opposite direction to the aforementioned studies (Dalley & Roiser, 2012). Therefore, subcortical sites are the likely candidates for this link between increased waiting (and choice) impulsivity with reduced serotonin.

Action cancellation: stopping impulsivity

Whilst manipulating the serotonergic system has clear effects on choice and waiting impulsivity, there is little evidence for a role of this neurotransmitter in cancelling an ongoing action. Studies using the stop-signal reaction time task have found no change in reaction time with serotonergic manipulations in both humans and animals (Clark et al., 2005; Eagle et al., 2009). Thus, like dopamine, serotonin seems to contribute little to non-affective simple motor inhibition, instead enabling organisms to wait and inhibit action for reward.

7.1. Vigour and effort

Whilst there is much evidence that serotonin gates action such that decreased serotonin levels increases the likelihood of an action taking place, whether the *vigour* of that action once engaged is also dependent on the 5-HT system is a different question. Indeed, these are separable processes that have been carefully dissociated in recent work that demonstrates the unique contributions of dopamine to the initiation and vigour of action (Alves et al., 2018).

It appears that the same can be said for serotonin. Both human (Crockett, Clark, Apergis-Schoute, Morein-Zamir, & Robbins, 2012; Den Ouden et al., 2015) and animal (Fletcher, Tampakeras, Sinyard, & Higgins, 2007) studies using different behavioural paradigms have demonstrated decreased response latencies and thus an increased vigour of responding when serotonergic receptors are stimulated, though this is not necessarily associated with a general increase in motoric activity (Hadamitzky, Feja,

Becker, & Koch, 2009). In the case of Fletcher et al. (2007), the effect was revealed by antagonism of the 5-HT_{2C} receptor specifically.

A recent study elucidated the mechanism by which 5-HT_{2C} receptors may mediate vigour and goal-directed activity by applying a putative antagonist of this receptor to animals as they completed a variety of tasks involving physical effort (Bailey et al., 2018). Consistent with previous findings (Denk et al., 2005), animals' choice for different forms of work was unaffected by this antagonism. However, response vigour increased significantly, with an escalation in lever pressing bout length and press rate. This reflects previous work from the same authors that also showed animals will increase number of responses and duration of effort for reward with this same 5-HT_{2C} ligand (Bailey et al., 2016).

8. Summary

The mesolimbic dopamine system, particularly the VTA-nucleus accumbens projection, is physiologically and theoretically well-placed to translate motivation into action; the 'limbic-motor interface' of the brain, it is strongly responsive to unpredicted reward and reward-related stimuli, but dopamine release in this region also appears to vary dependent on animals *moving* for reward.

The consequence of this efflux is to activate dopamine receptors located on striatal medium spiny neurons. The two key dopamine receptor families have traditionally been thought to play opposing roles in behaviour, permitting and inhibiting action

dependent on the rewards and costs involved. Indeed, pharmacology studies of impulse control seem to suggest that activation of D1 receptors predisposes animals to carrying out action even when it is inappropriate to do so, though the role of D2 receptors in waiting for reward is less well-established. And in paradigms investigating responsivity to reward-associated cues, both D1 and D2 receptors in the ventral striatum appear to contribute to the expression of goal-directed behaviour.

The serotonergic system also interacts closely with the mesolimbic dopamine system, both physiologically and also to influence behaviour in response to reward; whilst traditionally it has been thought of as an opponent to dopamine, more recent work suggests that the two neurotransmitters may work together to promote reward-related behaviours. Additionally, serotonin and the 5-HT_{2C} receptor in particular has also been strongly implicated in the ability to *wait* for reward – in other words, a form of patience – as well as the vigour of action associated with reward obtainment.

9. Thesis aims

Although the concepts of motivation, reward prediction, learning, and action have all been associated with dopamine, there are many unanswered questions that reflect a current inability to tie these aspects of its function together. Such questions include: how does mesolimbic dopamine both encode an apparently homogeneous reward prediction error whilst also strongly influencing task engagement, vigour, and general levels of motivation? How separable are the contributions of the different timecourses of dopamine release to these functions? How might such dopamine

release differentially affect the receptor subtypes, and do they go on to influence action or inaction? Finally, what is the specific role of serotonin in the expression of goal-directed action for reward, and how does this neurotransmitter interact with dopamine in such behaviours?

In this thesis, a single novel behavioural paradigm, a Go/No-Go task, has been developed to attempt to answer these questions. This task was designed to exploit some of the commonly attributed functions of dopamine and serotonin; having trials with different reward sizes allows both for elicitation of reward prediction errors, as well as allowing for the direct examination of motivational levels on behaviour. And, importantly, contrasting instances where animals must make action – Go trials – against trials where animals must inhibit all action to receive reward – No-Go trials – allows for a degree of control over behaviour that other traditional impulse control studies lack: an explicit parametric manipulation of when and whether action needs to be initiated as a function of the potential future gain (or the opportunity cost of acting inappropriately). By utilising methods that allow for either the recording of dopamine release in the NAc core, or the specific manipulation of the dopaminergic receptor subtypes both in the NAc core and across the brain, in conjunction with a well-controlled behavioural paradigm, this thesis will attempt to bridge the two perspectives of dopamine as a teacher and as a motivator.

2 Methods

Chapter abstract:
Here the general methods used in subsequent experimental chapters are outlined, including a description of the behavioural paradigm, statistical analyses, and techniques used for the recording and manipulation of neurotransmitters in the brain.

1. Behaviour

1.1. The Go/No-Go task

The fundamental purpose of the rodent Go/No-Go task as described here is to be able to manipulate and control whether a particular action should be made or withheld in order to gain different amounts of reward within a single paradigm. This allowed for measurements and manipulation of dopamine in situations with different reward expectations *and* different action requirements using a factorial design.

The task was run in standard operant chambers, with a central nose poke and levers on either side on one wall and the food magazine on the opposite wall (Figure 1). A typical trial proceeds as follows. After voluntarily commencing a trial by poking into a central nose port, one of four auditory cues was presented after a short delay. The cue signalled two pieces of information: first, whether the animal had to make ('*Go*') or withhold ('*No-Go*') an action to receive reward, and second, the size of the reward on offer for doing this correctly.

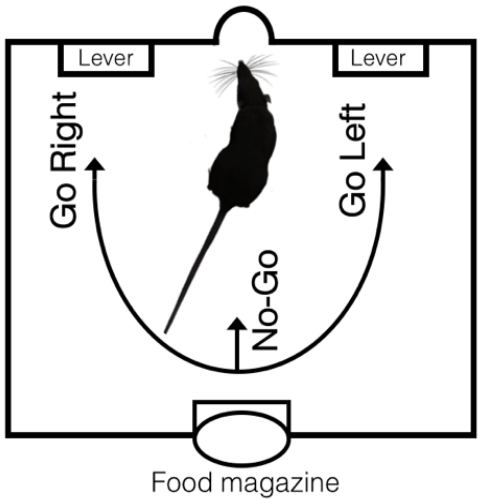


Fig. 1. Illustration of operant chamber layout.

In trials where animals had to perform an action for reward, the action required was a response on a lever situated either to the left or right of the nose poke, depending on the cue presented (Figure 1). The cue sounded for a maximum of 5s, or until animals made a lever response. Conversely, on No-Go trials animals had to withhold

action and remain in the central nose poke for a further 1.5 – 1.7 seconds. Here, the cue sounded for the duration of the required No-Go period, though if animals exited the nose poke prematurely the cue ceased. In both cases, after a 1s delay, reward was delivered for a correct response into a magazine situated on the opposite wall of the operant chamber.

The amount of reward on offer varied, with either *'small'* reward trials in which a single 45mg sucrose pellet was on offer, or *'large'* reward trials, which offered animals two such pellets if they made or withheld an action successfully (Figure 2). Within either Go or No-Go trials, the requirements for success were matched regardless of the reward size on offer.

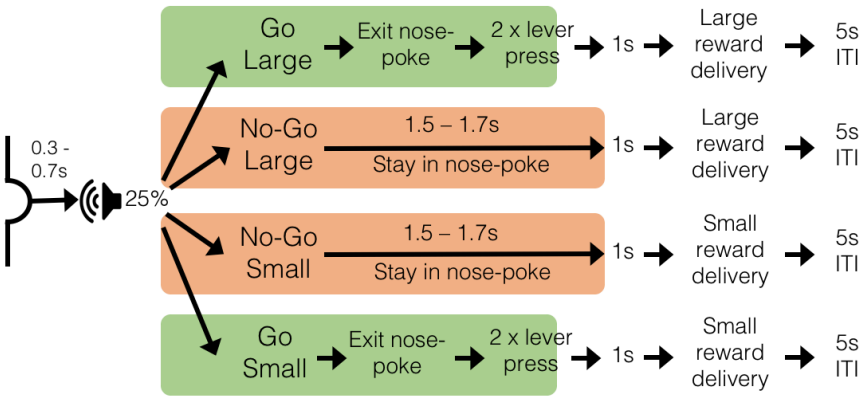


Fig. 2. Schematic of the trial types in the Go/No-Go task.

1.2. Behavioural apparatus

Testing was carried out in three identical operant chambers with an area of 30.5 x 24.1 x 29.2cm and containing stainless steel grid floors (Med Associates VT, USA). They were housed within a custom-built sound-attenuating cabinet ventilated with a fan, which provided a background noise of ~64 dB. Each chamber contained two

retractable levers 9.5cm either side (left and right) of a central nose-poke, which was fitted with an infrared beam signalling when animals entered the receptacle (Figure 1). A food magazine was fitted on the wall opposite into which sucrose pellets were dispensed. Each chamber was also fitted with a house-light and a speaker for delivering auditory stimuli.

1.3. Behavioural training

Animals were initially introduced to the operant chambers before beginning magazine training, during which 45mg sucrose pellets (Test Diet, Sandown Scientific, UK) were delivered every 60s until 21 pellets were delivered. If the animal had consumed at least 18 of the 21 pellets from the food magazine, it then progressed to training on the No-Go trial type.

The nose poke duration required by the No-Go trial type consisted of a pre-cue period and a cue period. The required nose-poke duration was incrementally increased across training sessions, dependent on whether the individual had reached a criterion of ≥ 60% success rate, until animals were able to sustain a randomly jittered pre-cue period of 0.3-0.7s and a maximum cued hold period of 1.5 – 1.7s (or, in one case, 1.7 – 1.9s in one study in Chapter 7, as described there) in order to receive a single sucrose pellet. To distinguish between genuine nose poke exits and small shifts in posture or movement that may have caused the nose poke detector to stop being activated, a short 'buffer' period of a maximum of 0.1s was also introduced. If animals ceased activating the nose poke for more than this buffer period, the trial was signalled as incorrect. Throughout the No-Go portion of the training, the levers were retracted.

Exiting the nose poke before the pre-cue period was complete resulted in an aborted trial, with the only consequence being that the nose poke timer would reset to zero.

Each cue consisted of a single auditory stimulus, which could either be a tone, buzz, white noise, or clicker sound. Cues were counterbalanced across animals. The cue sounded until the end of the hold period (No-Go Correct) or, if the rats exited the nose poke prematurely, at the time of exiting the nose poke (No-Go Fail). After a successful trial there was a 5s inter-trial interval (ITI) during which the nose poke remained inactive. The end of the ITI was not signalled by any external cue. A premature head exit caused the house light to be illuminated for the duration of a 5s time-out immediately as the animal exited the nose poke, after which the house light turned off and a standard 5s ITI commenced. After animals had achieved a success rate of ≥ 60% the second No-Go cue, requiring the same hold time but eliciting a 'large' reward (two sucrose pellets) was introduced.

Once animals had reached criterion for both No-Go trial types, they were then trained on the Go trial type. Animals were first lever trained on a fixed ratio 1 schedule until they had made at least 20 presses on the left and right levers respectively. They were then trained on a simplified version of the full task. During the entire session the two levers were extended, and a nose poke entry sustained for the pre-cue period would again elicit one of two auditory cues, but in this instance the rats had to press either a left lever or a right lever twice depending on the cue. One lever was rewarded with a single pellet ('small' reward), whilst the other was rewarded with two pellets ('large' reward); the side was counterbalanced across animals. The cue sounded either

until the rat pressed the correct lever (Go Correct), the incorrect lever (Go Incorrect), or they failed to press either lever within 60s of cue onset (Go Fail). A 5s ITI followed successful Go trials. Failing to press the correct lever or either lever resulted in the house light illuminating for a 5s time-out period, before the house light turned off and the 5s ITI commenced.

During training, an error-correction procedure was used so that the next trial after an error would always be of the same cue and trial type, and in addition the wrong lever was withdrawn so that the animals could not press it. Once a criterion of ≥ 60% successful Go responses was reached, the cue duration (and therefore maximum reaction time) was lowered to 5s.

At this point interleaved No-Go trials were re-introduced. Each of the trial types occurred with equal probability. All rewards were delivered 1s after successful completion of a trial, and after reward delivery there was a 5s inter-trial interval (ITI) during which animals were unable to initiate a new trial. No cue indicated the end of the ITI and animals were free to initiate the subsequent trial whenever they chose after this time. In the case of an error or failed trial, the house-light would illuminate 1s after the erroneous response and would stay on for a 5s time-out period. After this period, the house-light turned off and the usual ITI commenced. A session ended after the animals had either gained 100 rewards or had worked for 60 minutes, whichever came first.

1.4. Measures of performance

Accuracy measures in the Go/No-Go task are more nuanced than simply describing 'success' or 'failure'. Whilst only one type of error could be made in No-Go trials (leaving the nose poke prematurely, before the end of the cue period), in Go trials animals could either press the incorrect lever ('Wrong Lever' Go) or fail to make a response during the 5s cue time ('Response Omission' Go).

Finally, animals sometimes entered the nose poke but did not remain for the full pre-cue period before leaving, therefore failing to initiate the next trial. These instances were termed 'aborted' trials, and were counted independently after reward was retrieved from the food magazine subsequent to a successful trial and after failure, and also dependent on whether the animal had aborted the trial during or after the ITI period.

1.5. Measures of latency

Whilst success rate and types of errors made are critical measures for understanding the likelihood of action initiation and the selection of specific actions, changes to behavioural latencies in a task can allow for a deeper understanding of the role of dopamine transmission both in terms of motoric and, importantly, motivational changes. The Go/No-Go paradigm used here has a rich set of latencies that can be analysed for this purpose. These are outlined below, and also in the schematics in Figure 3 and Figure 4.

Time in nose poke: nose poke exit times for both successful and failed trials. On both Go and No-Go trials, this is aligned to cue onset. In Go trials, this does not include trials where animals remained in the nose poke > 1.7s.

***Travel time*:** on Go trials, time taken after exiting the nose poke to press a lever.

Lever response latency*:* time taken to complete the required 2 lever presses on successful Go trials.

Magazine latency: time taken from successful trial completion to entry in the food magazine. On Go trials this is aligned to the second lever press, whilst on No-Go trials this is from nose poke exit. On rare occasion the food magazine sensor that indicated when animals had entered the magazine became blocked partway through a session, and so to ignore these erroneous values any latencies that were ≤ 200ms were excluded from analyses.

Re-engagement latency: time taken to first re-enter the nose poke at the end of a trial. For successful trials, this is from leaving the food magazine, and for unsuccessful trials, this is from 1s after making the error, as latencies < 1s occurred after failed No-Go trials where animals immediately re-entered the nose poke and therefore are not representative of animals wanting to re-engage in the task as such. In both cases this does not take into account whether the ITI period (or timeout period, in the case of an unsuccessful trial) has been completed. This metric will be influenced both by

general task engagement but also by timing, as animals cannot initiate the next trial until the ITI period is complete.

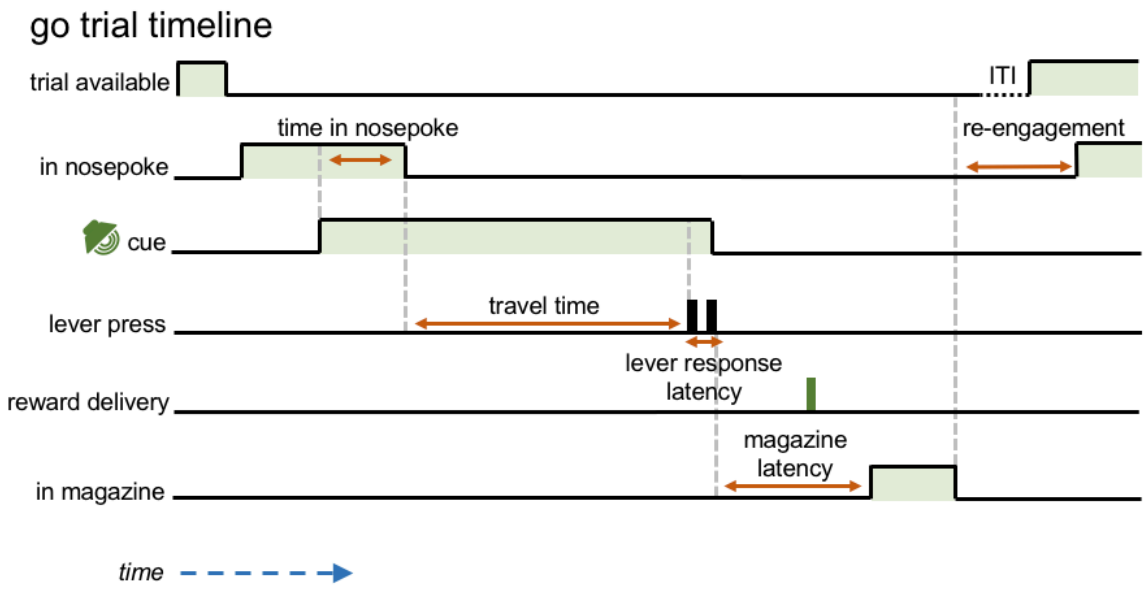


Fig. 3. Schematic of latencies of events that could occur within a Go trial. Lines indicate the passage of time and arrows indicate time differences between important events.

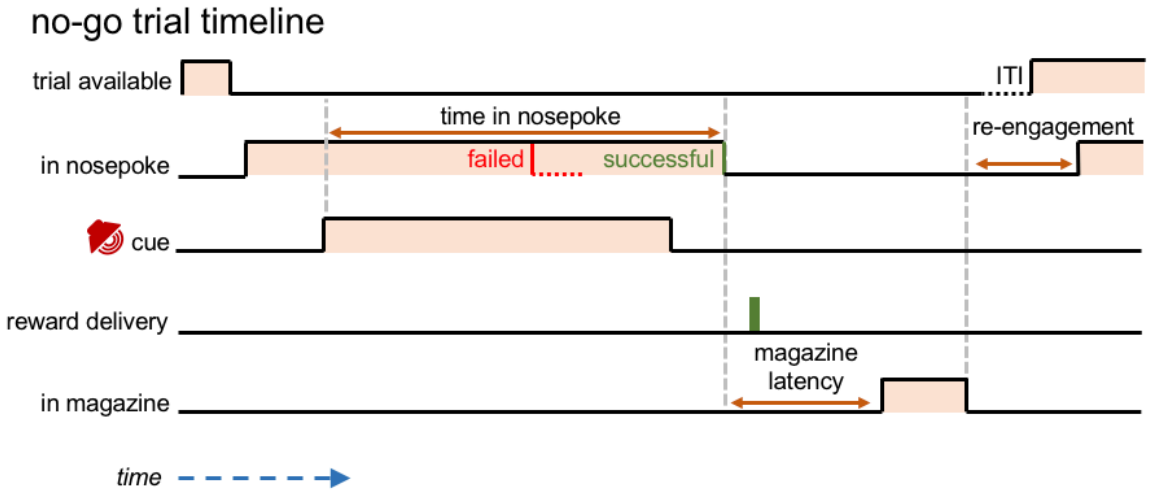


Fig. 4. Schematic of latencies of events that could occur within a No-Go trial.

2. Statistical approaches

2.1. Analysis of mean and median measures

Whenever appropriate, measures were investigated using an analysis of variance (ANOVA) approach. Analysis of means are reported throughout, but if median analyses differed in effect then they too were reported. In brief, ANOVA procedures evaluate mean differences between levels of a manipulation as a proportion of the variance between these levels to give the test statistic of the F-ratio (or Fisher-ratio) (Gravetter & Wallnau, 2009). The F-ratio can be generally conceived of as follows:

$$F = \frac{variance\ between\ sample\ means\ or\ manipulations}{variance\ within\ sample\ means\ or\ manipulations}$$

In practice this involves calculating the deviations of all observations from their associated means – the Sum of Squares – for both between and within treatment effects. Due to the experimental design used throughout most of this thesis, the majority of ANOVAs applied consider within-subject manipulations unless stated otherwise.

Whilst the ANOVA approach is considered robust against violations of the normal distribution assumption (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010), the within-subject equivalent to the assumption of homogeneity of variance – sphericity – is more crucial as this can lead to an increase in the Type I error rate (Haverkamp & Beauducel, 2017). For the assumption of sphericity to be met in the population being sampled from, the variances of the *differences* between levels of a manipulation

need to be equal or similar. This can be assessed using Mauchley's test (Mauchley, 1940). For all within-subjects ANOVAs conducted here, if Mauchley's test was violated, a Huynh-Feldt correction was applied (Huynh & Feldt, 1976). This correction estimates the degree of deviation from sphericity in order to correct the degrees of freedom of the ANOVA's *F* distribution. When a significant interaction was found, post-hoc pairwise comparisons were conducted in order to understand the differences driving the interaction effects.

Although the sample sizes used in the studies here are similar to many similar studies that have been published in the literature, it is still useful to have a measure of the magnitude of an effect that is relatively independent of sample size. One such measure is eta squared, which indicates the proportion of variance accounted for by each of the main effects and interactions (and the error term) in an ANOVA, calculated simply by dividing the Sum of Squares of an effect by the total Sum of Squares (Gravetter & Wallnau, 2009). However, one issue with using eta squared is that the proportion accounted for by each effect depends to some extent on the number of other effects and their significance, and so partial eta squared can be used as an alternative (Cohen, 1973). Partial eta squared is calculated by dividing the Sum of Squares of an effect by itself plus the Sum of Squares for the error term associated with that effect, and therefore its degree is not influenced by the significance of the other terms. Because of this it is the preferred test of effect size for a repeated measure design and is therefore reported throughout.

2.2. *Analysis of latency distributions*

To gain a more nuanced understanding of the behavioural effects of the pharmacological manipulations applied in this thesis that might not become apparent by a simple analysis of variance, the latency measures described in section 1.5 were also examined when possible.

All values in the histograms plotted were calculated as a function of number of total observations for a particular level of a manipulation for a subject, and thus the y-axis is normalised count, or alternatively can also be thought of as the discrete probability of a data point falling in a specific bin. This is averaged over multiple subjects to give the average probability as plotted on the y-axis of the histograms presented.

Indeed, it is *because* probability is averaged across subjects for the histograms presented that a Vincentization procedure was then used for any subsequent latency analyses. Averaged distributions, particularly when the number of observations in a bin are low or variable, can misrepresent the underlying representative distribution of a data set (Whelan, 2008). Vincentization is a method of linear interpolation which overcomes this issue by calculating quantiles of the individual distributions, and then averaging these to produce a group distribution (Ratcliff, 1979).

In more detail, this method involves first ordering each subject's reaction times by magnitude. A number of quantiles (or equally-sized groups) is chosen and each quantile is calculated from these ordered reaction times as follows: if wanting to divide an individual's distribution of n observations into q quantiles, each observation

is first repeated by *q*, and so the first quantile can be calculated by summing the first *n* latencies and dividing this total by *n* (Ratcliff, 1979). The same procedure is applied to the remaining quantiles, and the values of each quantile are then averaged horizontally across participants to give group quantiles. Throughout this thesis 0.1 quantiles are used.

The original proposer of this method, Ratcliff (1979), showed that Vincentized distributions retain a similar mean, variance, and shape to the individuals' distributions whilst 'weighing' the data more fairly across participants, and it is because of this that Vincentizing has become a popular tool for examining full reaction time distributions. However, the approach is not without its critics, and it is recognised that averaging parameter estimates for individuals remains a more accurate and potentially powerful tool (Rouder & Speckman, 2004). Nevertheless, Vincentizing a distribution is advantageous particularly when either the number of subjects or the number of observations per subject are low, hence its application here where the use of animals meant that it was not possible to attain subject numbers more comparable to those in human reaction time studies.

In order to visualise changes to the Vincentized latency distributions, cumulative density functions (CDFs) of these distributions were plotted. A CDF shows the cumulative probability of responding as a function of response speed (Ridderinkhof, 2002). In other words, the ordinate of a CDF plot is simply a measure of cumulative quantile, and the abscissa is response time. In all CDF plots shown throughout, the 0.05 to 0.95 quantiles are presented in 0.1 quantile intervals, smoothed with a 3-point

moving average. Due to the relatively limited number of observations (in comparison to studies of reaction time latencies in humans in which a similar procedure would be used), latency distributions were analysed by comparison of the 0.05, 0.45, and 0.95 quantiles for each dose of a drug and each reward size in a within-subjects ANOVA, and any significant main effects or interactions were also investigated with appropriate Sidak-corrected pairwise comparisons.

3. Fast-scan cyclic voltammetry

3.1. Principles of fast-scan cyclic voltammetry

The experiments described in the first experimental chapter of this thesis employ a method for the measurement of *in vivo* chemical neurotransmission, fast-scan cyclic voltammetry (FCV), whilst animals engage in the behavioural task described in *1.1.* of this chapter.

Our ability to accurately assess levels of chemical neurotransmission has depended on the evolution of suitable techniques. One such method is microdialysis, which involves the direct sampling of analyte concentrations in the extracellular fluid of the region of interest (Chefer, Thompson, Zapata, & Shippenberg, 2010). Whilst both highly sensitive and selective for measuring levels of specific neurotransmitters, microdialysis is limited by its temporal resolution, with samples often collected every five to twenty minutes (though more recent studies have been able to reduce this interval to allow for a sub-minute resolution (Wang, Roman, Schultz, Jennings, & Kennedy, 2008). Additionally, the size of the microdialysis probes used can cause

substantial damage to the tissue in or adjacent to the region of interest, with some probes reaching a diameter of up to 0.4mm.

As an alternative method, electrochemical monitoring allows for a much greater temporal resolution and increased spatial precision. The basic principle behind methods such as amperometry, chronoamperometry, and the approach employed here, FCV, involves the measurement of electrical currents generated by the oxidation-reduction (redox) reaction of molecules that are in contact with a probe. In FCV specifically, a potential sweep is applied to an electrode probe at a rapid scan rate; the sweep involves applying a triangular waveform that ramps up to an oxidising potential and then back down to a reduction potential (Figure 5a; Fox & Wightman, 2017). The application of this voltage waveform causes the redox reaction – in other words, loss (oxidation) and subsequent gain (reduction) of electrons – of any electroactive species present on the surface of the electrode at the time. Considering that electric current at its simplest conception is the movement of electrons, in all cases this redox reaction results in a measurable change of current that is directly proportional to the concentration of electroactive molecules on the surface of the electrode at the time the scan is applied (Bunner & Rebec, 2016).

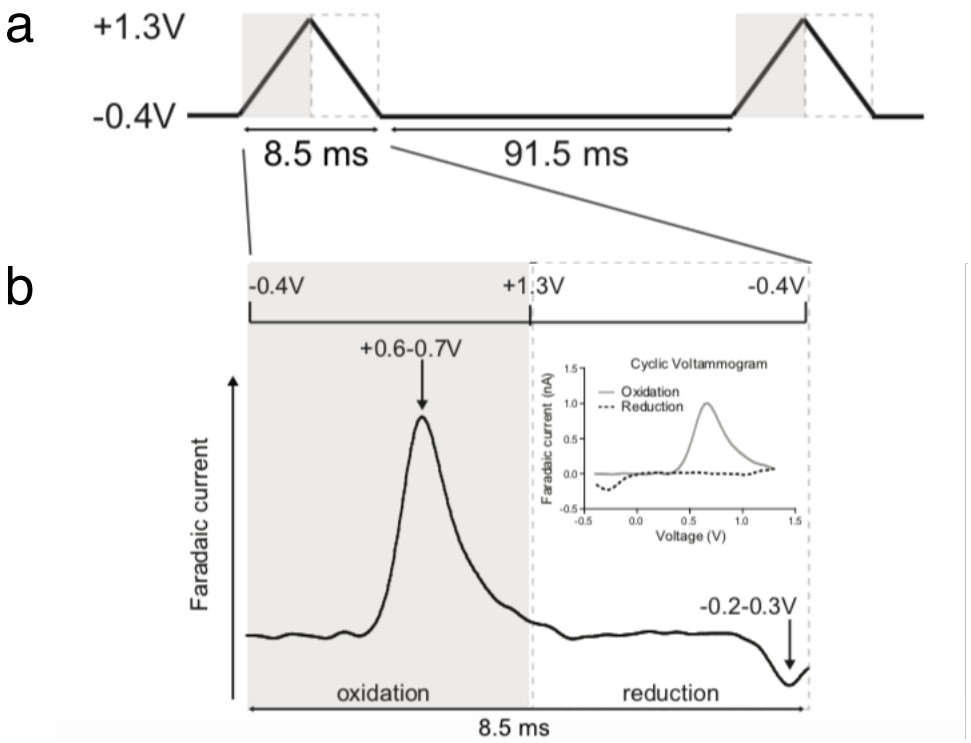


Fig. 5. Schematic of key FCV principles adapted from Arnold, Burgeno, & Phillips (2015). a) Triangular waveform applied in FCV when attempting to measure levels of dopamine, ramping from -0.4V to +1.3V and back down across the course of 8.5ms, every 100ms. b) Changes in current across this 8.5ms sweep period due to the oxidation and reduction of dopamine and dopamine-o-quinone respectively. Peak oxidation and reduction currents are indicated with arrows. b inset) Example of a dopaminergic cyclic voltammogram, showing current vs. voltage.

Modern implementations of FCV involve the insertion of a reference electrode and a working electrode (Rodeberg, Sandberg, Johnson, Phillips, & Wightman, 2017). The silver/silver chloride (Ag/AgCl) reference electrode allows maintenance of an absolute potential at the working electrode, and is usually implanted in a region away from the recording site – here, over the posterior sensorimotor cortex. The working electrode is a carbon-fibre microelectrode that is implanted into the region of interest. For all of the experiments in this thesis, these electrodes consist of a carbon fibre that has been threaded through fused-silica tubing for chronic implantation

(Clark et al., 2010). The exposed end of the carbon fibre, typically 150 – 200 microns in length, is the surface upon which electroactive molecules can accumulate (Rodeberg et al., 2017). Therefore, in addition to its excellent temporal resolution, the spatial resolution of FCV is also such that specific regions or areas of interest can be precisely targeted. Additionally, the small size of the working electrode minimises the extent of tissue damage at the recording site.

However, because of its reliance on the occurrence of a redox reaction, one hurdle that experimenters using voltammetry have to overcome is that of identifying the substance being detected (Arnold, Burgeno, & Phillips, 2015). There are many substances in the brain that can undergo oxidation, including other neurotransmitters outside of the molecule of interest. FCV overcomes this issue partly by applying the aforementioned triangular waveform which produces a cyclic voltammogram – current against voltage potential (Figure 5b, inset). As the exact potential at which a molecule will undergo oxidation will vary dependent on its molecular structure, the cyclic voltammogram can act as a 'fingerprint' of a specific molecule, aiding its identification. In the case of dopamine, the potential applied to the electrode typically ramps from -0.4V (vs Ag/AgCl) to a maximum voltage of +1.3V, and then back down to the holding potential, and the scan can be completed in 8.5ms before the potential is held until the next voltage scan (Arnold et al., 2015). Scans are often applied every 100ms, yielding the excellent temporal resolution that FCV is known for. Dopamine undergoes a peak oxidation reaction at around +0.6V to +0.7V and dopamine-o-quinone undergoes peak reduction at -0.2 to -0.3V, so maximal current changes at these points of the voltage scan are observed in the presence of

dopamine (Figure 5b). As the scans are applied rapidly – ten scans every second – colour plots can be used to show the changes in current at each point in the waveform over time (Figure 6; Michael, Travis, & Wightman, 1998). Importantly, the vast majority of the actual recorded signal is in fact capacitance current – to be able to identify behaviourally relevant changes in current (and thus dopamine level) due to a behavioural event or cue, current at the time of the event is subtracted from a baseline period. This means that the measurements taken are reflective of *relative*, rather than absolute changes in dopamine level.

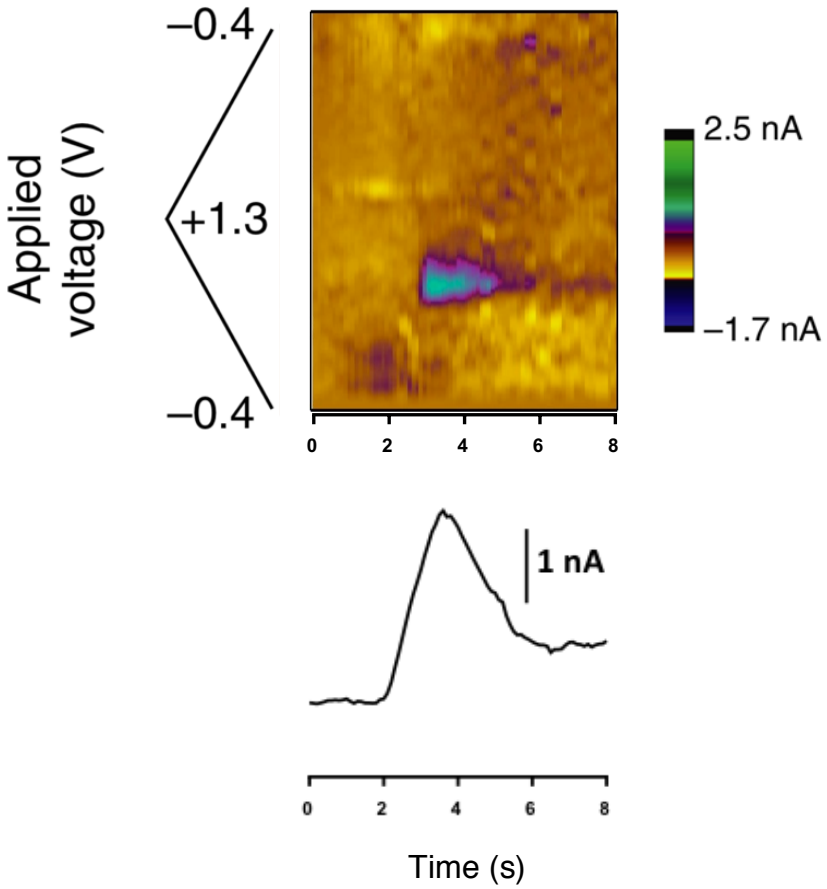


Fig. 6. Example FCV colourplot, showing time against applied voltage (V), with changes in current due to applied voltage indicated in colour – here in response to unsignalled reward at 2s. The colourplots are comprised of hundreds of consecutive voltammetric scans to show changes in levels of dopamine – change in current – over time, as shown extracted in the second plot. The subplot illustrates the cyclic voltammogram for this trace. Adapted from Panagi (unpublished).

Because of its temporal and spatial precision and its high chemical sensitivity, FCV enables the correlation of levels of dopamine fluctuation with on-line behaviour, and has therefore allowed for a greater understanding of the role of dopamine in a host of behavioural situations, including its role in learning (Wassum et al., 2012) and decision-making (Gan, Walton, & Phillips, 2010).

3.2. Surgical procedure

Once they had achieved stable baseline performance on the Go/No-Go task, rats were anaesthetised using inhaled isoflurane (4% vol/vol in O_2 induction and 1.5% for maintenance, delivered via facemask) and given buprenorphine (Vetergesic, 0.03 mg/kg, subcutaneous) and meloxicam (Metacam, 2 mg/kg, subcutaneous) at the start of surgery. To combat dehydration, animals were also given 3ml of glucosaline (0.5% in 0.9% saline, Aquapharm), divided across two subcutaneous injection sites. Body temperature was maintained at 37 ± 0.5 °C with the use of a homeothermic heating blanket.

Once animals were secured in a stereotaxic frame (Kopf Instruments) and their scalp shaved and cleaned with dilute hibiscrub and 70% alcohol, a local anaesthetic (bupivacaine, 2 mg/kg) was administered to the site of future incision. Eye gel (Lacri-Lube, Allergan) was applied to the eyes for protection. The skull was then exposed and bregma and lambda taken. The anterior/posterior tilt of the head was adjusted until the dorsal/ventral difference between bregma and lambda was 0.1mm or less.

Craniotomies were then drilled for the Ag/AgCl reference electrode (AP: -3.7, ML: -1.4), 4 anchoring screws (Precision Technology Supplies) and a voltammetric recording electrode in each hemisphere. After the screws were secured and the reference electrode inserted, custom-made carbon fibre microelectrodes were then lowered into the NAc core, at co-ordinates of +1.4mm anterior/posterior, ±1.3mm medial/lateral, and -7.0mm dorsal/ventral from the surface of the skull relative to bregma (Franklin & Paxinos, 2007). The carbon fibre microelectrodes and reference electrode were attached to a 6-pin headstage connector, which was secured in place along with a headpost with dental cement. Following surgery, animals were again administered buprenorphine, as well as another 3ml of glucosaline. Animals were also given palatable food. Meloxicam was given up to a further three days post-surgery dependent on the individual animal's recovery trajectory. Animals had on average 2 weeks of post-surgery recovery with food and water *ad libitum,* before food restriction and further behavioural training.

3.3. Fast-scan cyclic voltammetry recordings

Voltammetric scans were performed at a frequency of 10Hz throughout the session. Prior to a scan, the carbon fibre was held at a potential of -0.4V (versus Ag/AgCl) and then, during the scan, was ramped up to +1.3V and then back down to -0.4V at 400V/s, and held at -0.4V between scans. As described in section 3.1., the application of this waveform induces redox reactions in electrochemically active species, such as dopamine, at the surface of the carbon fibre, which are then recorded as changes in current over time. The average current from the scans obtained within the 0.5s before cue presentation was subtracted from the current across the scans within a trial to

give background-subtracted signals (P. E. M. Phillips et al., 2003; Roitman et al., 2004). As in Gan et al. (2010), to ensure that recording electrodes were reliably detecting behaviourally-evoked dopamine, the response to an unexpected sucrose pellet delivered to the magazine was measured at the start and end of each recording session. Delivery of an unexpected pellet has been shown to increase phasic dopamine release in the nucleus accumbens (Clark et al., 2010). A criterion for the inclusion of recordings from a session was that electrochemically verifiable dopamine release from the delivery of the unexpected pellets could be detected. From this, an extracted cyclic voltammogram was linearly regressed against a dopamine standard, with $R^2 \geq 0.75$ set as the criterion based on the discriminability of dopamine from other common neurochemicals in a flow cell (Heien, Phillips, Stuber, Seipel, & Wightman, 2003). Only sessions were sufficient discriminability was confirmed were included in subsequent analysis.

3.4. Voltammetric data analysis

Analysis of voltammetric data was initially carried out using software written in LabVIEW (National Instruments). Data were low-pass filtered at 2 kHz. As has been described previously (Flagel et al., 2011; Wanat, Kuhnen, & Phillips, 2010), principal component analysis using a standard training set of stimulated dopamine release detected by chronically implanted electrodes was used to isolate changes in dopamine concentration from other electrochemical signals as the first principal component among other unrelated electrochemical fluctuations, such as changes in pH (Heien, Johnson, & Wightman, 2004). Trials where the PCA failed to successfully extract dopamine current on > 50% of data points were excluded. Once dopamine-

related current changes were extracted, all further analysis was undertaken using MATLAB (Mathworks). Unless otherwise stated, all data were smoothed using a 0.5s moving window.

The discriminability of dopamine signals was analysed in each individual animal at each time point in a 5s period after an event of interest (in other words, time-locked to cue onset, head exit, or first lever press) using the auROC, an approach from signal detection theory (Green & Swets, 1966). In the first FCV experiment conducted, the Go trial type or successful trials were considered as positive cases. In the second FCV experiment, which introduced different reward sizes amongst other modifications, the large reward trial types were always considered as positive cases, except when comparing Go Large and No-Go Large trials, where the Go trial type was positive.

To quantify which factors affected dopamine levels, regression coefficients were estimated for each animal at each time point in a 6s window spanning from 1s previous to an event of interest to 5s following the event – here termed as a trial. A linear model was used with a constant term, representing an ordinary least-squares fit of the given regressors to the data over trials. The specific regressors used are outlined in Chapter 3. All regressors, whether continuous or categorical, were mean-centred. Regression coefficients in each animal were averaged and then transformed into absolute values.

3.5. Histology

After recordings were completed, animals were deeply anaesthetized with sodium pentobarbitone (200mg/kg of body weight, i.p.). Microlesions were then made at the electrode locations via current stimulation to aid subsequent histological localisation. The animals were then transcardially perfused with 0.9% saline followed by a 10% formalin solution (vol/vol). The brains were kept in 10% formalin solution until being sectioned. Brains were cut into 50 μm-thick coronal sections using a vibrating-blade microtome (Leica). Sections were then mounted on glass slides and stained with cresyl violet (Sigma Aldrich) to confirm electrode locations, before being mounted in DePeX mounting medium onto 1.5% gelatin-coated slides and enclosed with coverslips.

4. Systemic pharmacology

4.1. Drug solutions

The specific drug compounds and doses used are described in the relevant chapters, but in all cases bar one the drugs were dissolved in 0.9% sterile saline. The exception was SB 242084, a serotonergic 5-HT_{2C} ligand; due to its low solubility, this compound was instead prepared in 0.9% sterile saline containing 8% 2-hydroxypropyl-ß-cyclodextrin and 25mM citric acid. The use of citric acid necessitated adjustment of the pH of this vehicle solution, which was carried out by adding 100 μL volumes of 5M sodium hydroxide dissolved in de-ionised water. All drugs were made in batch, aliquoted, and frozen at -20°C. Individual aliquots were defrosted for use on testing days.

4.2. Drug application

All drugs were injected intraperitoneally in a volume of 1 ml/kg bodyweight. The side of the injection site was counterbalanced within animals across injections to reduce damage to the area. The amount of time between injection and behavioural testing varied between drugs as detailed in the relevant chapters, but was usually either 10 or 20 minutes. Each drug was administered using a counterbalanced Latin square design in order to reduce the risk of a systematic carry-over effect. Test sessions were separated by at least one treatment-free training day to ensure a return to baseline performance and complete washout of the drug, with criteria of ≥ 60% successful trials across all trial types, and session completion within 60 minutes. If these criteria were not met, animals continued with treatment-free training days until performance was satisfactory, at which point the next testing day commenced. In practice, however, this occurred very rarely.

5. Local pharmacology

5.1. Surgical procedure

As with the FCV surgeries, once animals had achieved stable baseline performance on the Go/No-Go task, they underwent surgery. The procedure replicated that of the FCV surgeries in terms of anaesthesia, analgesia, placement in the stereotaxic frame, and skull preparation. However, from the point of craniotomies onwards the procedure differed; in this case, six craniotomies were made: two for the implantation of bilateral guide cannulae (Plastics One, UK) and four for anchoring screws

(Precision Technology Supplies) to ensure the headcap remained securely attached to the skull post-surgery. After insertion of the screws, the guide cannulae were lowered. The cannulae consisted of an 8mm plastic pedestal which held together two 26-gauge metal tubes with a centre-to-centre distance of 3.4mm and a length of 7.5mm. They were implanted 1.5mm above the target site of the nucleus accumbens core, at co-ordinates of AP: +1.4mm, ML: ±1.7mm, DV: -6.0mm from the surface of the skull relative to bregma (Franklin & Paxinos, 2007). A slightly more lateral co-ordinate was used here in comparison to the FCV recordings to try to restrict the action of each agent to the NAc core and avoid the adjacent NAc medial shell. Dental acrylic (Associated Dental Products Ltd.) was then applied to secure the cannulae to the skull and screws. Any loose skin around the headcap was brought together with stitches, both at the front and the back of the incision. At the end of surgery, bilateral dummy cannulae of the same length as the guide cannulae were inserted to ensure patency, and a dustcap was secured onto the pedestal to secure the dummy and avoid accumulation of debris.

Following surgery, as with the FCV surgeries, animals were again administered buprenorphine and meloxicam, as well as another 3ml of glucosaline. Meloxicam was given up to a further three days post-surgery dependent on the individual animal's recovery trajectory.

5.2. Infusion procedure

From two days post-surgery animals were first habituated to manipulation of their implants by being lightly restrained and having the dummy cannulae removed and a

fresh set inserted. Re-training of the animals commenced on average 2 weeks after surgery until they had returned to ≥ 60% performance on each of the trial types.

The day before the first drug infusion session, a mock infusion was carried out. This involved insertion of the full-length injectors, which had been back-filled with saline, but no infusion of any substance. This was performed to avoid potential tissue damage-related confounds occurring in the first experimental session, and also to habituate animals to the infusion procedure.

The following day, animals underwent their first experimental infusion session. Two 10 μl glass Hamilton syringes with cemented needles were back-filled with 0.9% sterile saline and placed in a Cole-Parmer infusion pump. Double connector assembly tubing, which had previously been cleaned by flushing through with ethanol and air, was filled with saline before being attached to the Hamilton syringes. 33-gauge bilateral injectors were then attached to the double connector assembly. The same tubing and injectors were always used for different drug compounds and doses to avoid cross-contamination. The injectors had also been cleaned by being sonicated for 60 minutes in aliquots filled with 70% ethanol, and this cleaning procedure was repeated between sessions to reduce the risk of infection.

The Hamilton pistons were then pushed and a small air bubble pulled before the drug itself was pulled into the injectors. After ensuring that the injectors were not blocked, the rats were gently restrained, their dummy cannulae removed, and the injectors inserted into the nucleus accumbens core. The length of the injectors was 9mm,

meaning that once inserted they protruded past the guide cannulae by 1.5mm in order to reach the target location with a dorsal/ventral of -7.5mm. During infusion, 0.5 µl of either vehicle or drug solution was injected per hemisphere at a rate of 0.25 µl/minute. The injectors were subsequently left in place for a further two minutes after cessation of the infusion to allow diffusion of solution from the injectors. The injectors were then removed and it was ensured that they had not become blocked during the infusion. Any blockages were noted and the session repeated if necessary (though in practice, this only occurred twice across all sessions). The dummy cannulae and dustcap were then replaced, and the rats were returned to their homecage for 10 minutes to reduce possible effects of injection stress on performance. Testing then commenced. No animal underwent more than 12 experimental sessions due to the increased risk of gliosis with repeated infusions.

5.3. Histology

At the end of data collection, animals were deeply anaesthetised with sodium pentobarbitone (200mg/kg, i.p. injection). They were then transcardially perfused with 0.9% saline followed by a 10% formalin solution (vol/vol). The brains were kept in 10% formalin solution until being sectioned. Brains were sectioned into 60 µm-thick coronal sections using a vibratome (Leica). The sections were then stained with cresyl violet (Sigma Aldrich) before being mounted in DePeX mounting medium onto 1.5% gelatin-coated slides and enclosed with coverslips.

The results of the histology reveal that in the vast majority of cases, the injectors successfully targeted the NAc core (Figure 8, adapted from Paxinos & Watson, 2009),

often with a ventromedial placement within the core. Although cannulae placement was sometimes more anterior than intended, animals were not excluded as long as there was evidence that the injectors reached the NAc core. The anterior commissure was also targeted on some occasions. However, in the absence of evidence to the contrary, we have assumed that this did not preclude diffusion of the drug into the target area and thus these animals were not excluded from analysis.

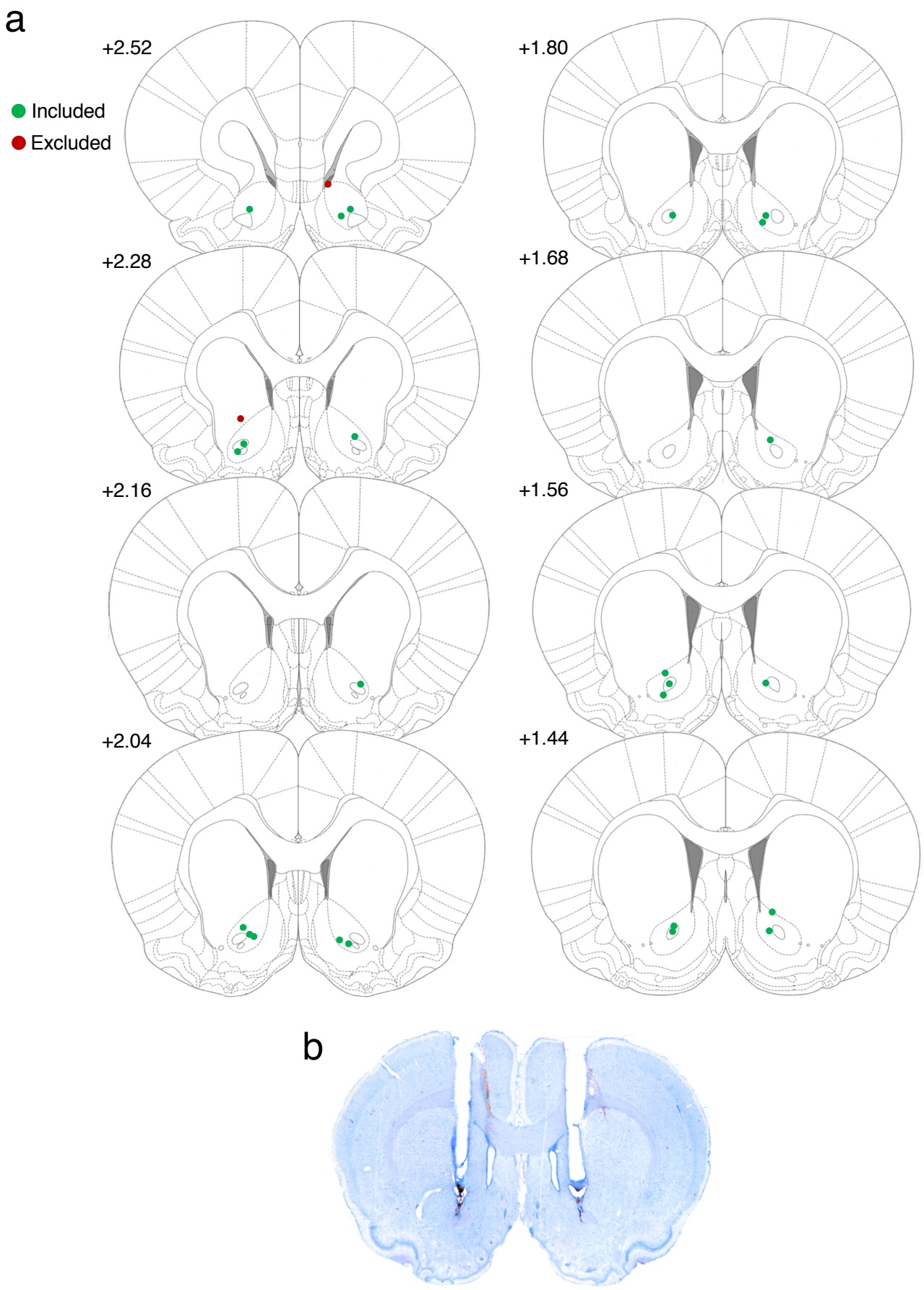


Fig. 8. Injector placement. (a) Locations of injector lesions from animals included (green) and animals excluded from analysis (red). Numbers to the left of coronal sections indicate distance anterior to bregma (mm). Adapted from Paxinos & Watson (2009). (b) Example of bilateral injector lesion and guide cannulae placement.

3 Action initiation shapes mesolimbic dopamine encoding of future rewards

Chapter abstract:
It is broadly accepted that the activity of midbrain dopamine neurons and consequent dopamine release in the nucleus accumbens core signals a reward prediction error: the discrepancy between expected and received reward. However, there is also evidence that such dopamine at both slow and fast timescales also plays an important role in driving behaviour for reward, and the relationship between these two functions is not well understood. In this chapter, by measuring subsecond dopamine release as animals either make or withhold action for either a small or large reward, it is revealed that dopamine release in this region is contingent both on reward prediction and correct action initiation.

The figures and data from this chapter are published in:

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*, 34 – 36.

For the data in this chapter, P.B., P.J.M., and M.E.W. conceived and planned the study, E. C. J. S. and L. L. G. performed surgery and collected the data, E. C. J. S., L. L. G. and M. E. W. conducted analyses.

1. Introduction

Dopamine has historically been linked to concepts of reward and reinforcement (Olds & Milner, 1954; Wise, 1978). More recently, subsecond changes in the activity of midbrain dopamine neurons has been hypothesised to encode the discrepancy between expected and received reward: a reward prediction error signal (Schultz et al., 1997). Since the proposal of midbrain dopamine neurons acting as an analogue for reward prediction error was first put forward, it has been confirmed using behavioural paradigms such as blocking (Waelti et al., 2001) and conditioned inhibition (Tobler, Dickinson, & Schultz, 2003), across rewards with different properties or dimensions (Lak et al., 2014), and across species (Cohen, Haesler, Vong, Lowell, & Uchida, 2012; Takahashi et al., 2016; O'Doherty, Dolan, & Schultz, 2006). This prediction error signal has also been identified in the downstream release of dopamine in the nucleus accumbens (NAc) core (Day et al., 2007; Hart et al., 2014). Thus the evidence in support of reward prediction error encoding by midbrain dopamine neurons and downstream striatal dopamine release is substantial, and correspondingly this theory has been highly influential and is well-accepted by most in the field.

Yet prediction error encoding may not adequately capture all aspects of the dopamine function. Dopamine is also known to be integral to movement; indeed, the motor disorder of Parkinson's disease is characterised by dopaminergic cell death and dopamine depletion in the striatum (as reviewed in Weisenhorn, Giesert, and Wurst, 2016). Whilst the physiology of Parkinson's Disease has predominantly been ascribed

to the nigrostriatal dopaminergic pathway, it has also been shown that elevating extracellular dopamine levels in the ventral striatum by repeated amphetamine administration potentiates movement (Robinson, Jurson, Bennett, & Bentgen, 1988). Specifically, pharmacologically enhanced levels of mesolimbic dopamine has been shown to promote *goal-directed* movement, in particular cue-driven instrumental responding for reward (Nicola et al., 2005; Robbins, Cador, Taylor, & Everitt, 1989).

Methods that allow for a subsecond readout of dopamine fluctuations in the NAc core have also intimated that reward-seeking behaviour might be an important component of this signal. Phasic dopamine release has been shown to increase coincident with the initiation of drug-seeking behaviours – lever pressing for cocaine – and electrical evocation of dopamine release at a similar timescale was able to recreate this behaviour, suggesting a causal relationship between this efflux and goal-directed action for drug reward (Phillips et al., 2003). Additionally, mesolimbic dopamine release peaks at the time of cued lever pressing for sucrose, even when the behavioural response to the cue is substantially delayed (Roitman et al., 2004).

Therefore, there is evidence that dopamine release even at shorter timescales may be shaped by action for reward, but the exact relationship between cue-evoked dopamine, movement, and reward processing has not yet been systematically examined. In a first experiment reported in Syed et al. (2016), we used fast-scan cyclic voltammetry (FCV) to measure dopamine release in the core of the NAc as well-trained animals performed a cued Go/No-Go task. Animals learnt cues that instructed them to either carry out action in 'Go' trials by lever pressing, or to

withhold all action in 'No-Go' trials by remaining in a nose poke for several seconds. Correct responses in both cases resulted in delivery of a single sucrose pellet. It was found that, whilst dopamine release sharply rose with presentation of a Go cue and correct action initiation, it was significantly attenuated throughout the action restraint period in successful No-Go trials, only rising once animals initiated action by exiting the nose poke to retrieve reward. However, mesolimbic dopamine release was nonetheless *also* mediated by reward expectation (calculated based on animals' performance levels), resulting in an interaction between this and the action requirements of a trial in shaping dopamine levels.

Whilst this first study progressed our understanding of the influence of goal-directed, cued movement on the subsecond dopamine signal, it has several limitations that we address in the experiment described here. The first is that reward expectation was not externally manipulated, but instead inherently varied as a function of average success rate of a trial type. The second is that a 3-cue design was used, with one cue indicating a No-Go trial and two cues (Go Left and Go Right) indicating Go trials. Therefore, No-Go trials were less likely to occur than Go trials and so may have been more surprising, and in conjunction with a difference in average success rates across these trial types, may have influenced the dopamine signal beyond the relevant independent variables of cue value and movement requirement.

1.1. Aims, approaches, and hypotheses

The aim of this study is to parse out the separable contributions of cue-driven reward prediction and action initiation for reward to the phasic mesolimbic dopamine signal.

To better understand the relationship between reward expectation and action initiation in the dopamine signal, in the experiment described in this chapter we again recorded mesolimbic dopamine release using FCV, but this time implementing a version of the Go/No-Go paradigm in which the cues now instruct both the action requirement *and* the size of the prospective reward – either one or two sucrose pellets – available.

We hypothesised that, contrary to a pure reward prediction signal and as revealed in the initial Go/No-Go experiment, mesolimbic dopamine release would increase in Go trials but be attenuated in No-Go trials. However, we also hypothesised that reward size on offer would also mediate dopamine levels such that there would be greater efflux in response to a cue predicting larger reward.

2. Methods

2.1. Animals

All procedures were carried out in accordance with the UK Animals (Scientific Procedures) Act (1986). A total of 12 naïve male Sprague-Dawley rats were used (Harlan). One rat was culled due to post-surgical complication, and out of the 22 remaining electrodes (2 per animal), 6 were broken/noisy, and 7 were misplaced. This left a total of 6 rats with 9 working electrodes that provided the data described here.

All animals were maintained on a twelve-hour light/dark cycle (lights on at 07:00). Rats were group-housed throughout initial training but were individually housed

following surgery. All testing was carried out during the light phase, and during training and testing periods, animals were food restricted such that their weights were kept between 85 – 90% of their free-feeding weight. Water was available *ad libitum* whilst animals were in their home cages.

Full details of the apparatus and behavioural training, as well as further details of the task, are described in Chapters 2 and 3. All recordings were carried out in well-trained rats who had achieved stable levels of performance.

2.2. Fast-scan cyclic voltammetry

The principles of FCV, the surgical procedure used, voltammetric data analysis, and histology are all described in Chapter 2. To reiterate briefly, carbon fibre microelectrodes were implanted bilaterally targeting the NAc core and, during experimental sessions, 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.

2.3. Voltammetric data analysis

Inclusion of data from a session was dependent on linear regression of an extracted cyclic voltammogram after delivery of an unsignalled reward against a dopamine standard, with $R^2 \geq 0.75$ set as the criterion based on the discriminability of dopamine from other common neurochemicals in a flow cell. If this criterion was passed, then PCA analysis using a standard training set or ground truth of stimulated dopamine release, with dopamine as the first principal component, was used to isolate changes in dopamine concentration from other electrochemical signals, such as changes in

pH. Trials where the PCA failed to successfully extract dopamine current from this training set on >50% of timepoints were excluded.

In order to quantify the degree of discriminability of the dopamine signals recorded in pairs of trial types in these studies, an auROC (area under the Receiver Operating Characteristic) approach was used (Green & Swets, 1966). After calculation of the auROC from each animal, this was averaged and significant discriminability at each time point was computed using 1,000 random permutations of the trial types and re-calculating the auROC to generate a null distribution. Permutation tests were considered significant at any time point when $p < .05$, corrected for multiple comparisons.

To quantify which factors affected dopamine levels, regression coefficients were estimated for each animal at each time point in a 6s window spanning from 1s previous to an event of interest to 5s following the event – here termed as a trial. A linear model was used with a constant term, representing an ordinary least-squares fit of the given regressors to the data over trials. The specific regressors used in this Experiment for analysis of cue-evoked dopamine on correct and valid Go and No-Go trials were: (1) action initiation (trial-by-trial time in nose poke between cue and head exit) and (2) reward prediction error (calculated for each given trial type as: [expected value – average(expected value)], where expected value = trial type success rate*reward size. The values used for (2) are also detailed in Table 2 in section 3.4. of the results section. All regressors, whether continuous or categorical, were mean-

centred. Regression coefficients in each animal were averaged and then transformed into absolute values.

2.4. *Measures of performance and latency*

Animals' performance on the Go/No-Go paradigm was analysed to understand their behaviour in the task. The measures of performance and latency measures are outlined in detail in sections 1.4 and 1.5 of Chapter 2, but those discussed in the current chapter include Go or No-Go trial success rate (Table 1), and time spent in the nose poke from cue onset in correct or failed Go and No-Go trials.

ERROR TYPES		
Go error types	Wrong lever 	Response omission 
No-Go error types	Premature head exit 	

Table 1. Error types in Go and No-Go trials.

2.5. *Behavioural data analysis*

These aforementioned measures were analysed using both parametric (within-subjects ANOVAs or paired samples T-Tests as appropriate, followed by pairwise comparisons – Sidak-corrected) and non-parametric (chi-squared test and follow-up Wilcoxon signed ranks tests) approaches. As some animals contributed FCV data

from two different sessions, behaviour from two recording sessions from all 6 animals were analysed to understand behaviour in the task.

3. Results

3.1. Animals were able to successfully make and withhold action in the Go/No-Go task

To understand how animals behaved in the Go/No-Go task, both their success rates and latencies to execute behaviour were analysed. Success rate was found to be highest in Go Large trials (Figure 1a; significant reward*action interaction: $F_{1,11} = 7.670$, $\eta_p^2 = .411$, $p = .018$; pairwise comparisons: small reward trials, Go vs. No-Go, $p = .990$; large reward trials, Go vs. No-Go, $p < .001$; Go trials, small vs. large reward, $p = .006$; non-parametric chi-squared test, [$\chi^2(3) = 13.62$, $p = .004$]; post-hoc Wilcoxon signed ranks test of Go Large against other trial types, [$W_6 = 0$, $p = .03$]). However, success rate of the other three trial types – Go Small, No-Go Large, and No-Go Small – was comparable.

By definition, animals were faster to initiate action in Go than No-Go trials, but were also quicker to leave the nose poke when the reward on offer was large compared to small in both Go and No-Go trials (Figure 1b; significant reward*action interaction: $F_{1,11} = 5.687$, $\eta_p^2 = .341$, $p = .036$; pairwise comparisons, Go trials: small vs. large reward, $p < .001$; No-Go trials: small vs. large reward, $p = .001$; non-parametric Wilcoxon signed ranks test, both [$W_6 = 0$, $p = .03$]). There was a tendency for animals to leave the nose poke earlier in failed small reward No-Go trials than when a large reward

was on offer, though this did not reach significance [Figure 1c; paired samples T-Test: t_{11} = -1.813, p = .097]. This suggests that animals had learnt both the cue-action and cue-reward size contingencies.

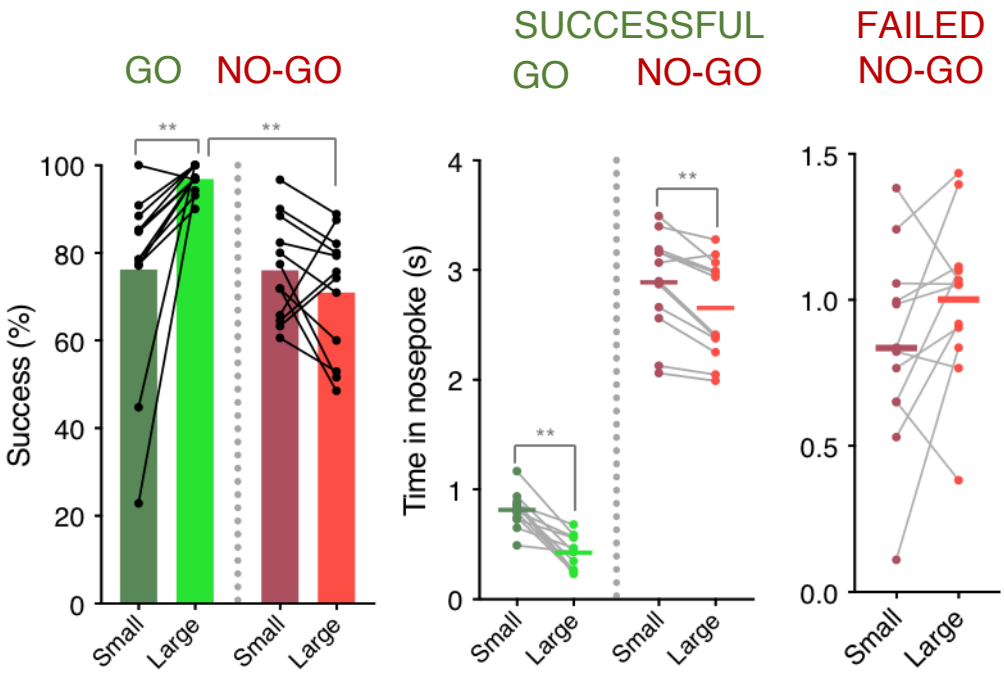


Fig. 1. Behavioural performance in the Go/No-Go task. (a) Mean success rate for Go (left of dotted line) and No-Go (right of dotted line) trials split by reward size on offer. Bars indicate overall mean, dots indicate within-subject means. (b) Time in nose poke from cue for successful Go (left of dotted line) and No-Go (right of dotted line) trials split by reward size on offer. (c) Time in nose poke from cue for failed No-Go trials split by reward size on offer. Significance bars indicate pairwise comparisons, ** $p < .01$.

3.2. NAc core dopamine release reflects reward prediction error encoding

After presentation of the Go Large cue, dopamine release rapidly increased and was sustained as animals performed the action required to garner reward (Figure 2a and 2b). However, presentation of the Go Small cue in fact led to a transient *dip* in dopamine levels that was highly distinguishable from the Go Large transients. This is consistent with reward prediction error encoding; as animals experienced an equal

number of small and large reward trials, their long-running expectation of reward would have been the average of these. Thus a cue indicating that the upcoming trial presented the opportunity of gaining a large reward would elicit a positive reward prediction error. Likewise, a cue for a small reward trial would elicit a negative prediction error for the same reason.

As with the Go trials, phasic dopamine release transiently increased immediately following a No-Go Large cue, and transiently decreased following a No-Go Small cue (Figure 2a and 2c). However, the degree of this change was many times smaller than in Go trials, and as animals went on to withhold action in No-Go trials, dopamine release remained significantly attenuated. This is particularly apparent when directly comparing Go and No-Go trials offering the same reward size (Figure 2a).

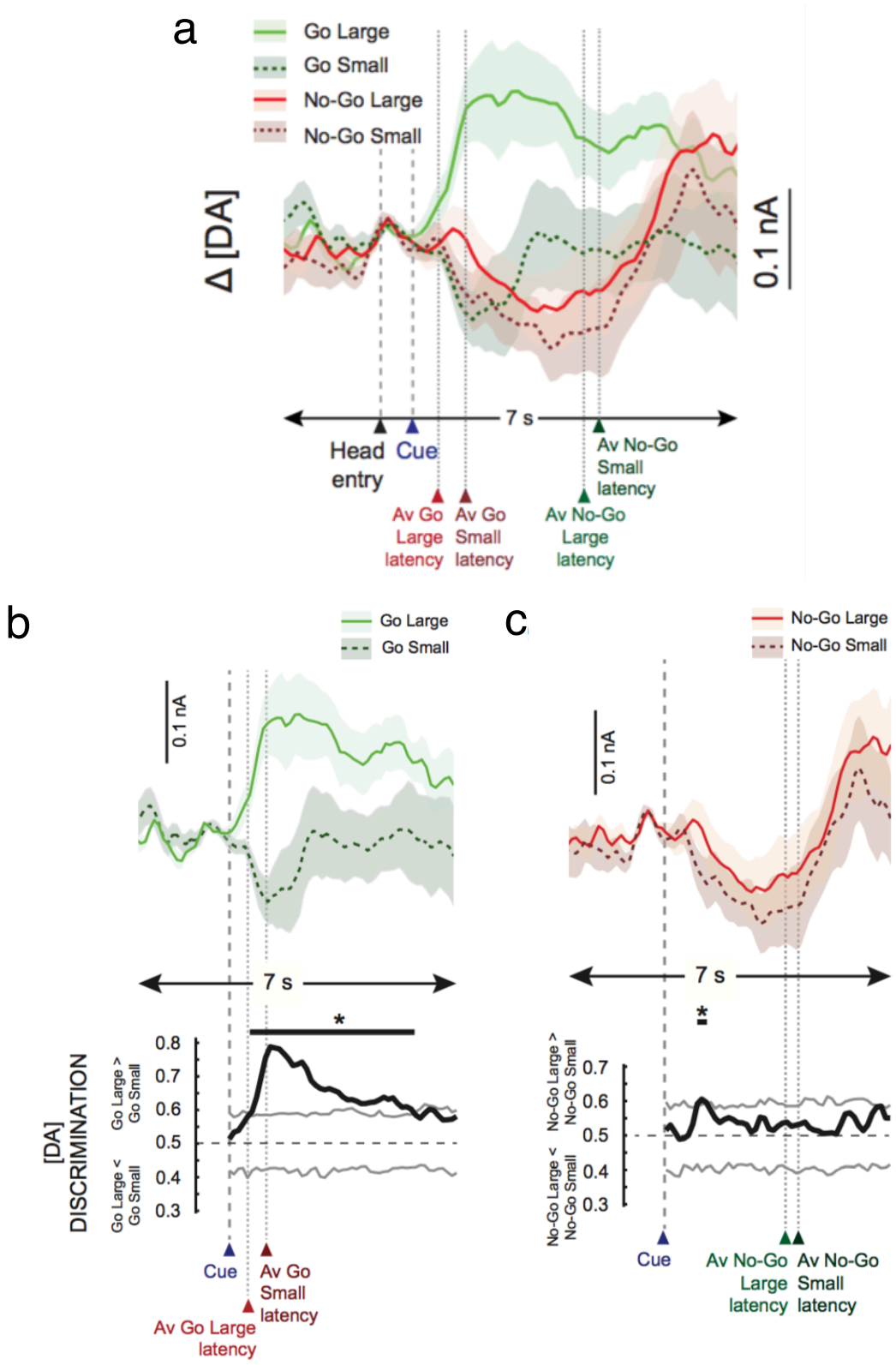


Fig. 2. Average dopamine traces across all four trial types. (a) Mean (lines) and S.E.M.s (shaded areas) of dopamine traces in the Go/No-Go task when aligned to cue onset. (b) Dopamine traces for Go Large and Go Small trials, aligned to cue onset. Lines indicate mean, shaded areas indicated S.E.M. Lower panel shows average discriminability between Go Large and Go Small trials for each timepoint against the permuted null distribution (grey lines representing maximum or minimum, * $p < .05$, Bonferonni-corrected). (c) Same as in (b) but for No-Go Large and Small trials.

3.3. NAc core dopamine release is shaped by correct action initiation

Re-alignment of these traces from both correctly executed Go and No-Go trials shows a sharp increase in dopamine release at point of first action initiation – nose poke exit (Figure 3). Importantly, replicating the results of the original Go/No-Go study, this was contingent on *correct* action initiation; if animals exited the nose poke prematurely in a No-Go trial, dopamine levels dipped before the error was explicitly signalled by the houselight (although the cue did indeed stop sounding as soon as the error was made), even though in both cases qualitatively the same movement was being executed (Figure 4b; shown only for No-Go trials as there were too few failed Go trials for the same analysis).

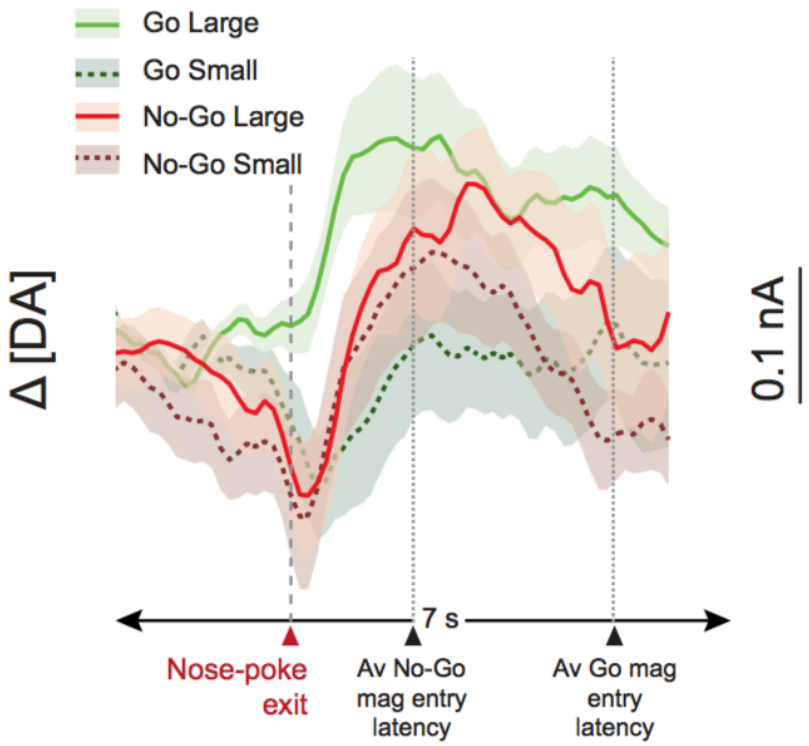


Fig. 3. Average dopamine traces across all four trial types aligned to point of first action initiation nose poke exit, showing mean (lines) and S.E.M.s (shaded areas).

Further evidence linking the dopaminergic response to action can be found by dividing traces from correctly performed No-Go trials by the speed of nose poke exit: either fast (< 600ms) or slow (> 1600ms). When aligned to cue offset, dopamine in fast No-Go head exit trials rises rapidly before responses in slow trials (Figure 4c), whilst re-alignment to nose poke exit again ties the moment of action initiation to an increase in the dopaminergic response (Figure 4d).

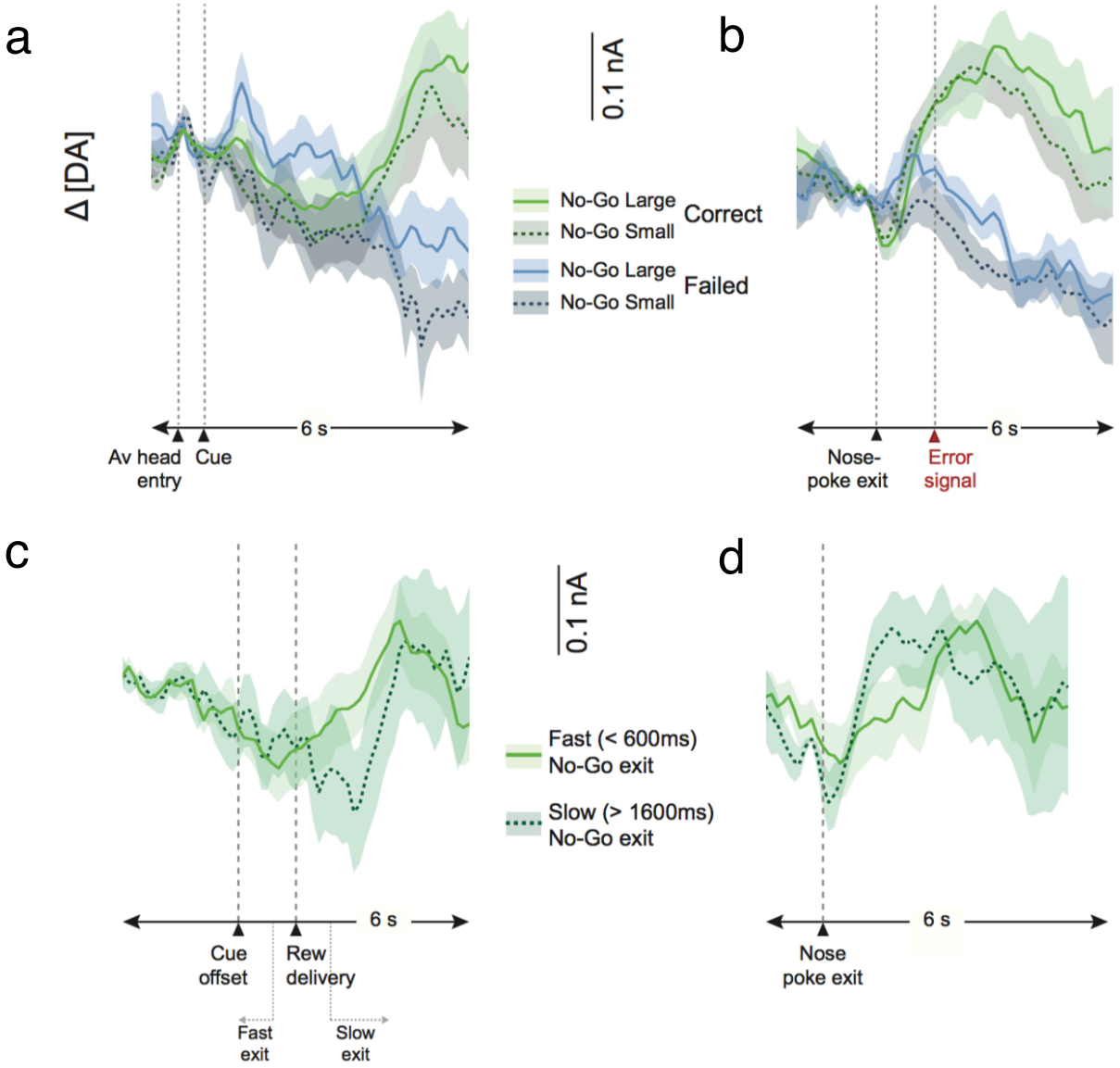


Fig. 4. Average dopamine traces for No-Go trials divided by success and speed. (a) Mean (lines) and S.E.M.s (shaded areas) of dopamine traces for correct No-Go trials (green) and failed No-Go trials due to premature head exit (blue), aligned to cue onset. Large reward trials have a solid line, small reward trials have a dotted line. (b) Same as in (a) but aligned to nose poke exit. (c) Average dopamine release on a subset of correctly performed No-Go trials where the rats' exit from the nose poke was either 'fast' (< 600ms after cue offset, filled green line) or 'slow' (> 1600ms after cue offset, dashed green line), aligned to cue offset (1s before reward delivery). (d) Same as in (c) but aligned to nose poke exit. Therefore, the fast head exit times occur > 400ms before reward delivery, and the slow ones > 600ms after reward delivery.

3.4. NAc core dopamine release reflects an interaction between reward prediction error encoding and correct action initiation

Finally, to understand the dynamics of these effects within a trial, a general linear model approach was used to understand the degree of combined contributions of reward prediction error and action initiation to dopamine release in a 2.5s post-cue period in the Go/No-Go task. Here, the regressor for reward prediction error was calculated by multiplying success rate of a trial type by reward size on offer (either 1 or 2 sucrose pellets) to give an expected value, and then by subtracting the average expected value from the expected value of each trial type (Table 2). Regressors for time of action initiation and an interaction term were also included.

This analysis solidified the conclusions made earlier; as well as main effects of reward prediction error and action initiation, these factors interacted to best predict the dopaminergic response in this time period (Figure 5). While the reward prediction error is the first significant predictor in the post-cue period, shortly thereafter the interaction term and then action initiation come into play. This suggests that, once cue-action-reward contingencies are fully learnt, NAc core dopamine levels in response to such cues is shaped not only by reward prediction error encoding, but *also* by the action required to obtain such reward.

PREDICTION ERROR CALCULATION				
Trial type	**Reward size**	**Average success rate**	**Expected value**	**Prediction error**
Go Small	1	0.82	0.82	-0.42
Go Large	2	0.98	1.96	0.72
No-Go Small	1	0.78	0.78	-0.45
No-Go Large	2	0.69	1.38	0.15
		Average expected value:	**1.23**	

Table 2. Prediction error calculation for use as a GLM regressor.

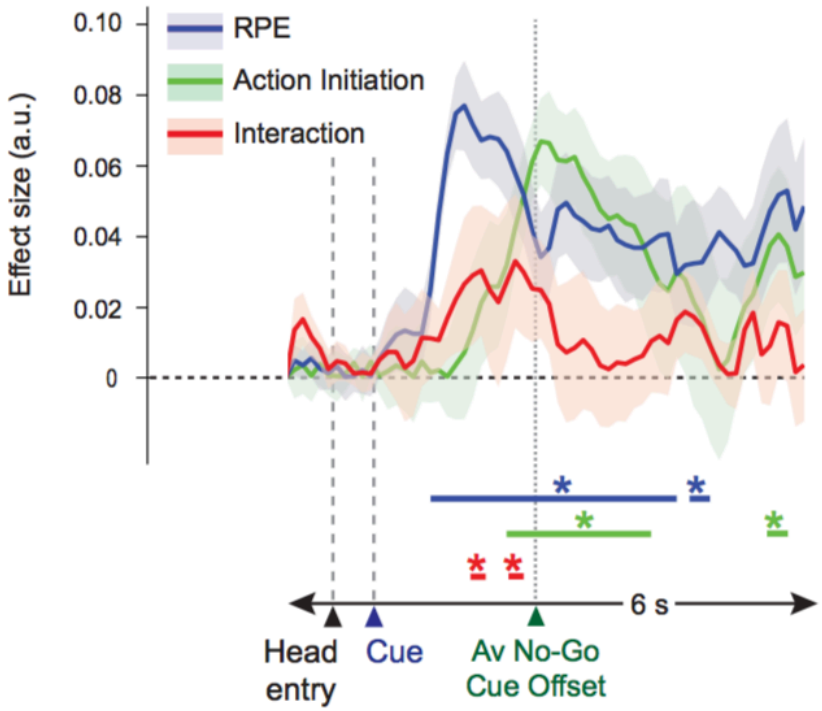


Fig. 5. A general linear model with regressors for reward prediction error (as calculated in Table 1), action initiation time, and an interaction term was applied to all dopamine responses during a 2.5s post-cue period in successful Go and No-Go trials. Both factors and an interaction were significantly predictive of dopamine levels at different points in this post-cue period. Average absolute effect sizes are plotted (lines indicate means, shaded areas are S.E.M.) Significance indicated when permutation testing indicated a significant difference, $p < .05$, Bonferonni-corrected, a.u., arbitrary units.

4. Discussion

Phasic dopamine release in the NAc core has been thought to reflect a similar signal to that found in the activity of midbrain dopaminergic neurons: the encoding of a reward prediction error, or the discrepancy between expected and received reward (Day et al., 2007; Schultz et al., 1997). However, there is also a substantial body of research that ventral striatal dopamine is key in facilitating behaviour in response to stimulus-reward associations, providing the energetic impetus for operant responding (Robbins & Everitt, 2007; Robbins & Everitt, 1992). Whilst this conclusion has often been drawn from manipulations of longer-term changes in dopamine levels,

such as in pharmacological approaches, studies measuring phasic, short-term changes in mesolimbic dopamine release in the NAc core have also demonstrated fluctuations coincident with reward-seeking action (Phillips et al., 2003; Roitman et al., 2004).

These previous studies did not directly manipulate action requirements and so only provide indirect evidence of this effect. Thus here, for the first time, the contribution of goal-directed action to mesolimbic dopamine release was understood in the context of animals having to either make or withhold action for reward. It was discovered that phasic dopamine release is contingent on correct action initiation, whilst also varying as a function of future potential reward. This is a first step towards uniting the two predominant literatures on the reward prediction error and purposeful movement aspects of dopamine.

4.1. Understanding the relationship between VTA cell firing and NAc dopamine release

One important aspect to note about these data is that they exclusively speak to fluctuations in downstream dopamine release in the NAc core; no direct measurements of the cell firing of midbrain dopamine neurons was taken in this study, and it was that approach in that locus that yielded the initial discovery of the reward prediction error signal (Schultz, Apicella, & Ljungberg, 1993). Whilst there is evidence of movement mediating midbrain cell firing, such as from the key proponent of the prediction error theory himself (Schultz, Ruffieux, & Aebischer, 1983) – it cannot be taken for granted that the activity of these neurons would directly

reproduce the dynamics seen in dopamine release in the accumbens during the Go/No-Go task.

Indeed, the relationship between VTA cell firing and NAc dopamine release has become an issue of interest; whilst any differences seen might be because of different experimental conditions – electrophysiological experiments in the midbrain often use headfixed approaches, limiting movement in contrast to the techniques used in the striatum such as that applied here – a recent direct comparison of dopamine cell firing and release in the same task shows a dissociation between the two, with VTA neurons indicating prediction error-like responses and NAc dopamine instead reflecting reward expectation (discussed further in the subsequent section; Hamid et al., 2016). This discrepancy may be due to local striatal mechanisms that are able to modulate dopamine release outside of the activity of VTA neurons (Glowinski, Chéramy, Romo, & Barbeito, 1988; Threlfell et al., 2012). Of course, Hamid et al. (2016) do not directly speak to the situation addressed here, where action must be withheld for reward, but their findings suggest the possibility that the activity of midbrain dopamine neurons may not be suppressed in the way that dopamine release has shown to be in the current experiments.

4.2. Dopamine ramps for reward – contingent on action and certainty

A point of note is that the sustained dopaminergic release to action initiation described in this chapter is strikingly reminiscent of 'ramping' seen in other studies as animals execute a series of behaviours whilst working towards reward (Hamid et al., 2016; Howe, Tierney, Sandberg, Phillips, & Graybiel, 2013; Wassum, Ostlund, &

Maidment, 2012). Such ramps do not correspond to surprise or unexpected reward or information, and are sustained even after learning – as seen here. Some have suggested that, rather than directly indicating goal proximity, such ramps might be indicative of expected value which increases as the animal approaches its target (Hamid et al., 2016). Others have claimed that ramping can indeed fit into a temporal difference learning framework, dependent on the transformation of the proximity to goal value (Gershman, 2014), or that ramping can arise from phasic prediction error coding if the decay of learned values is incorporated into the model (Kato & Morita, 2016). But regardless of the proposed mechanism or computation by which such ramping arises, the experiments conducted here demonstrate that it is dependent on action, as in No-Go trials no steady increase of dopamine is seen until the time of action initiation – even though, throughout the correct withholding of action in the cued No-Go period, animals are becoming more and more temporally proximal to reward.

In addition, the fact that such ramping is *not* seen here when animals are about to execute an incorrect action, even before it is explicitly indicated that the action is incorrect, is intriguing. One explanation is that dopamine only rapidly rises in conjunction with highly certain or confident actions, and in the case of a (soon-to-be) incorrect action, levels of uncertainty are greater (Fiorillo, Tobler, & Schultz, 2003) – although interestingly some have hypothesised that greater uncertainty should in fact lead to, rather than prevent, ramping (Niv, Duff, & Dayan, 2005). This potential influence of uncertainty on the dopaminergic signal might also translate into a bias in action selection towards either 'riskier' or 'safer' options (Stopper, Tse,

Montes, Wiedman, & Floresco, 2014), even in situations where the riskiness of a choice is subjective rather than objective.

Therefore, whilst uncertainty was not explicitly manipulated in the tasks described here, it appears that the large phasic increase in dopamine seen in correctly executed Go trials is contingent both on action initiation and possibly on the certainty with which it is carried out.

4.3. Phasic dopamine release: a driver of reward-guided behaviour?

Whilst the experiments described in this chapter are an important and necessary step towards a more holistic understanding of the NAc core dopamine response to reward prediction and movement, despite an excellent temporal resolution, FCV does not allow causality to be ascertained. In other words, it is not yet known whether the phasic dopamine release seen here, corresponding to action initiation in all trial types both in Experiment 1 and 2, actively *drives* behaviour, or if it is a response to the execution of behaviour – possibly either to sustain it, or to act as a downstream value signal to be updated upon reward receipt. Thus the experiments conducted in the subsequent empirical chapters attempt to address this issue by directly manipulating the degree of activation of dopamine receptors in this same Go/No-Go task, bringing together evidence from pharmacology and physiology studies to further understand the role of dopamine in reward-guided movement.

4 D1-like receptors shape initiation, execution, and selection of reward-guided action

Chapter abstract:
Dopamine is thought to play an important role in learning about reward and reward-predictive cues, but also in motivated behaviour for reward. The previous chapter demonstrated that dopamine release in the nucleus accumbens core reflects an interaction between both reward prediction and action initiation, but did not elucidate whether dopamine *causally* promotes reward-seeking behaviour. Here, by using systemic pharmacological manipulations of D1Rs when animals had to either make or withhold action for reward, it was revealed that this receptor family both influences cue-driven action initiation and ongoing action execution, and biases action selection. This lends further support to the hypothesis that dopamine is not merely passively performing as an expectancy discrepancy signal, but in fact plays a key role in *actively* guiding behaviour.

For the studies presented in this chapter, Laura Grima and Mark Walton planned the experiments, Laura Grima collected the data, and Laura Grima analysed and presented the data with input from Mark Walton.

1. Introduction

Historically, dopamine in the mammalian brain has been considered important both for reinforcement and learning about reward, and for driving movement. Yet the specific relationship between these functions is not yet fully understood. In Chapter 3, a novel behavioural paradigm was used in conjunction with fast-scan cyclic voltammetry (FCV), a method for measuring subsecond dopamine release, and we revealed for the first time that dopamine levels in the nucleus accumbens (NAc) core reflect an interaction between encoding of a reward prediction error *and* the goal-directed action required to garner reward (see also Syed et al., 2016). This is an important step in uniting these disparate literatures.

However, FCV only reports changes in dopamine levels – it does not perturb or alter them. Therefore, in the previous chapter it was not possible to ascertain whether an increase in dopamine was directly promoting or reporting purposeful action initiation. Here, by manipulating D1-like receptors (D1Rs) – those which are likely most sensitive to the phasic changes in dopamine levels as recorded by FCV (Dreyer et al., 2010; Richfield et al., 1989) – when animals must either make or withhold cue-driven action for reward, we have been able to tackle the question of whether dopamine transmission *causally* influences action initiation and restraint, as well as the degree to which this is reward- or cue-mediated.

1.1. D1Rs in prediction and action

Both the phasic activity of midbrain dopaminergic neurons and the downstream release of dopamine in the NAc core are purported to vary as a function of reward prediction error, responding to unexpected discrepancies from expectation in both the timing and amount of reward delivery, and to informative cues that predict upcoming reward (Day, Roitman, Wightman, & Carelli, 2007; Schultz, 1998; Schultz, Dayan, & Montague, 1997). In this way, phasic dopamine acts as a teaching signal, updating cue value in response to reward. Additionally, cue-evoked excitation of MSNs consequent to this phasic release has been shown to be dependent on D1R function (du Hoffmann & Nicola, 2014).

However, the FCV recordings in Chapter 3 unequivocally demonstrated the importance of action initiation to the phasic dopamine signal, previously seen by many as solely encoding reward prediction error. Nevertheless, it may still be the case that, although dopamine appeared to fluctuate with action, it is *reporting* rather than driving animals' behaviour, akin to updating state values as the rat moves through a trial (Hamid et al., 2016; Howe, Tierney, Sandberg, Phillips, & Graybiel, 2013).

In parallel, however, the literature on D1R function suggests a possible causal influence on behaviour. In addition to the sensitivity of these receptors to the phasic dopamine release seen in Chapter 3, the classical view of basal ganglia function would also suggest that it is this receptor subtype – predominantly expressed as part of the direct pathway – that enables the organism to initiate movement in the face of D2R indirect pathway activation, which simultaneously inhibits movement (Albin et al.,

1989; Frank, 2005). Both phasic stimulation of direct pathway MSNs and tonic stimulation of D1Rs promotes locomotor activity, lending support to this hypothesis (Dreher & Jackson, 1989; Kravitz et al., 2010).

Importantly, these phasic dopaminergic increases observed using FCV were not found in response to all forms of motoric behaviour, such as erroneous action initiation during the cue in a No-Go trial, but instead only in the context of *reward-seeking* action. Impulse control paradigms support the role of D1Rs in the initiation of goal-directed behaviours; whilst there is little evidence that these receptors (particularly in the ventral striatum) play a role in action cancellation (Dalley et al., 2011), a rodent's ability to postpone action is increased with D1R antagonism in the 5-choice serial reaction time task (Pattij et al., 2007), and decreased with their stimulation (Pezze et al., 2007). In the latter case, the erroneous responses are arguably reward-seeking, motivated actions, as instead of generally increasing motoric output the animal is drawn to respond in a port – an action that has previously led to reward. These studies would therefore suggest that D1Rs play a role in promoting reward-guided action. However, it has not yet been tested whether their manipulation would have a similar effect when having to withhold action for reward, and how trial-by-trial changes in reward expectation, causing positive or negative reward prediction errors, influences these effects.

1.2. D1Rs in cue-driven action and ongoing performance

Correct action initiation and execution in the 5-choice task is not only reward-seeking, but is *cue-triggered*. Such cue-triggered behaviour has been shown to be

dependent on D1R activation (Pattij et al., 2007 - although this is region-dependent) and studies specifically designed to investigate cue-driven responding have also demonstrated the necessity of this receptor subtype to such behaviour (Yun et al., 2004), coherent with the hypothesis that dopamine acts to facilitate the ability to respond to reward-predictive information (Robbins & Everitt, 2007). This is also consistent with the phasic dopamine release observed at the point of correct cue-driven action initiation in the FCV studies in Chapter 3.

Yet there is also evidence that self-initiated action is also dopamine-dependent, and Salamone, Correa, Farrar, & Mingote (2007) have proposed a role of dopamine in the invigoration of effortful responding for reward (Floresco, St. Onge, Ghods-Sharifi, & Winstanley, 2008; Phillips et al., 2007). This can be distinguished from the *ongoing performance* of action for reward; whilst initially engaging in a behaviour is dependent on D1R activation, its continued engagement – for example, execution of a sequence of nose pokes – is unaffected by D1R antagonism (Keeler, Pretsell, & Robbins, 2014). Yet in Chapter 3, with animals that were very well-trained in a deterministic cued task requiring relatively stereotyped behaviours in Go trials – either moving left or right in response to a cue – we found sustained dopamine release as animals moved *throughout* the trial until reward retrieval. It is therefore not yet known whether disrupting the downstream effects of this dopamine signal within the same task, and with similarly well-trained animals, will alter the ongoing performance of reward-directed behaviour or only action directly promoted by value-laden cues.

1.3. Aims, approaches, and hypotheses

The experiments in this chapter aim to resolve whether dopamine transmission, following the changes in release seen in conjunction with action initiation in Chapter 3, acts to directly promote behaviour, or instead reports it as a behavioural readout.

In order to address this, we examined the effects of systemic administration of either a D1R agonist or antagonist to rodents performing the Go/No-Go paradigm. If phasic dopamine release purely signals the discrepancy between expected and actual cue value, then stimulation or blockade of D1Rs might increase or decrease respectively the value of all cues presented, and thereby modulate performance similar to a change in reward size.

However, if dopamine levels causally promote reward-seeking *actions*, then D1R stimulation should result in animals being less able to withhold action for reward, and more likely to carry out reward-seeking action when it is inappropriate to do so. The antagonist, on the other hand, might have limited effects on No-Go performance – as no dopamine release was seen in such trials – but, if goal-directed action is dependent on the dopamine efflux, should reduce performance in Go trials.

Finally, if dopamine not only promotes the initiation of reward-seeking actions, but is also necessary for the *ongoing performance* of those actions, then a D1R manipulation will both alter the speed and engagement throughout Go trials.

2. Methods

2.1. Animals

All procedures were carried out in accordance with the UK Animals (Scientific Procedures) Act (1986). Two separate cohorts of animals were used in the two experiments described in this chapter. For the D1R agonist experiment, a total of 11 male Sprague-Dawley rats that had previously been used for FCV recordings in the same Go/No-Go task (see Chapter 3) were used. For the D1R antagonist experiment, a total of 14 male Sprague-Dawley rats that had not undergone any other procedures were used. One animal had to be excluded due to completing < 20% of possible correct trials in one drug session, resulting in an n of 13 for this second experiment.

All animals were maintained on a twelve-hour light/dark cycle (lights on at 07:00). Rats in the D1R agonist experiment were single housed due to previous surgery. Rats in the D1R antagonist experiment were grouped housed throughout training and the drug study. All testing was carried out during the light phase, and during training and testing periods, animals were food restricted such that their weights were kept between 85 – 90% of their free-feeding weight. Water was available *ad libitum* whilst animals were in their home cages.

Full details of the apparatus and behavioural training, as well as further details of the task, are described in Chapters 2 and 3. All pharmacological challenges were carried out in well-trained rats who had achieved stable levels of performance.

2.2. *Measures of performance and latency*

The measures of performance and latencies recorded in the Go/No-Go task are outlined in detail in sections 1.4 and 1.5 of Chapter 2, but are reiterated briefly in Table 1 below. In addition to the measures of performance described in Table 1 that influence success rate, animals could also 'abort' trials by not remaining in the nose poke for full pre-cue period between trials, therefore failing to initiate the next trial.

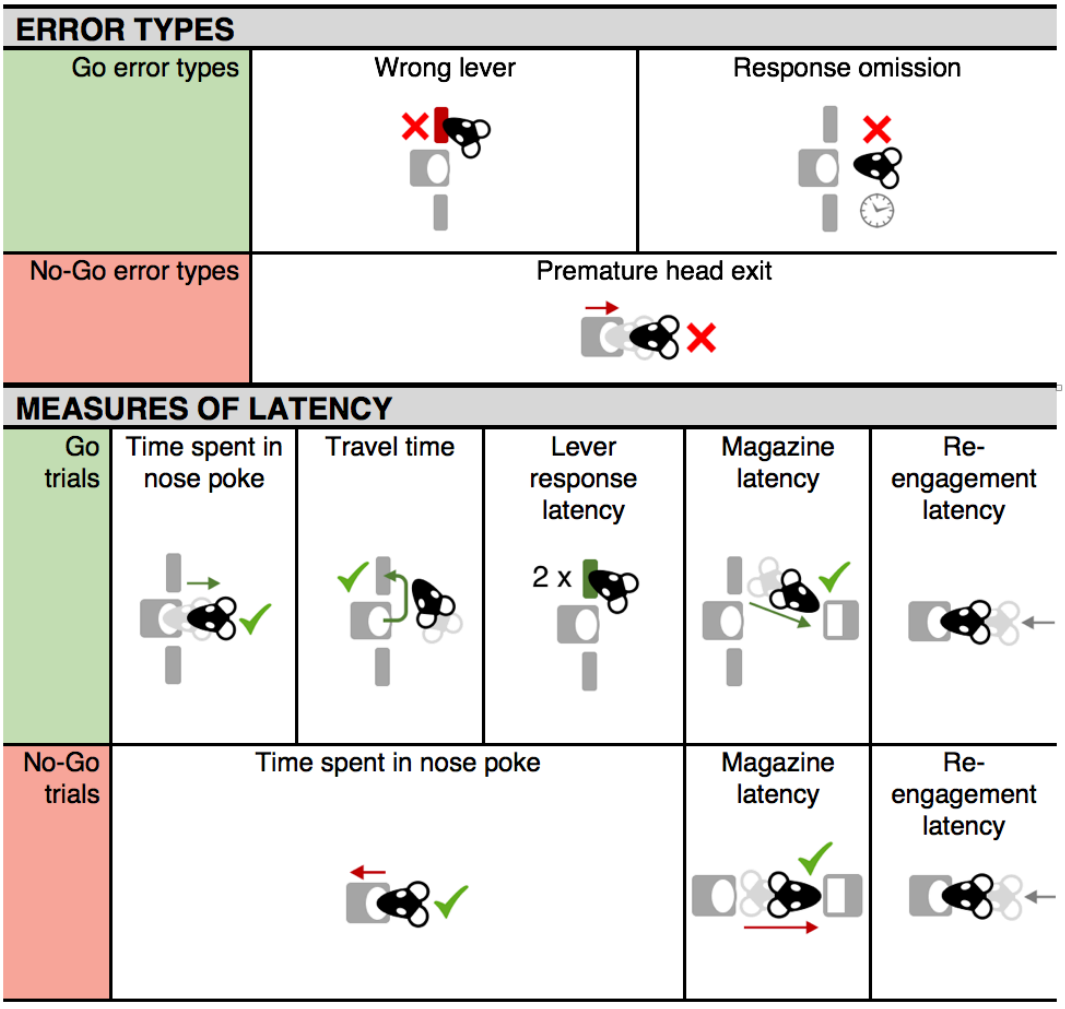


Table 1. Types of errors and measures of latency analysed in Go and No-Go trials.

2.3. Pharmacological compounds

SKF 81297 hydrobromide (Tocris Bioscience) was used as the D1-like agonist. It is a full, not partial, agonist, meaning that it produces a response at the receptors that has an efficacy comparable to endogenous dopamine. It shows around a 500 times greater affinity for D1-like over D2-like receptors (Neumeyer, Kula, Bergman, & Baldessarini, 2003). The D1-like antagonist was SCH 23390 (Tocris Bioscience). It is highly selective for both D1 and D5 receptors, with around an 800-fold greater affinity for these than for D2-like receptors (Bourne, 2001).

2.4. Pharmacological challenges

Both SKF 81297 and SCH 23390 were administered systemically via i.p. injection. SKF 81297 was administered 10 minutes before the behavioural session began, whilst SCH 23390 was administered 20 minutes before. All drugs in all experiments were applied in separate counterbalanced Latin square designs to reduce the influence of order effects.

In the first experiment, SKF 81297 was applied systemically at doses of 0.3 and 1.0 mg/kg, calculated as the salt. These doses were chosen to reflect studies that have used similar dose ranges to show the role of stimulating D1Rs in cue learning (Alleweireldt, Weber, Kirschner, Bullock, & Neisewander, 2002), waiting in a delay discounting task (Koffarnus et al., 2011), and increased risky choices (St Onge & Floresco, 2009). In addition to drug sessions, there was also a baseline vehicle (saline) session.

This second cohort was also used for the D1R antagonist experiment. These animals received SCH 23390 systemically at doses of 0.005 and 0.01 mg/kg, calculated as the salt. A higher dose of 0.02 was also used in these same animals in a separate Latin square, but affected behaviour such an extent that in 7 drug sessions, animals did not complete all trials within 60 minutes. Therefore, the data presented here are exclusively from the experiment using the first two lower doses. These doses were reflective of previous studies showing increased impulsivity in delayed discounting, (Koffarnus et al., 2011; van Gaalen et al., 2006), increased risk aversion (St Onge & Floresco, 2009), and decreased choice of high physical effort options (Bardgett, Depenbrock, Downs, Points, & Green, 2009; Hosking, Floresco, & Winstanley, 2014).

2.5. Data analysis approaches

A full discussion of the statistical approaches used in the analyses of the pharmacology data presented here can be found in section 2 of Chapter 2. To reiterate briefly, unless otherwise stated, mean values were analysed using 2 (reward: small or large) x 3 (drug dose: saline, low dose, and high dose) within-subjects ANOVAs. Median values were also analysed using the same approach for all measures reported, but were only explicitly discussed if the results of the ANOVA differed from testing of the mean values. Significant main effects or interactions were further investigated with appropriate pairwise comparisons, corrected for multiple comparisons using the Sidak approach.

Due to the relatively limited number of observations (in comparison to studies of reaction time latencies in humans in which a similar procedure would be used), latency distributions were analysed by comparison of the 0.05, 0.45, and 0.95 quantiles representing early, mid, and late latencies. These were then also included in a within-subjects ANOVA, along with reward and drug dose, and any significant main effects or interactions were also investigated with appropriate pairwise comparisons (Sidak-corrected).

3. Results

D1R Agonist

3.1. Systemic D1R stimulation reduces success rate in both Go and No-Go trials

If stimulation of D1Rs acts as a proxy positive reward prediction error, cue value and thus success rate – ability to execute correct and accurate action in Go trials, and ability to withhold action in No-Go trials – should be increased. However, if it acts to activate or increase the value of action, Go success rate should increase whilst No-Go performance should decrease. In fact, systemic D1R stimulation decreased success rate in a dose-dependent manner on *both* Go (Figure 1a) and No-Go (Figure 1b) trials.

In Go trials, animals' success rate was influenced by reward, with higher performance when a large reward was on offer [significant main effect of reward: $F_{1,10} = 15.498$, $\eta_p^2 = .608$, $p = .003$]. The D1R agonist reduced Go success rate at the highest dose of the drug, particularly on *small* reward Go trials [significant main effect of drug: $F_{1.440,14.400}$

= 4.535, η_p^2 = .312, p = .04, and drug*reward interaction: $F_{2,20}$ = 4.135, η_p^2 = .293, p = .031; pairwise comparisons: small reward, low vs. high dose p = .033, saline vs. high dose p = .077 (uncorrected)].

Overall success rates in No-Go trials were also reward-mediated, with average success rates greater on high reward trials [significant main effect of reward: $F_{1,10}$ = 5.130, η_p^2 = .339, p = .047]. Systemic administration of the D1R agonist dose-dependently and markedly reduced No-Go success rate, again with a tendency for this to be most prominent in *small* reward trials [significant main effect of drug: $F_{2,20}$ = 14.911, η_p^2 = .599, p < .001; drug*reward interaction: $F_{2,20}$ = 3.467, η_p^2 = .257, p = .051; pairwise comparisons in small reward trials: vehicle vs. low dose, p = .022, vehicle vs. high dose, p = .002; significant pairwise comparisons in large reward trials: vehicle vs. high dose, p = .061, low dose vs. high dose, p = .020].

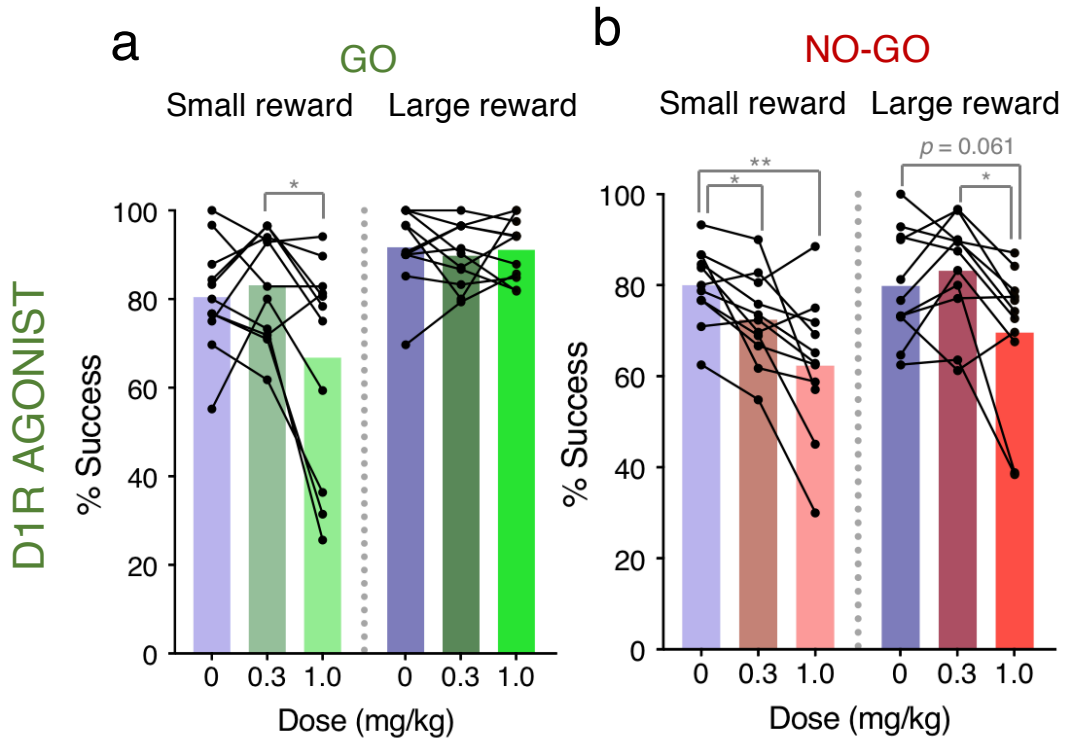


Fig. 1. Effect of systemic D1R stimulation (SKF 81297) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) D1R stimulation reduced small reward Go trial success rate at the highest dose of the drug. (b) D1R stimulation reduced No-Go success rate in both large and small reward trials, but particularly for small reward trials. Significance bars indicate post-hoc comparisons, * p < .05, ** p < .01, Sidak-corrected.

To further understand the detrimental effect of D1R stimulation in success rate in Go trials, the *types* of errors were analysed. In such trials, animals could either press the wrong lever from that indicated by the auditory cue (Wrong Lever error), or could omit responding at either lever during the 5s cue period (Response Omission error). Drug treatment increased the proportion of trials in which animals performed a Wrong Lever error dependent on the size of the reward on offer [Figure 2a; significant drug*reward interaction: $F_{2,20}$ = 4.179, η_p^2 = .295, p = .030; significant pairwise comparisons in small reward trials: vehicle vs. high dose, p = .029 (uncorrected); no main effect of drug, $F_{2,20}$ = 2.584, η_p^2 = .205, p = .100]. Therefore, when presented with a cue instructing a response to the small reward lever, D1R stimulation caused animals to be increasingly likely to select the option that had previously led to a larger reward. While Response Omission errors also increased numerically, particularly on small reward trials (Figure 2b), neither the main effect of drug or drug*reward interaction reached significance [$F_{2,20}$ < 3.31, p > .058].

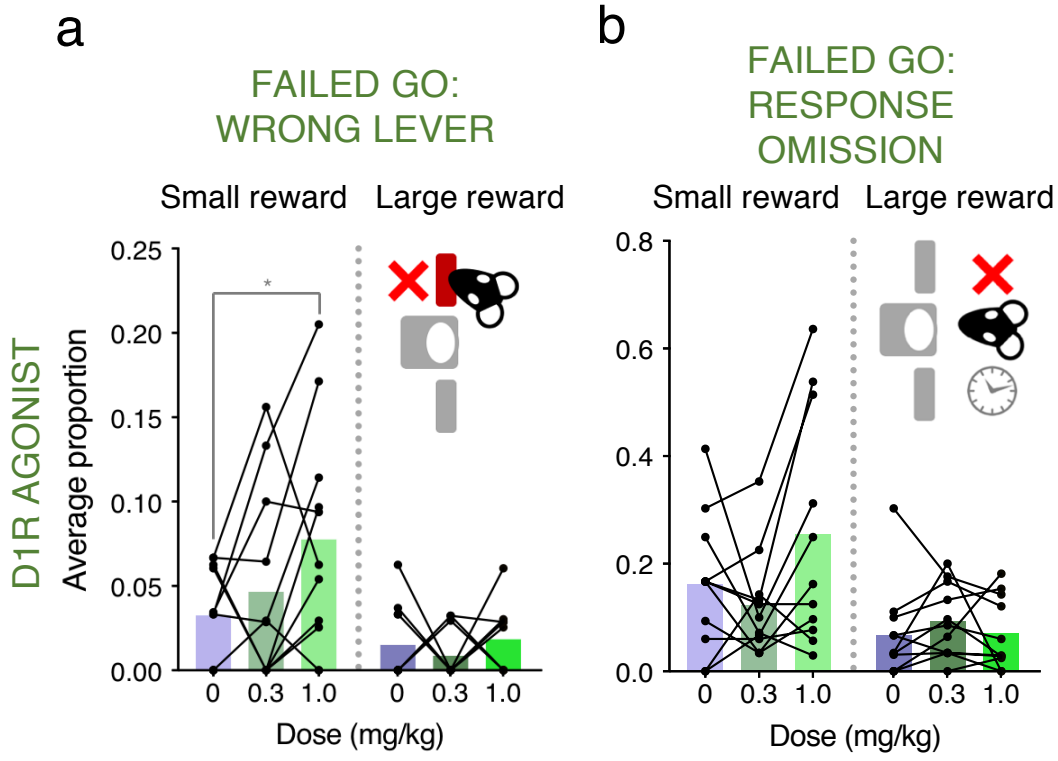


Fig. 2. Effect of D1R stimulation on the proportion of Go small or large trials where animals make (a) a wrong lever press or (b) omit response entirely during the 5s cue period. Significance bars indicate pairwise comparisons, * p < .05, ** p < .01, uncorrected.

We next investigated *when* animals were making errors on No-Go trials. Therefore, we plotted the distributions of nose poke exit latencies in failed No-Go trials as a proportion of all nose poke exit latencies. This showed an increased likelihood to leave only *early* in the No-Go cue period regardless of whether the size of the reward on offer was small (Figure 3a) or large (Figure 3b).

To more formally quantify this, average proportions of head exit times were divided into 'early' (< 800ms) or 'late' (> 800ms) errors (Figure 3c). A three-way within subjects ANOVA revealed that animals were more likely to leave prematurely in the No-Go period when the reward on offer was small [significant main effect of reward: $F_{1,10} = 5.134$, $\eta_p^2 = .339$, $p = .047$]. More importantly, the D1R agonist influenced when animals were likely to leave the nose poke in No-Go trials, with *early* head exits increasing significantly but late head exits remaining unaltered [significant drug*error period interaction: $F_{2,20} = 7.780$, $\eta_p^2 = .438$, $p = .003$; pairwise comparisons: early period, vehicle vs. low dose, $p = .008$, vehicle vs. high dose, $p < .001$, low dose vs. high dose, $p = .005$, late period, all $p > .5$]. This effect was also partially reward-mediated, with a greater increase in early head exits when a smaller reward was on offer [trend reward*drug interaction: $F_{2,20} = 3.459$, $\eta_p^2 = .257$, $p = .051$; post-hoc multivariate analyses: small reward, $F_{2,9} = 12.516$, $\eta_p^2 = .736$, $p = .003$; large reward, $F_{2,9} = 5.346$, $\eta_p^2 = .543$, $p = .030$]. However, there was no three-way interaction [$F_{2,20} = 1.003$, $\eta_p^2 = .091$, $p = .384$].

To summarise, as well as increased erroneous action selection in Go trials, errors of action initiation were increased with D1R stimulation. And when such errors of action

initiation were made, they were almost exclusively in the period just after cue onset, particularly when the reward to be gained (or the opportunity cost to be avoided) was small.

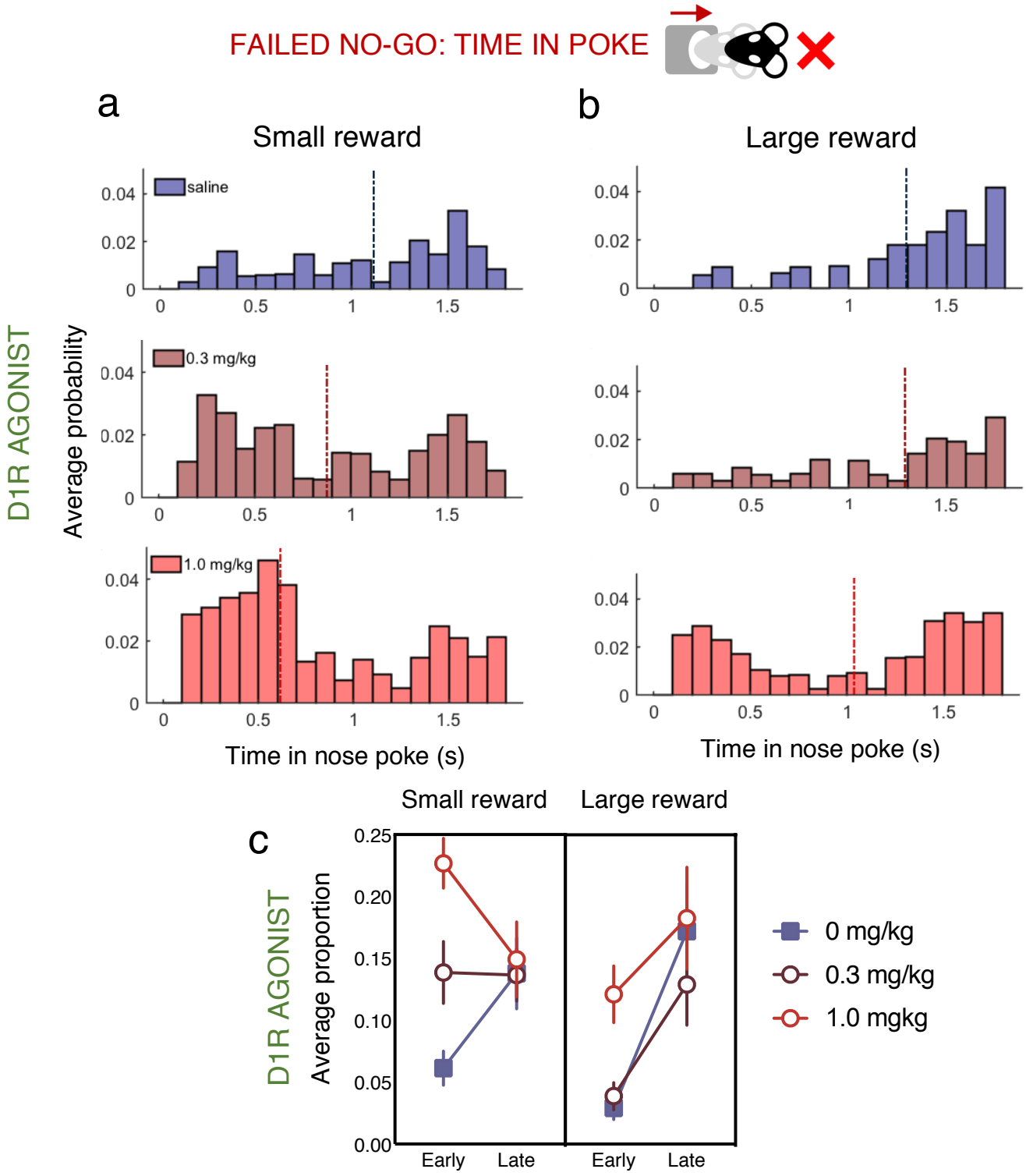


Fig. 3. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with D1R stimulation, calculated as probability over all head exit times, both successful and failed. Last bin of each histogram contains all values > 1.7s. Dotted line indicates median. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.3 mg/kg, red circle: 1.0 mg/kg).

3.2. *Systemic D1R stimulation does not alter the speed of cue-driven action initiation*

We previously showed that animals made more, and earlier, incorrect responses with application of a D1R agonist when it was inappropriate to initiate action. We next examined how this same drug affected the speed of *appropriate,* or purposeful, responses to initiate action. In correct Go trials, this is the latency to leave the nose poke after cue onset (Figure 4a). Whilst animals were quicker to leave the nose poke if a large reward was on offer [significant main effect of reward: $F_{1,10}$ = 13.105, η_p^2 = .567, p = .005;], the D1R agonist caused no reliable reduction in the mean time spent in the nose poke in successful Go trials regardless of reward on offer [no main effect of drug, nor drug*reward interaction: all Fs < 2.1, ps > .1].

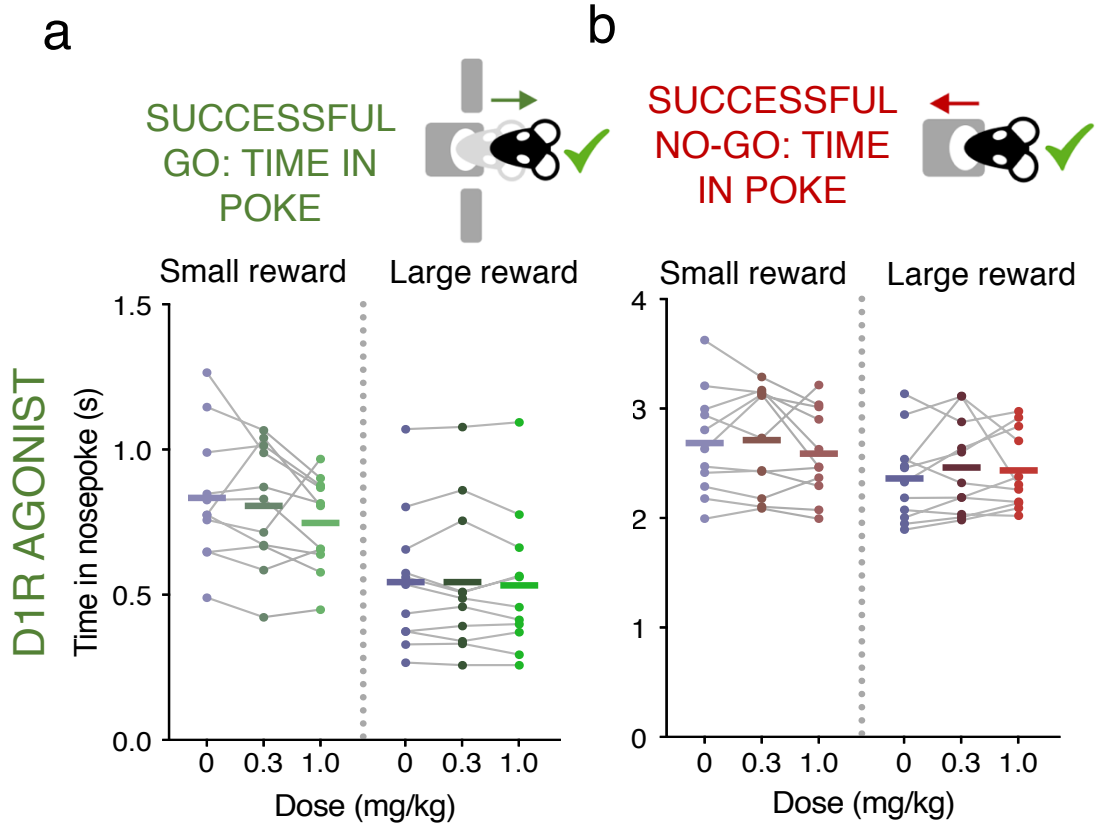


Fig. 4. Mean time spent in the nose poke (s) from cue onset in successful (a) Go or (b) No-Go trials. Solid coloured bars indicate overall mean time in nose poke, and individual dots are mean times for individual animals.

Moreover, even though visual inspection of the mean head exit time suggested there might be a speeding of latencies, the same result was found using a finer-grained analysis of the distribution of response times (Figure 5) [all *F*s < 2.9, *p*s > .08].

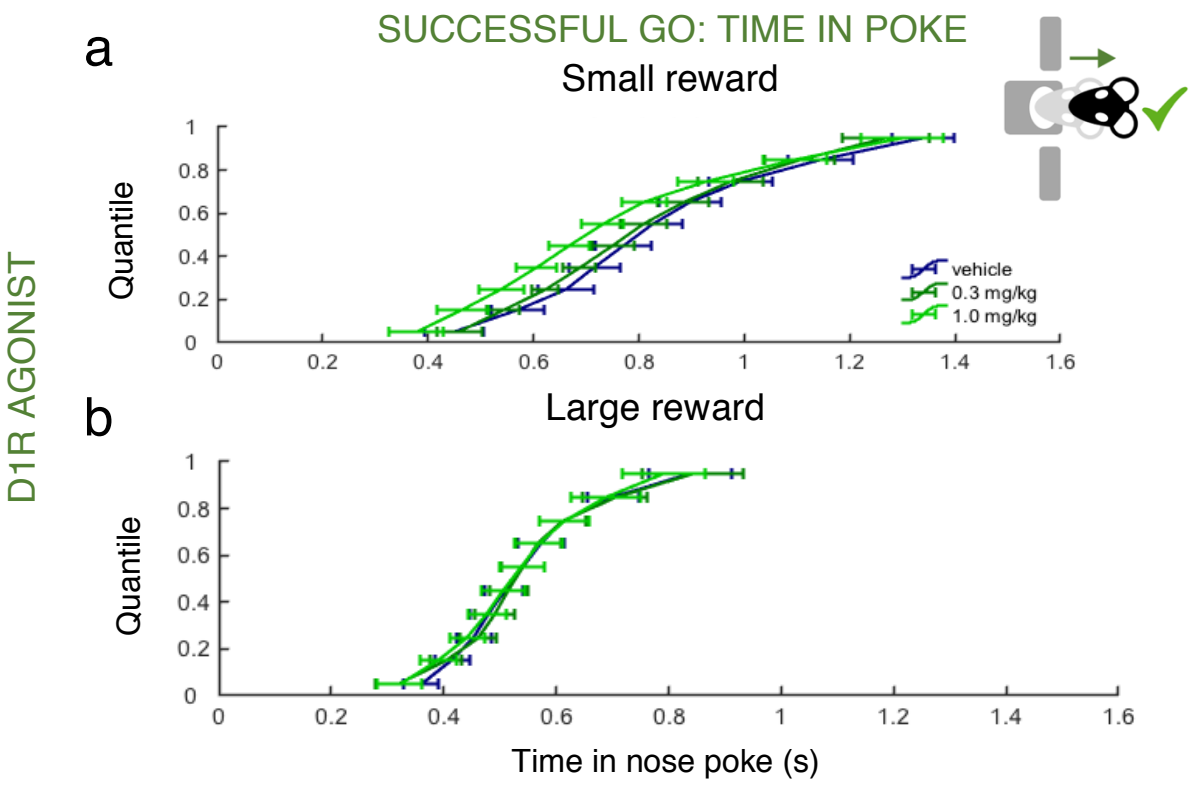


Fig. 5. CDFs for time in nose poke from cue in successful Go trials with D1R stimulation. (a) CDF for head exit time in small reward Go trials. (b) Same as in (a) but for large reward Go trials.

We also investigated whether the latency to initiate appropriate action in successful No-Go trials was similarly unaffected by D1R stimulation. Animals were again faster to exit the nose poke if a large reward was on offer [significant main effect of reward: $F_{1,10} = 9.556$, $\eta_p^2 = .489$, $p = .011$]. Additionally, the drug also had a reward-dependent effect on time spent in the nose poke in successful No-Go trials [significant drug*reward interaction: $F_{1.348,13.480} = 7.264$, $\eta_p^2 = .421$, $p = .015$], although no post-hoc pairwise comparisons between any of the doses were significant, and similarly there

was no significant main effect of drug on these latencies [$F_{2,20}$ = 0.473, η_p^2 = .045, p = .630].

3.3. Systemic D1R stimulation reduces energising of ongoing goal-directed behaviour

D1R stimulation promotes cue-elicited action initiation. However, we have also previously found that increases in dopamine were not just observed at cue presentation when an animal made a successful response, but were sustained throughout the trial to reward collection (see Chapter 3). Therefore, we hypothesised that the D1R agonist might also speed movement through a trial.

Surprisingly, though, the agonist in fact dose-dependently *increased* travel times between head exit and the correct lever on Go trials (Figure 6; significant main effect of drug: $F_{2,20}$ = 6.331, η_p^2 = .388, p = .007; significant pairwise comparisons: vehicle vs. high dose, p = .026]. However, this effect was not reward-mediated, and nor were animals on average faster to travel in large reward trials [no significant drug*reward interaction, nor main effect of reward: Fs < 2.4, ps > .1].

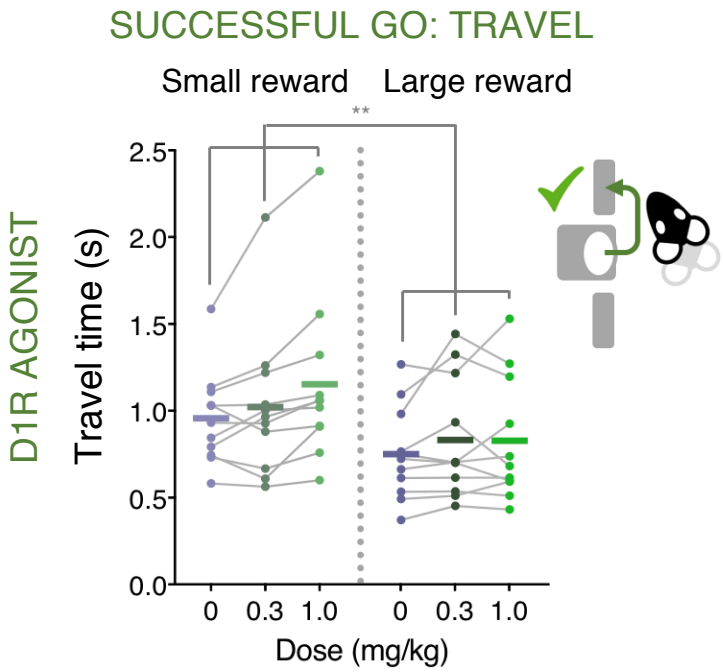


Fig. 6. Mean travel time from nose poke exit in successful Go trials was significantly increased by D1R stimulation when both a small or large reward was on offer. Significance bars indicate main effect of drug, ** p < .01

Importantly, it was not the case that all action latencies were affected. The D1R agonist did not affect time taken to complete the Go lever response requirement [no main effect of drug, reward, or drug*reward interaction: all *F*s < 1.0, *p*s > .3]. And whilst animals remained faster to collect large rewards in both Go [significant main effect of reward: $F_{1,10} = 12.954$, $\eta_p^2 = .564$, $p = .005$] and No-Go trials [significant main effect of reward: $F_{1,10} = 8.760$, $\eta_p^2 = .467$, $p = .014$], the D1R agonist had no effect on latency to reach the food magazine in either case [all *F*s < 0.2, *p*s > .6].

Travel times were not the only goal-directed behaviour that was slowed by the drug, however. Trials were self-paced, in that there was no external cue to signal when the inter-trial interval was complete and a new trial could begin. A simple incentive motivation account would suggest that increased dopaminergic activation of D1Rs should lead to increased speed of trial engagement. Yet systemic D1R stimulation in

fact *increased* the average latency to re-engage on the subsequent trial [Figure 7a; significant main effect of drug: $F_{1.097,10.965} = 4.879$, $\eta_p^2 = .328$, $p = .048$; although no pairwise comparisons were significant; similarly, no significant drug*trial outcome interaction: $F_{1.075,10.753} = 2.750$, $\eta_p^2 = .216$, $p = .126$].

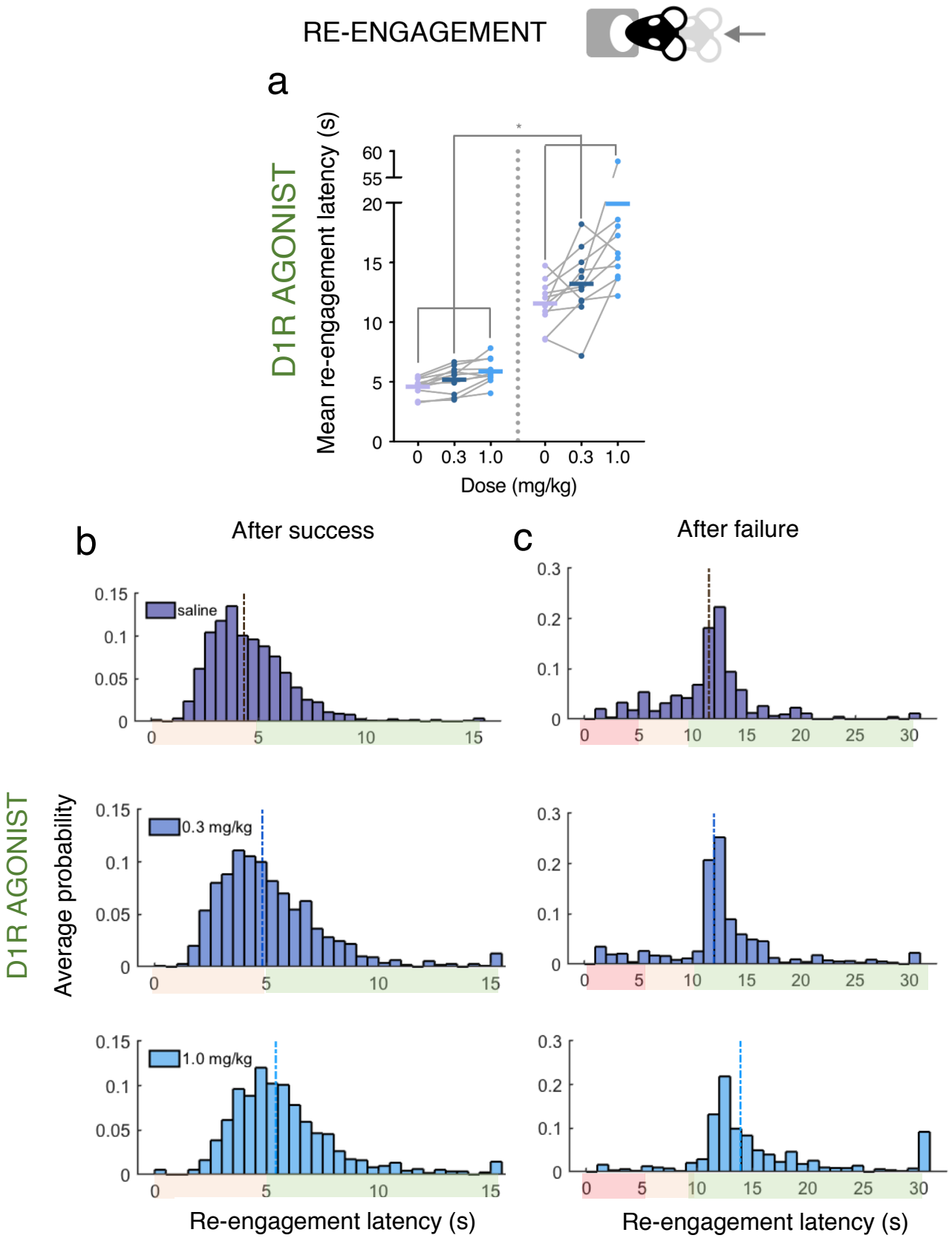


Fig. 7. Effect of D1R stimulation on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after both successful (left of dotted line) and failed (right of dotted line) trials increased with D1R stimulation. (b) Histograms of re-engagement latencies after success for each drug dose. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (0.5s bins; last bin contains all responses after 15s; dotted lines indicate medians). (c) Same as in (b) but after failure. Light red is the timeout period during which the houselight is on. (1s bins; last bin contains all responses after 30s). Significance bars indicate a main effect of drug, ** $p < .01$.

3.4. Systemic D1R stimulation promotes cued, not internally-driven, movements

We previously showed that the likelihood of an incorrectly initiated *cued* action was greatly increased by the D1R agonist. To understand whether this previous effect is specific to cued movement, we analysed aborted trials. These are defined as occasions when, after either successfully or unsuccessfully completing the preceding trial, animals voluntarily enter the nose poke to start a new trial when the nose poke is active, but do not remain for the duration of the precue period and instead prematurely leave the nose poke, thereby failing to initiate the subsequent trial.

Analysis of the proportion of aborted trials (Figure 8) showed that, whilst there was a higher rate of aborted trials after success than after failure [significant main effect of previous trial outcome: $F_{1,10} = 8.949$, $\eta_p^2 = .472$, $p = .014$], this measure was not altered by the drug [no main effect of drug, nor interaction: all Fs < 1.4, ps > .2]. This suggests that it is *cued* behaviour that is promoted by the D1R agonist, rather than an increase in the initiation of movement *per se*.

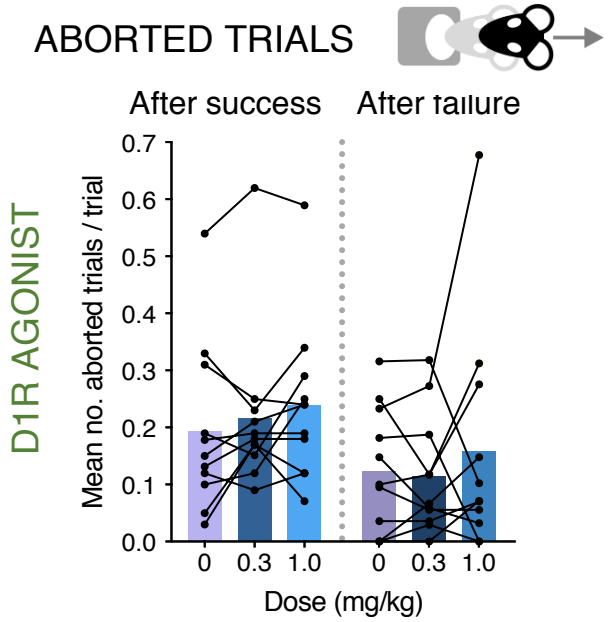


Fig. 8. Effect of D1R stimulation on mean rate of aborted trials (average number of aborted trials per trial) – when animals entered the nose poke to initiate a trial, but did not remain for the precue period and therefore failed to elicit a cue. The drug did not alter rate of aborted trials either after success (left of the dotted line) or after failure (right of dotted line).

D1R Antagonist

3.5. Systemic D1R blockade reduces success rate exclusively in Go trials

In a second experiment, we examined the effect of blocking D1Rs in the same Go/No-Go task. Converse to the effects seen with the D1R agonist, we predicted that blockade of D1Rs should prevent large effluxes of phasic dopamine from acting at the MSNs, leading to a reduction in goal-directed responding to such cues.

Indeed, as can be observed in Figure 9a, there was a significant, dose-dependent reduction in success rate after D1R antagonist administration in Go trials [significant main effect of drug: $F_{2,24} = 7.015$, $\eta_p^2 = .369$, $p = .004$; significant pairwise comparisons: vehicle vs. high dose, $p = .029$]. However, whilst animals' success rates were again

greater on large reward trials, the effect of the drug was not reward-mediated [significant main effect of reward: $F_{1,12} = 6.849$, $\eta_p^2 = .363$, $p = .023$; no drug*reward interaction: $F_{2,24} = 1.025$, $\eta_p^2 = .079$, $p = .374$].

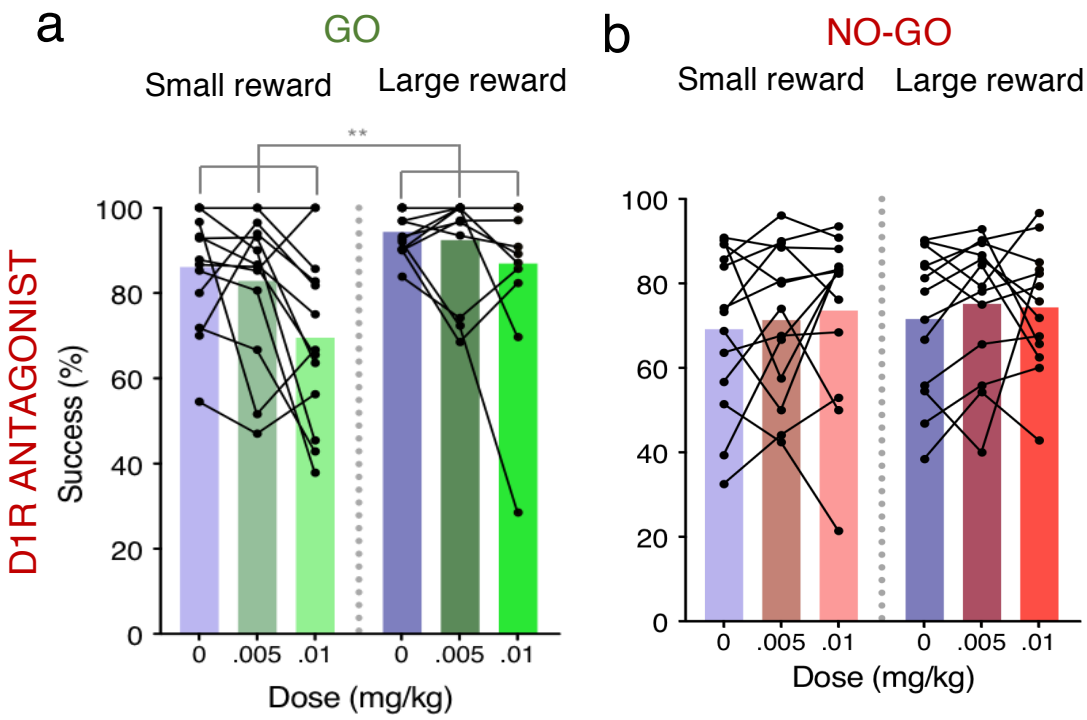


Fig. 9. Effect of systemic D1R blockade (SCH 23390) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) D1R blockade reduced success rate in both small and large reward Go trials. (b) D1R blockade had no effect on success rate in No-Go trials. Significance bars indicate main effect of drug, ** $p < .01$.

We also investigated whether the D1R antagonist altered success rate in No-Go trials. The FCV study in Chapter 3 found no phasic dopamine release to either of the No-Go cues, lending a prediction that blocking D1Rs should have little to no effect. This was the case, with No-Go success rate being unchanged at either dose of the D1R antagonist [Figure 9b; no significant main effect of drug or interaction with drug: $F < 0.6$, $p > .5$].

The reduced success rate in Go trials could be either due to an increase in Wrong Lever errors, or an increase in Response Omission errors. Whilst the probability of Wrong Lever errors did not change with the D1R antagonist on board [Figure 10a; no main effect or interaction with drug: all Fs < 1.1, p > .3], Response Omission errors dose-dependently increased substantially in both small and large reward Go trials [Figure 10b; significant main effect of drug: $F_{2,24}$ = 7.881, η_p^2 = .396, p = .002; significant pairwise comparisons: vehicle vs. high dose, p = .018; no drug*reward interaction: $F_{2,24}$ = 1.956, η_p^2 = .140, p = .163].

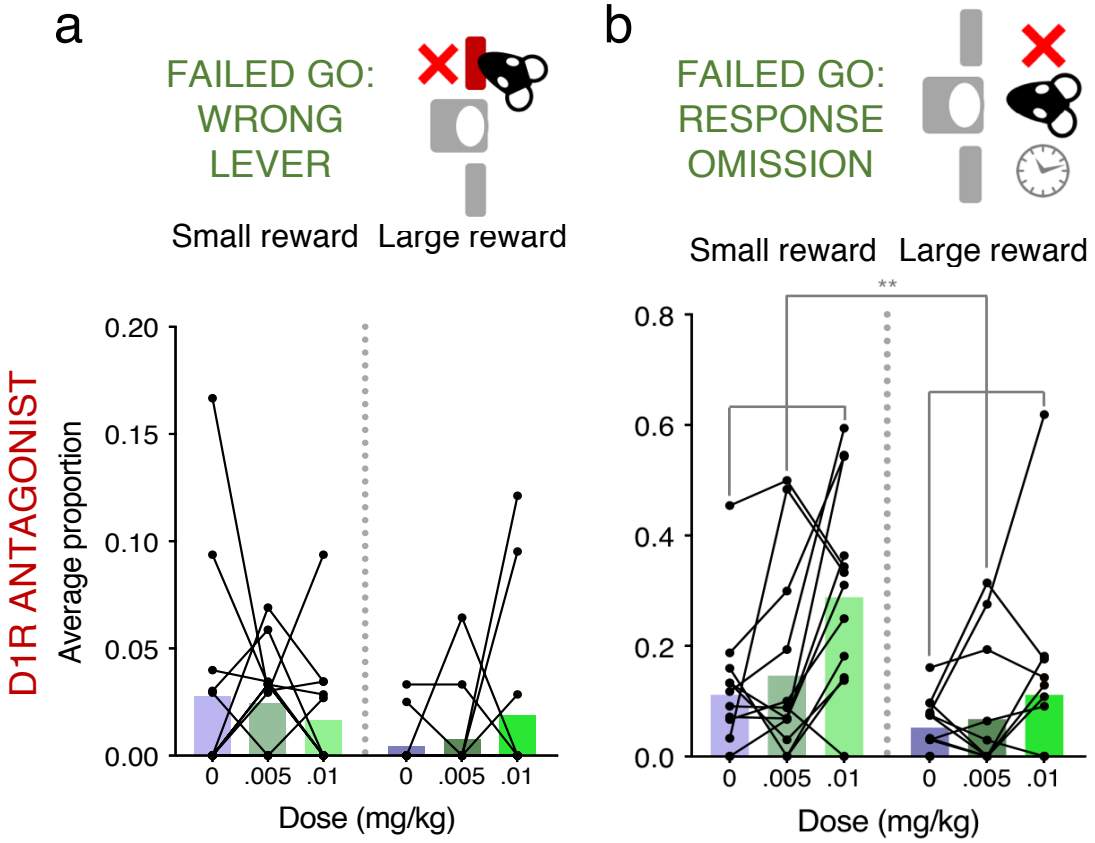


Fig. 10. Effect of D1R blockade on the average proportion of Go small on large trials where animals made (a) a wrong lever press or (b) omitted response entirely during the 5s cue period. Significance bars indicate a main effect of drug on the proportion of trials with response omissions, ** < .01.

Whilst the *number* of errors in No-Go trials did not change with the D1R antagonist, it might have been that the *distribution* of errors in terms of how quickly they were made was altered by the drug. However, this was not the case (Figure 11) [no main effect of drug, nor drug interactions with No-Go period or reward: all Fs < 0.5, ps > .6]. Therefore, animals' ability to withhold movement for reward was not impacted by D1R blockade, both in terms of number of errors and in terms of when errors were made.

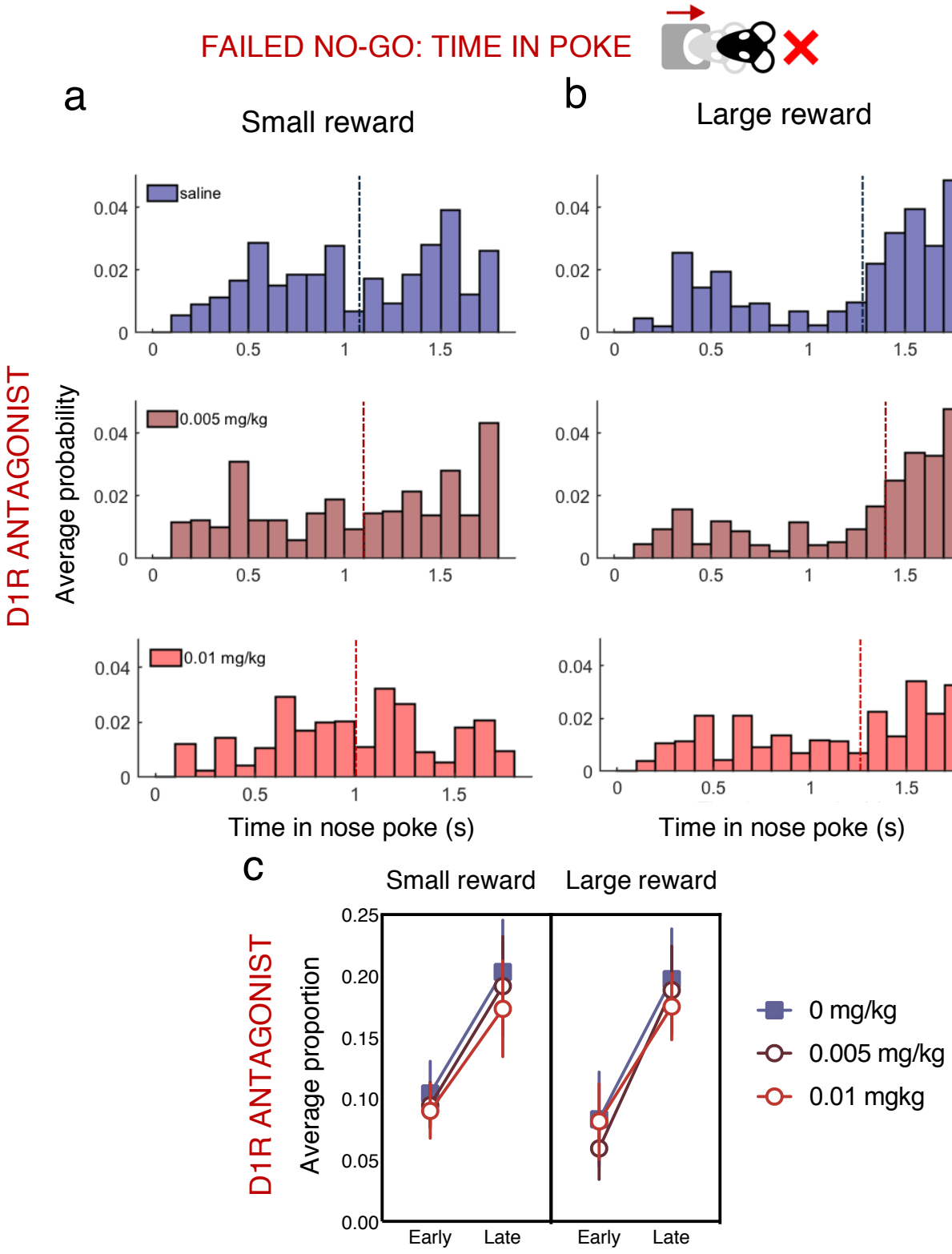


Fig. 11. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with D1R blockade, calculated as probability over all head exit times, both successful and failed, for each drug dose. Last bin contains all values > 1.7s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.005 mg/kg, red circle: 0.01 mg/kg. There was no effect of drug on these measures.

3.6. *Systemic D1R blockade slows appropriate cue-driven action initiation in Go trials*

We previously showed that animals had reduced Go trial success with a D1R antagonist on board. Similarly, in successful Go trials, the drug dose-dependently slowed rats' ability to initiate an action [Figure 12a; significant main effect of drug: $F_{2,24} = 8.607$, $\eta_p^2 = .418$, $p = .002$; drug*reward interaction: $F_{2,24} = 2.903$, $\eta_p^2 = .195$, $p = .074$]. Yet whilst appropriate action initiation was slowed by the D1R antagonist in successful Go trials, rats initiated action in successful No-Go trials after cue *offset* with similar latencies to baseline [Figure 12b; no main effect or interaction with drug, all Fs < 0.7, ps > .5]. Therefore, only cue-driven action initiation in Go trials was altered.

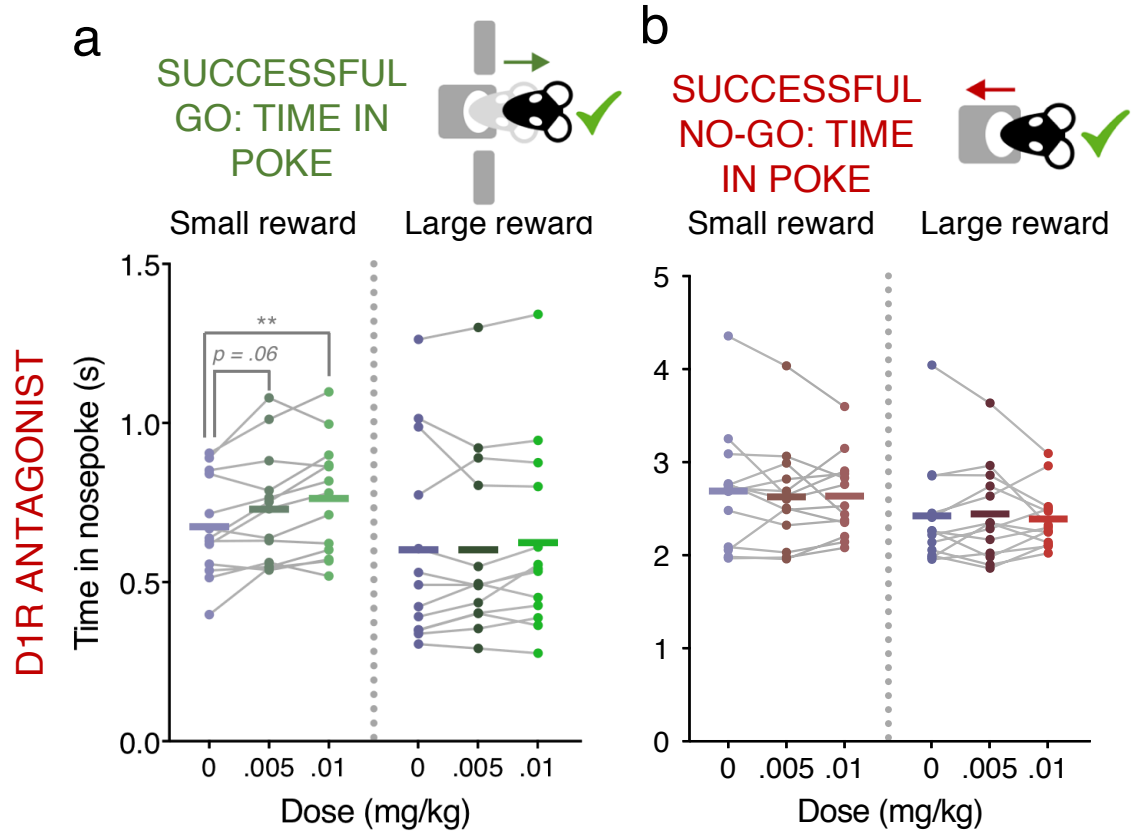


Fig. 12. Mean time spent in the nose poke (s) from cue onset in successful trials on the D1R antagonist. (a) Mean time in nose poke for successful Go trials, particularly when a small reward was on offer, was increased by the drug. (b) Mean time in nose poke for successful No-Go trials was unaltered. Significance bars indicate pairwise comparisons, * p < .05, ** p < .01, Sidak-corrected.

To further understand whether different Go trial head exit times were differentially affected by the D1R antagonist, we compared the 0.05, 0.45, and 0.95 quantiles from the CDF distributions (Figures 13a and 13b) alongside reward size and drug dose using a three-way within-subjects ANOVA. This revealed a near-significant three-way interaction between period, drug dose, and reward size on offer [$F_{4,44}$ = 2.544, η_p^2 = .188, p = .053] driven by a change in the *slowest* head exit times in small reward trials [pairwise comparisons: small reward late period, vehicle vs. low dose: p = .029, vehicle vs. high dose: p = .048; all other pairwise comparisons: p > .1]. This is further confirmed by separate analysis of small and large reward trials [small reward trials: significant drug*period interaction: $F_{4,44}$ = 2.904, η_p^2 = .209, p = .032; large reward trials: no significant drug*period interaction or main effect of drug: all Fs < 0.7, ps > .6].

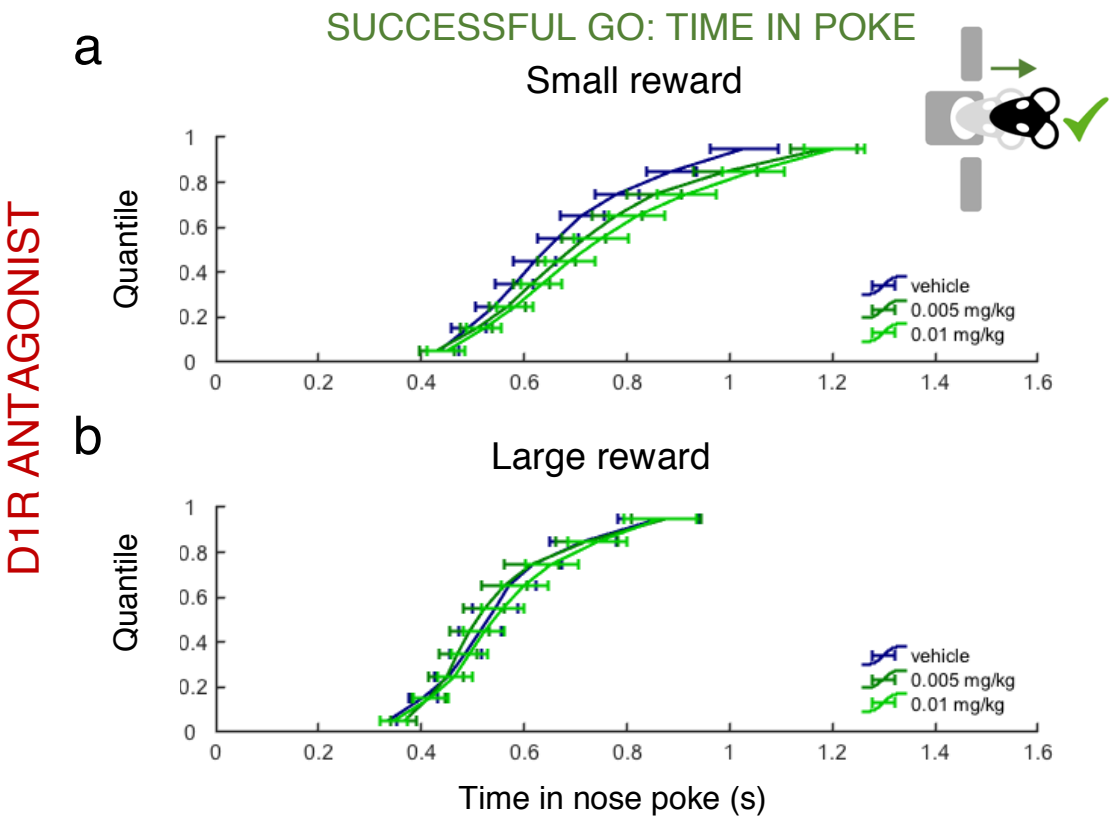


Fig. 13. CDFs for time in nose poke from cue in successful Go trials with D1R blockade. Lines on histograms show median time in nose poke (s), bin size 0.05s. (a) CDF for small reward trials, showing that the slowest half of responses were most increased by the high dose of the D1R antagonist. (b) Same as in (a) but for large reward trials, where there is little effect of the drug.

3.7. *Systemic D1R blockade reduces energising of ongoing goal-directed behaviour*

Earlier in this chapter, it was shown that stimulating D1Rs instead *slowed* travel times. Therefore, if the D1R antagonist works in the opposite direction to the agonist, then it should in fact speed these latencies.

Overall, animals were faster to travel with the promise of a larger reward [Figure 14; significant main effect of reward: $F_{1,12} = 18.416$, $\eta_p^2 = .605$, $p = .001$; small reward trials]. Blocking D1Rs at the highest dose of the drug slowed these latencies regardless of reward size on offer [significant main effect of drug: $F_{2,24} = 13.226$, $\eta_p^2 = .542$, $p = .173$; significant pairwise comparisons: vehicle vs. high dose, $p = .004$; low dose vs. high dose, $p = .001$; no drug*reward interaction: $F_{2,24} = 1.302$, $\eta_p^2 = .098$, $p = .290$].

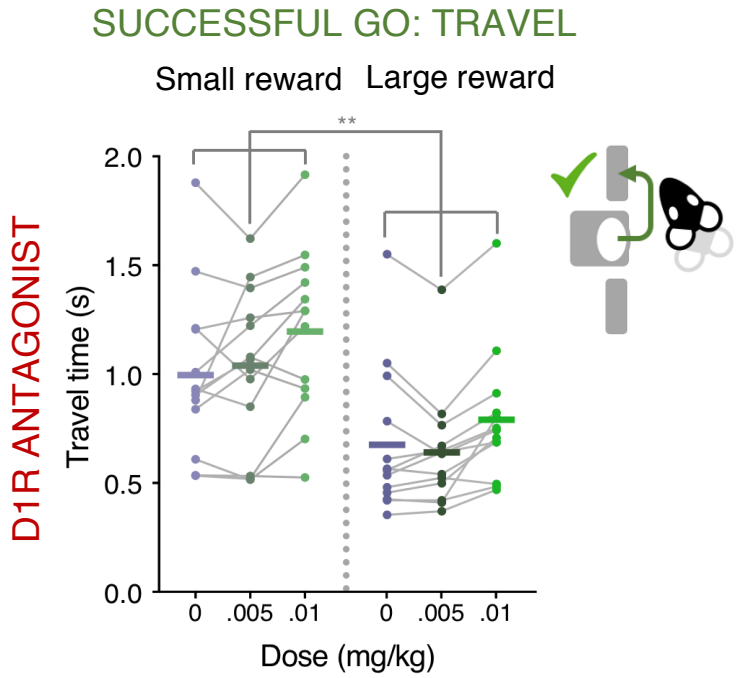


Fig. 14. Mean travel time from nose poke exit in successful Go trials was significantly increased by D1R blockade when both a small or large reward was on offer. Significance bars indicate main effect of drug, ** $p < .01$.

We again compared the 0.05, 0.45, and 0.95 quantiles of the CDF distributions (Figures 15a and 15b) using a three-way within-subjects ANOVA. We replicated the effect of drug on travel time latencies [significant main effect of drug: $F_{2,24}$ = 4.041, η_p^2 = .252, p = .031], but this effect was not mediated by either reward size on offer or latency period [no significant interactions: all Fs < 1.2, ps > .3].

Animals were not only slower to travel to the lever in successful Go trials with a D1R antagonist on board, but also took more time to complete the two lever presses required [significant main effect of drug: $F_{2,24}$ = 9.395, η_p^2 = .439, p = .001; significant pairwise comparisons: vehicle vs. high dose, p = .025; low dose vs. high dose, p = .008; no significant main effect of reward or interaction: all Fs < 3.2, ps > .08].

It is important to note that not *all* motoric behaviours were affected. Animals were able to retrieve reward from the food magazine at speeds no different to baseline levels in both Go and No-Go trials [no main effects of drug or reward, or interactions: all Fs < 3.2, ps > .1].

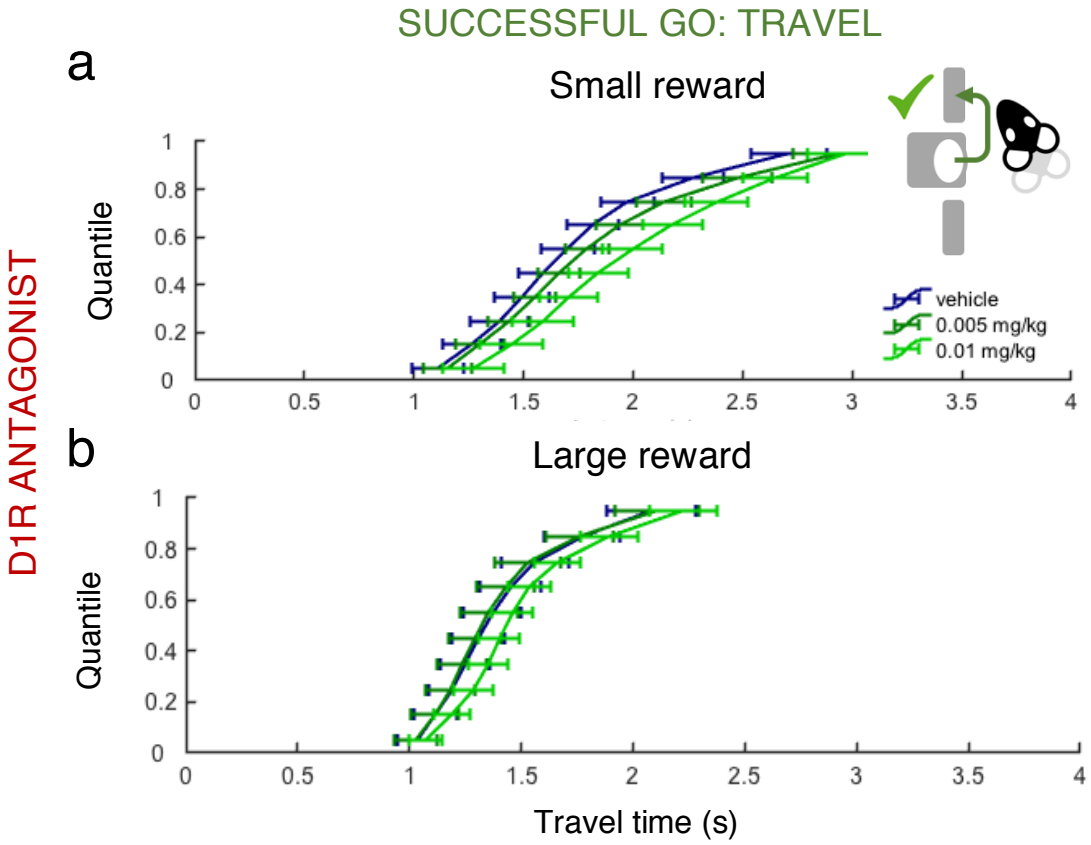


Fig. 15. CDFs for travel time from nose poke exit in successful Go trials with D1R blockade. (a) CDF showing slowing effect of SCH 23390 on travel times in small reward trials. (b) Same as in (a) but for large reward trials, where there again is a slowing effect of the drug.

As well as slowed travel times and lever response latencies, the D1R antagonist at both doses also increased mean latency to re-engage in the task [Figure 16a; main effect of drug, $F_{1.002,12.026} = 7.580$, $\eta_p^2 = .424$, $p = .003$; significant pairwise comparisons: vehicle vs. low dose, $p = .007$; vehicle vs. high dose, $p = .047$; low dose vs. high dose, $p = .057$]. The high dose of the drug had a particularly strong effect on re-engagement after success (Figure 16b), whilst both doses slowed this latency after failure [Figure 16c; significant drug*trial outcome interaction: $F_{1.002,12.018} = 5.332$, $\eta_p^2 = .308$, $p = .039$; after success pairwise comparisons: vehicle vs. low dose, $p > .1$; vehicle vs. high dose, $p = .01$; after failure: vehicle vs. low dose, $p = .005$; vehicle vs. high dose, $p = .068$].

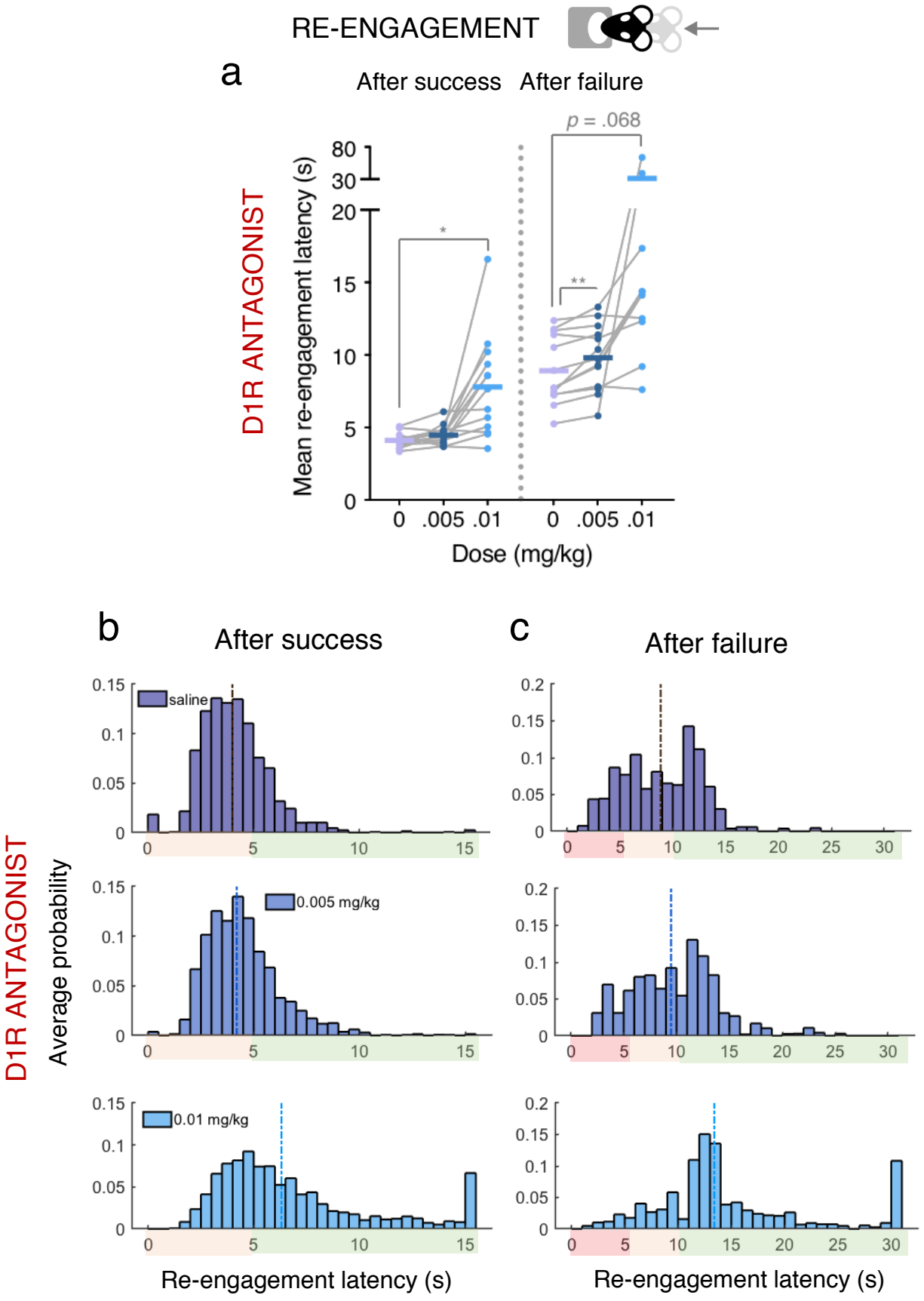


Fig. 16. Effect of D1R blockade on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after successful trials (left of dotted line) and after failure (right of dotted line) increased with D1R blockade. (b) Histograms of re-engagement latencies after success for each drug dose. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (0.5s bins; last bin contains all responses after 15s; dotted lines indicate medians). (c) Same as in (b) but after failure. Light red is the timeout period during which the houselight is on. (1s bins; last bin contains all responses after 30s). Significance bars indicate pairwise comparisons, * $p < .05$, ** $p < .01$, Sidak-corrected.

3.8. Systemic D1R blockade does not promote internally-driven, movements

We previously showed that the likelihood of an incorrectly initiated cued action on No-Go trials was unchanged by D1R blockade. Similarly, the D1R antagonist did not alter inappropriate *uncued* action initiation [Figure 17; no main effect of drug nor interaction with preceding trial outcome: all *F*s < 0.7, *p*s > .5].

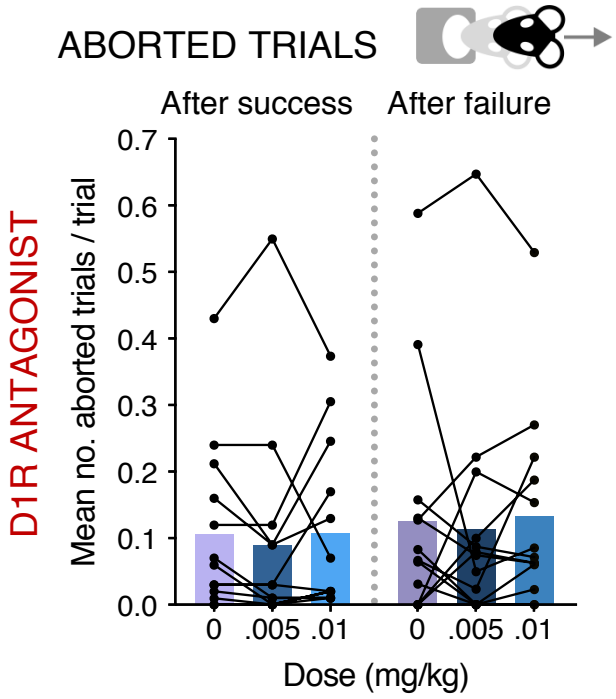


Fig. 17. Effect of D1R blockade on mean rate of aborted trials (average number of aborted trials per trial) – when animals entered the nose poke to initiate a trial, but did not remain for the precue period and therefore failed to elicit a cue. The drug did not alter rate of aborted trials either after success (left of the dotted line) or after failure (right of dotted line).

4. Discussion

Determining the respective contributions of reward processing and movement to the dopamine signal has been a challenge in the field. In Chapter 3, it was shown that phasic dopamine release in the core of the nucleus accumbens reflects an interaction between reward prediction *and* whether or not goal-directed action has to be initiated in order to garner reward. However, it was not possible to determine whether such fluctuations in dopamine levels were driving or reporting action. Therefore here, dopamine D1-type receptors, which are thought to be most sensitive to such phasic changes in dopamine, were either stimulated or blocked using systemic pharmacology in the same Go/No-Go task to address this question. In support of the idea that dopamine acting through D1Rs causally promotes action initiation, it was found that pharmacological stimulation of D1Rs caused increases in fast, inappropriate, impulsive responses on No-Go trials, and blockade slowed action initiation on Go trials. However, the agents' effects were not limited to cue onset, but also influenced the selection and ongoing execution of action for reward. Table 2 contains a summary of these findings, which are discussed in more detail subsequently.

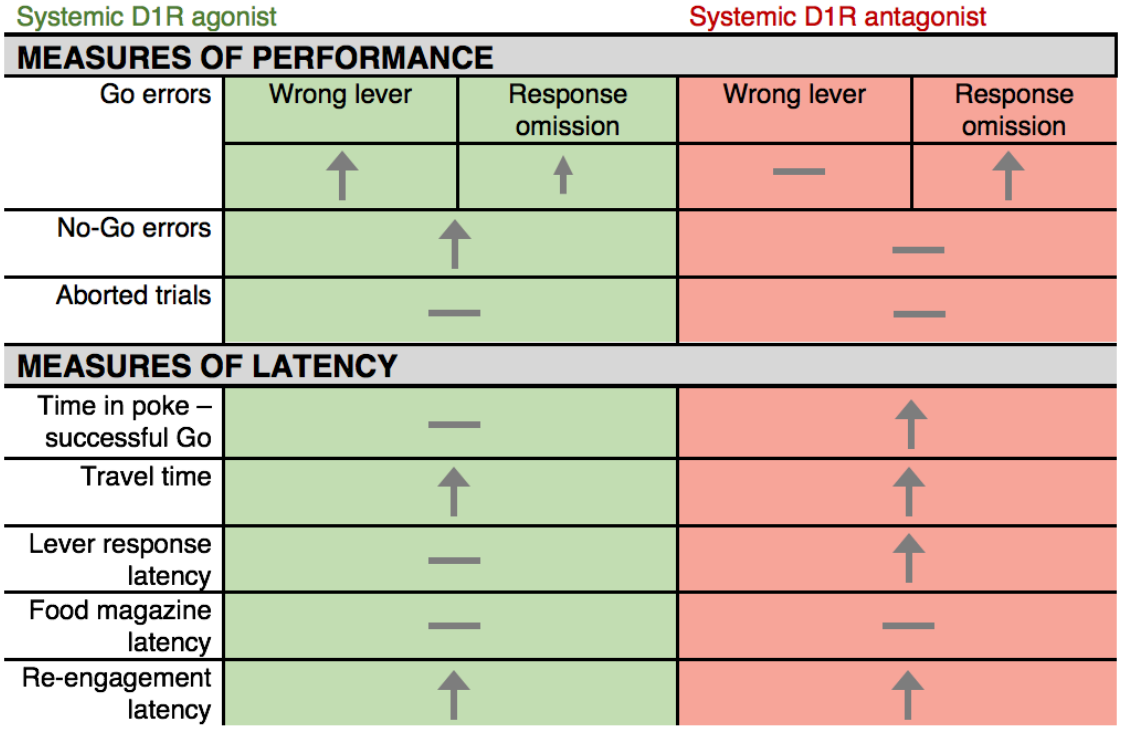


Table 2. Summary of effects of both the D1R agonist (green) or D1R antagonist (red) on measures of performance and measures of latency in the Go/No-Go task. Large arrows indicate a strong effect, small arrows indicate smaller or more specific effects.

4.1. D1Rs play a causal role in initiating action for reward

A pure reward prediction error account of the phasic dopaminergic signal might suggest that stimulating the receptor family purportedly most sensitive to such signals might either have little effect on performance, as the task is well learned, or may actually *improve* overall performance if it is signalling future reward. Instead, the D1R agonist *reduced* success rate in both Go and No-Go trials. In No-Go trials, this meant that the D1R agonist decreased animals' ability to withhold movement for reward. This was the case both when a small or large reward was on offer, although the effect was greatest in the latter trials. This suggests that, rather than passively receiving a signal of unexpected cue value, activation of D1Rs *promotes* cued action initiation, even if it is inappropriate. This is in agreement with findings using impulse control tasks such as the 5-choice that also show a decreased ability to inhibit goal-

directed action with D1R stimulation (Pezze et al., 2007). However, this effect was only seen in the period immediately after cue onset; if the rats were able to withhold movement initially, they were then no more likely to make an impulsive response in the second half of the No-Go period as control animals. Additionally, there was no increase in number of aborted trials (where the animal came out of the nose poke during the pre-trial period before a cue was presented) showing that the rats were able to restrain responding in the absence of external cues.

Of course, one drawback with the current study is that pharmacological manipulations are not brief but longer lasting, and so it is debatable whether a D1R agonist can 'mimic' a prediction error. Indeed, artificially prolonged dopamine signalling induced by pharmacological manipulation can have protracted effects on the neurons expressing the target receptors, for instance by causing homeostatic responses that constrain the plasticity mechanisms purported to encode learning via a reward prediction error (Beaulieu & Gainetdinov, 2011).

Therefore, it is useful to compare the effects of D1R stimulation with *blockade* of D1Rs, which – although also more prolonged manipulation – may prevent the post-synaptic effects of phasic dopamine release elicited by a prediction error in a comparable way to a briefer manipulation. As predicted, the D1R antagonist did not affect success rates in No-Go trials, and nor did it change the amount of time animals spent in the nose poke during such trials. This lack of an effect on No-Go trials is consistent with the FCV results from Chapter 3; there, the No-Go cue did not elicit a phasic dopaminergic response, and so blockade of the receptors that would be sensitive to

such a response should not alter performance as their activation is not necessary for success in such trials. Moreover, this voltammetry study also showed that animals are able to initiate actions without dopamine levels changing so long as they were not appropriate reward-seeking movements.

This is contrary to previous studies that have shown a reduction in premature responding with use of the same pharmacological agent, SCH 23390 (Pattij et al., 2007). However, that study used the 5-choice task which allows measurement only of a reduction in premature execution of one type of behaviour – nose poking – in contrast to the current study, in which movement of any kind during a No-Go trial is considered premature. Therefore, these are fundamentally different forms of behavioural inhibition that, presumably, would elicit different phasic dopamine signals such that less D1R activation is able to reduce premature responding in one but not the other.

4.2. D1Rs play a causal role in ongoing performance of action for reward

As seen in Chapter 3, phasic dopamine release is not only seen transiently at cue presentation, as has previously been observed during Pavlovian conditioning (Day et al., 2007), but is instead sustained during correct goal-directed action execution. Indeed, there is evidence that D1R blockade reduces the degree of vigour of such responding across several tasks, both cue-driven (du Hoffmann & Nicola, 2014a; Keeler et al., 2014; White & Rebec, 1994) and internally-driven (Yohn et al., 2015). Such evidence is also replicated here, with the D1R antagonist slowing not only action initiation, but also action execution for reward in Go trials – from the travel time, to

the lever response latency, and finally re-engagement latency, as well as reducing the overall likelihood of animals executing lever presses, consistent with evidence that dopamine is required for the control of vigorous movement for reward (Panigrahi et al., 2015).

This complements Bogacz's (2017) recent formulation of D1Rs or 'Go' neurons encoding payoffs, as blockade of D1Rs but not D2Rs results in the overall utility of all actions being reduced such that inaction is now more favourable over action. Berke (2018) has also argued that such effects can in fact parse with a reward prediction error framework, if dopamine antagonists are thought to act as a constant negative prediction error, reducing the value of reward-driven actions towards zero and producing 'extinction-like effects'. However, notably, it was not *all* action latencies that were slowed under D1R blockade as there was no change in reward collection latencies. This suggests that the ability of D1R activation to mediate the vigour of an action might be dependent on the predicted future gain of that action.

If this is the case, D1R stimulation should produce the converse effect in Go trials, increasing the value of goal-directed action and thus promoting their vigour. Yet here, the D1R agonist also *slowed* both travel time and re-engagement latencies. This slowing may have occurred for several reasons. First, sustained enhancement of dopamine receptor activity via pharmacological intervention may have longer-term, counteractive effects on dopamine release, although some have shown that this is predominantly a D2R-mediated effect (Boyar & Altar, 1987). Second, as discussed in more detail below, D1R stimulation also influenced the process of action selection,

which, if animals are taking longer paths to reach their goal, might also lead to slowed action execution latencies as a by-product. The latter could be determined by video analysis, which was not possible to achieve in the timeframe of this thesis. Thus whilst both stimulation and blockade of D1Rs had a similar effect on the ongoing performance of goal-directed actions, the mechanisms underlying these behavioural changes may differ.

4.3. D1R stimulation biases action selection towards a default preferred option

During learning, D1R activation has been shown to lead to the repetition of a specific action over other actions (Kravitz, Tye, & Kreitzer, 2012). However, dopamine is thought to play less of a role in action selection in well-learnt contexts as actions become more habitual (Wickens, Horvitz, Costa, & Killcross, 2007). Additionally, whilst there is building evidence linking dopamine with the initiation of action, the direct contribution of dopamine to action selection is less well established. Instead, both the activity of midbrain dopamine neurons (Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006) and dopamine release in the striatum (Hollon, Arnold, Gan, Walton, & Phillips, 2014) appear to report the value of an already selected action made elsewhere in the brain, rather than performing the necessary multiplexing across several options required of an action selection signal.

Yet in the studies described in this chapter, the D1R agonist had a surprising effect on action selection, causing animals to more often respond at the large reward lever

during small reward trials even though the paradigm here does not allow for free choice and therefore such action selection ultimately led to no reward. This is reflective of recent work recording *dorsal* striatal dopamine in a task where animals were required to internally switch from the selection of one action to another dependent on the duration of an interval (Howard et al., 2017). Of relevance to the current study, these authors showed a dip in dopamine levels when animals selected the contrary option to their default choice, whilst optogenetic stimulation of nigral dopamine neurons biased animals towards this default. Here, sustained activation of D1Rs could be acting in a similar manner, preventing any dips of dopamine from unlocking alternative actions, ultimately driving the animals to carry out a default movement that had previously led to a large reward. Indeed, the slowed travel time and re-engagement latencies with the D1R agonist on board may also be reflective of a disruption to action selection processes, as the range of possible actions at these timepoints is large, leading to increased competition and slowing of responses.

4.4. Pharmacological considerations

While the agents used here were selected both for their specificity and their widespread use in the literature, it is worth briefly discussing any effects they might have outside of those intended in this study. Along with its localisation to the D1R family (including both D1 and D5 receptors), the D1R agonist SKF 81297 has also been found to act as a partial agonist at D1-D2 receptor heteromers (Rashid et al., 2007) which is worth noting as up to 15% of dorsal striatal MSNs may co-express both receptor subtypes (Bertran-Gonzalez et al., 2008; Perreault et al., 2011). As the drug was applied systemically in this study, it therefore may have been active at these

heteromers in addition to D1 and D5 receptors. Whilst the behavioural implications of such activation are not yet known in detail, this issue is overcome in the subsequent chapter, in which this pharmacological agent is locally infused into the ventral striatum, where co-expression is less pervasive.

There is also *in vitro* evidence that the D1R antagonist used, SCH 23390, binds with relatively high affinity to the 5-HT_{2C} and 5-HT_{1C} serotonergic receptors (Bischoff, Heinrich, Sonntag, & Krauss, 1986), although particularly high doses are required to achieve similar affinity *in vivo*. However, this is an unlikely cause of the effects seen here as there should be a resultant increase rather than decrease in impulsive responding, which is inconsistent with the data presented (Winstanley, Theobald, et al., 2004).

4.5. Regional specificity of the presented effects

All of the data presented here were gathered from systemic manipulations. An obvious drawback to such an approach is that D1Rs across the whole brain were affected, making it difficult to ascertain the locus of the results seen. As discussed in Chapter 1, exactly *where* dopamine is acting can have a significant impact on how its effects are translated into behaviour.

For example, there are D1Rs in the prefrontal cortex that are purported to play a role in working memory function, as ascertained using delayed response tasks. Here, D1R antagonism has been shown to disrupt working memory performance (Sawaguchi & Goldman-Rakic, 1994). Similarly, Chudasama & Robbins (2004) used a modified

version of the 5-choice paradigm to include a memory-dependent component, and found that levels of D1R stimulation that could improve attentional accuracy could also either disrupt or facilitate short-term working memory dependent on the length of the delay. However, in the Go/No-Go paradigm, cues were presented continuously throughout the trial to ensure a low working memory load, and so the effect seen here are unlikely to be attributable to working memory modulation.

Perhaps more relevantly, the importance of D1R activation to attentional processes has also been demonstrated; infusion of D1R-specific pharmacological agents can either improve (in the case of an agonist) or worsen (in the case of an antagonist) performance in the 5-choice serial reaction time task (Granon et al., 2000). It is not possible to rule out attentional effects in the Go/No-Go task with the current data set, but this is overcome to some extent in the following chapter by use of a local infusion approach into a brain region not traditionally associated with attentional modulation. Moreover, the attentional load was again minimised by the long presentation times of the cues. By contrast, classic sustained attention tasks in rodents tend to use very short cue presentations.

Finally, and perhaps most importantly, a lack of regional specificity means that the effects here cannot be ascribed exclusively to dorsal or ventral striatum. This is likely an important distinction; for example, whilst vigour of action execution has been found to be modulated by D1R stimulation or blockade in the dorsomedial striatum (Eagle et al., 2011), this same effect has not always been replicated in the nucleus accumbens (Pattij et al., 2007). Similarly, the action selection effect seen here with

D1R stimulation may be mediated by the dorsal rather than the ventral striatum (Parker et al., 2016b), which has repercussions on the interpretation levied from such data when attempting to contrast prediction error and behavioural activation accounts of dopaminergic function (Leventhal et al., 2014). Due to this thesis' focus on *mesolimbic* dopamine, the experiments described in the following chapter overcome these concerns by locally infusing the same D1R receptor ligands used here into the nucleus accumbens core.

4.6. Summary

By manipulating the level of activity of D1Rs, we found that stimulation of these receptors promoted cue-potentiated action whilst their blockade slowed cue responding, consistent with evidence that they are integral for the expression of cue-driven incentive motivation (Yun, Nicola, & Fields, 2004). Surprisingly, we also found that D1R stimulation slowed the ongoing performance of action and re-engagement in the task, possibly related to disruption of action selection processes (Howard et al., 2017). For a greater understanding of these effects, we locally infused these D1R compounds into the NAc core in the following chapter.

5 Nucleus accumbens core D1-like receptors shape cue-promoted action initiation for reward

Chapter abstract:
We have previously shown that phasic mesolimbic dopamine release is shaped both by reward prediction and action required to garner reward, and that such action initiation and ongoing performance of action is causally mediated by the degree of activation of receptors thought to be sensitive to that phasic release. However, these causal effects were not anatomically localised, and so here we locally infused a D1R agonist or antagonist directly into the NAc core. We found that stimulation of these receptors drove cue-promoted action initiation and their blockade impeded instrumental action for reward, but that neither manipulation modulated ongoing performance of correct action. Therefore, NAc core D1Rs are integral specifically in initiating action in response to value-laden cues.

For the studies presented in this chapter, Laura Grima and Mark Walton designed and planned the experiments, Laura Grima and Marios Panagi conducted surgeries, and Laura Grima collected, analysed, and presented the data with input from Mark Walton.

1. Introduction

Dopamine's role both in reward prediction and in driving goal-directed behaviour is well established. However, understanding how a single neurotransmitter can convey two such different signals has been a challenge. One possible account is that dopamine might be performing different functions in different regions; for example, there is evidence that ventral striatal dopamine is more involved in learning of reward-relevant information, whilst dorsal striatal dopamine is central in movement, with depletion of dopamine in this region contributing to the pathology of Parkinson's Disease (Tritsch & Sabatini, 2012). More recently, the dorsal striatum has also been considered important in acting for reward (Parker et al., 2016). Yet in Chapter 3, using FCV to measure subsecond dopamine release in the NAc core, we found evidence that phasic dopamine was shaped by action initiation as well as by encoding of future reward.

However, whether dopamine was causally promoting goal-directed behaviour could not be determined using this approach. Therefore, in Chapter 4, we systemically activated or blocked D1-like receptors (D1Rs) – the receptor subtype likely most sensitive to phasic changes in dopamine levels (Dreyer et al., 2010; Richfield et al., 1989) – within the same Go/No-Go task in order to understand their contribution (if any) to motivated behaviour. We showed that level of D1R activation has profound influences on ability to initiate and perform ongoing action for reward, but that ability to *withhold* action is not dependent on D1R activation, reflecting the

attenuated dopamine release seen in No-Go trials in the FCV study. Intriguingly, action *selection* was also biased with overstimulation of these receptors.

Whilst these experiments in Chapter 4 further bridge the gap between the studies of reward prediction and movement for reward, their lack of regional specificity means that we do not yet know whether dopamine fluctuations in the NAc core as seen in Chapter 3 causally drives action for reward. Dopamine release in this region has been associated with the reward prediction error hypothesis (Day et al., 2007). A related reinforcement learning theory conceives of the striatum as an actor-critic model (Joel, Niv, & Ruppin, 2002), with the dorsal striatum aiding learning of actions to maximise future reward, whilst the ventral striatum predicts – or critiques – the consequences of this policy (O'Doherty et al., 2004). In this framework, the ventral striatum as critic is *passively* updating the weighted sum of future rewards by comparing between predicted and actual rewards received, and not *actively* promoting action. Therefore, one possibility is that the influence of action on reward predictions errors observed with voltammetry actually reflects a report of the behavioural state of the animal, and is not directly involved in the promotion of reward-seeking actions.

However, studies that have manipulated the activation of dopamine receptors at longer timecourses have suggested that D1Rs in the NAc core are causally involved in driving motivated behaviour, both in terms of (Pattij, Janssen, Vanderschuren, Schoffelmeer, & van Gaalen, 2007; Pezze, Dalley, & Robbins, 2007) and invigorating correctly executed behaviour (Nowend, Arizzi, Carlson, & Salamone, 2001). In

contrast, manipulations of the adjacent dorsal striatum have comparatively little effect on premature responding (Agnoli, Mainolfi, Invernizzi, & Carli, 2013).

Importantly, tasks that require the cancellation of action are also less dependent on the accumbens (Eagle et al., 2011), suggesting that it is action *restraint* that might be mediated by dopamine in the accumbens. Yet we found no reduction in impulsive behaviour in No-Go trials with the systemic D1R antagonist in Chapter 4, contrary to other action restraint studies in the accumbens (Pattij, Janssen, Vanderschuren, Schoffelmeer, & van Gaalen, 2007; Pezze et al., 2007). This suggests that the form of action restraint required here is qualitatively different. Thus, the contribution of the D1R family in the NAc core when needing use cues to *withhold* action for future reward is not yet known.

1.1. Aims, approaches, and hypotheses

The experiments in this chapter aim to localise the effects seen with systemic manipulation of D1Rs in Chapter 4 to the NAc core, a region strongly associated both with prediction error encoding and with movement for reward. In particular, it is asked whether D1Rs in the accumbens play a similarly important role in initiating movement as shown previously, and whether their over-activation in this region also slows ongoing performance of behaviour and biases it towards previously highly rewarding responses.

To address these questions, rodents well-trained in the Go/No-Go paradigm were implanted with cannulae directed at the NAc core to allow for local infusion of a D1R agonist or D1R antagonist – the same agents as those used in Chapter 4.

Based on our own FCV and pharmacology studies, as well as evidence from the impulse control literature, we predicted that dopamine in the NAc core promotes motivated action such that infusion of a D1R agonist in this should replicate the increase in early incorrect responses in No-Go trials seen with the same systemic manipulation. Similarly, the antagonist should have little effect on these same trials.

Whilst the FCV studies revealed sustained dopamine release *throughout* action execution in correct Go trials, the systemic studies provided mixed evidence for the necessity of D1R activation in the ongoing performance of such behaviours, with *both* the D1R agonist and D1R antagonist slowing travel times and increasing Response Omission errors. Other studies have suggested that the ability to execute action for reward is indeed NAc core D1R-mediated (Nowend et al., 2001; Soares-cunha et al., 2016b) and, as this is consistent with our own FCV data, it is hypothesised that infusion of a D1R agonist into the NAc core specifically should *speed* progression throughout a trial, with the D1R antagonist producing the converse effect.

Finally, it is hypothesised that the observed action selection effect of the D1R agonist seen in Chapter 3 was due to dorsal, rather than ventral striatal modulation of these receptors (Parker et al., 2016). Therefore, infusion of the D1R agonist into the NAc

core here should not reproduce this effect, with dopamine receptors in this region being more important for *driving* rather than directing behaviour.

2. Methods

2.1. Animals

One cohort of 14 male Sprague Dawley rats was used for both experiments. These animals had previously been used for several systemic pharmacological experiments, including some of the data for Chapter 4, before being implanted for the infusion study discussed here. Data from one animal were excluded from both experiments due to misplaced cannulae. Data from a second animal were excluded from the local D1R antagonist experiment, as on the high dose of the drug the animal did not complete any trials. Thus the presented local D1R agonist data include an n of 13 animals, and the D1R antagonist data include an n of 12 animals. Rats were group-housed throughout training and after adequate recovery post-surgery.

Full details of the apparatus and behavioural training, as well as further details of the task, are described in Chapters 2 and 3. All local infusion procedures were carried out in well-trained rats that had achieved stable levels of performance. It should be noted that the performance of several animals did not reach pre-surgery levels even after extensive training, and were >2 standard deviations from average performance of the group. To avoid arbitrary exclusion criteria, these animals have not been excluded from analyses. Nonetheless, we have highlighted any particular analyses when these individuals have an unduly large influence on the statistics.

2.2. *Measures of performance and latency*

The measures of performance and latencies recorded in the Go/No-Go task are outlined in detail in sections 1.4 and 1.5 of Chapter 2, and follow the structure used in Chapter 4.

2.3. *Pharmacological compounds*

The pharmacological compounds used in this chapter were again the selective D1-like agonist, SKF 81297 hydrobromide, and the selective D1-like antagonist, SCH 23390 hydrochloride. For a full description of their properties, see section 2.1. of Chapter 4.

2.4. *Pharmacological challenges*

In the experiments described in this chapter, both drugs and a saline control were infused locally into the nucleus accumbens (NAc) core, in separate counterbalanced Latin square designs for each drug. A description of the surgical and infusion procedures, as well as detailed histology, are provided in section 5 and associated subsections in Chapter 2.

SKF 81297 was infused at either 0.4 μg/μl (low dose) or 4.0 μg/μl. As a volume of 0.5 μl was infused per hemisphere, consistent with many other studies in the field, in practice this resulted in an infusion of 0.2 μg or 2.0 μg per side. These concentrations were chosen to reflect doses used in studies of action inhibition (Winstanley et al., 2010 - although it should be noted that the agonist was infused into the orbitofrontal

cortex in this study), risky decision-making (Stopper, Khayambashi, & Floresco, 2013), and reversal learning (Haluk & Floresco, 2009). Others have infused partial D1R agonists into the accumbens in the 5-choice task using similar doses (e.g. Pezze, Dalley, & Robbins, 2007).

SCH 23390 was infused at either 0.2 μg/μl (low dose) or 2.0 μg/μl. As with the D1R agonist, a volume of 0.5 μl was infused per hemisphere, thus either 0.1 or 1.0 μg of drug was infused either side. These concentrations were also chosen to reflect those used in studies of impulsivity (Pattij et al., 2007), action inhibition (Eagle et al., 2011), risky decision-making (Stopper et al., 2013) and reversal learning (Haluk & Floresco, 2009).

2.5. Data analysis approaches

A full discussion of the statistical approaches used in the analyses of the pharmacology data presented here can be found in section 2 of Chapter 2. The approach mirrored that taken in Chapter 4 for the systemic administration.

3. Results

NAc core D1R Agonist

3.1. NAc core D1R stimulation reduces success rate exclusively in No-Go trials

In Chapter 4, systemic administration of a D1R agonist decreased success rate in both Go and No-Go trials. Here, although animals were again better at executing a Go

response with promise of a larger reward [Figure 1a; significant main effect of reward: $F_{1,12}$ = 6.091, η_p^2 = .337, p = .030], the D1R agonist, when infused into the NAc core, had no effect on success rate in Go trials [no main effect of drug or interaction: all Fs < 2.3, ps > .1].

By contrast, when instructed to withhold action on No-Go trials, animals were no better at doing so with the promise of a larger reward [Figure 1b; no main effect of reward: $F_{1,12}$ = 0.539, η_p^2 = .043, p = .477]. However, replicating the systemic study, infusion of the D1R agonist dose-dependently reduced success rate in No-Go trials [significant main effect of drug: $F_{2,24}$ = 8.459, η_p^2 = .413, p = .002; significant pairwise comparisons: vehicle vs. high dose, p = .007]. Unlike the systemic study, the effect here was not reward-mediated [$F_{2,24}$ = 0.854, η_p^2 = .066, p = .438].

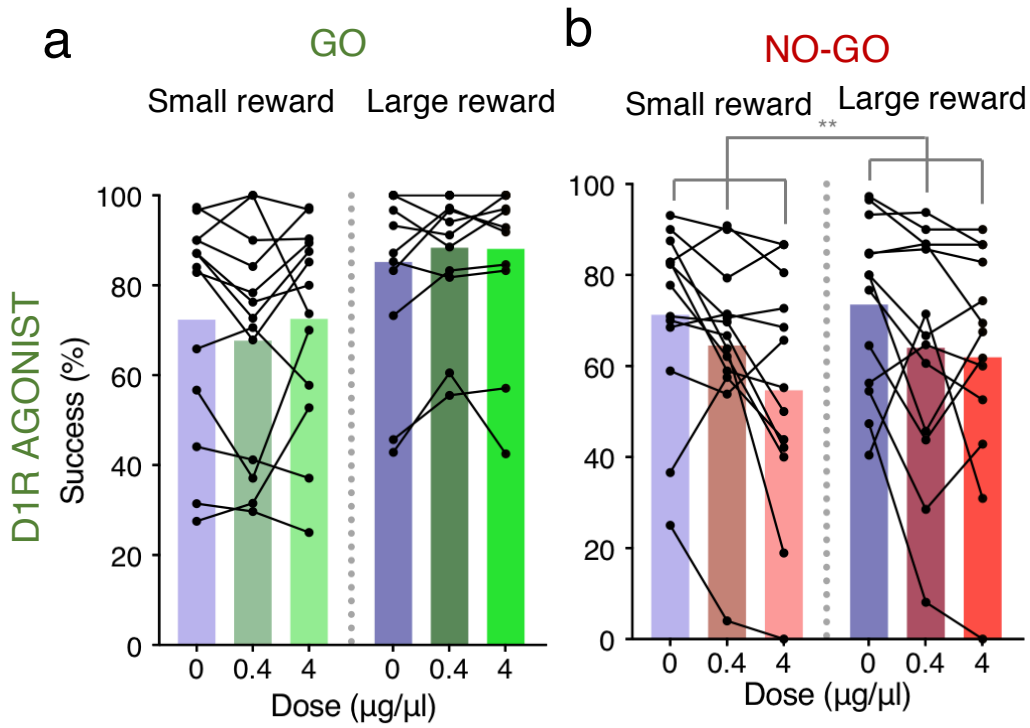


Fig. 1. Effect of local D1R stimulation (SKF 81297) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) D1R stimulation had no effect on success rate in Go trials. (b) D1R stimulation reduced No-Go success rate in both large and small reward trials. Significance bars indicate main effects of drug, ** p < .01.

To further understand the detrimental effect of the D1R agonist on success rate in No-Go trials, we again analysed when animals were leaving the nose poke during the No-Go period in failed trials. These distributions reflected the success rate, showing that animals were increasingly likely to leave early in the No-Go cue period regardless of whether the size of the reward on offer was small (Figure 2a) or large (Figure 2b).

To more formally quantify this, average proportions of head exit times were divided into 'early' (< 800ms) or 'late' (> 800ms) errors (Figure 2c). A three-way within subjects ANOVA with reward (small and large), drug (vehicle, low dose, and high dose), and error period (early and late) revealed that animals were no more likely to leave earlier in the No-Go period when the reward on offer was small [no significant main effect of reward: $F_{1,12}$ = 0.540, η_p^2 = .043, p = .476]. Additionally, animals were generally more likely to leave late in the error period [significant main effect of period: $F_{1,12}$ = 7.107, η_p^2 = .372, p = .372]. When including all animals, the D1R agonist increased the proportion of errors, both early and late [significant main effect of drug: $F_{2,24}$ = 8.465, η_p^2 = .414, p = .002; significant pairwise comparisons: vehicle vs. high dose, p = .007; no interaction with reward or period: all Fs < 1.1, ps > .3]. However, when excluding one animal whose No-Go success rate was > 2 standard deviations from average performance of the group, we found that the effect of the drug was mediated by period such that both doses increased *early* errors, whilst only the high dose had such an effect on *late* errors, similar to that found with systemic administration of the D1R agonist [significant drug*period interaction: $F_{2,22}$ = 3.613, η_p^2 = .247, p = .044; early errors pairwise comparisons: vehicle vs. low dose, p = .056,

vehicle vs. high dose, p = .063; late errors: vehicle vs. low dose, p > .9; vehicle vs. high dose: p = .027].

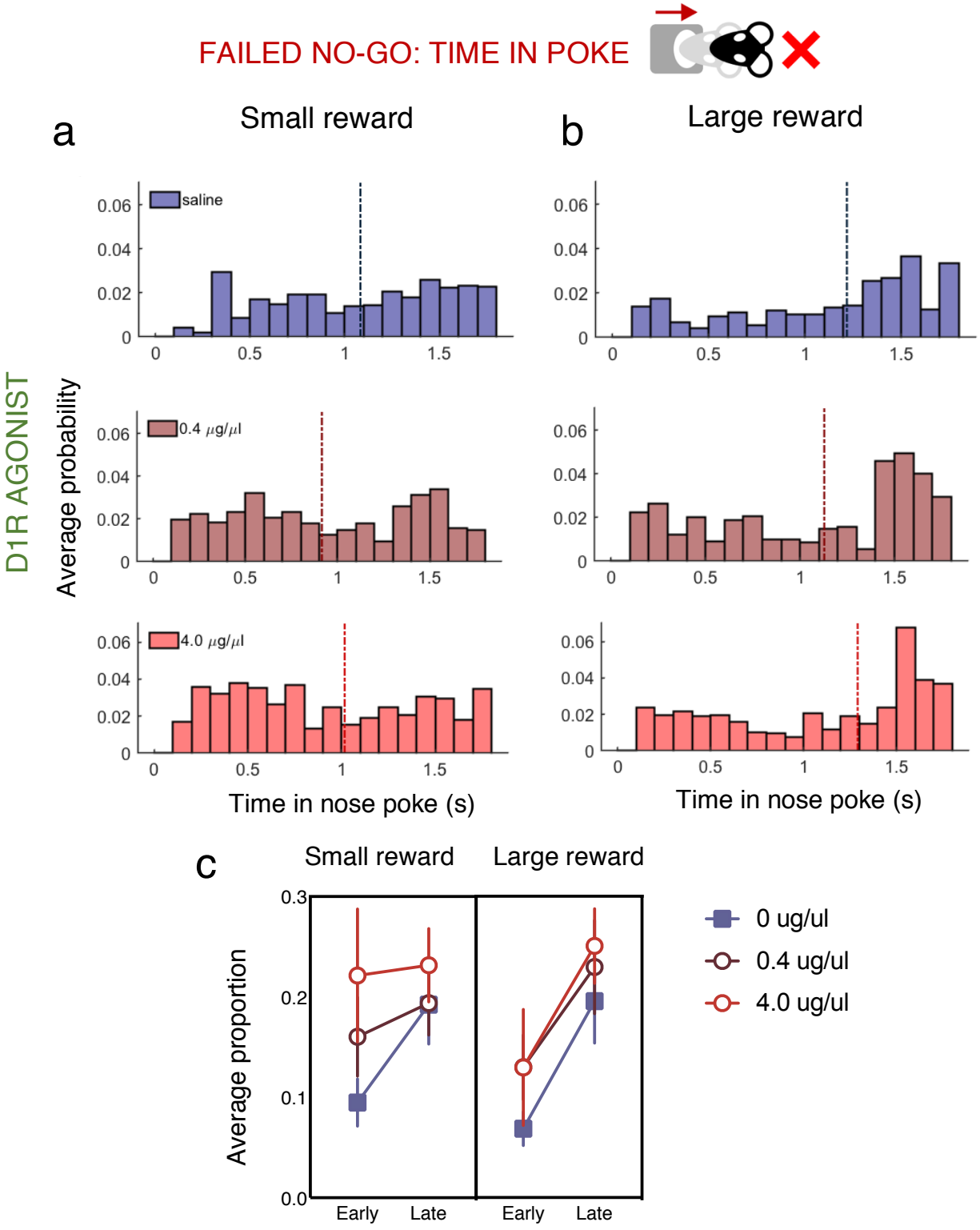


Fig. 2. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with local D1R stimulation, calculated as probability over all head exit times, both successful and failed, for each drug dose. Last bin contains all values > 1.7s. (c) Mean proportion of time spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (purple square: vehicle; brown circle: 0.4 ug/ul; red circle: 4.0 ug/ul.

As described above, local infusion of the D1R agonist did not alter average success rate in Go trials. However, given the effect of the systemic D1R agonist on failed Go trials, we therefore analysed the proportion of Go small or large trials where animals either carried out a Wrong Lever error, or a Response Omission error.

With regards to the latter type of error, animals were less likely to omit a response when a larger reward was on offer [Figure 3b; trend main effect of reward: $F_{1,12} = 4.514$, $\eta_p^2 = .273$, $p = .055$]. However, the drug had no effect on this metric [no main effect of drug or interaction: all Fs < 2.5, $p > .1$]. Yet whilst there was no significant main effect of reward or drug on Wrong Lever errors [Figure 3a; all Fs < 3.6, ps > .08], the drug did exert an effect dependent on the reward size on offer [trend drug*reward interaction: $F_{2,24} = 3.197$, $\eta_p^2 = .210$, $p = .059$]. However, no pairwise comparisons were significant after correction. Moreover, this effect appeared driven partly by the subject that made <40% correct Go trial responses (2 standard deviations from the mean) as when this animal was excluded, no effects were significant [all Fs < 2.6, ps > .1].

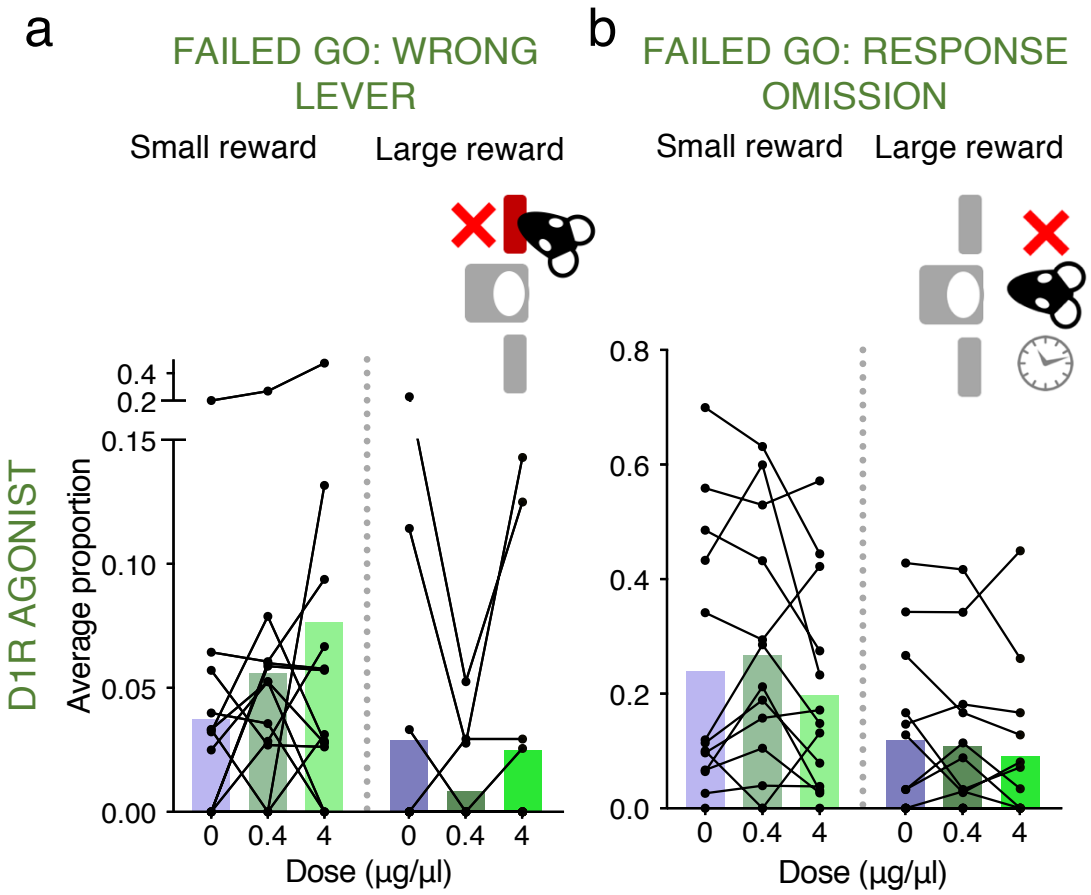


Fig. 3. Effect of local D1R stimulation on the average proportion of Go small or large trials where animals make (a) a wrong lever press or (b) omit response entirely during the 5s cue period. Bars are overall means across all animals, dots are individual animal means.

3.2. NAc core D1R stimulation promotes cue-driven action initiation in Go trials

We previously showed that stimulation of D1Rs in the NAc core increased earlier latencies to initiate action when it was inappropriate to do so in No-Go trials. We next examined how this drug affected the speed of appropriate responses to initiate action.

In correct Go trials, animals were no faster to leave the nose poke with promise of larger reward in such trials [Figure 4a; no significant main effect of reward: $F_{1,12} = 1.193$, $\eta_p^2 = .09$, $p = .296$]. However, the D1R agonist reduced time in the nose poke from

cue [significant main effect of drug: $F_{1.505,18.061} = 4.046$, $\eta_p^2 = .252$, $p = .05$; no significant pairwise comparisons; no reward*drug interaction: $F_{2,24} = 0.006$, $\eta_p^2 = .000$, $p = .994$].

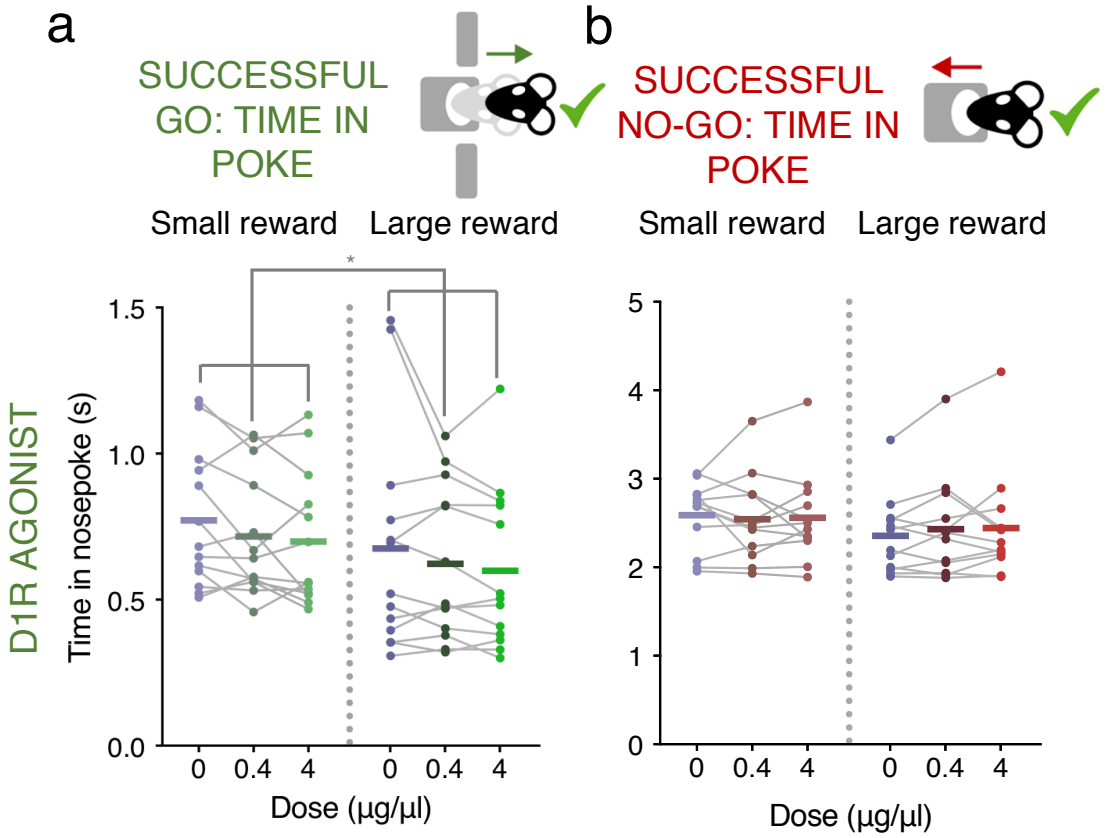


Fig. 4. Mean time spent in the nose poke (s) from cue onset in successful trials in (a) Go or (b) No-Go trials. Solid coloured bars indicate overall mean time in nose poke, and individual dots are mean times for individual animals. Significance bars indicate a main effect of drug, * p < .05.

Visual inspection of the distribution of head exit latencies in successful Go trials (Figures 5a and 5b) suggests a possible interactive effect of both reward and period with drug (0.05 [early], 0.45 [mid], and 0.95 [late] quantiles). To quantify this, a three-way within-subjects ANOVA comparing reward size (small or large), drug (vehicle, low dose, and high dose), and quantile (early, mid, and late) was carried out. This indeed revealed that the drug effect was mediated by both reward size and period of the distribution [significant reward*drug*period interaction: $F_{4,48} = 3.031$, $\eta_p^2 = .202$,

p = .026; significant pairwise comparisons: small reward mid period, vehicle vs. high dose, p = .055]. To tease this effect apart, we then analysed small and large reward trials separately, but found that main effects and interactions of drug with period were non-significant [all Fs < 2.6, ps > .09].

We then investigated whether the latency to initiate appropriate action in successful No-Go trials was similarly speeded by D1R stimulation (Figure 4b). Animals were faster to leave the nose poke with the promise of a larger reward [excluding 1 animal with 0% No-Go success rate in two sessions, significant main effect of reward: $F_{1,12}$ = 4.869, η_p^2 = .307, p = .050]. However, unlike at the time of the Go cue, the D1R agonist had no effect on these action initiation latencies, regardless of the reward size on offer [no main effect or interaction: Fs < 2.7, ps > .9].

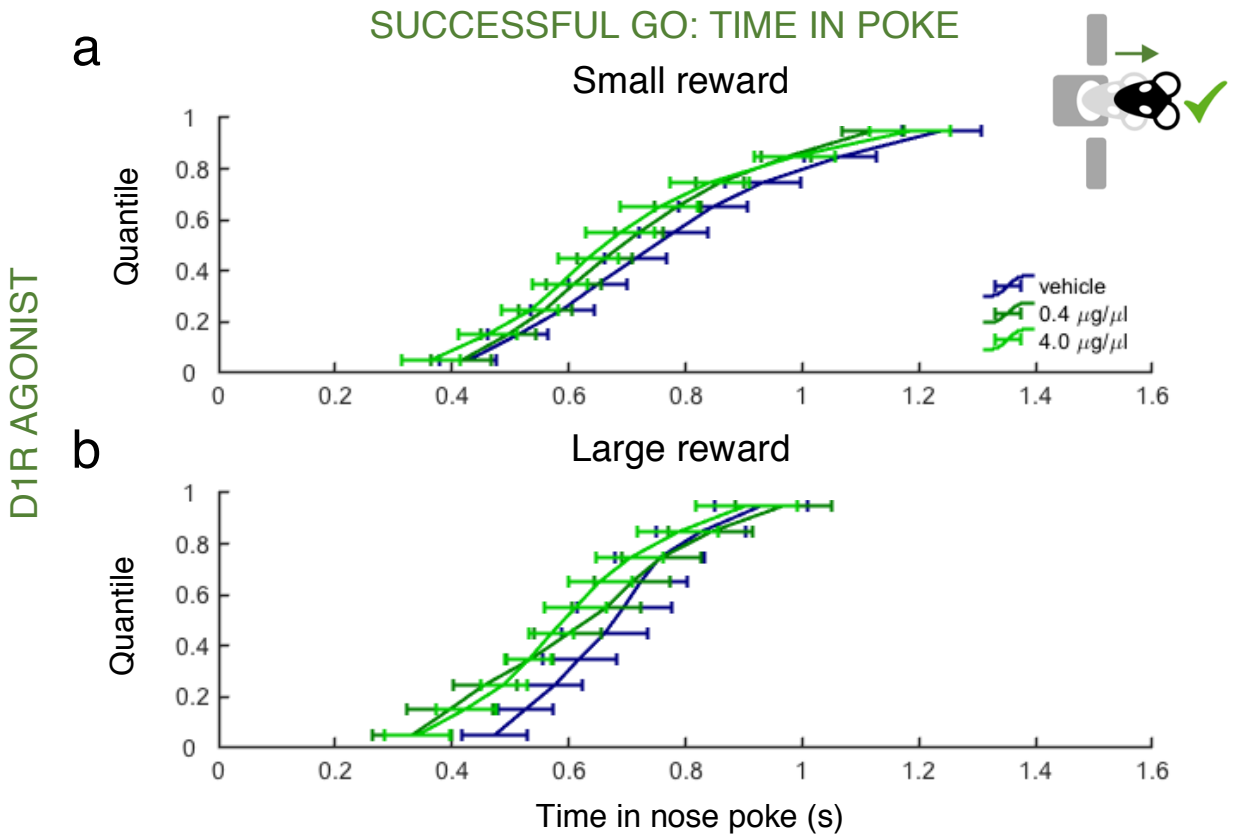


Fig. 5. CDFs for time in nose poke from cue in successful Go trials with local D1R stimulation. (a) CDF for small reward trials. (b) Same as in (a) but for large reward trials.

3.3. NAc core D1R stimulation does not alter energising of goal-directed behaviour

Local stimulation of D1Rs in the NAc core promoted cue-elicited action initiation, both when it was inappropriate – in No-Go trials – and when it was appropriate, in successful Go trials. However, in Chapter 4 we found that systemic D1R stimulation also *slowed* certain goal-directed behaviours, particularly travel time, contrary to an incentive motivation hypothesis of dopamine.

Here, animals were faster to travel with promise of a larger reward [Figure 6; significant main effect of reward: $F_{1,12} = 10.733$, $\eta_p^2 = .472$, $p = .007$]. However, the D1R

agonist neither slowed nor speeded mean travel time, regardless of reward size on offer [no main effect of drug or interaction: all *F*s < 0.9, *p*s > .4]. Analysis of the distributions of travel times similarly revealed no effect of drug or mediation by period of travel [all *F*s < 2.8, *p*s > .08].

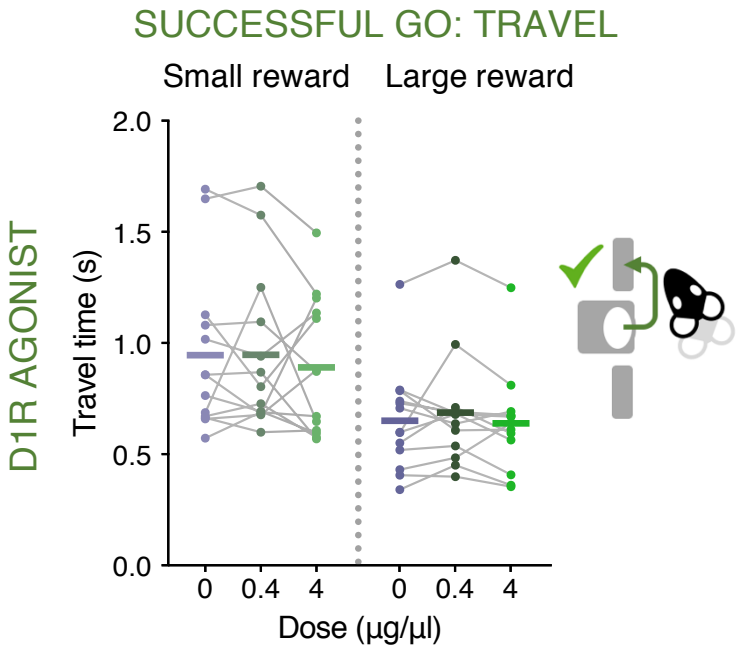


Fig. 6. Mean travel time (s) from nose poke exit in successful Go trials was unaltered by local D1R stimulation.

Other latencies to carry out action were similarly unaltered by the drug. Lever response latency was not affected by D1R stimulation [no main effects of reward or drug, nor interaction: all *F*s < 0.4, *p*s > .7]. Time taken to reach the food magazine after both successful Go and No-Go trials was reduced with promise of a larger reward [Go: significant main effect of reward: $F_{1,12} = 4.904$, $\eta_p^2 = .290$, $p = .047$; No-Go: significant main effect of reward: $F_{1,11} = 7.580$, $\eta_p^2 = .408$, $p = .019$], but was not altered by the drug in either case [no main effect of drug or interaction: all *F*s < 1.0, *p*s > .3].

Systemic administration of a D1R agonist as shown in Chapter 3 slowed latencies to re-engage in the task. Here, animals were generally quicker to re-engage after success than after failure [Figure 7a; significant main effect of trial outcome: $F_{1,12} = 43.095$, $\eta_p^2 = .782$, $p < .001$]. But no drug effect was found with local administration [no main effect of drug or interaction: all Fs < 1.9, ps > .1]. This is also apparent in the latency distributions both after success (Figure 7b) and after failure (Figure 7c).

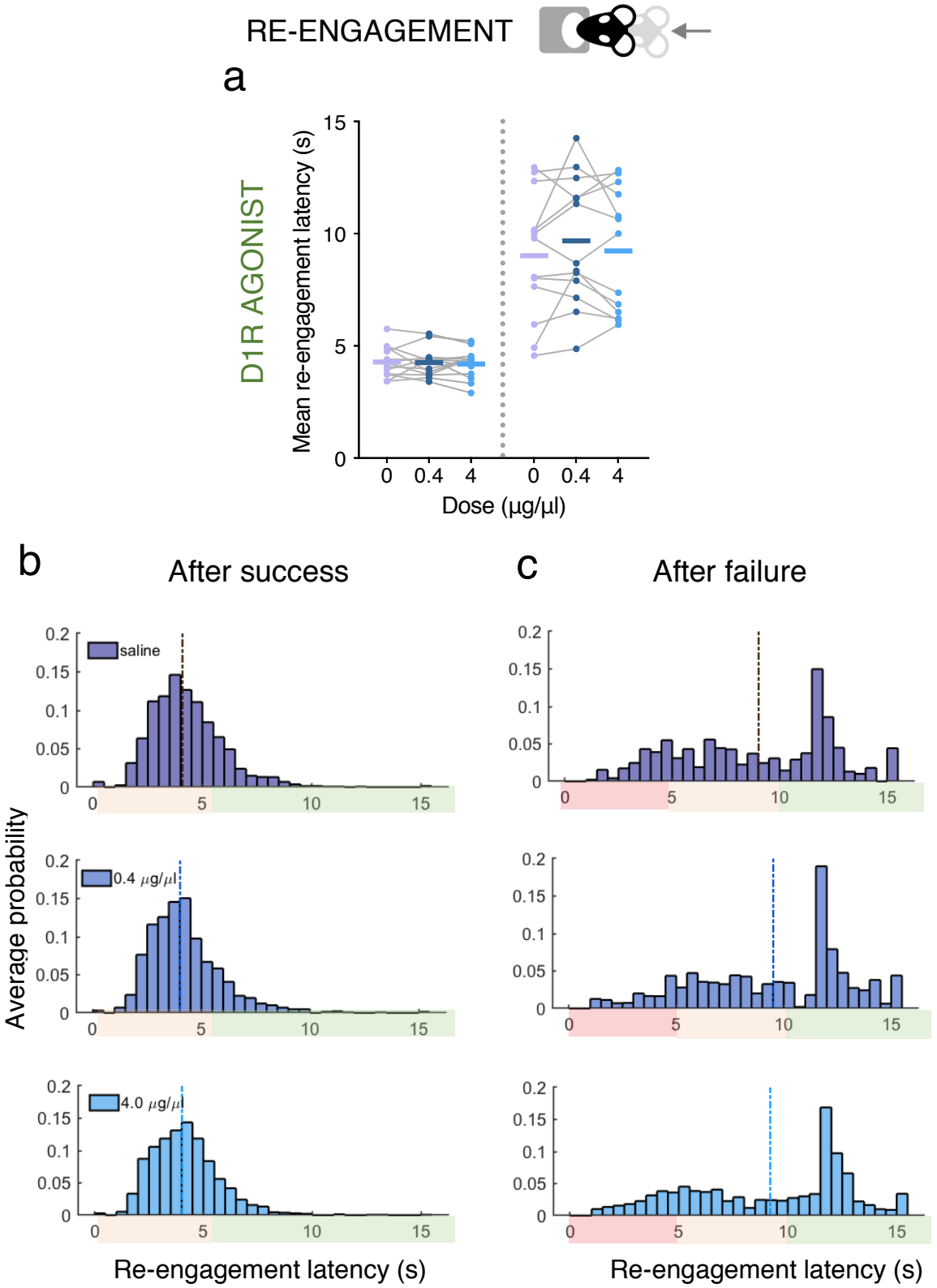


Fig. 7. Effect of local D1R stimulation on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after both successful (left of dotted line) and failed (right of dotted line) trials was unchanged. (b) Histograms of re-engagement latencies after success, split by drug dose. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (0.5s bins; last bin contains all responses after 15s; dotted lines indicate medians). (c) Same as in (b) but after failure. Light red is the timeout period during which the houselight is on. (0.5s bins; last bin contains all responses after 15s).

3.4. NAc core D1R stimulation does not promote uncued action initiation

Analysis of mean number of aborted trials (Figure 8) showed, in general, there was no difference in this measure dependent on previous trial outcome [no significant main effect of trial outcome: $F_{1,12} = 1.060$, $\eta_p^2 = .081$, $p = .324$]. The drug slightly increased number of aborted trials, although this was not a reliable effect and did not result in any significant pairwise comparisons [trend main effect of drug: $F_{2,24} = 2.980$, $\eta_p^2 = .199$, $p = .07$; no interaction with trial outcome: $F_{2,24} = 2.000$, $\eta_p^2 = .143$, $p = .157$]. Exclusion of an outlier > 2 standard deviations from the mean also lessened the effect [trend main effect of drug: $F_{2,23} = 2.719$, $\eta_p^2 = .198$, $p = .088$].

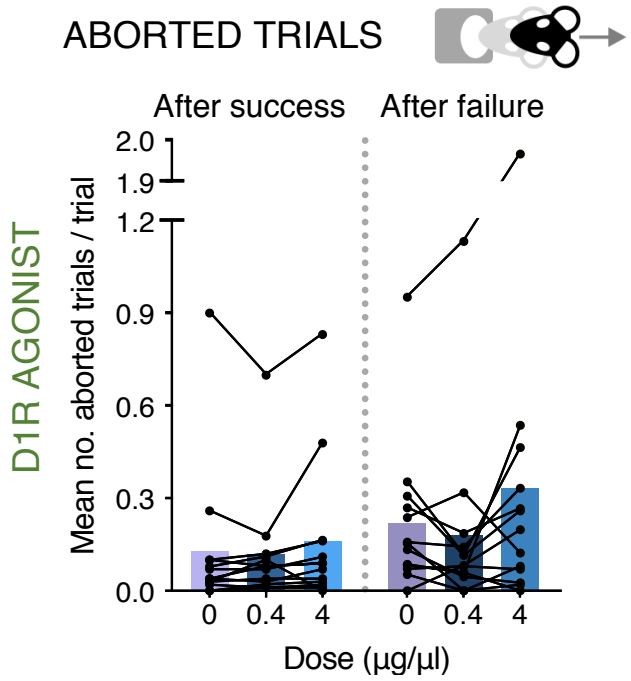


Fig. 8. Effect of local D1R stimulation on mean rate of aborted trials (average number of aborted trials per trial) – when animals entered the nose poke to initiate a trial, but did not remain for the precue period and therefore failed to elicit a cue. The drug did not significantly alter rate of aborted trials either after success (left of the dotted line) or after failure (right of dotted line).

NAc core D1R Antagonist

3.5. NAc core D1R blockade reduces success rate exclusively in Go trials

In a second experiment, we infused a D1R antagonist into the NAc core as the same animals performed the Go/No-Go task. As with systemic application of the drug, their ability to carry out action for reward was dose-dependently reduced by the local D1R antagonist [Figure 9a; significant main effect of drug: $F_{2,22} = 4.559$, $\eta_p^2 = .293$, $p = .022$; significant pairwise comparisons: vehicle vs. high dose, $p = .05$]. Yet here, this effect was also to some extent mediated by the size of the prospective reward, such that there was a loss of the distinction between reward sizes at the highest dose of the drug [trend drug*reward interaction: $F_{2,22} = 3.247$, $\eta_p^2 = .228$, $p = .058$; near-significant pairwise comparisons in large reward trials with correction: vehicle vs. high dose, $p = .069$; small reward trials, all $p > .1$; uncorrected values: large reward trials, vehicle vs. high dose, $p = .024$; low dose vs. high dose, $p = .042$; small reward trials, low dose vs. high dose, $p = .066$].

We then investigated whether success rate in No-Go trials was influenced by the local D1R antagonist. However, as with the systemic study, No-Go success was not altered by D1R stimulation in the NAc core [Figure 9b; no significant main effects of drug or reward, or interaction: all Fs < 1.6, ps > .2].

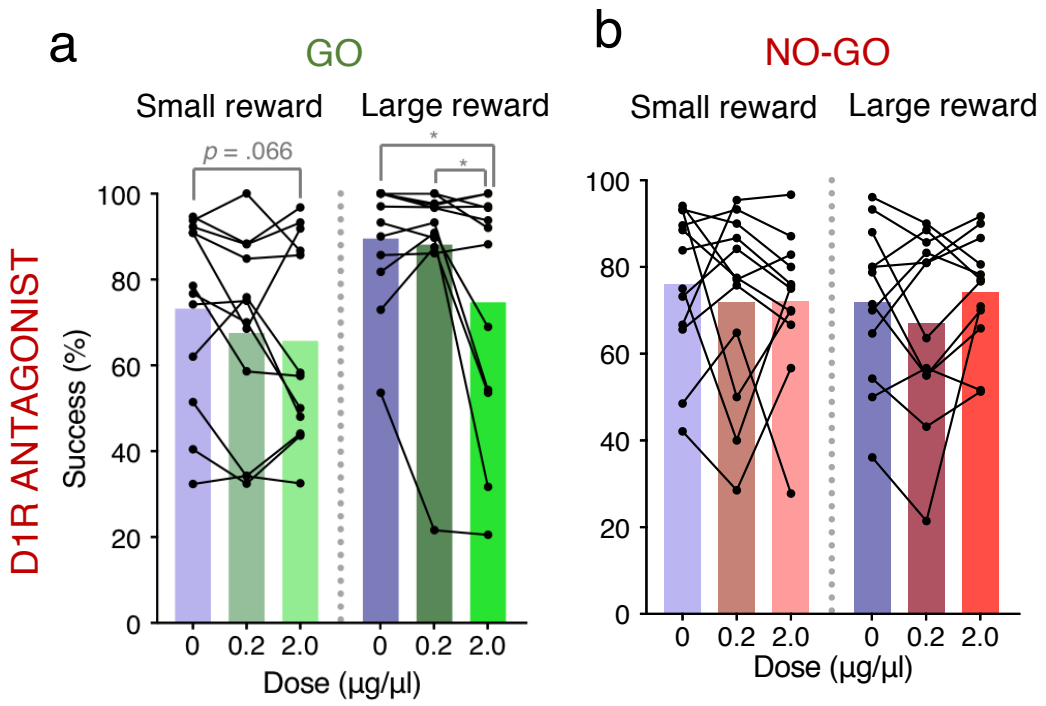


Fig. 9. Effect of local D1R blockade (SCH 23390) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) D1R blockade reduced success rate in Go trials, particularly when a large reward was on offer. (b) D1R blockade had no effect on success rate in No-Go trials. Significance bars indicate uncorrected significant pairwise comparisons, * $p < .05$, ** $p < .01$.

To understand what drove this reduced success rate in Go trials, the types of errors made – either Wrong Lever or Response Omission – were again analysed. The D1R antagonist had no effect on Wrong Lever errors [Figure 10a; no main effect of drug or interaction: all Fs < 1.9, ps > .1] and similarly did not change animals' propensity to lever press in No-Go trials [no main effect of drug or interaction: all Fs < 1.6, ps > .2]. However, it did dose-dependently increase the proportion of Response Omission errors regardless of the reward size on offer [Figure 10b; significant main effect of drug: $F_{2,22} = 4.591$, $\eta_p^2 = .294$, $p = .022$; significant pairwise comparisons: vehicle vs. high dose, $p = .03$; no drug*reward interaction: $F_{2,22} = 1.402$, $\eta_p^2 = .113$, $p = .267$].

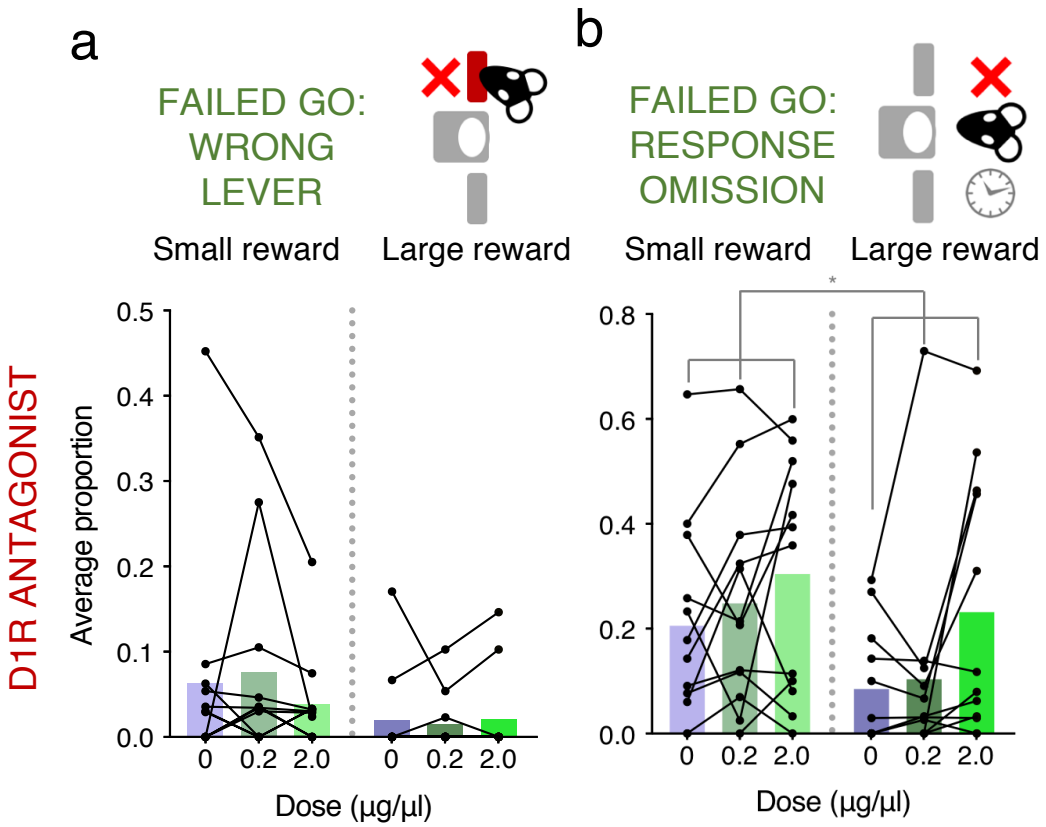


Fig. 10. Effect of D1R blockade on the average proportion of Go small or large trials where animals made (a) a wrong lever press or (b) omitted response entirely during the 5s cue period. Significance bars indicate a main effect of drug on the proportion of trials with response omissions, * p < .05.

Whilst the *number* of errors in No-Go trials did not change with local infusion of the D1R antagonist, it might have been that the *type* of errors in terms of how quickly they were made was altered. We quantified this by comparing average proportion of early (< 800ms) and late (> 800ms) head exits in the No-Go period (Figure 11). Whilst animals again generally made more late than early head exits [significant main effect of period: $F_{1,11}$ = 26.844, η_p^2 = .709, p < .001], neither reward nor drug had any effect on when animals left the nose poke in failed No-Go trials [no main effects of reward or drug, and no significant interactions: all Fs < 1.7, ps > .2].

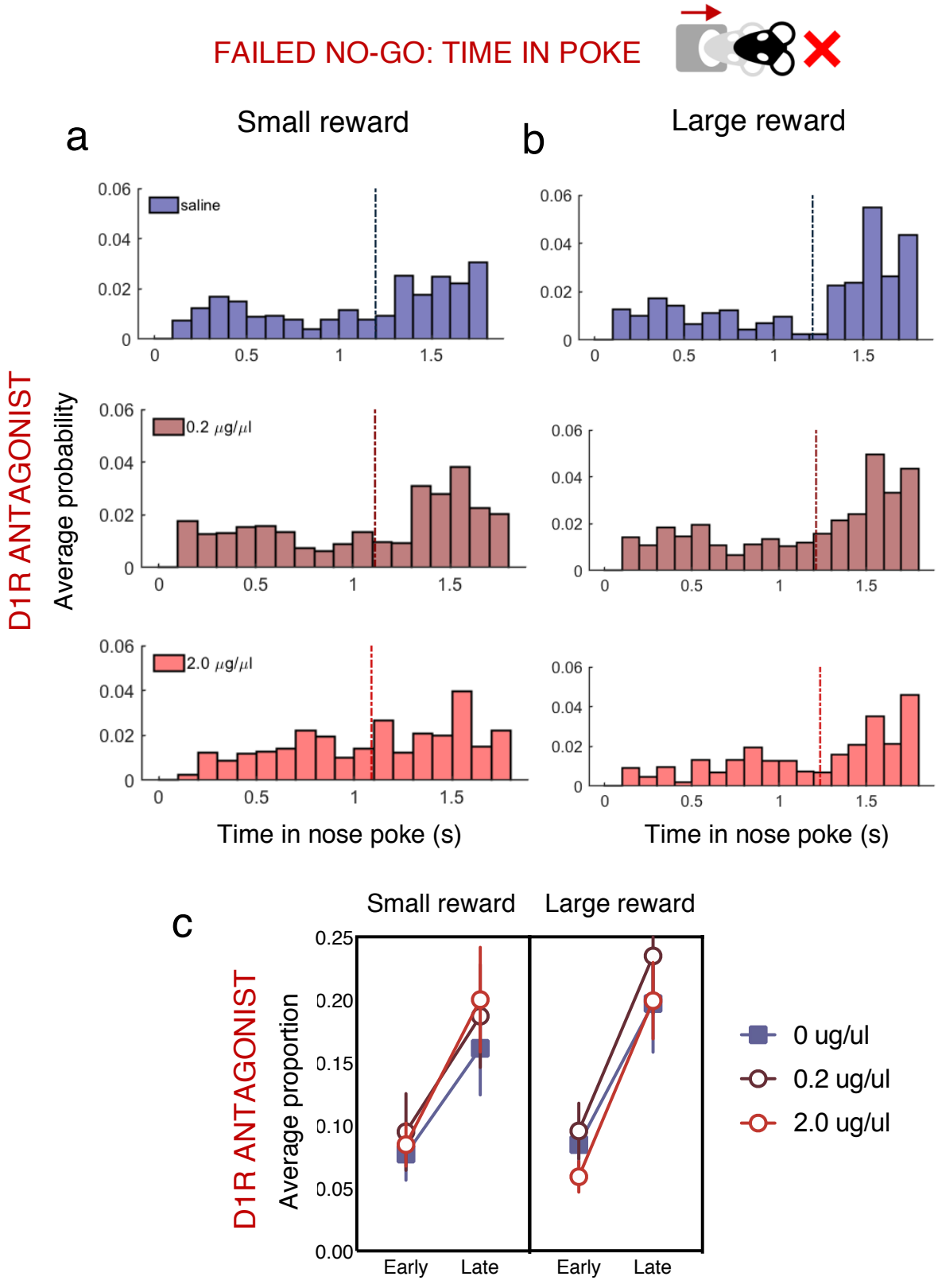


Fig. 11. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with local D1R blockade, calculated as probability over all head exit times, both successful and failed, for each of the drug doses. The last bin contains all values > 1.7s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.2 ug/ul, red circle: 2.0 ug/ul. There was no effect of drug on these measures.

3.6. Local D1R blockade alters cue-driven action initiation in Go trials

We previously showed that, when it was inappropriate to initiate action, animals' responses with a D1R antagonist were not altered. To investigate whether *appropriate* action initiation was slowed by the drug, we analysed time spent in the nose poke from cue onset in successful Go trials (Figure 12a).

Mean time spent in the nose poke was reduced by the promise of a larger reward [trend main effect of reward: $F_{1,11} = 4.305$, $\eta_p^2 = .281$, $p = .062$]. However, the drug did not reliably alter mean time spent in the nose poke in successful Go trials [no main effect of drug or interaction: all Fs < 2.6, ps > .09].

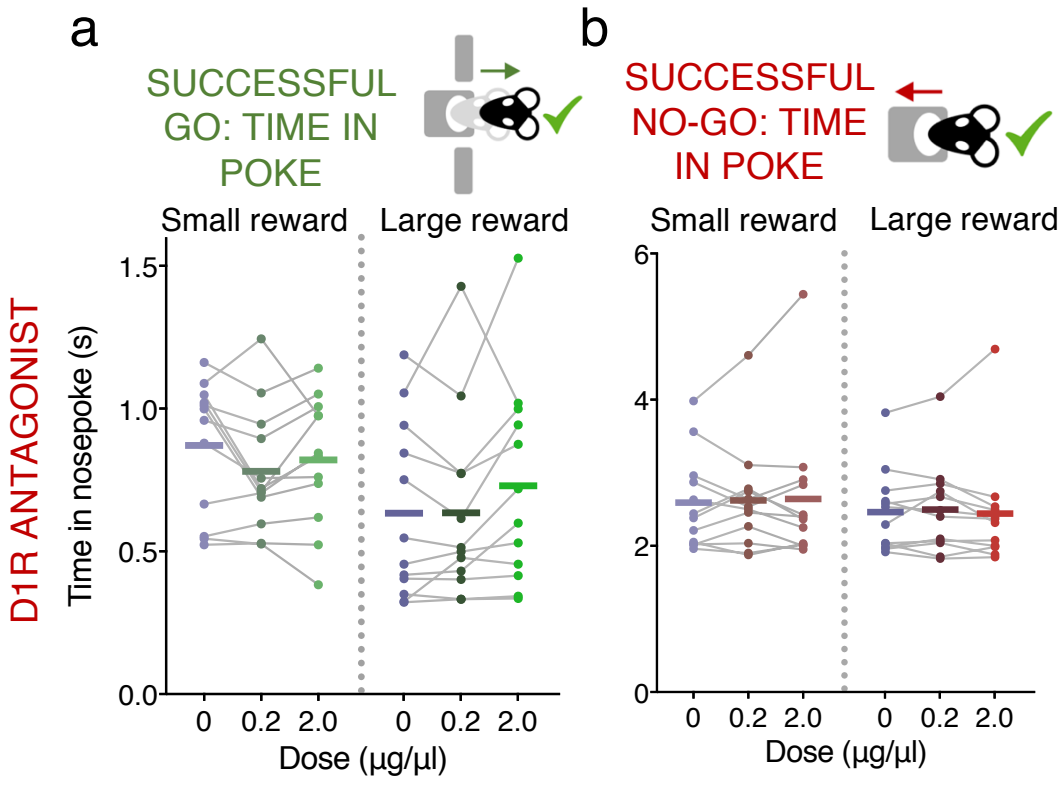


Fig. 12. Mean time spent in the nose poke (s) from cue onset in successful trials with a local D1R antagonist on board. (a) Mean time for successful Go trials in the nose poke was unaltered. (b) Similarly, mean time for successful No-Go trials in the nose poke was unaltered.

However, analysis of the Go trial head exit latency distributions (Figure 13) reveals an interactive effect of the drug with reward [significant drug*reward interaction: $F_{2,22}$ = 4.71, η_p^2 = .289, p = .023; pairwise comparisons: small reward, all p > .1; large reward, vehicle vs. high dose, p = .007] and with period [significant drug*period interaction: $F_{2.768,30.445}$ = 3.929, η_p^2 = .263, p = .020; pairwise comparisons: early period, all ps > .6; mid period; all ps > .2; late period, low dose vs. high dose, p = .003; no three-way interaction: $F_{2.240,24.642}$ = 1.840, η_p^2 = .143, p = .177]. In practice, this means that the drug exerted greatest effects on head exit times in *large* reward trials, particularly on the slowest head exit times.

We then investigated whether appropriate action initiation in successful No-Go trials, prompted by cue o*ffset*, was altered by local infusion of the D1R antagonist. As can be observed in Figure 12b, mean time spent in the nose poke in successful No-Go trials was not influenced by local blockade of D1Rs [no main effect of drug or interaction: all Fs < 0.7, ps > .4].

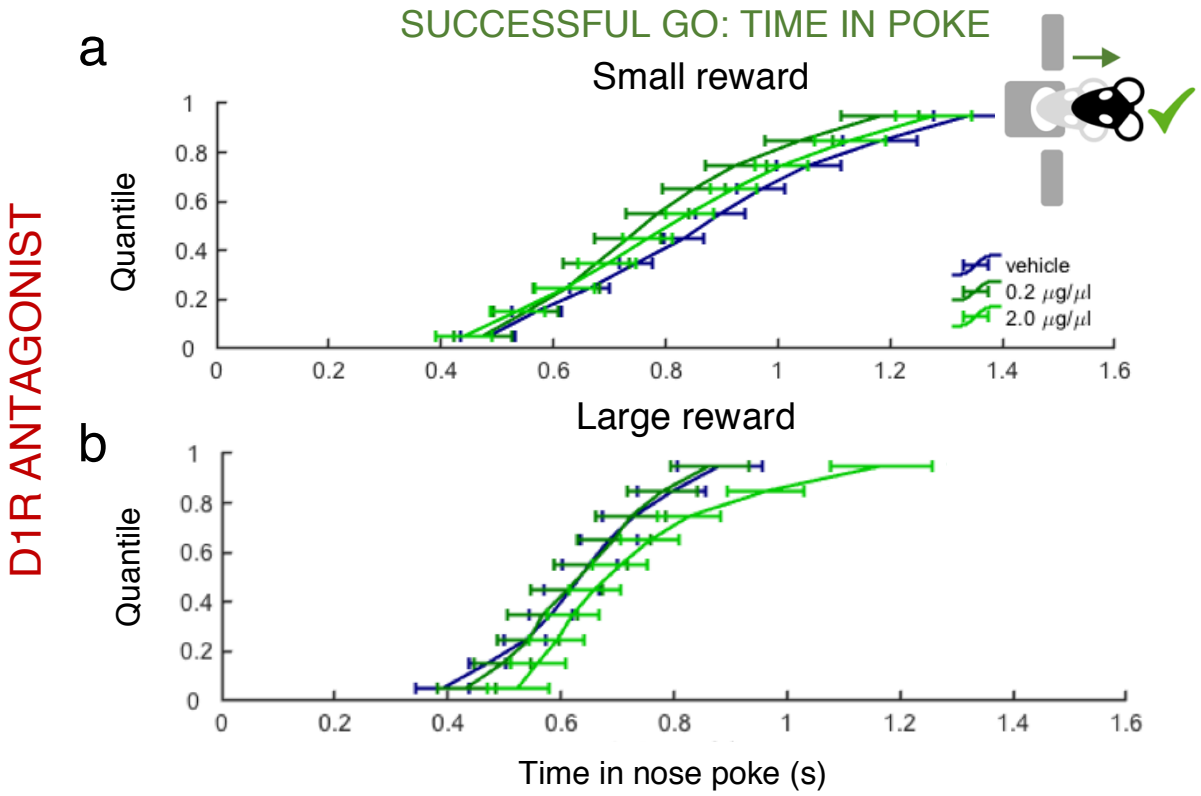


Fig. 13. CDFs for time in nose poke from cue in successful Go trials with local D1R blockade. (a) CDF of head exit latencies for successful small reward Go trials. (b) Same as in (a) but for successful large reward Go trials.

3.7. Local D1R blockade does not alter the energising of ongoing goal-directed action

We have shown that, although animals are less likely to execute a successful response in Go trials following local infusion of a D1R antagonist into the NAc core, the effects on time taken to initiate action in successful Go trials were only subtle, and both successful and failed No-Go trials were unaffected. We therefore next examined whether local blockade of D1Rs would instead influence latencies to execute goal-directed behaviour.

However, unlike the effect of systemic administration, the local D1R antagonist did not change the mean time taken to travel in Go trials [Figure 14a; no significant main effect or interaction; all Fs < 0.6, ps > .6].

The drug also had limited effects on mean lever press times (Figure 14b), with a small numerical slowing at the highest dose causing the main effect of drug to approach significance [$F_{2,22} = 3.361$, $\eta_p^2 = .234$, $p = .053$; no significant pairwise comparisons; no drug interaction with reward: $F_{2,22} = 0.016$, $\eta_p^2 = .001$, $p = .984$].

Equally, when retrieving reward after a successful Go trial, the D1R antagonist had no effect on this latency [no main effect of drug or interaction: all Fs < 1.8, ps > .2].

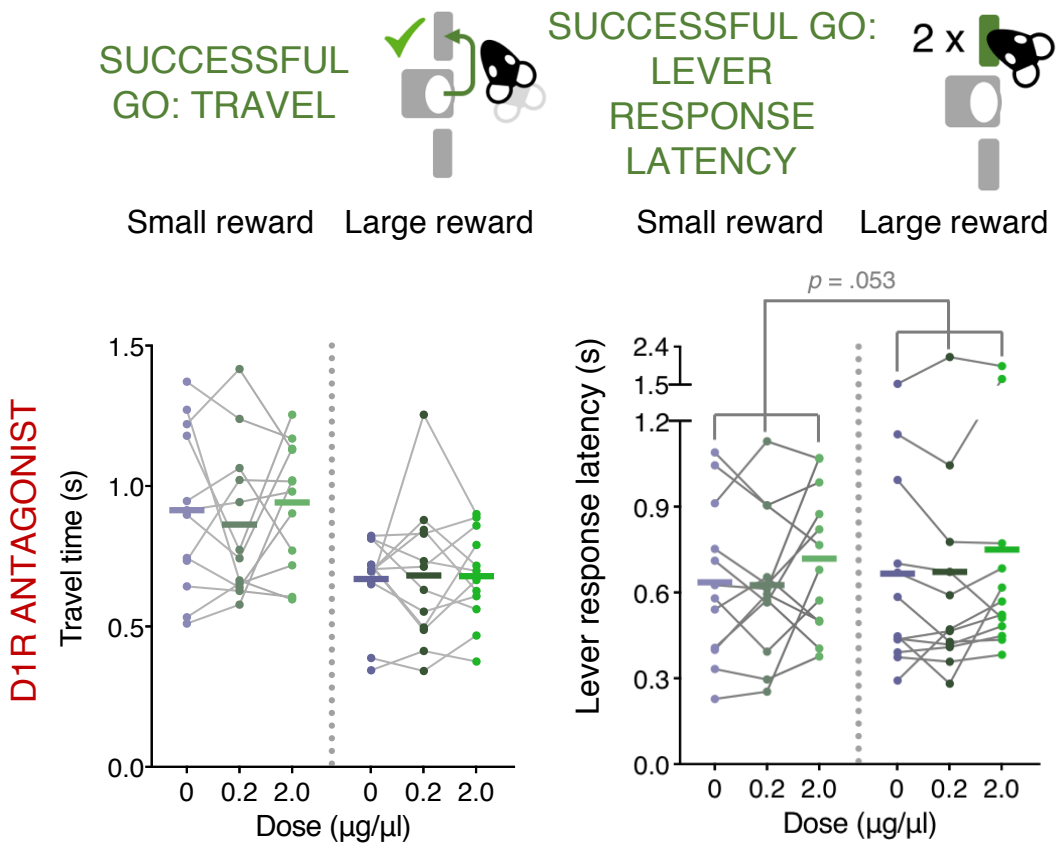


Fig. 14. Mean (a) travel time from nose poke exit and (b) lever-response latency in successful Go trials.

Finally, we analysed latencies to re-engage in the task. In Chapter 4, the systemic D1R antagonist had a profound slowing effect on these latencies. Here, animals were faster to re-engage after success than after failure [significant main effect of trial outcome: $F_{1,11} = 6.521$, $\eta_p^2 = .372$, $p = .027$]. However, the drug did not exert a reliable effect on mean re-engagement latencies [no significant main effect of drug or interaction: all Fs < 1.7, ps > .2].

3.8. Local D1R blockade does not promote uncued action initiation

Finally, to determine whether the D1R antagonist had any effect on the promotion or inhibition of *uncued* action, we analysed the rate of aborted trials (Figure 15) and found that the drug had no effect on animals' propensity to initiate action generally, neither after success nor after failure [no main effects of drug or trial outcome, nor interaction: all Fs < 1.1, ps > .3].

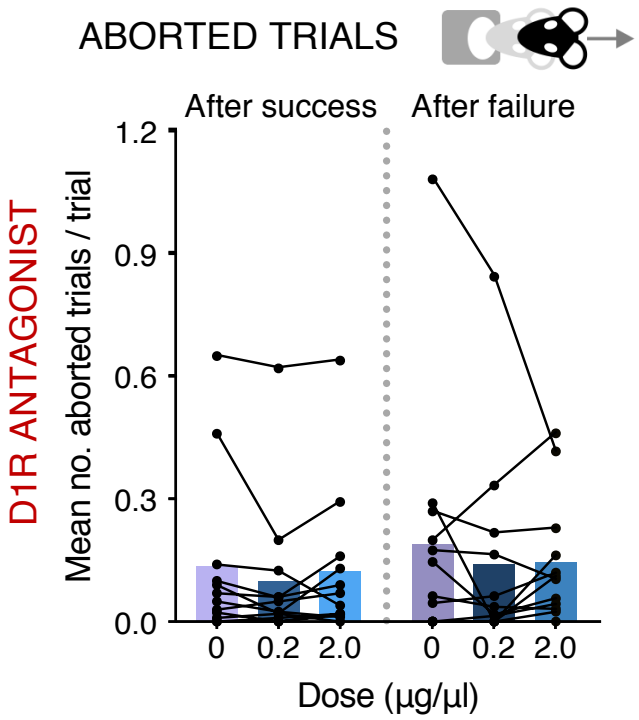


Fig. 15. Effect of D1R blockade on mean rate of aborted trials (average number of aborted trials per trial) – when animals entered the nose poke to initiate a trial, but did not remain for the precue period and therefore failed to elicit a cue. The drug did not alter rate of aborted trials either after success (left of the dotted line) or after failure (right of dotted line).

4. Discussion

Understanding dopamine's role in both reward prediction and in driving goal-directed behaviour may be dependent on regional differences, such that ventral striatal dopamine is more involved in representation of reward-relevant information whilst dorsal striatal dopamine encodes movement for reward (Kravitz et al., 2012). Yet in Chapter 3, we found that subsecond dopamine release in the NAc core is also shaped by goal-directed action initiation, meaning that it might instead be the case that this signal acts as a 'critic' for the ongoing action as 'acted' by the dorsal striatum, rather than purely representing reward-based information (Joel et al., 2002). If this is

the case, then D1R manipulation within the ventral striatum should not *causally* alter behaviour.

However, systemic manipulation of receptors thought to be most sensitive to phasic changes in this signal, D1Rs, in Chapter 4 revealed that level of D1R activation causally influenced ability to initiate and perform ongoing action for reward, and also drove action initiation. The lack of regional specificity of these manipulations made it challenging to link the variations in dopamine efflux seen with FCV in the NAc core specifically to these causal effects. Therefore, in this chapter the same D1R pharmacological challenges were applied to the NAc core, and, as with systemic manipulations of these receptors, their activation drove certain reward-oriented behaviours whilst their blockade impeded these (Table 2). Nevertheless, there were also some key differences between the whole brain and local manipulations that are discussed in more detail subsequently.

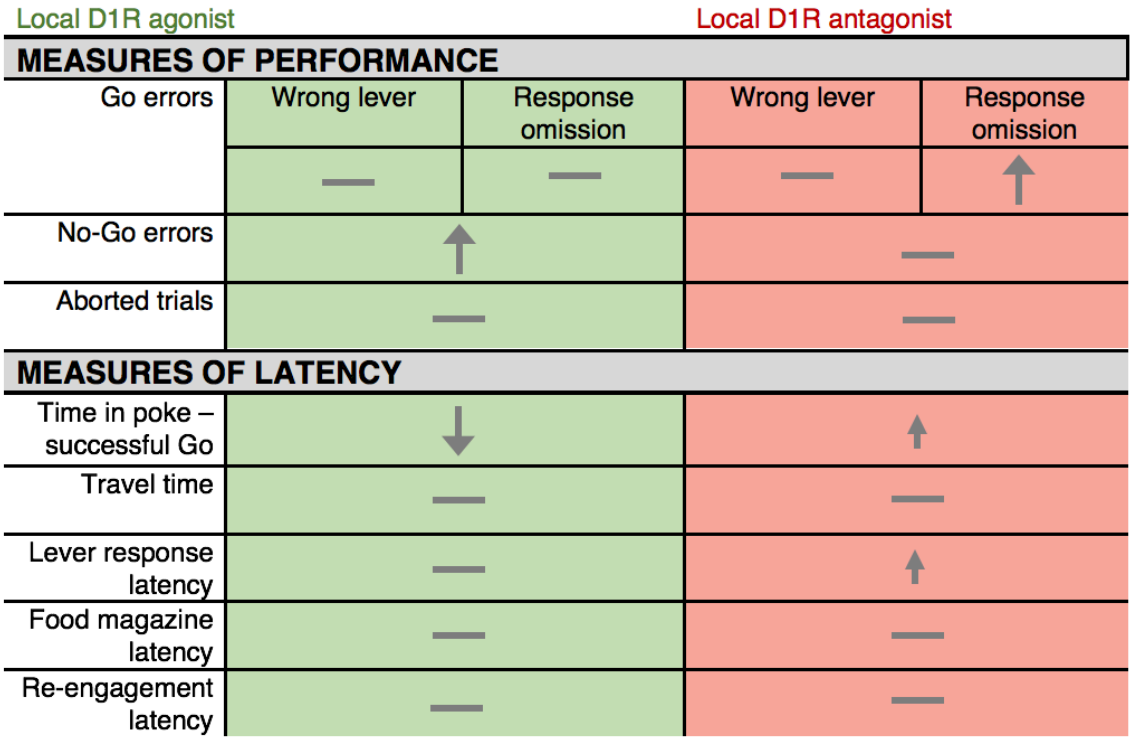


Table 2. Summary of effects of both the local D1R agonist (green) or D1R antagonist (red) on measures of performance and measures of latency in the Go/No-Go task. Large arrows indicate a strong effect, small arrows indicate smaller or more specific effects.

4.1. Cue-driven action for reward is causally mediated by D1Rs in the NAc core

We previously found that stimulating D1Rs systemically drove action initiation when it was inappropriate – in No-Go trials – contrary to a simple reward prediction error hypothesis of the dopaminergic signal. Here, we replicated this effect with local infusion of the D1R agonist into the NAc core. Such an effect is consistent with previous studies of action postponement, which showed that infusions of D1R agonists into the NAc core increased the number of premature responses made in the 5-choice task (Pezze et al., 2007).

However, in contrast to the systemic manipulations and to Pezze et al. (2007), we additionally found that animals were faster to initiate *appropriate* action initiation, suggesting that these receptors are key in promoting goal-directed movement as an expression of incentive motivation. Furthermore, and again in contrast with Pezze et al. (2007), this effect is specific to *cue-driven* actions, as animals were not more likely to abort trials with local infusion of the D1R agonist, and similarly time to leave the nose poke after successful No-Go trials was unaltered. In addition, local infusion of the D1R antagonist *decreased* the speed of cue-driven action initiation, particularly the tail end of the distribution, and also did not alter rate of aborted trials, again suggesting that levels of D1R activation are fundamental in the translation of reward processing to cue-dependent action.

This is consistent with Nicola's premise that dopamine in the NAc core is integral to the invigoration of reward-seeking actions by value-laden cues (McGinty, Lardeux, Taha, Kim, & Nicola, 2013). On the other hand, the authors go further to propose that this invigorating effect should only be apparent on 'flexible', rather than stereotyped, behaviours, citing evidence from reaction time tasks in which disruptive manipulation of the NAc only subtly alters the ability of cues to increase vigour (Brown & Bowman, 1995). This study also parallels ours in that animals must withhold movement for a brief period in a poke before leaving at cue presentation. Yet here, we found a speeding effect of the drug on a similar measure. This discrepancy may be due to the fact that, although it is most *efficient* for animals to leave as soon as they identify the cue for a Go trial in our task, they are not punished for remaining

for a second or two longer – as long as they ultimately press the correct lever – unlike in Brown & Bowman (1995), where any responses > 100ms were considered erronerous. Indeed, here it may be more optimal to remain in the poke for a slightly longer duration to be more certain that the current trial type is not a No-Go trial (for which leaving the nose poke early would be costly). Because of this, time to leave the nose poke in our task may well require more cognitive 'flexibility' than other simple reaction time tasks, and posits the intriguing suggestion that Nicola's flexible approach model could well incorporate this alongside behavioural flexibility as a function of NAc dopamine.

Ability to withhold action was unaltered by the D1R antagonist, consistent both with the systemic study and with the FCV results of Chapter 3. The latter which demonstrated an attenuation of dopamine efflux throughout the No-Go cue period, suggesting that dopamine is not required when having to withhold action for reward. Yet this is a different conclusion from that drawn by studies of impulse control such as the 5-choice, which have previously demonstrated a reduction of impulsivity with NAc core D1R blockade (Pattij et al., 2007). Key to this difference is that in the 5-choice, a reduction in impulsivity may coincide with either a decrease in locomotion or alternatively expressed behaviours, whereas here, it requires a sustained inhibition that has already been initiated. Therefore, for the first time it is demonstrated that D1R activation is *not* required when the form of inhibition categorically involves no movement.

4.2. Likelihood, but not speed, of instrumental responding is mediated by D1Rs in the NAc core

Our FCV studies in Chapter 3 revealed sustained dopamine release as animals progressed throughout Go trials, and in Chapter 4, it was shown that blockade of D1Rs indeed led to a slowing of ongoing performance. Infusion of the D1R antagonist into the NAc core here replicated an increasing inability to execute instrumental action for reward, leading to a large increase in Response Omission errors.

However, there is little evidence that antagonism of these same receptors within the ventral striatum mediates the speed of progression throughout *correct* Go trials, with travel time and food magazine latencies left unaltered by the manipulation. This is consistent with studies showing a reduction in appropriate cue responding with similar D1R blockade (Yun et al., 2004) and suggests that, rather than acting as a locomotor effect, D1Rs uniquely mediate the incentive motivational properties of a stimulus or operandum (Robbins & Everitt, 2007) – although that is specifically with regards to the decision to *initiate* a well-learned action sequence, but not its ongoing execution. This is also consistent with the flexible approach hypothesis (Nicola, 2010) as these latencies were garnered from successful Go trials and were therefore relatively stereotyped, such that NAc dopamine may not have a modulatory effect. However, the overall reduction in Go success rate with local infusion of the D1R antagonist suggests that the activation of these receptors is necessary when there is ambiguity as to the correct response, whilst the lack of effect on re-engagement

latencies again suggests that it is *cue-guided* action that was impeded in failed Go trials, rather than action more generally.

Interestingly, we previously found that systemic stimulation of D1Rs resulted in a *slowing* of travel times and re-engagement latencies. This was not replicated with local infusion of the D1R agonist, but neither did this manipulation speed progression through Go trials. Other impulse control studies have similarly found a lack of an invigorating effect on such latencies with NAc core D1R stimulation (Pezze et al., 2007), suggesting that degree of D1R activation in this region critically modulates whether or not an action is carried out, but not necessarily the vigour with which it is executed. Instead, there is a wealth of evidence that the dorsal striatal dopamine controls movement vigour (Alves et al., 2018; Panigrahi et al., 2015).

4.3. Action selection is not biased by level of D1R activation in the NAc core

Systemic stimulation of D1Rs in Chapter 4 caused animals to increasingly select the *large* reward lever in small reward Go trials. Whilst the direct contribution of mesolimbic dopamine to biasing action selection in well-learnt contexts is less well-established, there is evidence of *dorsal* striatal dopamine playing an important role in switching from the selection of one action to another (Howard et al., 2017). It was therefore hypothesised that this previous effect was due to stimulation of D1Rs outside of the accumbens, and so would not be replicated with local infusion of the same ligand. Indeed, we did not find this biasing effect in Go trials, confirming our

hypothesis that dopamine in the NAc core is critical in mediating the ability to make or withhold action, but not in determining the specific action chosen. This perhaps unsurprising considering NAc circuitry, which lacks segregation between D1R and D2R-expressing populations and project directly to the output nuclei of the basal ganglia (the ventral mesencephalon) or indirectly via the ventral pallidum, with neither region traditionally considered a nucleus for action selection processes (Kupchik et al., 2015).

4.4. Summary

By locally infusing a D1R antagonist or agonist into the NAc core during the Go/No-Go task, here we found that these receptors in this region specifically alter cue-driven action initiation, with little effect on ongoing performance of correct action. However, D1R blockade also decreased likelihood of making an operant response, again without affecting latencies such as travel time when animals were able to lever press, suggesting that these receptors are integral to the *initial* drive to perform an action – dependent on the size of the reward on offer – but not to complete a series of actions for reward. This is consistent with other evidence that early sequence components, but not terminal actions, are mediated by these receptors (Keeler et al., 2014). Finally, the lack of an improvement in success when having to withhold action for reward with the D1R antagonist on board suggests that these effects are critically expressed in the face of *action*, but not *inaction*.

6 D2-like receptors gate goal-directed action initiation and execution

Chapter abstract:
One proposed mechanism by which dopamine is able to simultaneously influence reward processing and motivation for reward is for the former to be predominantly mediated by D1-like receptors, and the latter by D2-like receptors. More specifically, tonic dopamine levels – which D2Rs are thought to be sensitive to – may encode *reward rate*, thereby driving response vigour. Alternatively, activation of these receptors as part of the indirect pathway may instead *inhibit* behaviour. Here, by using systemic pharmacological manipulations of D2Rs as animals made or withheld action for reward, it was revealed that activation of this receptor family impeded action initiation and execution for reward but without altering ability to withhold action, whilst blockade of these receptors had little influence on behaviour. This suggests that D2Rs gate goal-directed behaviour, but only when *movement* is required.

For the studies presented in this chapter, Laura Grima and Mark Walton planned the experiments, Laura Grima collected the data, and Laura Grima analysed and presented the data with input from Mark Walton.

1. Introduction

Whilst it is well-established that dopamine is important for both reward prediction and motivated behaviour, it is not well understood how this single neurotransmitter is able to convey these two quite disparate signals. One possibility that has been investigated thus far in this thesis is that dopamine's effects are region-dependent. Yet both when recording dopamine release and when manipulating the activation of D1Rs in the same region – the NAc core, purported to be solely important for reward prediction – we have found an *interaction* between prediction and action for reward.

An alternative that has been proposed in the literature is that the two well-described timescales of dopamine release – phasic and tonic – might convey different forms of information, with the former more sensitive to reward prediction and the latter more involved in activation of behaviour for reward. The focus in this thesis thus far has been on understanding the phasic signal and its consequences on D1-like receptors (D1Rs) that are thought to be sensitive to it, but D2-like receptors (D2Rs) may also play an important role in conveying reward or movement signals.

An important feature of striatal D2Rs is that they are expressed both post-synaptically on MSNs, but also presynaptically as autoreceptors on dopamine neurons (Ford, 2014a). The activation of these autoreceptors results in a decrease in excitability of dopamine neurons and associated dopamine release into the synaptic cleft, whilst their inhibition results in the converse (De Mei, Ramos, Iitaka, & Borrelli, 2009), and dependent on drug dose used, a pharmacological agent can preferentially activate or

block one or the other; high doses of D2 drugs tend to most impact postsynaptic receptors, whilst autoreceptor effects are found at lower doses (Tanaka et al., 1992). Importantly, autoreceptors have been found to play an important role both in locomotor activity (Eilam & Szechtman, 1989) and, more recently, in *reward-driven* behaviour, with their ablation resulting in an enhanced motivational phenotype (Bello et al., 2011).

Additionally, due to their higher inherent affinity, these receptors are thought to be more readily activated by changes to slow, or *tonic* fluctuations in dopamine release, as well as to dips in this release (Richfield, Penney, & Young, 1989; although see section 4.1 of the disucssion for caveats to this viewpoint).

1.1. Tonic dopamine and reward rate

It has long been thought that a rise in tonic levels of dopamine leads to an increase in motivation (Salamone & Correa, 2012), which can be expressed in terms of increased activation (Missale et al., 1998; Robbins & Everitt, 2007; Taylor & Robbins, 1984), vigour (Beierholm et al., 2013), energy expenditure, (Beeler et al., 2012), or willingness to exert effort for reward (Salamone, Correa, Farrar, & Mingote, 2007). Hence one model of how the dopaminergic signal can simultaneously convey a reward prediction error, or 'teaching' signal, whilst also influencing drive for reward is that the former is encoded by phasic dopamine release, and the latter varies as a function of tonic dopamine.

One particular proposal is that tonic dopamine may encode reward rate, or the *opportunity cost* of not engaging; in other words, if reward rate is high, the cost of not acting is higher than that when reward rate is low, meaning that it is more worthwhile in such circumstances to perform action vigorously (Niv et al., 2007). If average reward rate is reported by tonic levels of dopamine, then this lends a prediction that the reduction of such levels – or its effects via D2Rs – would artificially reduce the opportunity cost of inaction such that any action carried out would be less vigorous.

1.2. Inhibition of behaviour: the 'No-Go' receptor

In addition, the classical model of basal ganglia circuitry instead proposes that D2Rs, as part of the indirect pathway, impede conflicting actions such that their overactivation can inhibit behaviour entirely (Albin, Young, & Penney, 1989; Frank, 2005). Indeed, blockade of these receptors speeds response inhibition in a stop-signal task (Eagle et al., 2011), leading to a conception of D2R activation as putting the 'brakes on behaviour'.

However, more recent models suggest a more cooperative effect of D1Rs and D2Rs, particularly in ventral striatum (Kupchik et al., 2015). This complements studies showing the necessity of activation of both receptor subtypes in goal-directed behaviour (Cui et al., 2013; Soares-Cunha et al., 2016). Whilst these two proposals – D2Rs as inhibitors vs. as drivers of behaviour for reward – are in opposition, neither have been examined in the light of explicitly having to withhold *all* action for reward.

1.3. Behavioural switching and action selection

As described above, due to their low affinity property, D2Rs are particularly sensitive to tonic fluctuations in dopamine – but particularly to *reductions* of such levels, as at baseline they are highly occupied. It has therefore been posited that these receptors may be sensitive to the suppression of activity, and associated decrease in dopamine release, in response to omission of expected rewards (Tobler et al., 2003). This may be related to a facilitation of behavioural flexibility due to corticostriatal interactions (Meck & Benson, 2002), such that inactivation of D2Rs increases set-shifting behaviours (Goto & Grace, 2005), whilst their stimulation causes behavioural perseveration and inflexibility (Haluk & Floresco, 2009).

Keeler, Pretsell, & Robbins (2014) have suggested a mechanism for this by which MSNs expressing D2Rs encode reward association strength via internalisation of those receptors (Bartlett et al., 2005), meaning that application of a D2 agonist would inhibit pathways with a weak reward association strength and lead to selection of the same, previously high-value action. In the experiments described throughout this thesis, the animals have already fully learnt the task and there is no sudden switching of action-reward contingencies requiring new learning. Yet according to Keeler et al.'s (2014) formulation, alteration of D2R activity may regardless have an effect on action selection in the Go/No-Go task, biasing action towards the high value response.

1.4. Aims, approach, and hypotheses

The aim of the experiments in this chapter is to understand the contributions of D2Rs to reward processing and goal-directed action initiation and restraint. Therefore a selective D2R agonist, as well as a selective D2R antagonist, were applied systemically as animals performed the Go/No-Go task. Chapters 4 and 5 elucidated the effect of D1R stimulation and blockade in this same task, both across the brain and locally within the core of the NAc.

If D2Rs act in opposition to D1Rs, as suggested by some prominent models, then their stimulation should in fact *reduce* motivated behaviour, whilst their blockade should do the converse. This would also be consistent with a more nuanced but relevant conception of D2Rs as encoding the costs of acting (Bogacz, 2017). However, if, as more recent evidence might suggest, D2Rs act in synergy with D1Rs to promote behaviour for reward, then the direction of effects should be consistent with those found in Chapters 4 and 5, whilst their blockade should reduce goal-directed behaviour.

Consideration of a reward rate, or opportunity cost, hypothesis also lends itself to a specific prediction: if D2Rs are indeed encoding reward rate, then an increase in their activation should *increase* the vigour of behaviour – evidenced by a change in Go trial latencies – but without necessarily altering the allocation of behaviour itself. Alternatively, if this reward history learning is used specifically to bias action selection, then D2R stimulation should reduce behavioural flexibility and cause animals to repeat actions previously associated with larger rewards.

2. Methods

2.1. Animals

The same cohort of 11 male Sprague Dawley rats was used for both of the experiments described in this chapter. These animals had previously been used for FCV recordings in the same Go/No-Go task (see Chapter 3) and for the systemic D1R agonist study (see Chapter 4). For the D2R agonist study, two animals were excluded due to computer failure, resulting in an n of 9 for this experiment. For the D2R antagonist study, one animal was also excluded due to computer failure, resulting in an n of 10 for this second experiment.

Full details of the apparatus and behavioural training, as well as further details of the task, are described in Chapters 2 and 3. All pharmacological challenges were carried out in well-trained rats who had achieved stable levels of performance.

2.2. Measures of performance and latency

The measures of performance and latencies recorded in the Go/No-Go task are outlined in detail in sections 1.4 and 1.5 of Chapter 2.

2.3. Pharmacological compounds

The agonist chosen for this experiment was the selective D2/3R agonist quinpirole hydrochloride (Tocris Bioscience). Quinpirole is the most widely used agonist of D2 receptors both in vivo and in vitro, with greatest densities of binding sites found in

the striatum, nucleus accumbens, and olfactory tubercles (Levant, Grigoriadis, & DeSouza, 1992).

The antagonist chosen was eticlopride hydrochloride (Tocris Bioscience), a D2/D3 antagonist with an 200-fold greater affinity to D2 receptors than other similar antagonists such as sulpiride (Martelle & Nader, 2008). However, as with quinpirole – and, indeed, all atypical antipsychotics – it is also a potent D3 receptor antagonist with a binding affinity of 8.8 pK_i, possibly even acting as an inverse agonist at these sites (Griffon, Pilon, Sautel, Schwartz, & Sokoloff, 1996), as well as having an affinity of 7.0 pK_i for the D4 receptor (Durcan, Rigdon, Norman, & Morgan, 1995) However, it still has greatest affinity at D2 receptors at 9.2 pK_i, (Tang, Todd, & O'Malley, 1994), with highest binding in the striatum, as well as a very low affinity for D1 receptors (Hall, Kohler, & Gawell, 1985).

2.4. Pharmacological challenges

Both quinpirole and eticlopride were administered systemically via i.p. injection 10 minutes before the behavioural session began. In all cases drugs were applied in a Latin square design, although here only one vehicle (saline) session was used by which the effects of both drugs have been compared.

Quinpirole was applied at 0.0125 and 0.0375 mg/kg based on studies showing that similar doses reduce impulsivity in the 5-choice serial reaction time task (Fernando et al., 2012) and modulates reward expectancy in a gambling task (Winstanley, Cocker, & Rogers, 2011).

Eticlopride was applied at 0.01 and 0.03 mg/kg also based on dosages applied in studies of risky decision-making (St Onge & Floresco, 2009; Winstanley et al., 2011), impulsive decision-making (van Gaalen et al., 2006), and waiting for reward (van Gaalen et al., 2006).

2.5. Data analysis approaches

A full discussion of the statistical approaches used in the analyses of the pharmacology data presented here can be found in section 2 of Chapter 2. The approach used here is the same as that used in Chapters 4 and 5.

3. Results

D2R Agonist

3.1. Systemic D2R stimulation exclusively reduces success rate in Go trials

If the Frank (2005) Go/No-Go model is accurate, then the D2R agonist should reduce Go trial success rate whilst improving success rate in No-Go trials. However, if D2Rs act in tandem with D1Rs to promote goal-directed behaviour, then their stimulation should in fact increase Go trial success rate and reduce No-Go trial success rate. In fact here, we found that Go trial success rate was profoundly reduced by D2R stimulation (Figure 1a), whilst No-Go success rate was unaltered by the drug (Figure 1b).

As shown in previous chapters, animals' success rate was higher for large reward than small reward Go trials [significant main effect of reward: $F_{1,8} = 8.375$, $\eta_p^2 = .511$, $p = .02$]. This was dose-dependently and linearly reduced by the D2R agonist regardless of the size of the reward on offer [significant main effect of drug: $F_{2,16} = 45.299$, $\eta_p^2 = .850$, $p < .001$; significant pairwise comparisons: vehicle vs. high dose, $p < .001$; low dose vs. high dose, $p < .001$; no significant drug*reward interaction: $F_{2,16} = 0.290$, $\eta_p^2 = .035$, $p = .752$].

Overall success in No-Go trials was not mediated by reward size on offer [no significant main effect of reward: $F_{1,8} = 0.571$, $\eta_p^2 = .067$, $p = .472$]. In addition, the D2R agonist had no effect on animals' ability to withhold action for reward [no main effect of drug or interaction: all Fs < 0.6, ps > .5].

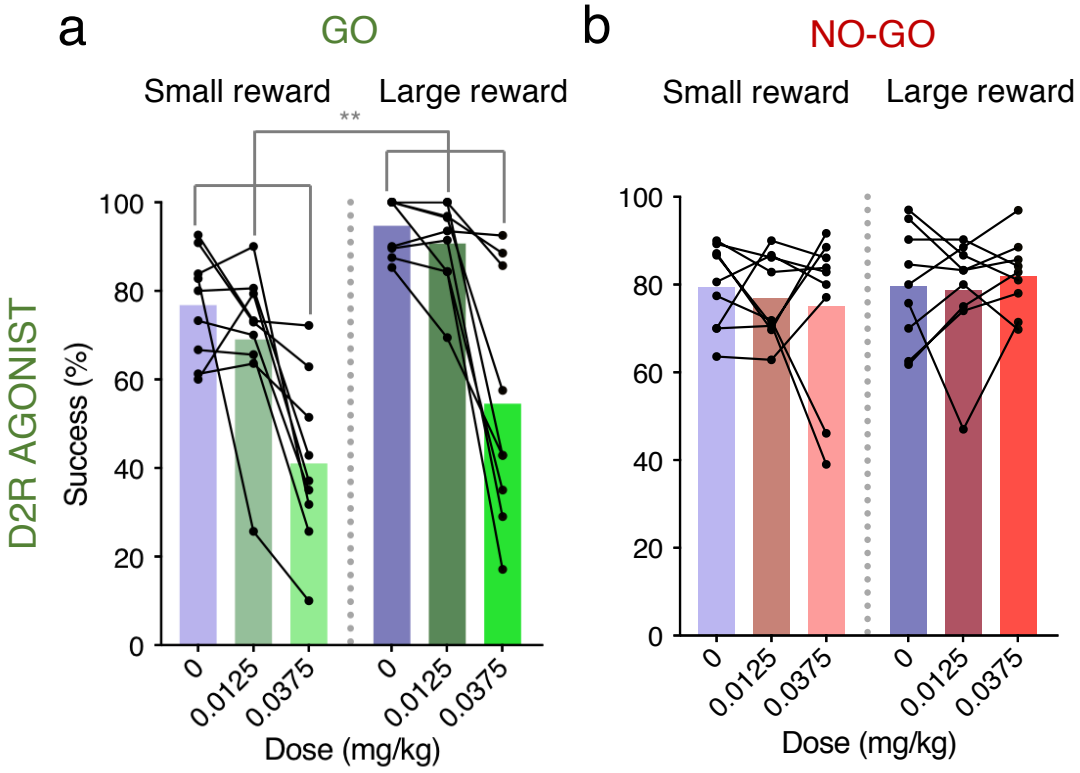


Fig. 1. Effect of systemic D2R stimulation (quinpirole) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) D2R stimulation reduced success rate in both small and large reward Go trials. (b) D2R stimulation had no effect on success rate in No-Go trials. Significance bars indicate a main effect of drug, ** p < .0.1.

To further understand the detrimental effect of D2R stimulation on success rate in Go trials, the types of errors were analysed. The reduction of Go trial success rate caused by the D2R agonist was due to a large, dose-dependent increase in Response Omission errors at both levels of reward [Figure 2b; significant main effect of drug: $F_{2,16} = 47.519$, $\eta_p^2 = .856$, $p < .001$; significant pairwise comparisons: vehicle vs. high dose, $p < .001$; low dose vs. high dose, $p < .001$; drug*reward interaction: $F_{2,16} = 0.216$, $\eta_p^2 = .026$, $p = .808$]. While there was a visual trend for Wrong Lever errors to be asymmetrically influenced by the drug (Figure 2a; reduced on small reward and increased on large reward trials), this was largely driven by two animals and overall the group was not reliably affected by D2R stimulation [no significant main effect of drug or interaction: all Fs < 4.0, ps > .07].

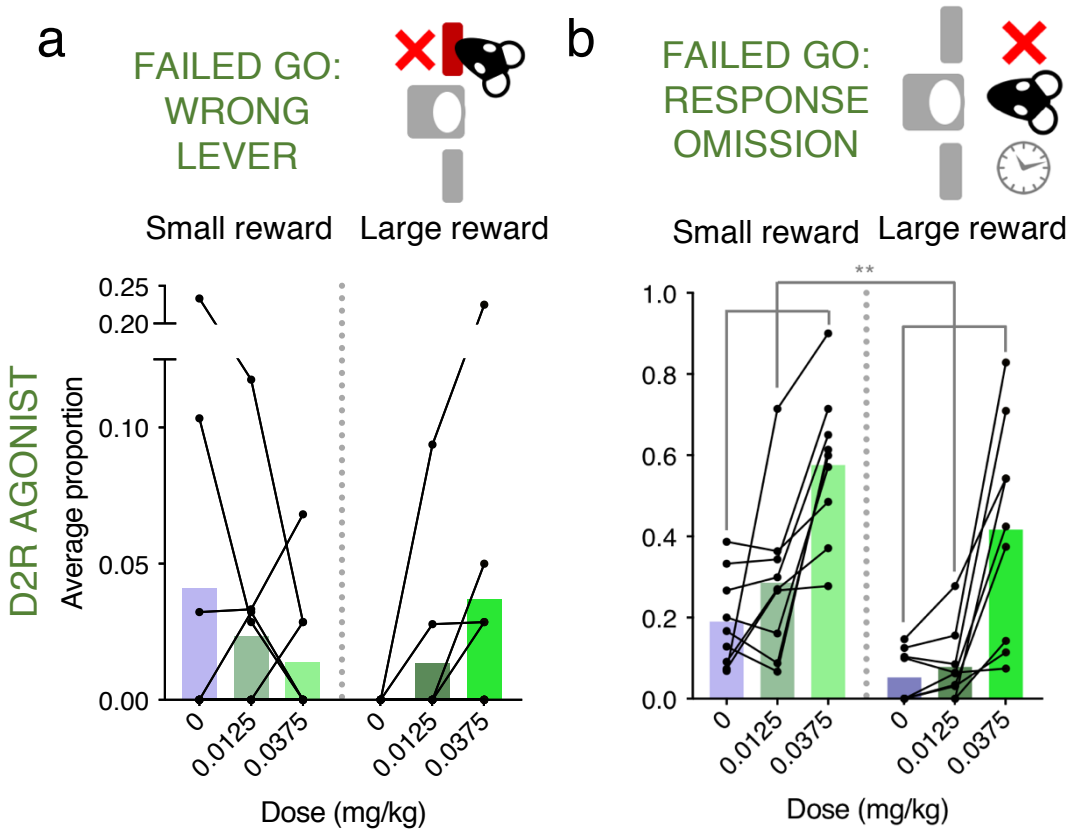


Fig. 2. Effect of D2R stimulation on the proportion of Go small or large trials where animals make (a) a wrong lever press or (b) omit response entirely during the 5s cue period. Significance bars indicate a main effect of drug, ** p < .01.

We next investigated No-Go trial errors; although there was no change in animals' likelihood of making an error with the D2R agonist on board, the *distribution* of such errors may have been altered by the drug. Again, however, a three-way within subjects ANOVA looking at 'early' or 'late' No-Go errors showed no effect of the D2R agonist on these errors [Figure 3; all *F*s < 2.5, *p*s > .1].

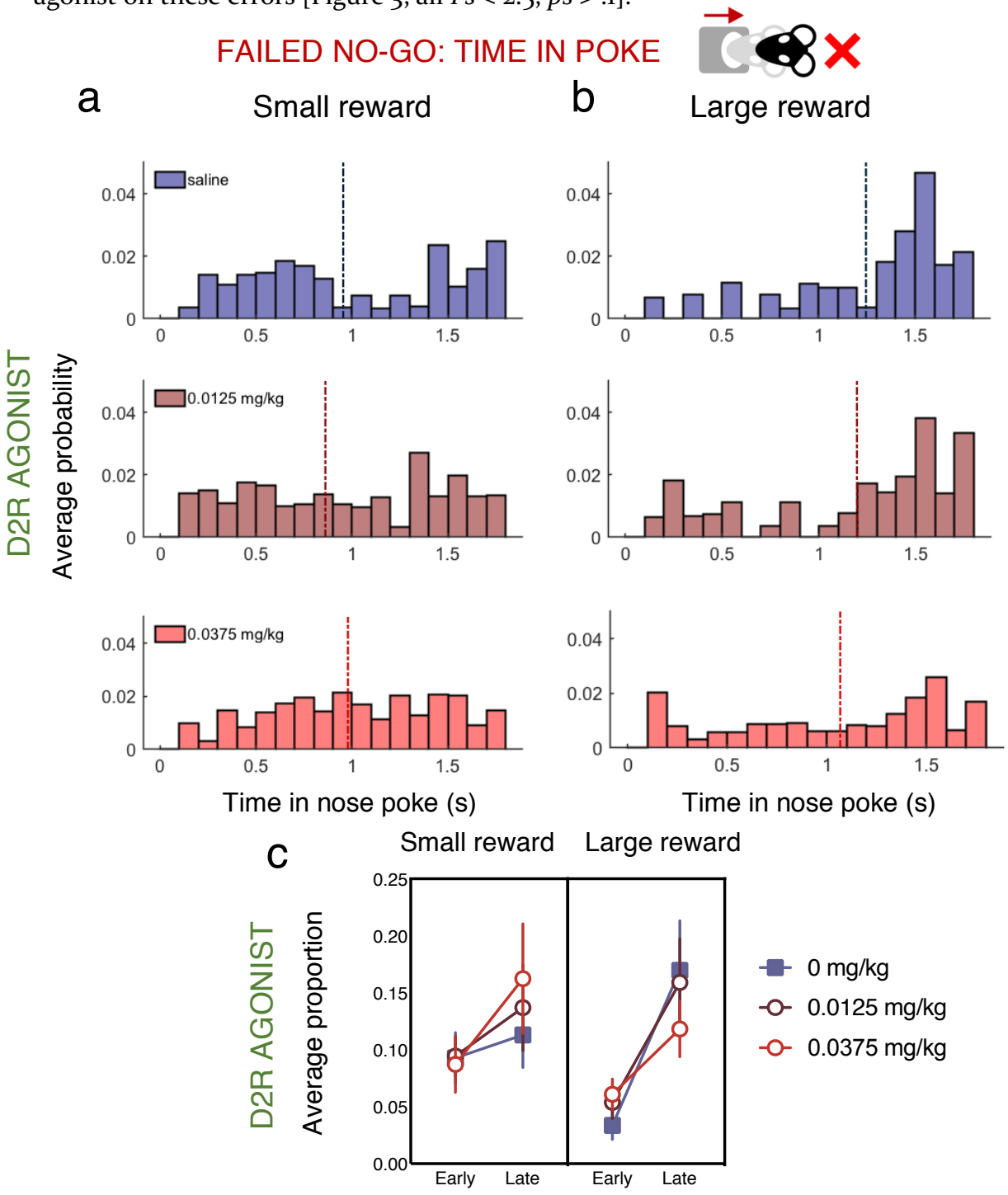


Fig. 3. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with D2R stimulation, calculated as probability over all head exit times, both successful and failed, for each drug dose. Last bin contains all values > 1.7s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.0125 mg/kg, red circle: 0.0375 mg/kg. The D2R agonist had no significant effect on early vs. late head exits.

3.2. *Systemic D2R stimulation slows cue-driven action initiation latencies*

We next examined how this same drug affected the speed of *appropriate* action initiation. As shown previously, rats were generally faster to leave the nose poke on correct Go trials with promise of a large reward [Figure 4a; significant main effect of reward: $F_{1,8} = 6.428$, $\eta_p^2 = .446$, $p = .035$]. The D2R agonist increased the latency to initiate action in successful Go trials regardless of the reward size on offer, though only at the highest dose [$F_{2,16} = 7.785$, $\eta_p^2 = .493$, $p = .004$; pairwise comparisons: vehicle vs. high dose, $p = .074$, low dose vs. high dose, $p = .015$; no drug*reward interaction: $F_{2,16} = 0.304$, $\eta_p^2 = .037$, $p = .742$].

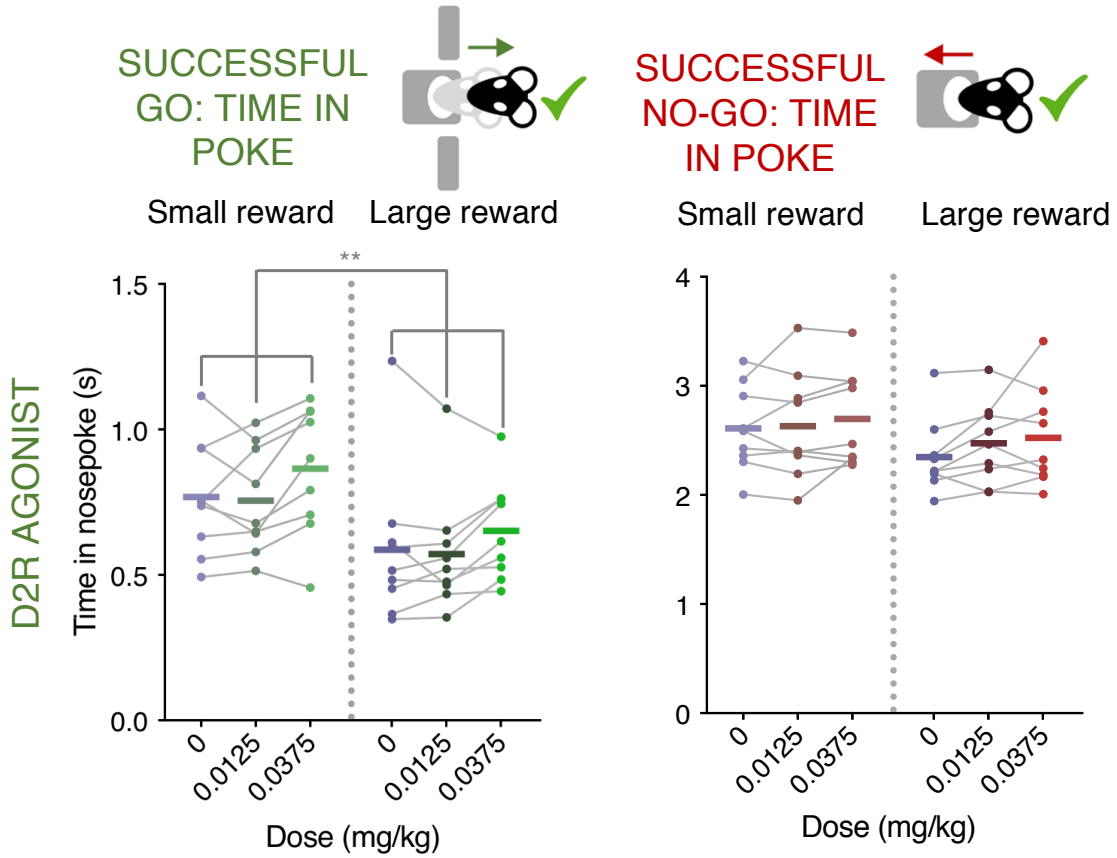


Fig. 4. Mean time spent in the nose poke (s) from cue onset in successful (a) Go or (b) No-Go trials. Significance bars indicate a main effect of drug, ** p < .01.

A finer-grained analysis of the distribution of response times by comparing the early, mid, and late points of the CDF distributions of successful small (Figure 5a) and large (Figure 5b) reward trials replicated the effect of the D2R agonist [significant main effect of drug: $F_{2,16} = 4.926$, $\eta_p^2 = .381$, $p = .022$]. However, while there were visual tendencies for the D2R agonist to have effects at different timepoints, this drug effect did not significantly interact with either reward size on offer or period of the distribution [no significant interactions: all Fs < 1.7, ps > .2].

We also investigated whether the latency to initiate appropriate action in successful No-Go trials, at point of cue *offset*, was similarly slowed by D2R stimulation (Figure 4b). However, this head exit latency was not altered by the drug [no main effect of drug or interaction: all Fs < 1.8, ps > .2]. Therefore, only cued action initiation in Go trials was impeded by the D2R agonist.

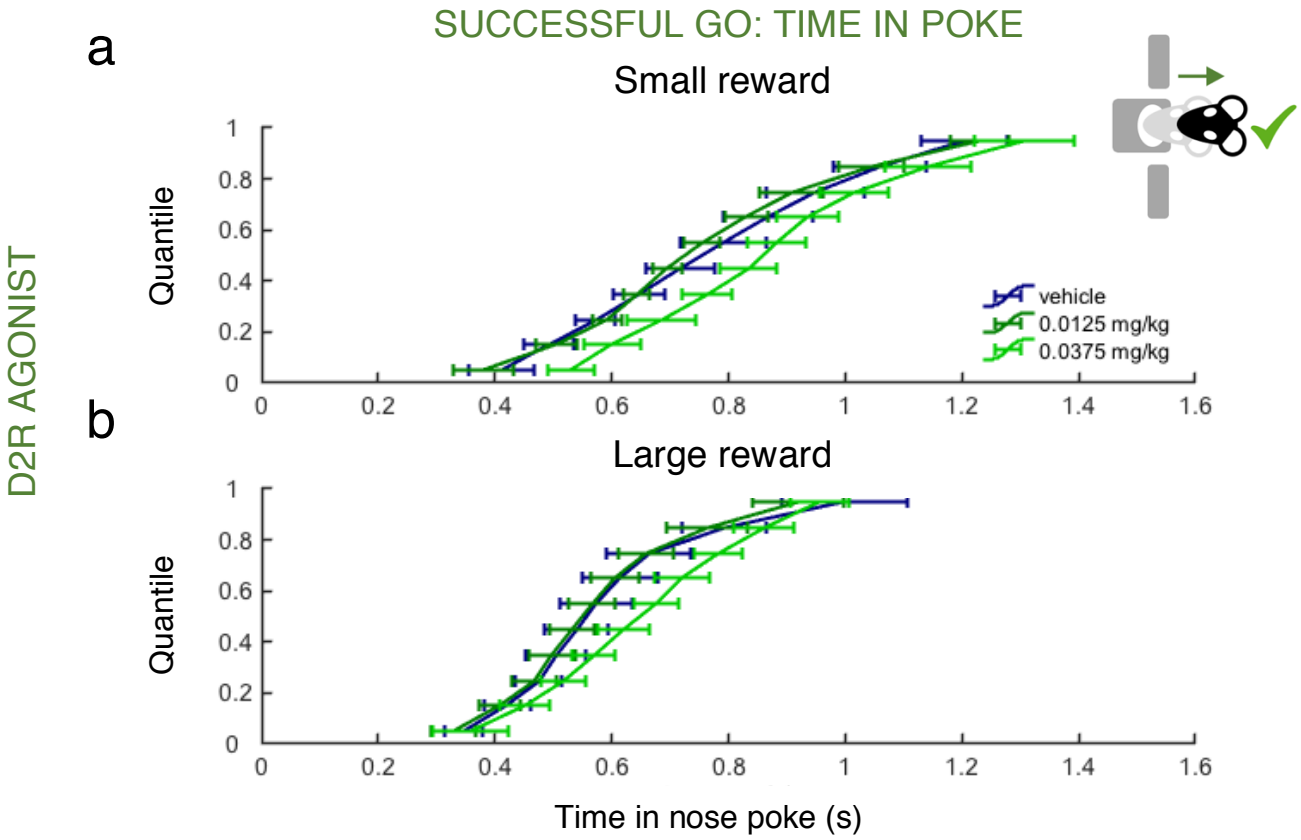


Fig. 5. CDFs for time in nose poke from cue in successful Go trials with D2R stimulation. (a) CDF showing slowing effect of quinpirole on head exit distributions for small reward trials. (b) Same as in (a) but for large reward trials, where there is again a slowing effect of the drug.

3.3. Systemic D2R stimulation reduces energising of ongoing goal-directed behaviour

D2R stimulation reduced cue-elicited action initiation. However, if activation of these receptors transmits information about reward rate, then their stimulation should *speed* movement throughout a Go trial.

This was not the case; whilst animals were faster to travel with promise of a larger reward [significant main effect of reward: $F_{1,8} = 5.929$, $\eta_p^2 = .426$, $p = .041$], the agonist dose-dependently increased travel times between head exit and the correct lever on successful Go trials, regardless of the size of the reward on offer [Figure 6a; significant

main effect of drug: $F_{2,16} = 14.173$, $\eta_p^2 = .639$, $p < .001$, linear ($p < .001$) but not quadratic ($p = .743$); pairwise comparisons: vehicle vs. low dose, $p = .093$ Sidak-corrected, vehicle vs. high dose, $p = .001$; no drug*reward interaction: $F_{2,16} = 1.467$, $\eta_p^2 = .155$, $p = .260$].

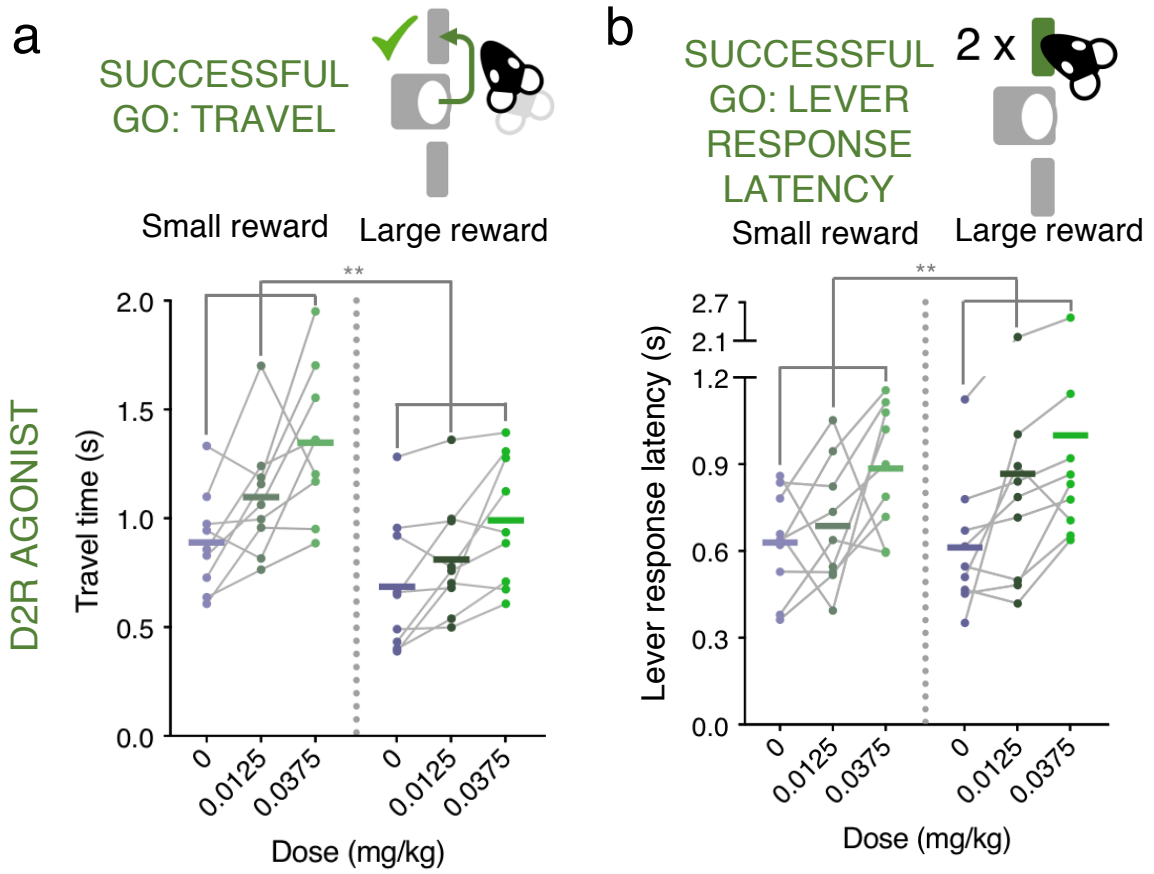


Fig. 6. Mean action execution latencies with the D2R agonist on board. (a) Mean travel time from nose poke exit in successful Go trials was significantly increased by D2R stimulation when both a small or large reward was on offer. (b) Mean lever response latency was also increased by the drug. Significance bars indicate main effect of drug, ** p < .01.

To understand whether all latencies across the distribution were similarly slowed, or whether this was a more selective effect, we compared drug effects throughout the travel time CDFs for small (Figure 7a) and large (Figure 7b) reward Go trials. Whilst the slowing effect of the drug was replicated [significant main effect of drug: $F_{2,16} =$

7.745, η_p^2 = .492, p = .004], this did not interact with either reward or latency period [no interactions: all Fs < 1.1, ps > .4].

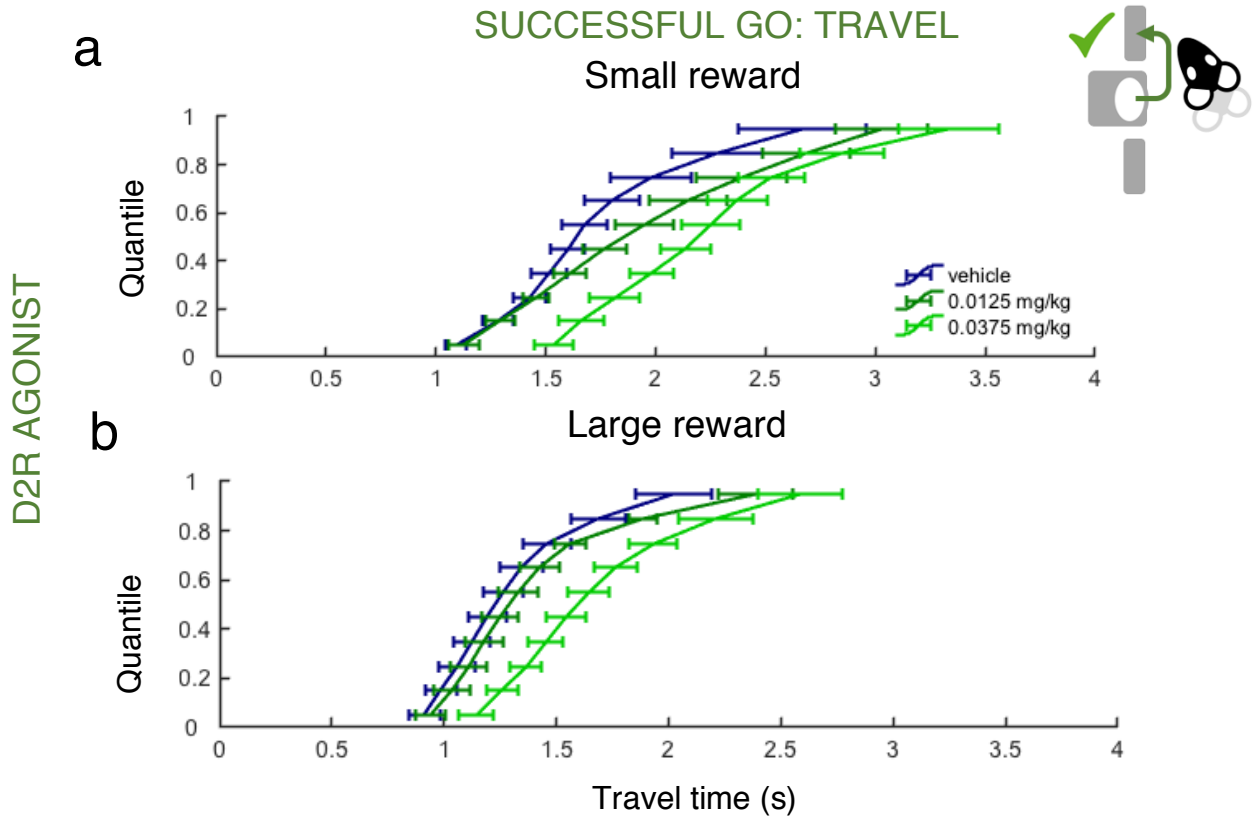


Fig. 7. CDFs for travel time from nose poke exit in successful Go trials with D2R stimulation. (a) CDF illustrating slowing effect of quinpirole on travel times for small reward trials. (b) Same as in (a) but for large reward trials, where there again is a slowing effect of the drug.

Similarly, whilst animals were no faster to complete the lever response requirements when a larger reward was on offer [Figure 6b; no significant main effect of reward: $F_{1,8}$ = 0.297, η_p^2 = .036, p = .601], this latency was also profoundly and dose-dependently slowed by the D2R agonist [significant main effect of drug: $F_{2,16}$ = 12.563, η_p^2 = .611, p = .001; within-subject contrasts; pairwise comparisons: vehicle vs. high dose, p = .004; low dose vs. high dose, p = .033; no drug*reward interaction: $F_{2,16}$ = 0.870, η_p^2 = .098, p = .438].

Unlike with pharmacological manipulations of D1 receptors, we even found that latencies to collect reward on successful trials were increased. On Go trials, the drug reward-dependently increased this latency, particularly when moving to collect a *small* reward [significant drug*reward interaction: $F_{2,14} = 7.962$, $\eta_p^2 = .532$, $p = .005$; significant small reward pairwise comparisons: vehicle vs. high dose, $p = .002$, low dose vs. high dose, $p = .001$; large reward pairwise comparisons, all $p > .1$; significant main effect of drug: $F_{2,14} = 34.459$, $\eta_p^2 = .831$, $p < .001$]. Similarly, on No-Go trials, latencies were dose-dependently slowed by the D2R agonist, although in this case the degree of the effect was not reward-mediated [significant main effect of drug: $F_{1.177,8.236} = 31.237$, $\eta_p^2 = .817$, $p < .001$; pairwise comparisons: vehicle vs. low dose, $p < .001$, vehicle vs. high dose, $p = .001$, low dose vs. high dose, $p = .013$; no significant drug*reward interaction: $F_{1.238,8.663} = 1.535$, $\eta_p^2 = .180$, $p = .256$].

Stimulating D2Rs also had profound effects on time taken to re-engage in the task (Figure 8). However, the highest dose of the drug slowed these latencies regardless of previous trial outcome [significant main effect of drug: $F_{1.336,10.685} = 30.459$, $\eta_p^2 = .792$, $p < .001$; pairwise comparisons: vehicle vs. high dose, $p < .001$, vehicle vs. low dose, $p = .374$, low dose v high dose, $p = .002$; no drug*trial outcome interaction: $F_{1.151,9.211} = 0.741$, $\eta_p^2 = .069$, $p = .564$]. Therefore, the D2R agonist slowed progression through all stages of trials.

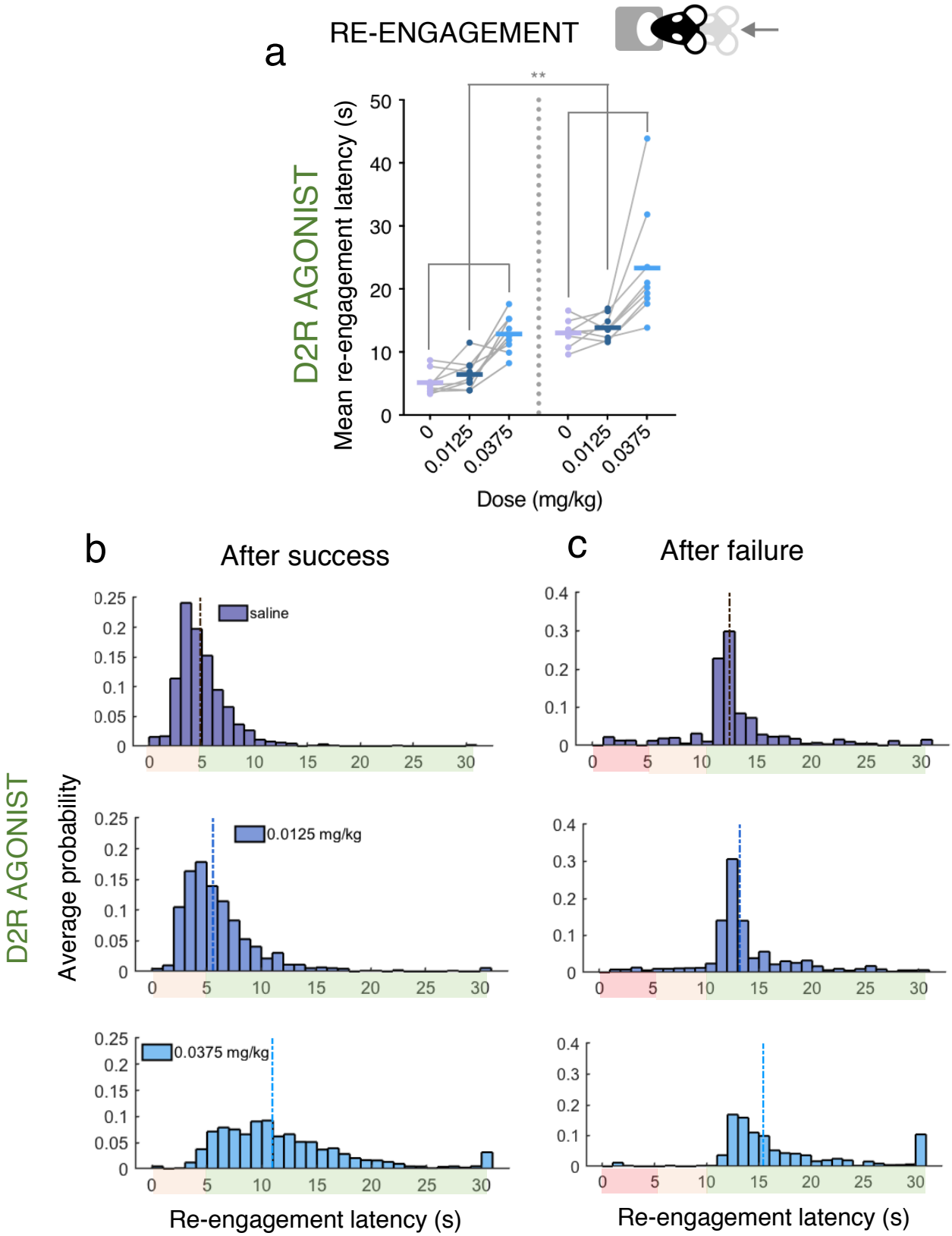


Fig. 8. Effect of D2R stimulation on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after successful trials increased with D2R stimulation, both after success (left of dotted line) and after failure (right of dotted line). (b) Histograms of re-engagement latencies after success for each drug dose. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (1s bins; last bin contains all responses after 30s; dotted lines indicate medians). (d) Same as in (b) but after failure. Light red is the timeout period during which the houselight is on. (1s bins; last bin contains all responses after 30s). Significance bars indicate main effect of drug, ** $p < .01$.

3.4. Systemic D2R stimulation promotes uncued movements

Analysis of mean rate of aborted trials (Figure 9) revealed that the D2R agonist in fact *increased* animals' propensity to make uncued movements, regardless of the outcome of the previous trial [significant main effect of drug: $F_{2,16} = 5.817$, $\eta_p^2 = .421$, $p = .013$; pairwise comparisons: vehicle vs. low dose, $p = .059$, vehicle vs. high dose, $p = .007$; no drug*trial outcome interaction: $F_{2,16} = 2.274$, $\eta_p^2 = .221$, $p = .135$]. This suggests that whilst cued behaviour is slowed by D2R stimulation, uncued behaviour can be potentiated by the drug.

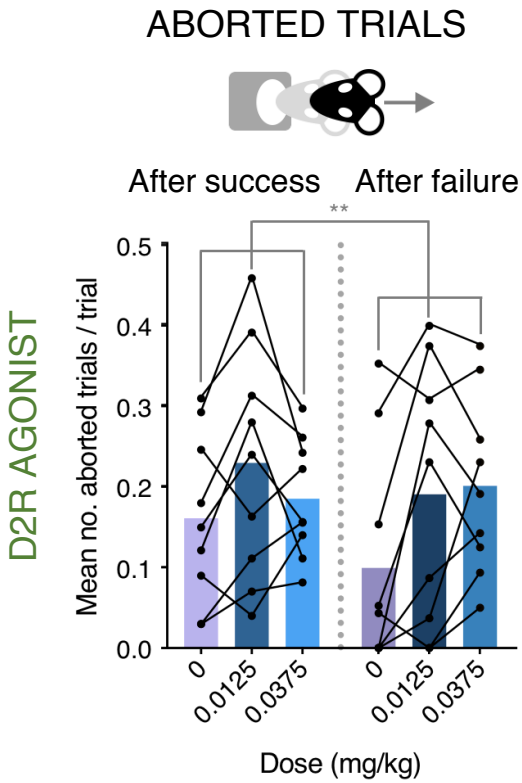


Fig. 9. Effect of D2R stimulation on mean rate of aborted trials (average number of aborted trials per trial). Significance bars indicate main effect of drug, ** $p < .01$.

D2R Antagonist

3.5. Systemic D2R blockade does not alter success rate in the Go/No-Go task

In a second experiment, we examined the effect of blocking D2Rs in the same Go/No-Go task. Converse to the inhibitory effects of D2R stimulation, we predicted that D2R antagonism should drive goal-directed behaviour. In fact, while there was a tendency for lower performance on Go trials at the highest dose, we observed no reliable effect of the drug on success rate [Figure 10a; no main effect of drug or interaction: all Fs < 2.9, p > .08; significant main effect of reward: $F_{1,9}$ = 62.391, η_p^2 = .874, p < .001] or similarly No-Go trials [Figure 10b; no main effects of drug or reward or interaction: all Fs < 0.4, ps > .6].

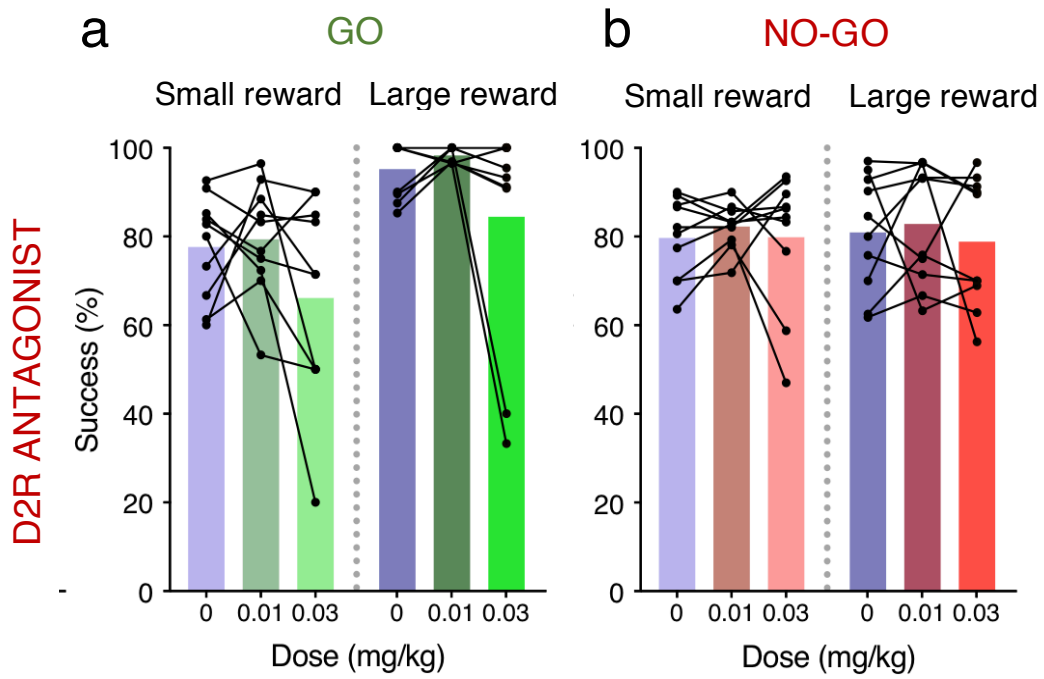


Fig. 10. Effect of systemic D2R blockade (eticlopride) on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Both (a) success rate in Go trials and (b) success rate in No-Go trials was unaffected by the drug.

The drug also did not have any consistent influence on the proportion of trials where animals made either error type [Wrong Lever error: no main effect of drug or interaction: all *F*s < 0.4, *p*s > .7; Response Omission error: no main effect of drug or interaction: all *F*s < 3.5, *p*s > .08].

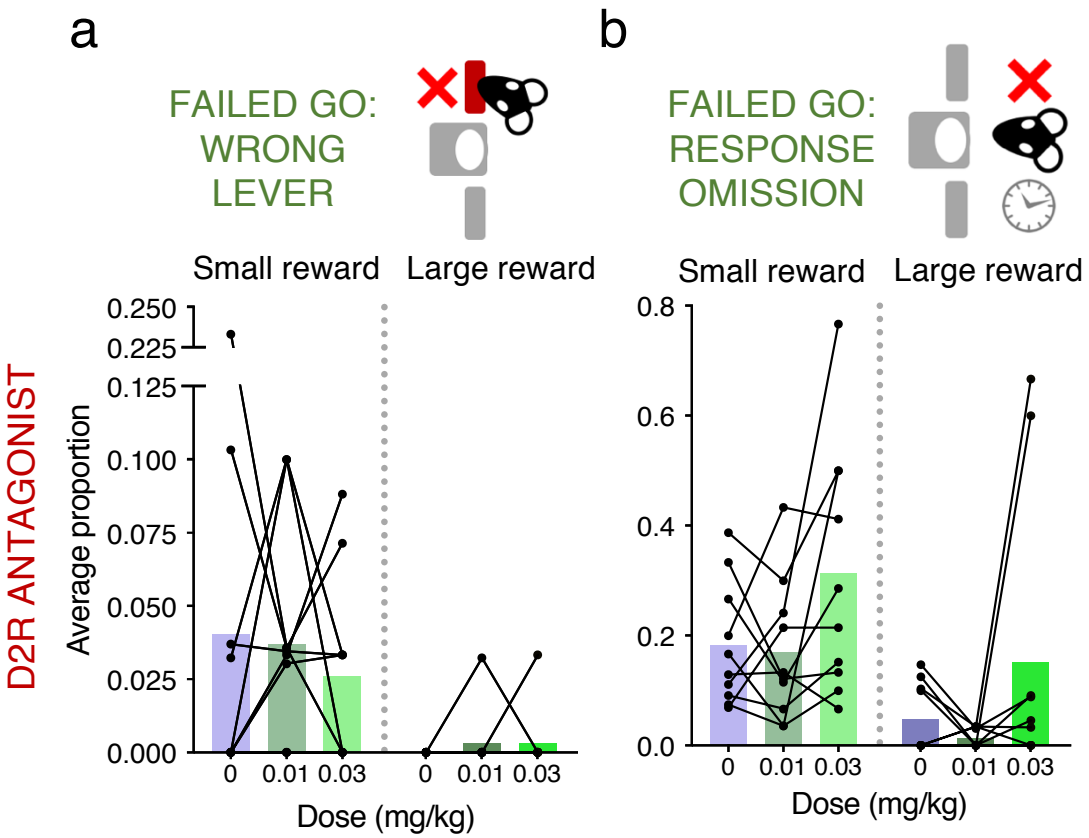
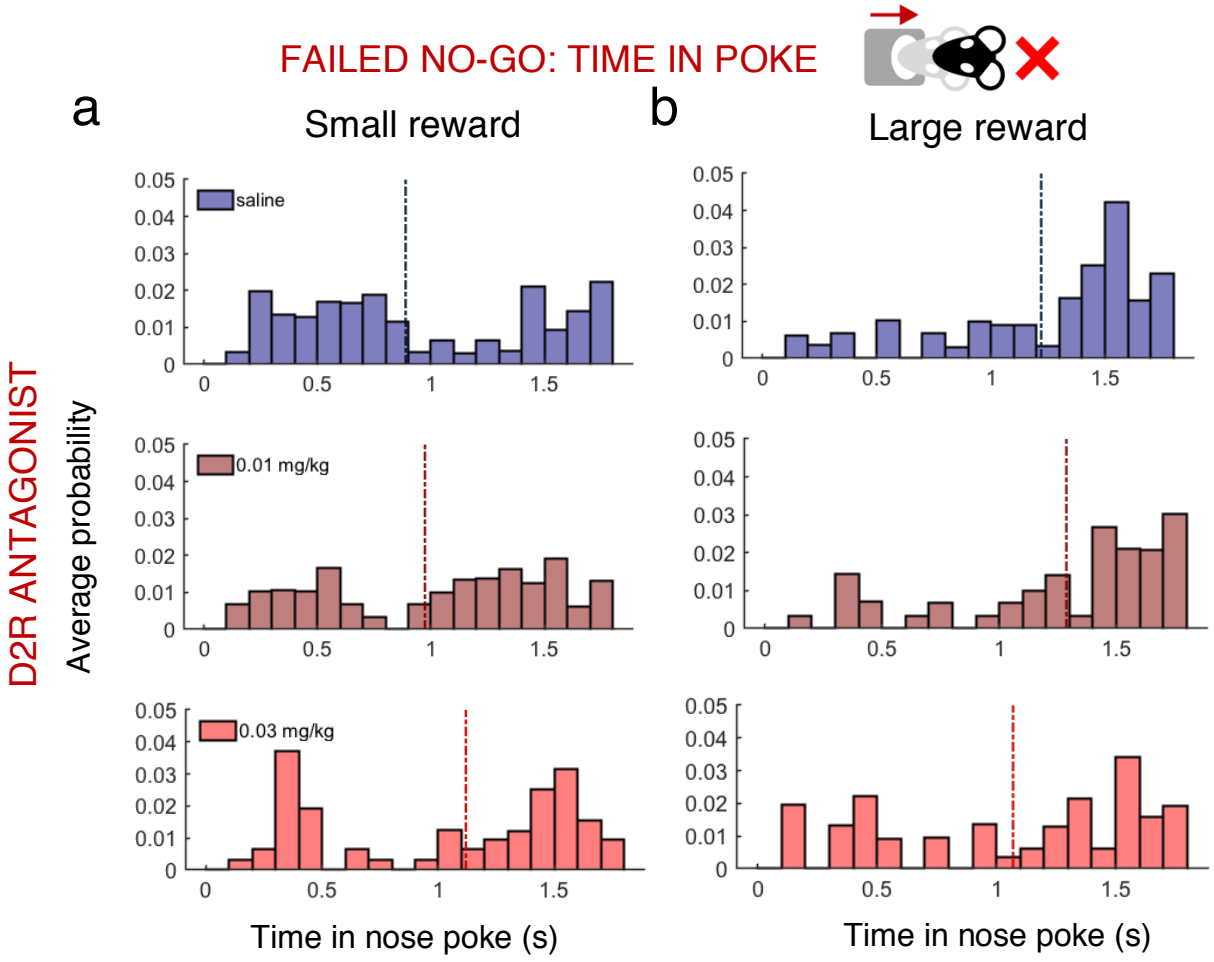


Fig. 11. Effect of D2R blockade on the average proportion of Go small (left of dotted line) or large (right of dotted line) trials where animals made (a) a wrong lever press or (b) omitted response entirely during the 5s cue period. The drug had no significant effect on either measure.

Similarly, distribution latencies for both small (Figure 12a) and large (Figure 12b) reward trials also show no effect of the drug [no main effect of drug or interaction with period or reward: all *F*s < 2.6, *p*s > .1]. Therefore, animals' ability to withhold movement for reward was not impacted by D2R blockade, both in terms of number of errors and when those errors were made.



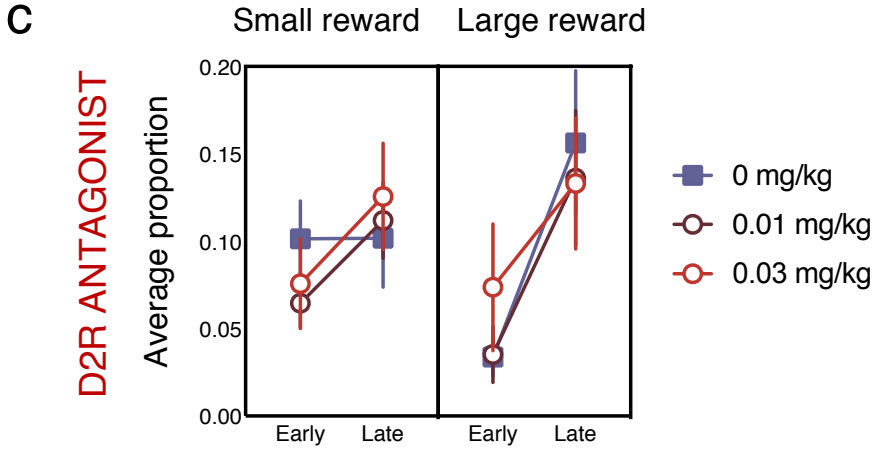


Fig. 12. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward failed No-Go trials with D2R blockade, calculated as probability over all head exit times, both successful and failed, for each drug dose. Last bin contains all values > 1.7s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.01 mg/kg, 0.03 mg/kg.

3.6. Systemic D2R blockade does not alter appropriate cue-driven action initiation

We next analysed time spent in the nose poke from cue in successful Go (Figure 13a) and No-Go (Figure 13b) trials. However, the drug did not alter this latency [no main effect of drug or interaction: all *F*s < 1.3, *p*s > .3].

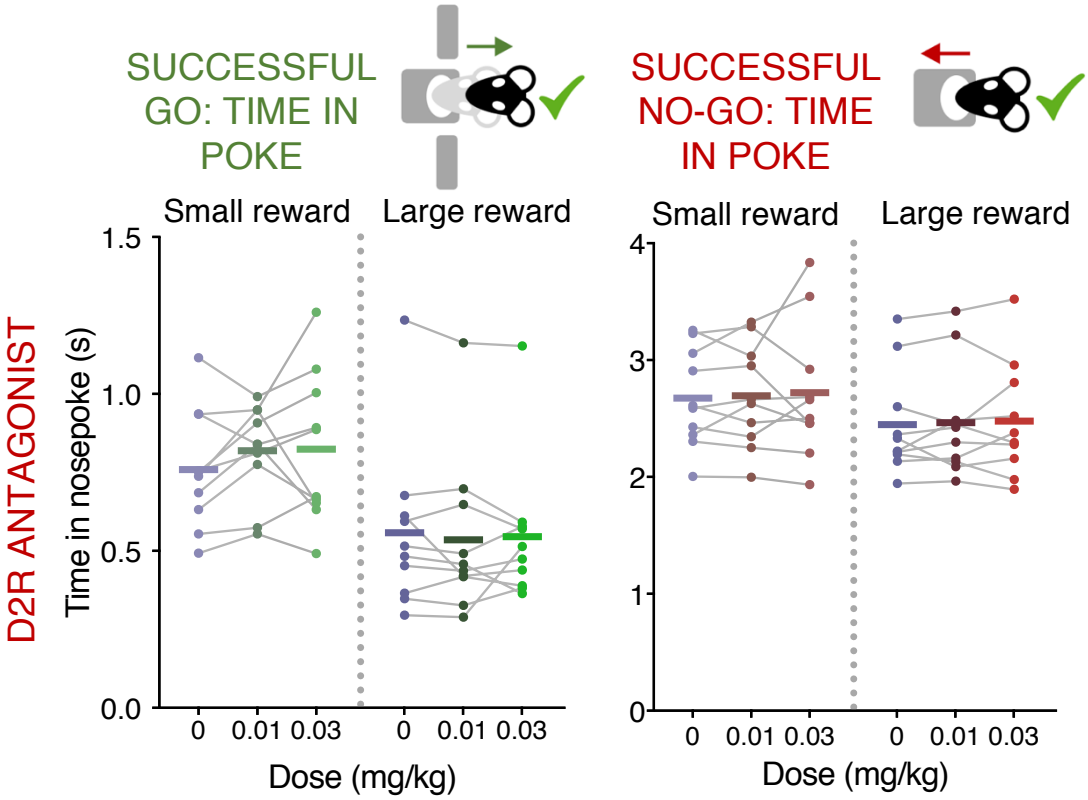


Fig. 13. Mean time spent in the nose poke (s) from cue onset in successful (a) Go or (b) No-Go trials with a D2R antagonist on board.

This was confirmed by a finer-grained analysis of time spent in the nose poke in successful Go trials, which also showed no effect of drug [Figure 14; no main effect of drug or interactions with reward or period: all *F*s < 2.0, *p*s > .1].

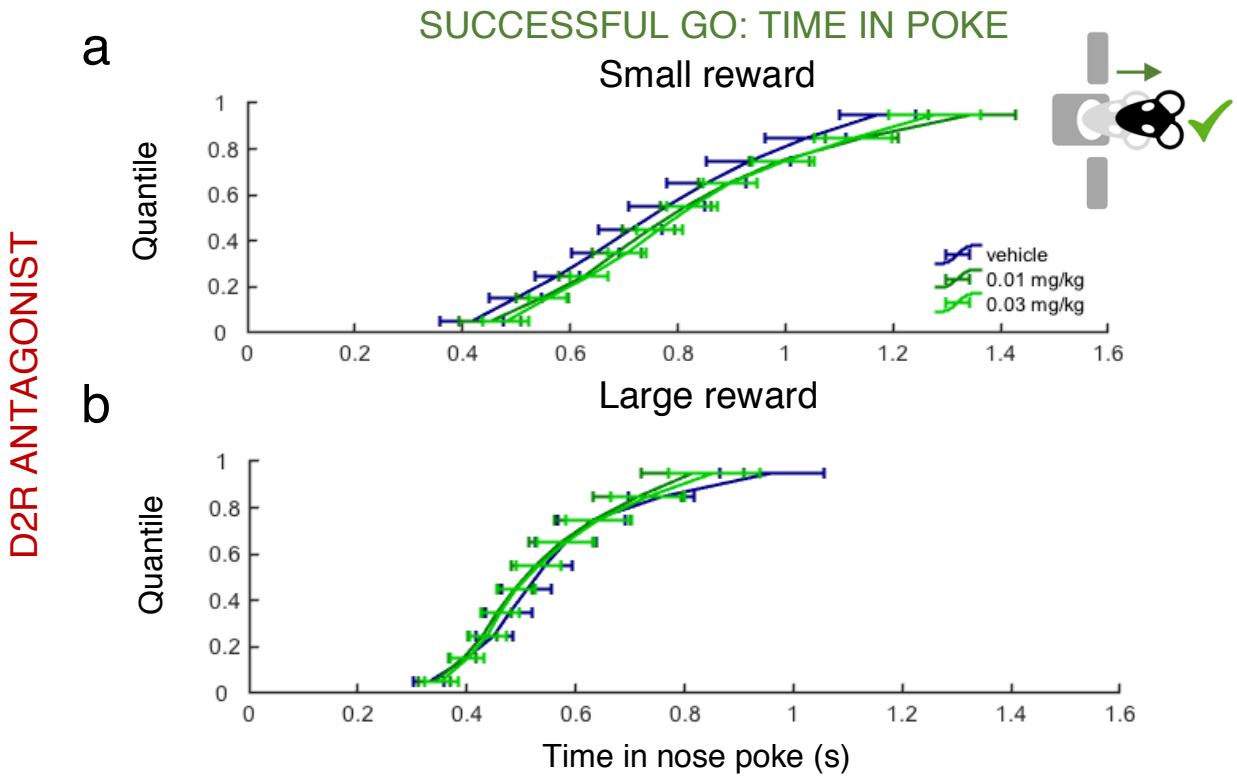


Fig. 14. CDFs for time in nose poke from cue in successful Go trials with D2R blockade. (a) CDF illustrating no effect of eticlopride on head exit distributions for small reward trials. (b) Same as in (a) but for large reward trials.

Similarly, the drug again had no effect on whether animals stayed in the nose poke for > 1.7s in successful Go trials [no main effect of drug or interaction: all *F*s < 2.7, *p*s > .09]. It also did not significantly affect No-Go head exit latencies after successfully withholding action for the cue duration [no main effect of drug or interaction: all *F*s < 0.2, *p*s > .8]. In summary, eticlopride did not reliably alter cue-driven action initiation in the Go/No-Go task.

3.7. Systemic D2R blockade does not alter the energising of goal-directed behaviour

The D2R antagonist had no effect on cue-elicited action initiation in Go trials. However, if these receptors encode average reward rate then their blockade should slow latencies to execute goal-directed behaviour.

Instead, mirroring the previous results, D2R blockade did not influence mean Go travel times [Figure 15a; no main effect of drug or interaction: all *F*s < 1.6, *p*s > .2]. This was also the case with a more detailed analysis of the first and second halves of the CDFs, with the drug having no effect on either early or late latencies [no significant main effect of drug or interactions with reward or period: all *F*s < 2.7, *p*s > .09]. Animals were also not faster to lever press for a larger reward [Figure 15b; no main effect of reward: $F_{1,9} = 0.145$, $\eta_p^2 = .016$, $p = .712$]. Similarly, there again was no effect of D2R blockade on this latency [no main effect of drug or interaction: all *F*s < 1.8, *p*s > .1].

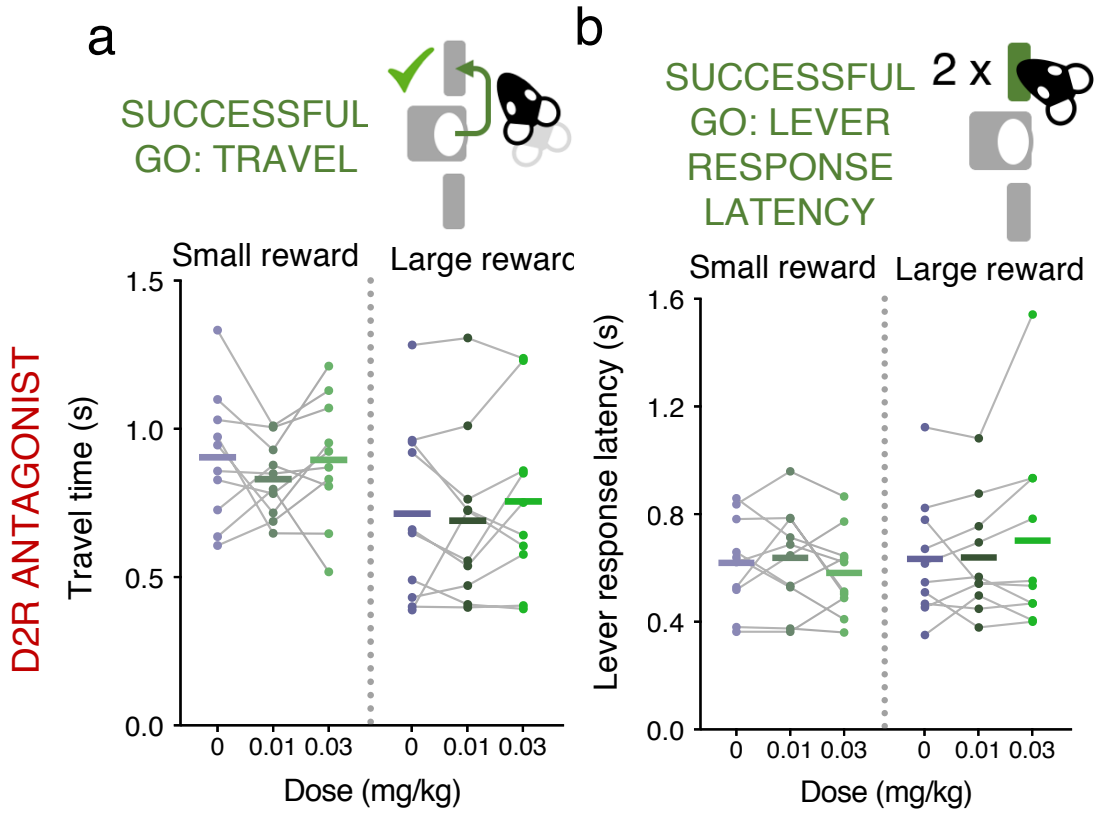


Fig. 15. Mean action execution latencies with the D2R agonist on board. (a) Overall (bars) and individual (dots) mean travel time from nose poke exit in successful Go trials. (b) Mean lever response latency in successful Go trials.

Time to collect reward from the food magazine was similarly unaffected by the D2R antagonist [no significant main effect of drug or interaction on either Go or No-Go trials: all Fs < 1.1, ps > .3].

Finally, animals' re-engagement latencies were also unaltered by the D2R antagonist [no significant main effect of drug or interaction: all Fs < 2.5, ps > .1]. This was also the case with analysis of median re-engagement, as plotted here due to several outliers [Figure 16; no significant main effect or interaction: all Fs < 2.4, ps > .1].

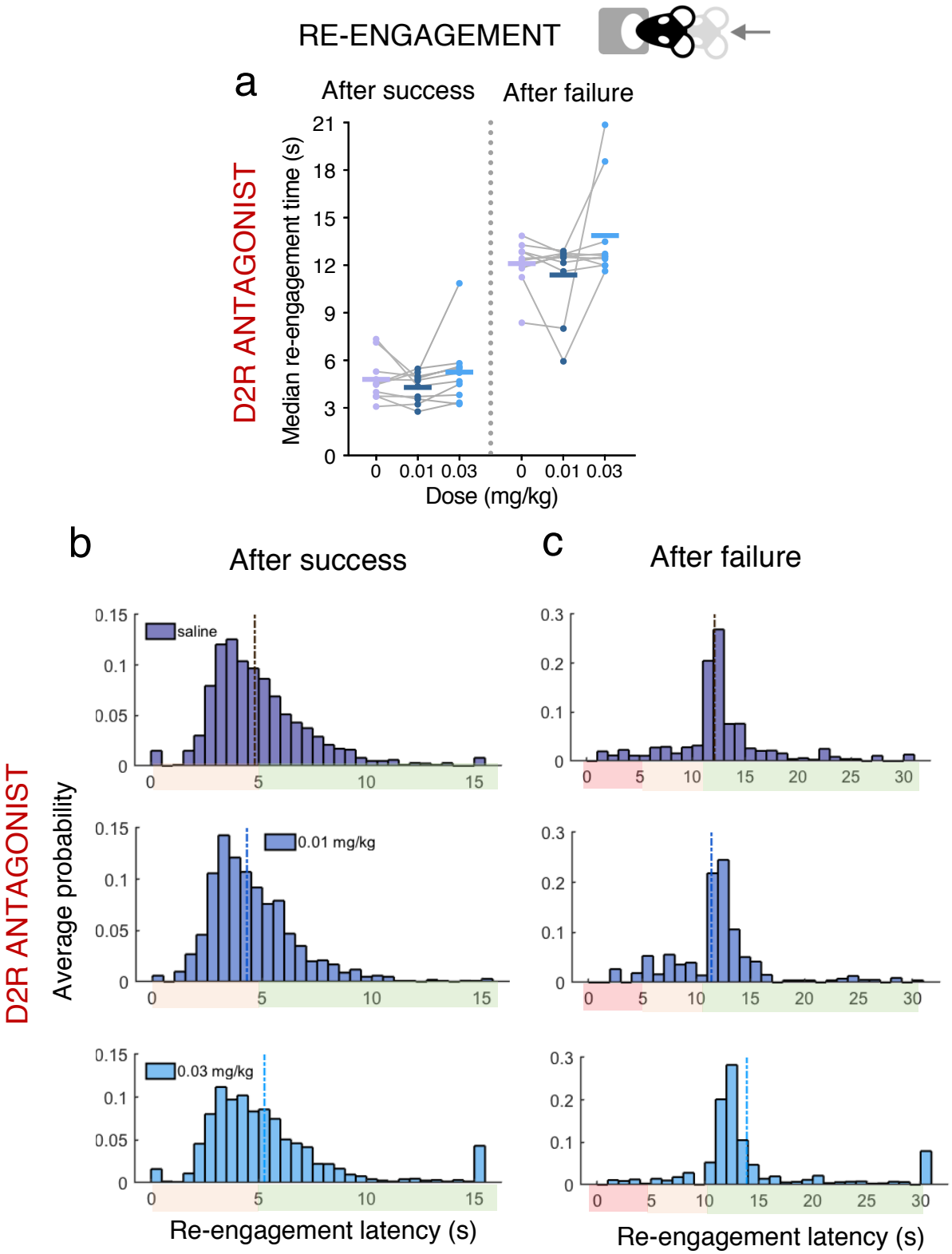


Fig. 16. Effect of D2R blockade on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after successful trials was unaltered with D2R blockade, both after success (left of dotted line) and after failure (right of dotted line). (b) Histograms of re-engagement latencies after success. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (0.5s bins; last bin contains all responses after 15s; dotted lines indicate medians). Dotted lines indicate median. (c) Same as in (b) but after failure, with 1s bins. Last bin contains all responses after 30s. Light red is the timeout period during which the houselight is on. (1s bins; last bin contains all responses after 30s).

3.8. Systemic D2R blockade does not promote uncued action initiation

Overall, animals aborted more trials after success than after failure [near-significant main effect of trial outcome: $F_{1,9} = 4.009$, $\eta_p^2 = .308$, $p = .076$]. However, the drug did not alter the degree to which animals initiated uncued action initiation [no significant main effect of drug or interaction: all Fs < 1.3, ps > .3].

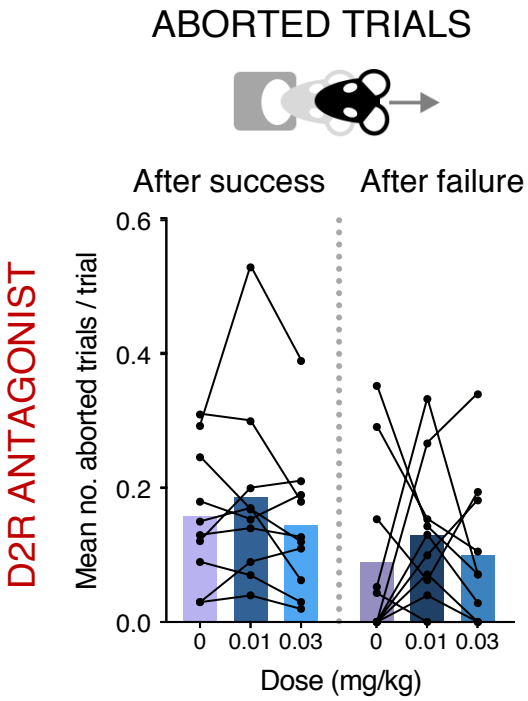


Fig. 17. Effect of D2R blockade on mean rate of aborted trials (average number of aborted trials per trial).

4. Discussion

Understanding how a single neurotransmitter, dopamine, is able to convey both reward and movement signals has become a topic of intense interest in the field. Thus far in this thesis, we have demonstrated the importance of phasic dopamine release and its potential consequences at D1Rs to both processes. However, the other principal form of the dopamine receptor family, D2Rs, may also play a unique role in

action for reward. In particular, these receptors have been linked both to calculation of average reward rate via fluctuations in levels of tonic dopamine, and to the inhibition of movement as part of the indirect pathway of the basal ganglia.

To parse these functions, here we either pharmacologically stimulated or blocked D2Rs systemically in the same Go/No-Go task used previously. In support of the hypothesis that D2R stimulation inhibits goal-directed action, we found that application of a D2R agonist slowed movement throughout trials where action was required, whilst having no impact when *withholding* of cue-evoked action was needed. In contrast, D2R blockade had few reliable effects on task metrics, suggesting that normal transmission at these receptors may not be necessary for successful performance in this task. Table 2 contains a summary of these findings, which are discussed in more detail subsequently.

	Systemic D2R agonist		Systemic D2R antagonist	
MEASURES OF PERFORMANCE				
Go errors	Wrong lever	Response omission	Wrong lever	Response omission
	—	↑	—	—
No-Go errors	—		—	
Aborted trials	↑		—	
MEASURES OF LATENCY				
Time in poke – successful Go	↑		—	
Travel time	↑		—	
Lever response latency	↑		—	
Food magazine latency	↑		—	
Re-engagement latency	↑		—	

Table 2. Summary of effects of both the D2R agonist (green) or D2R antagonist (red) on measures of performance and measures of latency in the Go/No-Go task. Large arrows indicate a strong effect, small arrows indicate smaller or more specific effects. For measures of performance, green arrows indicate an improvement and red arrows indicate a deterioration.

4.1. D2R activation inhibits action, but may not be necessary for goal-directed behaviour

The classical model of basal ganglia circuitry suggests that D2Rs act to inhibit competing actions as part of the indirect pathway, such that their overactivation can inhibit behaviour entirely (Albin, Young, & Penney, 1989; Frank, 2005). However, more recent data suggests that the direct and indirect pathways act *concurrently* to achieve behaviour, with stimulation of either driving goal-directed behaviour (Cui et al., 2013; Soares-cunha, Coimbra, Sousa, & Rodrigues, 2016a).

Here, we found evidence of the former: systemic D2R stimulation profoundly impeded action throughout Go trials, from action initiation (at the high dose of the drug) and then particularly through Go trial execution and reward retrieval at both low and high doses. These effects were indiscriminate, inhibiting action regardless of the size of the prospective reward. However, it was not the case that ability to *withhold* action was improved by the drug. In addition, latency to initiate action after a No-Go trial was similarly unaffected. This suggests that *movement* is an integral component of the effects of D2R stimulation in terms of inhibition of behaviour, such that a behaviour that does not require movement remains uninfluenced.

Yet whilst D2R stimulation had profound slowing effects on Go trials in the Go/No-Go task, interestingly blockade of these same receptors did little to alter behaviour, both in terms of performance and latency to complete trials. This was not predicted; whilst systemic application of the same D2R antagonist at the same doses has been shown not to modulate impulsive choice (van Gaalen et al., 2006) or responding (Blokland, Sik, & Lieben, 2005), it has been repeatedly shown that this drug slows latency to respond across different tasks both when applied systemically (White & Rebec, 1994; Winstanley et al., 2011) and when locally infused into the NAc core (Haluk & Floresco, 2009; Pattij et al., 2007). This may in part be due to the current study being underpowered, as both time in nose poke and proportion of Response Omission errors in Go trials both numerically increase with increasing drug dose, consistent with the literature.

In addition, whilst there was no statistically significant effect of the D2R antagonist on re-engagement latencies, the highest dose of the drug did result in a numerical increase in this latency. Inspection of the latency distributions suggest that, if animals re-engage shortly after trial completion, they are able to do so without any adverse effect. However, if they *disengage* from the task, they are likely to remain disengaged for longer periods of time. Similar effects have been found with other studies using the same D2R antagonist (Pattij et al., 2007), suggesting that activation of D2Rs is needed for sustained working for reward. And whilst this increase in re-engagement latency might seem counterintuitive in view of the direct/indirect model – inhibiting an inhibitory pathway should excite behaviour – it in fact makes more sense in light of proposals that the direct and indirect pathways act concurrently to achieve behaviour (Cui et al., 2013; Soares-cunha et al., 2016).

4.2. Presynaptic vs. postsynaptic effects of the D2R ligands

Whilst others have found that D2R activation is in fact *necessary* for goal-directed behaviour, as aforementioned we found that the D2R agonist slowed action initiation and execution in Go trials. One reason for this might be differential effects of the D2R ligands on autoreceptors and postsynaptic receptors in the striatum (Ford, 2014). It has been shown that lower doses of agonists such as quinpirole, typically up to 0.1 mg/kg, preferentially stimulate such D2 autoreceptors, causing decreased dopamine release and thus reduced movement (Lomanowska, Gormley, & Szechtman, 2004; Passetti, Levita, & Robbins, 2003). This is in contrast to effects found with stimulation of postsynaptic D2Rs, which results in eventual hyperactivity (Eilam & Szechtman, 1989). Therefore, the slowing of action latencies found with the D2R agonist here may

be locomotor in nature, and this is also implied by the lack of reward size sensitivity of the effect.

If the D2R agonist doses here were indeed preferentially acting at autoreceptors, this would result in a *decrease* of dopamine release into the synaptic cleft, including decreased amplitude of phasic transmission (Benoit-Marand, Borrelli, & Gonon, 2001; Escobar et al., 2015) and therefore decreased D1R activation. This is also broadly consistent with our findings from Chapter 4, which show that systemic D1R blockade slows many of the latencies that are also slowed with the D2R agonist here. However, there are also some important differences between these data. In particular, latency to retrieve reward from the food magazine was unaffected by the D1R antagonist, unlike here where the D2R agonist grossly slowed this movement. Similarly, effects on those latencies of the D1R antagonist, such as head exit time in Go trials, were sometimes reward-mediated, which was not the case here. Therefore, whilst some of the behavioural effects we have observed with D2R stimulation may well be caused by decreased phasic dopamine transmission, this is not able to explain all of the data.

In addition, the lack of such slowing with *local* administration of the D1R antagonist in Chapter 5 suggests that the effects seen both here and in Chapter 4 may be mediated by dorsal, rather than ventral, striatal neurons. Indeed, more recently it has been found that the direct and indirect pathway model of dopamine receptor function does not hold in the nucleus accumbens, as D2-MSNs in this region – unlike in the dorsal striatum – project to an output nucleus directly (Kupchik et al., 2015). In fact, activation of *ventral* striatal D2-MSNs has been shown to drive motivated

behaviour (Soares-cunha, Coimbra, Sousa, & Rodrigues, 2016). Thus the increase of action execution latencies with a systemic D2R agonist on board here is likely localised to autoreceptors in the dorsal striatum.

Lower doses of D2 antagonists also tend to preferentially act at autoreceptors. Indeed, other studies that have used eticlopride at the same doses applied here have found a reduction in response latencies at 0.01 mg/kg but an *increase* in those same latencies at 0.3 mg/kg (Rogers, Wong, McKinnon, & Winstanley, 2013). The authors in this case suggested the former might reflect presynaptic D2 activity and thus enhanced dopamine release, but postsynaptic D2 activity at the higher dose, consistent with other studies showing a similar biphasic effect of D2R-targeted ligands (Geyer, Russo, Segal, & Kuczenski, 1987). If this is indeed the case, then at least the lower dose of eticlopride used here should consequently have induced an *increase* in striatal dopamine levels to therefore mimic the effects of the D1R agonist. Yet whilst both drugs did increase re-engagement latency, their effects were otherwise disparate. This, in combination with the literature, suggests that whilst the D2R agonist may be predominantly mediating the effects of presynaptic receptors, the D2R antagonist – particularly the high dose, but perhaps both doses – may in fact also be acting at postsynaptic receptors, allowing for an interpretation of such effects in the context of reward rate and action selection hypotheses.

4.3. Levels of D2R activation do not encode reward rate

One aim of the experiments in this chapter was to understand whether hypotheses that tonic dopamine levels reflect reward rate, and therefore increases or decreases

the vigour of goal-directed behaviour as appropriate, hold true (Niv et al., 2007). This was studied by manipulating levels of D2R activation, as these receptors are thought to be more sensitive to slower, tonic changes in dopamine release due to their high affinity at baseline (Richfield, Penney, & Young, 1989). This resulted in the hypothesis that stimulating such receptors should mimic an increase in reward rate and thereby increased action vigour, whilst blockade of these same receptors should cause the converse.

In fact, we found the opposite effect with D2R stimulation, and, importantly, no effect on such latencies with D2R blockade. However, before refuting this hypothesis, there are several aspects of the manipulation that should be considered. First, we have based this hypothesis on the purported sensitivity of D2Rs to tonic dopamine release. However, their physiology is likely more nuanced, with evidence that these receptors are also sensitive to pauses in dopamine neuron firing (Dreyer et al., 2010) and even to *phasic* increases of dopamine release in the face of elevated basal levels of dopamine (Marcott et al., 2014). Whilst neither of these possibilities can directly explain the reduction in goal-directed behaviour we observed with D2R stimulation here, this is an added complication when attempting to understand the consequences of altered tonic dopamine levels due to reward rate alterations.

Second, the hypothesis that it is tonic, not phasic, dopamine that alters degree of motivation for reward itself has come under question. First, there is evidence that alteration of tonic dopamine levels can affect the vigour of response outside of opportunity cost (Zenon, Devesse, & Olivier, 2016). Second, it is not possible to

differentiate in such pharmacological studies between tonic and phasic contributions due to inherent temporal limitations of the method. This is particularly the case given it is likely that phasic signalling influences overall dopaminergic tone to mediate goal-directed behaviour (Phillips & Wightman, 2004). Hamid et al. (2016) more directly tested this by using both methods with a high and a low temporal resolution in order to capture tonic and phasic fluctuations in dopamine respectively, and found that although reward rate was indeed encoded by tonic levels of dopamine and related to increased motivation to work for reward, it was relative changes in *subsecond* fluctuations of dopamine that carried the motivation signal.

The caveat that D2Rs may not be exclusively sensitive to tonic dopamine release makes it problematic in drawing strong conclusions with regards to encoding of average reward rate by slow changes in dopamine levels. Additionally, as discussed above, it is likely that only the D2R antagonist manipulation acted postsynaptically, meaning that only evidence from this experiment can be used to refute or support this hypothesis. Finally, the lack of anatomical specificity of these manipulations adds further difficulty. However, with these caveats in mind, in our hands D2R blockade did not alter action vigour, suggesting that reward rate encoding is not reflected in tonic dopamine levels.

4.4. D2Rs in behavioural switching and action selection

Whilst D2R activation is broadly thought to impede execution of goal-directed behaviour, a more nuanced view is that they play an important role within the basal ganglia in the *selection* of one response over another via suppression of alternative

responses, and therefore in behavioural flexibility (Keeler et al., 2014). Indeed, NAc core D2R stimulation during the 5-choice has been shown to increase perseverative responding (Pezze et al., 2007), and in a gambling task has been shown to increase erroneous reward collection responses (Winstanley et al., 2011). Whilst this has been shown in the contexts of tasks in which animals are well-trained, similar to the experiments described here, we did not find such an effect on response selection in Go trials when stimulating D2Rs. This may be because here, there is no decision-making component (as in the gambling task) or response ambiguity (as in the 5-choice). Instead, all required responses are deterministic and clearly indicated to the animal throughout the trial. Therefore, whilst D2Rs may play a role in behavioural switching in the context of making decisions about which action to select, it is unlikely that they have such an influence in an unambiguous setting.

4.5. Summary

Here, D2R stimulation substantially reduced success rate in Go trials by decreasing the likelihood that animals would make an operant response, and also significantly slowed the latencies of both goal-directed action initiation and ongoing action, regardless of reward size on offer, most likely due to mediation of autoreceptors in the dorsal striatum. This suggests that dopamine is necessary to invigorate and sustain goal-directed actions and for task engagement. However, blockade of these receptors only subtly altered behaviour on this task by slowing latencies. Whilst this is consistent with an account of D2Rs driving goal-directed behaviour in synergy with D1Rs (Cui et al., 2013; Soares-cunha et al., 2016a), these effects were not consistent,

Therefore, whilst normal D1R function is necessary for performance in the Go/No-Go task, this appears not to be the case for D2-like receptors.

7 5-HT_{2C} receptors modulate ability to wait and enact action for reward

Chapter abstract:
Serotonin is integral to the ability to wait, or inhibit behaviour, for reward but it is not yet known whether this mediation of 'patience' is tied to movement *per se*. In addition, disruption of the serotonergic system, particularly of the 5-HT_{2C} receptor subtype, can in fact drive goal-directed behaviour. Here we show that effective systemic 5-HT_{2C} antagonism causes animals to be more impulsive, but that this is not a cue-driven effect. This ligand also raised incentive motivation as demonstrated by an increase in ability and latency to carry out action for reward. Moreover, the degree of these effects was determined by reward size on offer, but were not mediated by 5-HT_{2C} receptors in the NAc core. Together, these results demonstrate a role for this receptor subtype in action for reward, but in contrast to our manipulations of the dopaminergic D1 receptors, these are largely cue-independent.

For the studies presented in this chapter, Laura Grima and Mark Walton planned the experiments, Laura Grima and Oliver Harmson collected the data, and Laura Grima analysed and presented the data with input from Mark Walton.

1. Introduction

Serotonin has long been conceived of as the 'opponent' neurotransmitter to dopamine; whilst dopamine drives behaviour for reward, serotonin is thought to inhibit that drive (Cools et al., 2011; Soubrie, 1986). Indeed, whilst there is little evidence for a role of serotonin in action cancellation (Clark et al., 2005; Eagle et al., 2009, 2008), this neurotransmitter has been implicated in impulsive decision-making (Denk et al., 2005) but also particularly in enabling response *restraint* for reward; global depletion of 5-HT increases premature responding in the 5-choice task (Harrison, Everitt, & Robbins, 1997) and in a traditional Go/No-Go task (Harrison, Everitt, & Robbins, 1999). Similarly, tonic firing of dorsal raphe serotonergic neurons increases during the delay period whilst waiting for reward (Miyazaki, Miyazaki, & Doya, 2012) and optogenetic stimulation of these same neurons enables mice to withhold inappropriate action for longer periods (Fonseca, Murakami, & Mainen, 2015). Some of these effects have also been localised to transmission at the 5-HT_{2C} receptor (Winstanley, Theobald, et al., 2004). Therefore, the ability to *wait* for reward is strongly modulated by the serotonergic system, and the 5-HT_{2C} receptor in particular.

However, whether this alteration in 'patience' is tied to inhibition of movement *per se* is less well-established. Some studies have shown a slowing of waiting and movement times, but no effect on reaction times with photostimulation of 5-HT neurons (Fonseca et al., 2015). More direct studies of motor behaviour have shown a strong suppression of spontaneous locomotor activity with the same manipulation

(Correia et al., 2017). Yet these and the aforementioned pharmacological studies have been performed in the context of uncued action inhibition, with animals usually having to wait for a delay period of unsignalled length, whilst less is known about the effects of serotonergic manipulations on *cued* motor inhibition for reward.

Interestingly, in tasks of goal-directed *action* rather than inaction, disruption of the 5-HT system can in fact improve performance. Ligands that act broadly antagonistically on the 5-HT_{2C} receptor specifically have been shown to enhance incentive motivation (Browne & Fletcher, 2016; Simpson et al., 2011) and to increase willingness to work for reward (Bailey et al., 2016). Conversely, stimulating these receptors in the VTA *reduces* incentive motivation as measured by operant responding for sucrose (Valencia-Torres et al., 2017).

A possible mechanism by which this receptor may influence both action restraint and motivation is through control of mesolimbic dopamine transmission. Application of a 5-HT_{2C} antagonistic ligand in the ventral tegmental area leads to a dose-dependent increase in the basal firing rate of these neurons, as well as an increase in bursting activity and, consequently, an rise in mesolimbic dopamine levels in the accumbens (Matteo et al., 1999b). In addition, the nucleus accumbens itself is heavily innervated by the central serotonergic nucleus of the dorsal raphe (Brown & Molliver, 2000) and high levels of the 5-HT_{2C} receptor can be found within the striatum, most likely localised to GABAergic projection neurons (Eberle-Wang, Mikeladze, Uryu, & Chesselet, 1997). However, despite studies on their behavioural effects (Robinson et al., 2008b), less is known about these striatal 5-HT_{2C} receptors in terms of modulation

of accumbens dopamine directly, but several pathways – from the serotonergic and dopaminergic midbrain – are able to mediate dopamine levels in this region.

Outside of a role in behavioural inhibition, the serotonergic system has also been linked to value signalling; reward signals in 5-HT neurons of the dorsal raphe (Bromberg-Martin, Hikosaka, & Nakamura, 2010; Li et al., 2016; Nakamura, Matsumoto, & Hikosaka, 2008), and responses to reward-predicting cues (Cohen, Amoroso, & Uchida, 2015) have been shown. However, whether this is equivalent to dopaminergic prediction error encoding *per se* is less well-established. One study found that 5-HT neurons reflected mainly unsigned prediction errors, or surprise (Matias, Lottem, Dugué, & Mainen, 2017), and another has found differences in responses of these neurons dependent on reward size, with an increase in activity in response to a large reward as compared to a small reward – again not as a typical reward prediction error per *se*, rather in response to the absolute value of the reward regardless of whether it was expected or unexpected (Li et al., 2016). Therefore, whilst it is known that these neurons respond to reward, the behavioural relevance of this signal and how it interacts with behavioural inhibition or excitation is unclear.

1.1. Aims, approach, and hypotheses

The aim of the subsequent studies was to understand the role of the serotonergic system, particularly the 5-HT_{2C} receptor, in cued response restraint or execution and its interactions with reward size on offer. Additionally, due to known interactions with the dopaminergic system, we aimed to understand whether these effects could be localisable to the ventral striatum.

To study this, we applied a 5-HT_{2C} ligand, SB 242084, both systemically and locally into the NAc core as animals performed the Go/No-Go task. We predicted that, consistent with the literature linking the 5-HT system to ability to wait for reward, functional blockade of 5-HT_{2C} receptors would increase impulsive behaviour. In tandem, we also hypothesised that animals' performance in Go trials would increase as 5-HT_{2C} blockade drives motivated behaviour, and that this alteration in performance would be mediated by size of potential reward.

2. Methods

2.1. Animals

Two separate cohorts of animals were used in the experiments described in this chapter. For the first systemic 5-HT_{2C} study, a total of 12 male unimplanted Sprague Dawley rats were used that had previously been administered the D1R antagonist systemically (though these D1R data were ultimately not included in this thesis due several animals doing little with the antagonist on board).

For the local NAcC 5-HT_{2C} study, a total of 12 male Sprague Dawley rats that had previously been used for the D1R manipulations (described in Chapters 4 and 5) were used. One animal was excluded due to misplaced cannulae based on histological observations. Two further animals were excluded due to evidence that their cannulae were no longer patent, as ascertained by a positive control study using local NAc core infusion of amphetamine described in section 2.5 of the methods and 3.8 of the results

of this chapter, resulting in an n of 9 for this experiment (note that the pattern of results was unchanged if these two animals were included).

Two further systemic 5-HT_{2C} studies were conducted for replication purposes, one before implantation of the cannulae and one after the local infusions had been performed. Both used the same cohort of animals as the local infusion study, but also included two additional animals that had not taken part in the local infusion study due to time constraints, resulting in an n of 14. All rats were group housed throughout training and data collection.

Full details of the apparatus and behavioural training, as well as further details of the task, are described in previous chapters. One difference to note is that here, the animals in the first systemic study experienced a required No-Go duration of 1.7 - 1.9s, rather than the standard 1.5 - 1.7s used throughout the other experiments in this thesis (and the local and second systemic studies in this chapter). All pharmacological challenges were carried out in well-trained rats who had achieved stable levels of performance.

2.2. Measures of performance and latency

The measures of performance and latencies recorded in the Go/No-Go task are identical to those in previous Chapters.

2.3. Pharmacological compounds

SB 242084 (Tocris Biosciences) was used as the 5-HT_{2C} ligand in the experiments described here due to its high affinity – it is the most specific 5-HT_{2C} receptor agent developed to date (Kennett et al., 1997), with 100- and 160-fold selectivity over the closely related 5-HT_{2B} and 5-HT_{2A} receptor subtypes receptor subtypes (Di Matteo, Di Giovanni, & Esposito, 2000). It is typically referred to as a selective 5-HT2c antagonist, although its downstream effects are more complicated (see Bailey et al. 2016)

As SB 242084 is not easily dissolvable in liquid, 8% cyclodextrin in sterile saline combined with 25mM of citric acid was used as the vehicle solution in all studies in this chapter. The pH was tested and adjusted to between 6 and 7 using 5M NaOH (sodium hydroxide) when necessary.

In a positive control experiment, *d*-amphetamine hemisulphate (obtained from another lab) was used (Elverfors & Nissbrandt, 1992).

2.4. Pharmacological challenges

In the first systemic experiment, SB 242084 was administered systemically via i.p. injection 20 minutes before testing. The doses applied were 0.1 and 0.5 mg/kg calculated as the salt, as well as a separate vehicle session. These doses were based on studies that have shown both increased premature responding and decreased correct response latencies in the 5-choice serial reaction time task (Fletcher, Tampakeras, Sinyard, & Higgins, 2007b; Higgins, Enderlin, Haman, & Fletcher, 2003; Winstanley, Theobald, et al., 2004). In a second systemic experiment in a separate cohort of

animals, this same pharmacological agent was administered using the same procedure at the 0.1 mg/kg dose only.

In the local infusion experiment, SB 242084 was administered to the NAc core at either 0.2 μg/μl (low dose) or 1.0 μg/μl (high dose), in addition to a separate vehicle session, 10 minutes before behavioural training. As with the dopaminergic manipulations from Chapter 5, 0.5 μl was infused per hemisphere, resulting in an infusion of 0.1 μg or 0.5 μg per side. These concentrations were chosen to reflect other results that have shown increased impulsivity using this compound in the 5-choice serial reaction time task (Robinson et al., 2008). Additionally, in the positive control *d*-amphetamine local infusion experiment, the drug was administered at a dose of 10 μg/μl, as similar doses previously used in the ventral striatum have been shown to exert profound locomotor effects (Ikemoto, 2002). Due to the 5-HT_{2C} vehicle differing from the *d*-amphetamine vehicle, we used sessions where saline was locally infused as baseline for the D1R agonist and antagonist studies with which to compare the effect of *d*-amphetamine.

In all experiments, the ligand was applied in separate counterbalanced Latin square designs to reduce the influence of order effects.

2.5. Local infusion positive control study

At the time of local infusion of the 5-HT_{2C} ligand, animals had already experienced between 6 and 8 infusions as part of the experiments reported in Chapter 5. Therefore, to confirm ongoing patency of the cannulae, *d*-amphetamine was

subsequently infused as a positive control. Any animals in which the amphetamine infusion did not have an effect were excluded from analyses of the 5-HT_{2C} local experiment presented here.

Degree of change in two behavioural measures were used as criteria for exclusion. The first concerned change in **No-Go performance**, and the second concerned **time in nose poke in successful Go trials**. These measures were chosen as there is a clear prediction as to the direction of change due to *d*-amphetamine, and also because there is enough variance in both cases to allow for the degree of change to be detectable.

The degree of change required in order to be included in the final data set was at least half of the average change in at least one of these measures (at either reward size). In other words, as the average change in No-Go performance in low reward trials was a decrease in accuracy of 21%, if an individual animal's performance did not decrease by at least 10.5%, it may have been excluded. However, it was excluded only if the animal's average reaction time *also* did not decrease by at least half of the average. The direction of change was also incorporated; an increase in both of these measures would also lead to exclusion of an animal.

In practice, this lead to two animals being excluded, resulting in an n of 9 for the 5-HT_{2C} local infusion study here. In both cases, their No-Go performance and successful Go trial reaction times were unchanged with local infusion of *d*-amphetamine (as shown in section 3.8 of the results). All of the local 5-HT_{2C} data discussed in this

chapter only considered animals who *did* show an effect of the *d*-amphetamine infusion, and thus it is likely that any lack of effect of the 5-HT_{2C} ligand was genuine, rather than a byproduct of blocked cannulae or other cause of lack of diffusion of the drug such as gliosis or other accumulation of biological matter.

2.6. Data analysis approaches

A full discussion of the statistical approaches used in the analyses of the pharmacology data presented here can be found in section 2 of Chapter 2. The approach mirrored that taken in Chapters 4 for the systemic administration. For the systemic replication, data from the pre- and post-infusion experiments were both included, with experiment treated as a within-subjects factor.

3. Results

Systemic 5-HT_{2C} receptor manipulation

3.1. Systemic manipulation of 5-HT_{2C} receptors reduces success in No-Go trials and increases success in Go trials

If the 5-HT_{2C} receptor is integral to the ability to wait for reward, then its disruption should reduce success rate in No-Go trials. Additionally, if disruption of this receptor enhances goal-directed behaviour as shown by improved performance in other tasks, then the 5-HT_{2C} ligand should also improve Go trial success rate.

As predicted, the 5-HT_{2C} ligand reduced No-Go trial success rate, particularly when the prospective future reward was small [significant drug*reward interaction: $F_{2,22}$ =

5.179, η_p^2 = .320, p = .014; significant small reward trial pairwise comparisons: vehicle vs. low dose, p = .015; large reward trials, all p > .8; significant main effect of drug: $F_{2,22}$ = 4.159, η_p^2 = .274, p = .029].

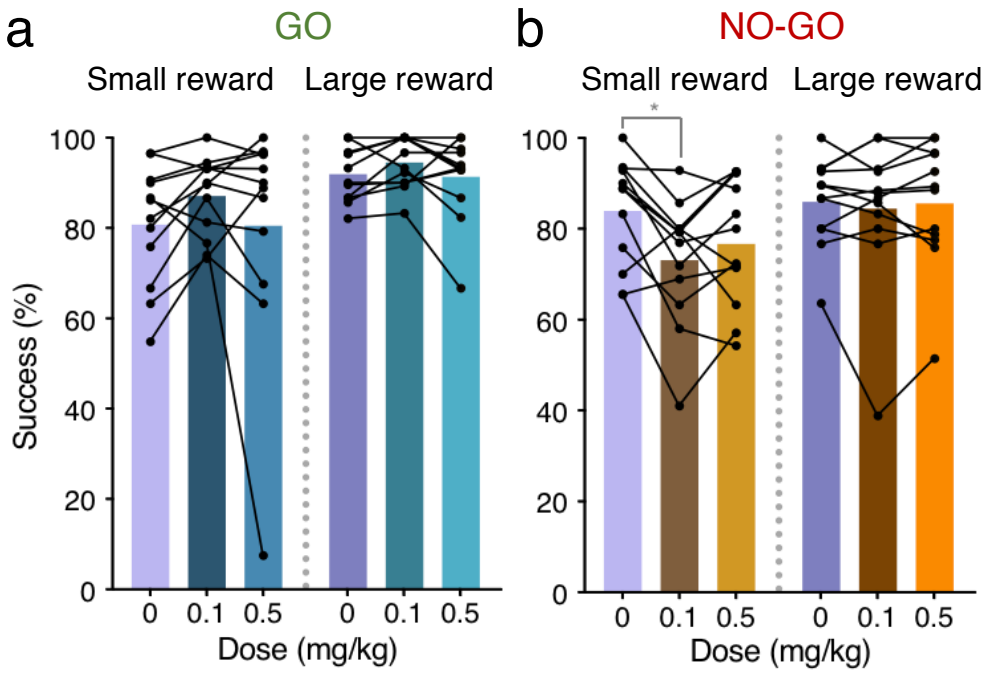


Fig. 1. Effect of the systemic 5-HT$_{2C}$ ligand on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials. Bars indicate overall mean, and dots are averages for individual animals. (a) 5-HT$_{2C}$ manipulation had no effect on Go trial success rate. (b) 5-HT$_{2C}$ manipulation reduced No-Go success rate in small reward trials. Significance bars indicate post-hoc comparisons, * p < .05, Sidak-corrected.

In Go trials, whilst there was no interaction between drug and reward [$F_{1.403,15.437}$ = 0.232, η_p^2 = .021, p = .718], there was a significant quadratic main effect of drug, caused by an increase in success rate at the low, but not the high dose, of the 5-HT$_{2C}$ ligand [$F_{1,11}$ = 11.147, η_p^2 = .503, p = .07; significant pairwise comparisons: vehicle vs. low dose, p = .012].

To understand whether the errors made in No-Go trials were driven by an increase in cue responding, we investigated *when* animals were leaving in failed No-Go trials. In contrast to the effect observed following D1 agonist administration, this showed an increased likelihood to leave *later* in the No-Go period, particularly in small reward trials (Figure 2a).

As before, to more formally quantify this, we analysed average proportions of head exits times in terms of whether they were 'early' or 'late' (Figure 2c; although here, due to the longer required No-Go holding period of 1.7 – 1.9s, 'early' was defined as < 900ms and 'late' defined as > 900ms). This revealed that, in general, animals were more likely to leave late in the No-Go period [significant main effect of period: $F_{1,11}$ = 8.347, η_p^2 = .431, p = .015]. However, this propensity was altered by the 5-HT_{2C} ligand; whilst early erroneous head exits remained unchanged by the drug, *late* head exits significantly increased, particularly at the low dose [near-significant drug*period interaction: $F_{2,22}$ = 3.404, η_p^2 = .236, p = .052; early pairwise comparisons: all ps > .7; late pairwise comparisons: vehicle vs. low dose, p = .052, vehicle vs. high dose, p = .150]. In addition, the effects of the drug were greatest in small reward trials, although this did not result in a three-way interaction [significant drug*reward interaction: $F_{2,22}$ = 5.179, η_p^2 = .320, p = .014; small reward pairwise comparisons: vehicle vs. low dose, p = .015; large reward pairwise comparisons: all ps > .8; no significant three-way interaction: $F_{2,22}$ = 1.286, η_p^2 = .105, p = .296].

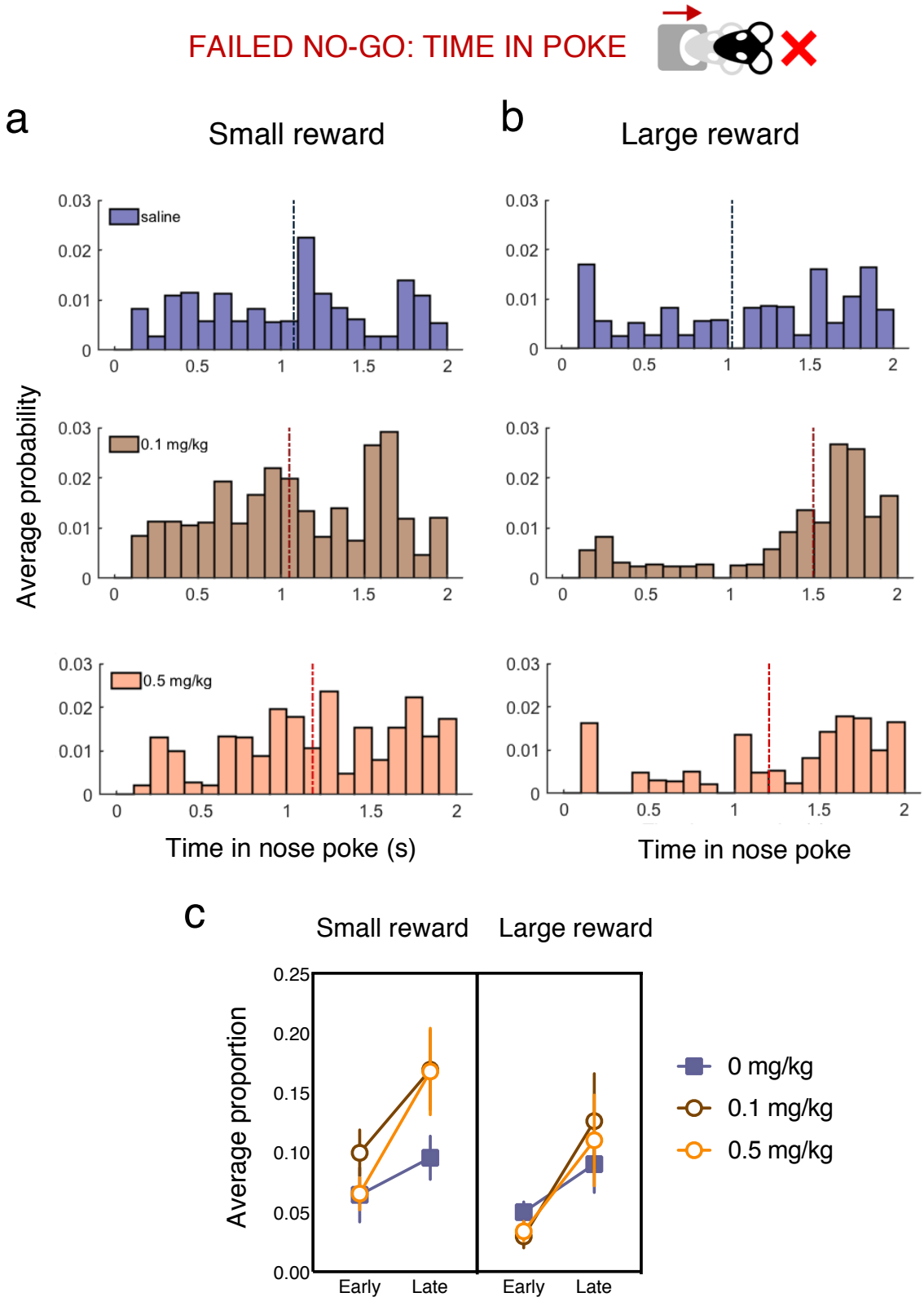


Fig. 2. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with 5-HT_{2C} manipulation, calculated as probability over all head exit times, both successful and failed, for each level of drug dose. Last bin contains all values > 1.9s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 900ms) or late (> 900ms) dependent on drug dose applied (square: vehicle, brown circle: 0.1 mg/kg, orange circle: 0.5 mg/kg.

The 5-HT_{2C} ligand increased success rate in Go trials at the low dose. To understand whether this was related to a decrease in a specific type of Go trial error, Wrong Lever (Figure 2a) and Response Omission (Figure 2b) errors were analysed.

Animals were no more likely to incorrectly lever press when a small reward was on offer [no main effect of reward: $F_{1,11} = 1.395$, $\eta_p^2 = .113$, $p = .263$], and this propensity was also unaltered by the drug [no effect of drug or interaction: all Fs < 1.5, ps > .2]. In contrast, animals were less likely to omit responding at the low dose of the 5-HT_{2C} ligand [quadratic main effect of drug: $F_{1,11} = 7.670$, $\eta_p^2 = .411$, $p = .018$; significant pairwise comparisons: $F_{1,11} = 7.670$, $\eta_p^2 = .411$, $p = .018$; no significant main effect or interaction: all Fs < 1.2, ps > .3]. Therefore, the drug improved success rate in Go trials by increasing animals' propensity to respond.

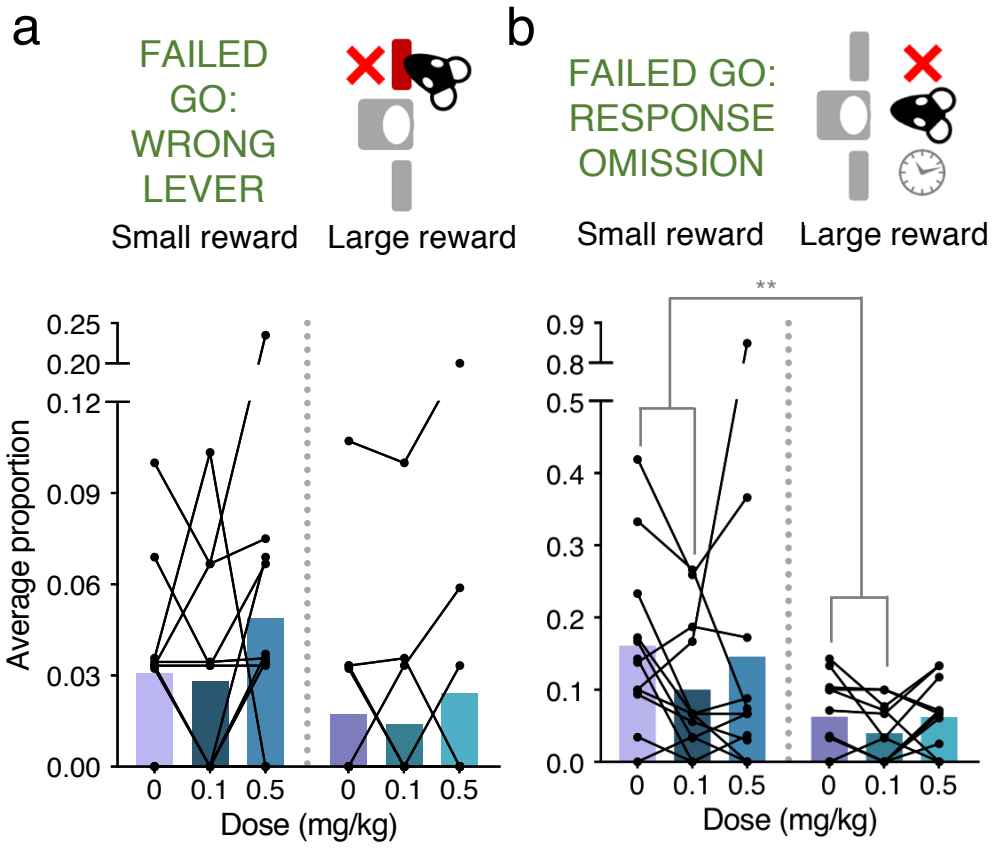


Fig. 3. Effect of the 5-HT$_{2C}$ ligand on the proportion of Go small or large trials where animals make (a) a wrong lever press or (b) omit response entirely during the 5s cue period. Significance bars indicate pairwise comparisons of main effect of drug dose, ** $p < .01$.

3.2. *Systemic manipulation of 5-HT$_{2C}$ receptors does not alter speed of cue-driven action initiation*

We previously showed that animals made more incorrect responses in No-Go trials when it was inappropriate to initiate action, but that they were more likely to leave *late* in the No-Go period of such error trials. We next examined how this drug affected the speed of appropriate cue-driven action initiation.

In both correct Go (Figure 4a) and No-Go (Figure 4b) trials, there was no effect of the 5-HT$_{2C}$ ligand, regardless of whether animals were leaving the nose poke at cue onset

in Go trials [no main effect of drug or interaction: all *F*s < 2.4, *p*s > .1] or leaving at cue offset in No-Go trials [no main effect of drug or interaction: all *F*s < 0.9, *p*s > .4].

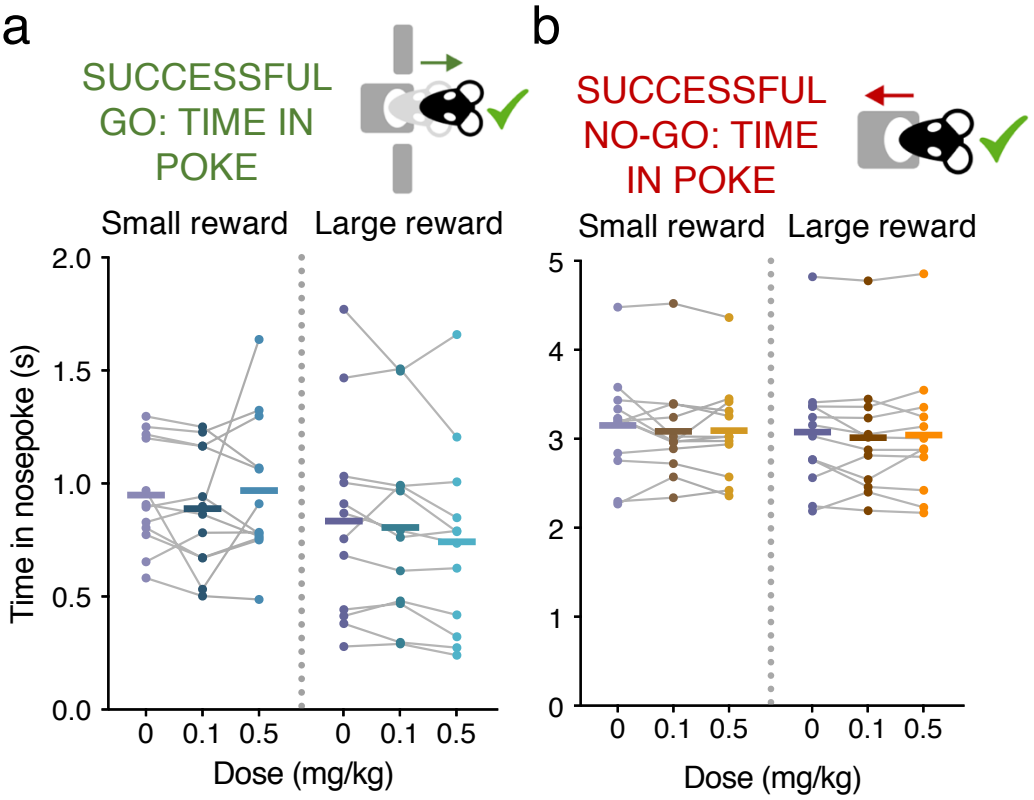


Fig. 4. Mean time spent in the nose poke (s) from cue onset in successful (a) Go or (b) No-Go trials. Solid coloured bars indicate overall mean time in nose poke, and individual dots are mean times for individual animals. The 5-HT_{2C} ligand had no effect on these measures.

3.3. *Systemic manipulation of 5-HT_{2C} receptors energises ongoing goal-directed behaviour*

If disrupting 5-HT_{2C} receptors results in an improvement in goal-directed behaviour, then it can be hypothesised that here we should expect to see animals' progression throughout Go trials to be speeded by the drug. Indeed, the drug did in fact speed mean travel times regardless of the reward size on offer [significant main effect of drug: $F_{2,22} = 7.161$, $\eta_p^2 = .394$, $p = .004$; pairwise comparisons: vehicle vs. low dose: $p =$

.002, vehicle vs. high dose: p = .049; no significant drug*reward interaction: $F_{2,22}$ = 0.327, η_p^2 = .029, p = .724].

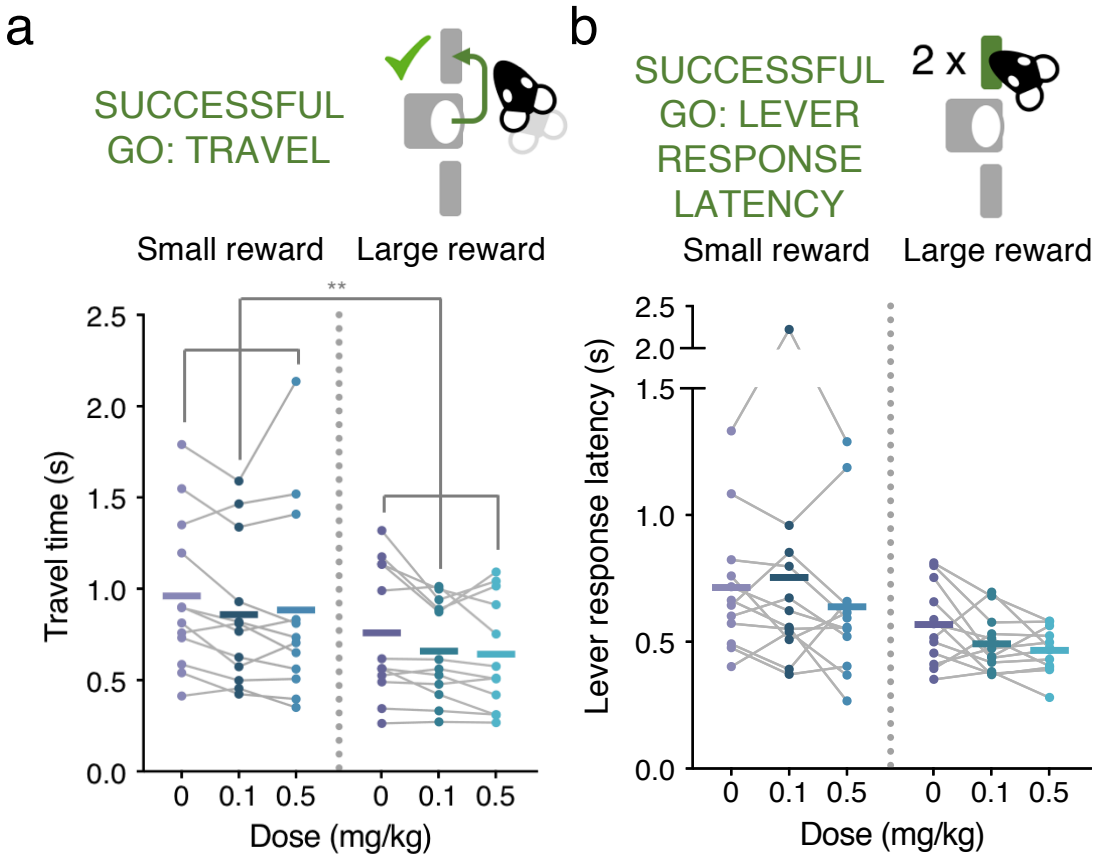


Fig. 5. Mean action execution latencies with systemic administration of the 5-HT$_{2C}$ receptor ligand. (a) Mean travel time from nose poke exit in successful Go trials was significantly reduced by the drug regardless of reward size on offer. (b) Mean lever response latency was also reduced by the drug. Significance bars indicate main effect of drug, ** p < .01.

To further understand whether different travel times were differentially affected by the 5-HT$_{2C}$ ligand, we again compared the 0.05, 0.45, and 0.95 quantiles of the CDFs (Figure 6b and 6d) at different reward sizes and drug doses. We replicated the drug effect [significant main effect of drug: $F_{2,22}$ = 5.367, η_p^2 = .328, p = .013] but also found that the drug exerted greatest effects in the mid or *slow* travel times, leaving fast travel times unaltered [significant drug*period interaction: $F_{2.264,24.899}$ = 4.248, η_p^2 = .279, p = .022; pairwise comparisons: early period, all ps > .2; mid period, vehicle vs.

low dose, p = .002; late period, vehicle vs. high dose, p = .013]. However, this effect was not significantly mediated by reward size on offer [no drug interaction with reward or three-way interaction: all Fs < 2.6, ps > .1].

Lever response latency was also numerically reduced by the drug (Figure 5b), resulting in a trend significant main effect of drug [$F_{2,22}$ = 2.578, η_p^2 = .190, p = .099; when excluding an outlier > 2 S.D.s from the mean $F_{2,20}$ = 4.558, η_p^2 = .313, p = .023; pairwise comparisons: vehicle vs. high dose, p = .007].

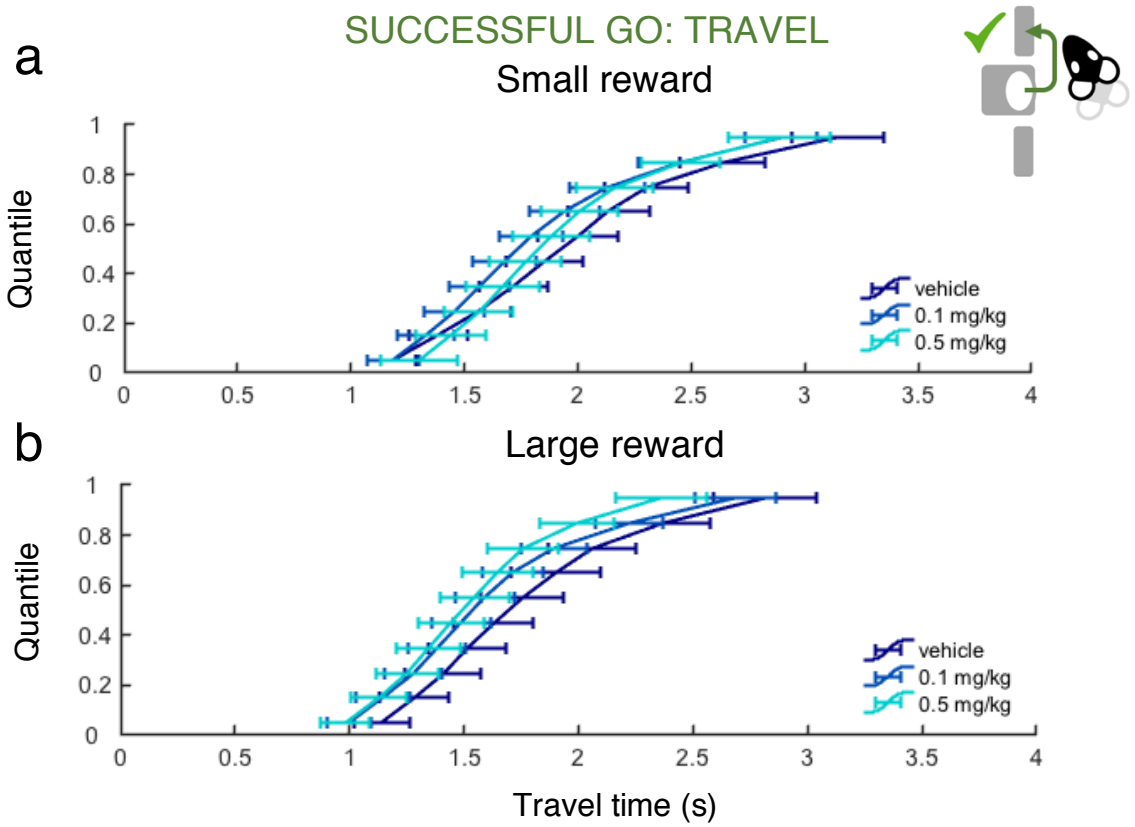


Fig. 6. CDFs for travel time from nose poke exit in successful Go trials with the systemic 5-HT2C ligand. (a) CDF for small reward trials, showing speeding effect of the drug on travel times. (b) Same as in (a) but for large reward trials, where there again is a speeding effect of the drug.

The 5-HT_{2C} ligand again dose-dependently reduced latency taken to retrieve reward [Go trials: significant main effect of drug: $F_{2,18}$ = 3.957, η_p^2 = .305, p = .038, no significant pairwise comparisons; No-Go trials: significant main effect of drug: $F_{1.356,12.202}$ = 17.343, η_p^2 = .658, p = .001; pairwise comparisons: vehicle vs. low reward, p = .032, vehicle vs. high reward: p = .001]. However, the degree of this effect was not mediated by the size of the reward on offer [no drug*reward interactions: all Fs < 2.3, ps > .1].

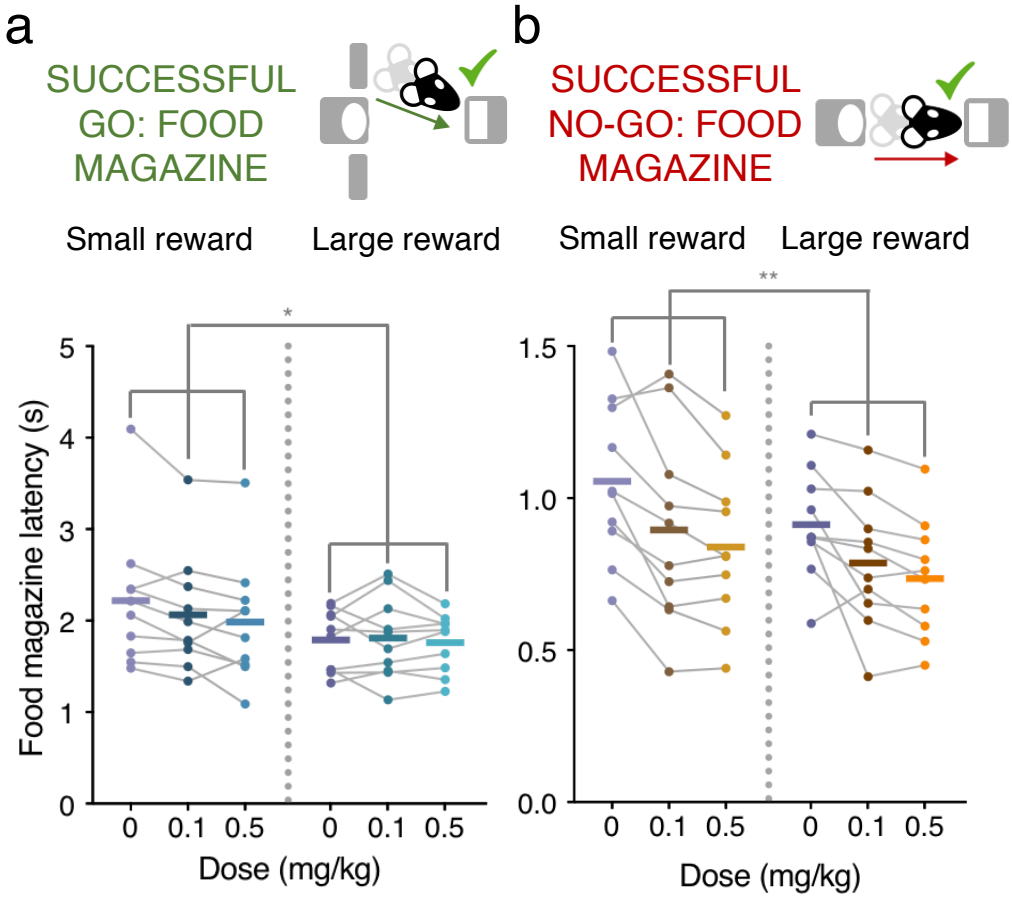


Fig. 7. Mean time taken to reach the food magazine (a) from lever pressing after a successful Go trial or (b) from nose poke exit after a successful No-Go trial. In both cases the 5-HT_{2C} ligand reduced this latency.

Whilst other action execution latencies were speeded by the 5-HT_{2C} ligand, mean time to re-engage was unaltered by the drug [Figure 8; no main effect of drug or interaction: all *F*s < 1.6, *p*s > .2].

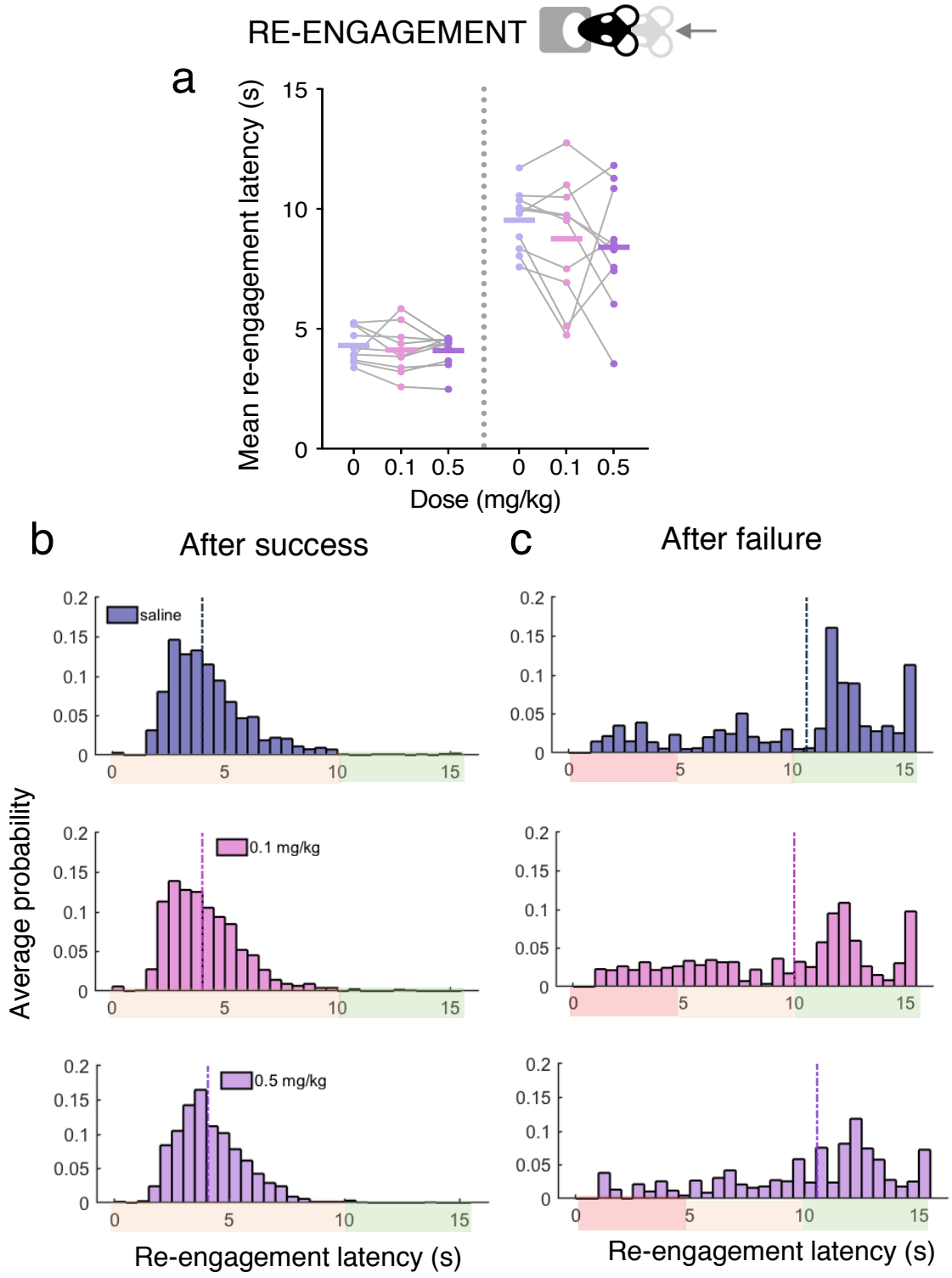


Fig. 8. Effect of the 5-HT_{2C} ligand on re-engagement latency. (a) Overall mean (bars) and mean individual (dots) re-engagement latency after both successful (left of dotted line) and failed (right of dotted line) trials were unaltered with the 5-HT_{2C} ligand. (b) Histogramx of re-engagement latencies after success for each drug dose. Light orange indicates ITI period, light green indicates the nose poke is active and a trial can be initiated (0.5s bins; last bin contains all responses after 15s; dotted lines indicate medians). (c) Same as in (b) but after failure. Light red is the timeout period during which the houselight is on. (0.5s bins; last bin contains all responses after 15s).

3.4. Systemic manipulation of 5-HT$_{2C}$ receptors does not alter internally-driven action initiation

We previously showed that the 5-HT$_{2C}$ ligand increased animals' propensity to make incorrectly initiated action, but that this did not seem to be a specifically cue-driven effect as, in such incorrect trials, animals in fact increasingly *delayed* action initiation. However, analysis of the rate of aborted trials does not reveal a significant change in animals' propensity to initiate action outside of cue presentation [Figure 9; no main effect or interaction with drug: all Fs < 0.3, ps > .7; no significant main effect of trial outcome: $F_{1,11} = 0.422$, $\eta_p^2 = .037$, $p = .529$].

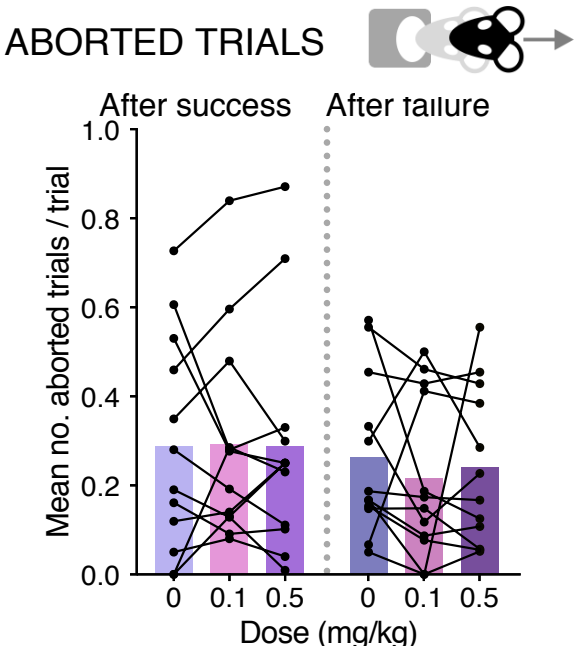


Fig. 9. Effect of D1R stimulation on mean rate of aborted trials (average number of aborted trials per trial) – when animals entered the nose poke to initiate a trial, but did not remain for the precue period and therefore failed to elicit a cue. The drug did not alter rate of aborted trials either after success (left of the dotted line) or after failure (right of dotted line).

Local 5-HT$_{2C}$ receptor manipulation

3.5. Local manipulation of 5-HT$_{2C}$ receptors in the NAc core has no effect on performance in the Go/No-Go task

Systemically administering the 5-HT$_{2C}$ ligand resulted in several profound effects on performance in the Go/No-Go task, including a reduction in No-Go success rate but an *increase* in Go success rate, with animals both able to respond more accurately and at decreased latencies on the drug. To understand whether these effects were caused by manipulation of 5-HT$_{2C}$ receptors in the NAc core specifically, we infused the same pharmacological agent into this region. However here, the drug had little effect on success rate in the Go/No-Go task [Figure 10; all *F*s < 0.8, *p*s > .4].

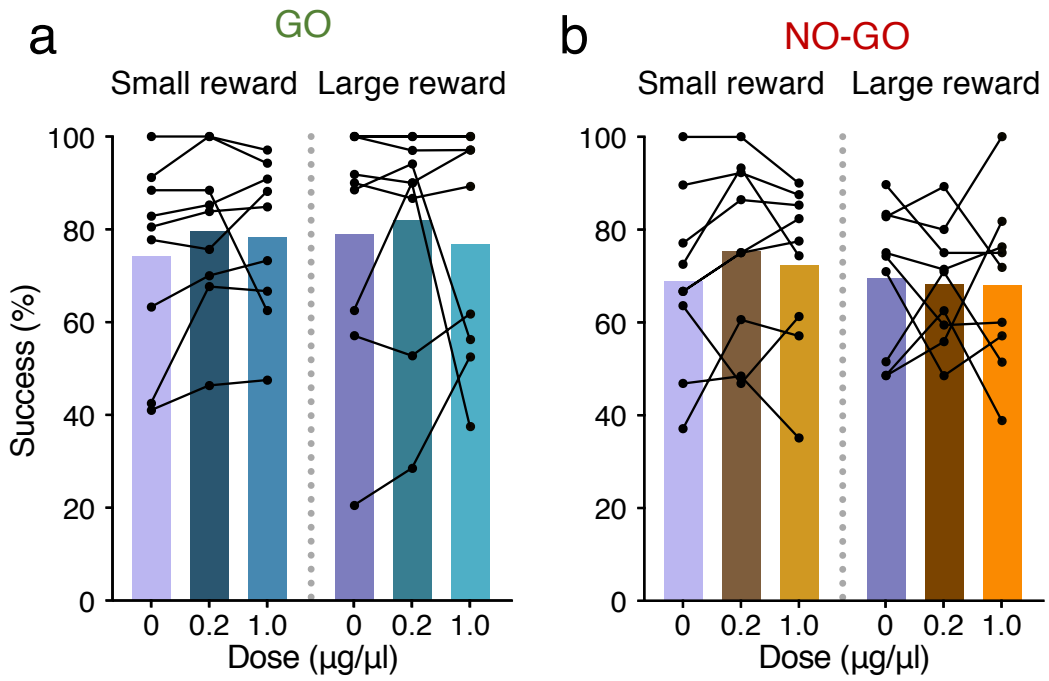


Fig. 10. Mean performance in the Go/No-Go task with local administration of SB 242084. (a) Overall mean performance (bars) and individual mean performance (dots) for low reward (left of dotted line) and high reward Go trials was not altered by the drug. (b) Same as in (a) but for No-Go trials.

This lack of effect is also reflected in Go trial errors. Both Wrong Lever errors and Response Omission errors were unaltered by the drug [no significant main effect of drug or interaction: all *F*s < 0.9, *p*s > .4]. Similarly, the distribution of No-Go trial errors (Figure 11) was also unaltered by the drug [all *F*s < 0.7, *p*s > .5].

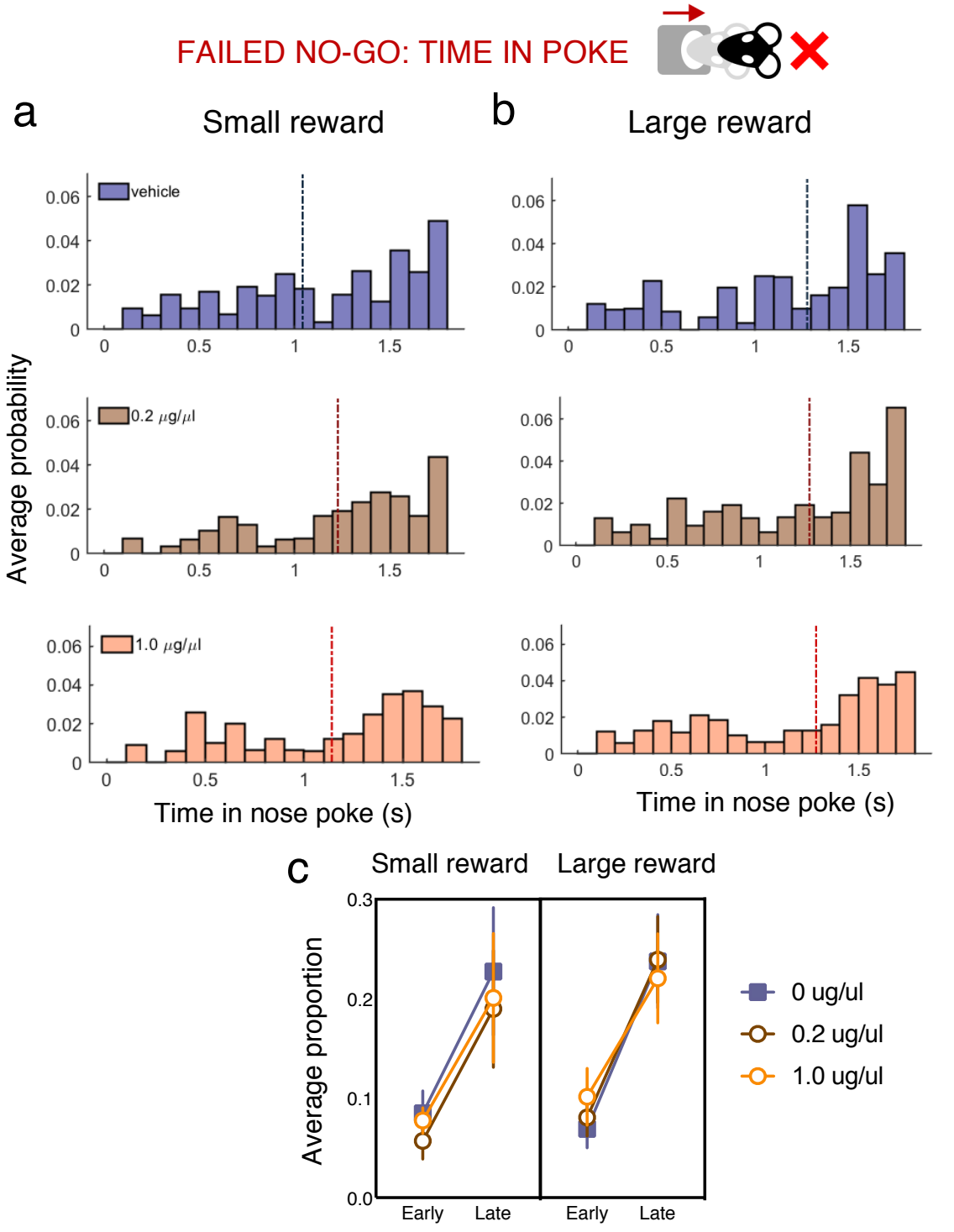


Fig. 11. Average probability histograms (0.1s bins) for time to leave nose poke in (a) small reward and (b) large reward trials with local 5-HT$_{2C}$ administration, calculated as probability over all head exit times, both successful and failed, for each drug dose. Last bin contains all values > 1.7s. (c) Mean proportion of times spent in the nose poke across trials that were early (< 800ms) or late (> 800ms) dependent on drug dose applied (square: vehicle, brown circle: 0.2 ug/ul, red circle: 1.0 ug/ul).

3.6. Local manipulation of 5-HT$_{2C}$ receptors in the NAc core has no effect on action initiation latencies in the Go/No-Go task

We next investigated whether action initiation latencies were altered by local infusion of the 5-HT$_{2C}$ ligand. Systemic administration of the drug did not affect mean speed of action initiation in successful Go or No-Go trials, and this was also the case here [Figure 12a; no significant main effect of drug or reward, or interaction: all *F*s < 3.0, $p > .07$].

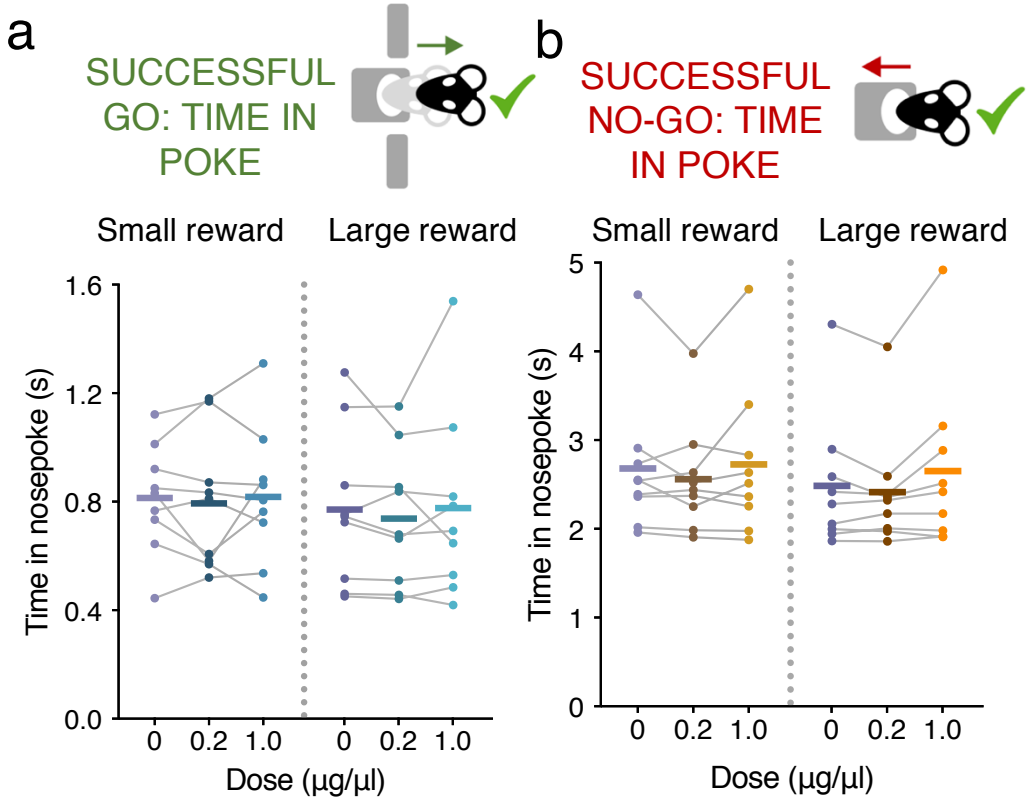


Fig. 12. Mean time spent in the nose poke (s) from cue onset in successful (a) Go or (b) No-Go trials with local administration of the 5-HT$_{2C}$ ligand. Solid coloured bars indicate overall mean time in nose poke, and individual dots are mean times for individual

3.7. Local manipulation of 5-HT$_{2C}$ receptors in the nucleus accumbens has no effect on action execution latencies in the Go/No-Go task

The drug again had no reliable effect on travel time latencies, lever response latencies or food magazine latencies [Figure 13; no main effects of drug or reward, or interaction: all *F*s < 1.5, *p*s > .2] .

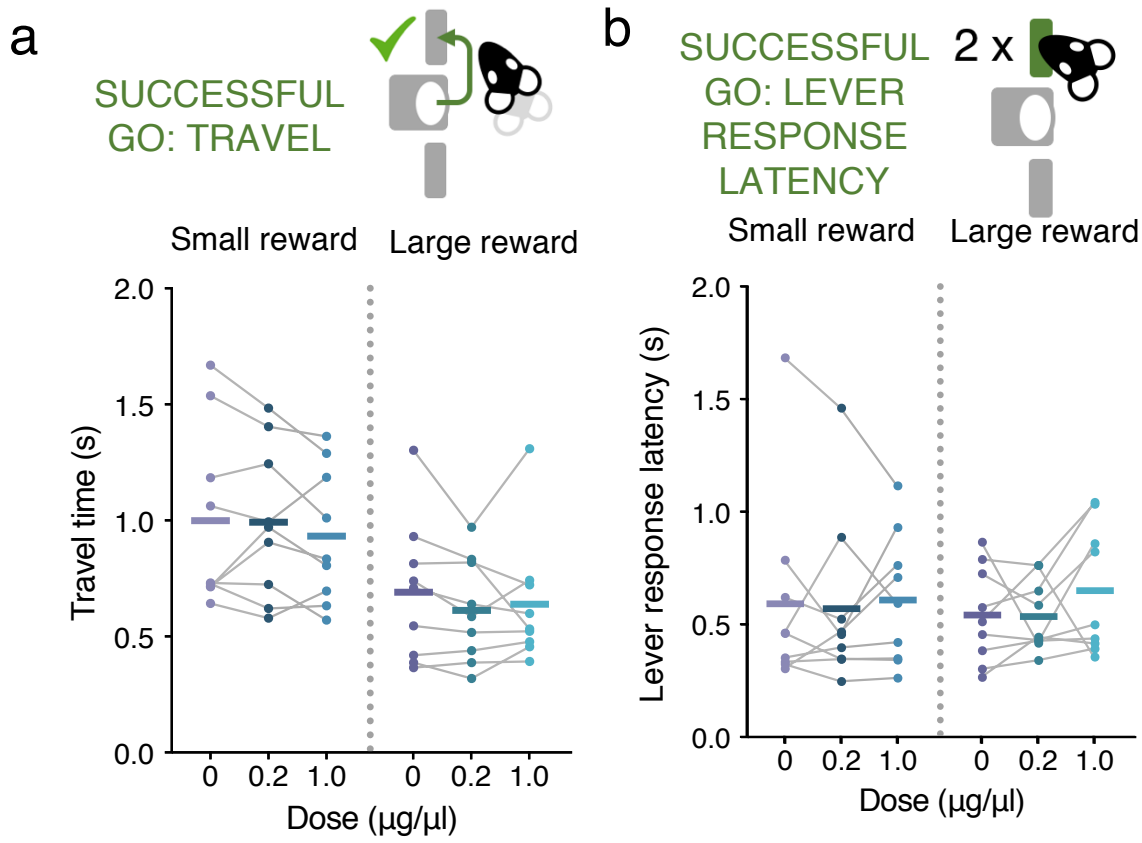


Fig. 13. Mean action execution latencies with local infusion of the 5-HT$_{2C}$ ligand. (a) Mean travel time from nose poke exit in successful Go trials. (b) Mean lever response latency from first to second lever press.

Animals were faster to re-engage in the task after success than after failure [significant main effect of trial outcome: $F_{1,8} = 9.781$, $\eta_p^2 = .550$, $p = .014$]. However, as with all other latencies, the drug exerted no effect [no significant main effect of drug

or interaction: all Fs < 1.2, ps > .3]. Similarly, rate of aborted trials was also unaffected by the drug [no main effect of drug or trial outcome, or interaction: all Fs < 0.5, ps > .5].

Local amphetamine infusion

3.8. Lack of effect of NAc core 5-HT$_{2C}$ manipulation is not due to failed infusions

At the time of local infusion of the 5-HT$_{2C}$ ligand, animals had already experienced between 6 and 8 infusions as part of the experiments reported in Chapter 5. Therefore, to confirm ongoing patency of the cannulae, *d*-amphetamine was subsequently infused as a positive control.

d-amphetamine significantly reduced mean No-Go success rate [Figure 14a; significant main effect of drug: $F_{1,12} = 14.047$, $\eta_p^2 = .539$, $p = .003$] and mean head exit latencies in Go trials [Figure 14b; significant main effect of drug: $F_{1,12} = 12.739$, $\eta_p^2 = .515$, $p = .004$] in most animals. In addition, two animals who showed no effect of the drug on these latencies were excluded from the previous analyses. Therefore, this demonstrates that the lack of effect of local infusion of the 5-HT$_{2C}$ ligand was not due a lack of patency of the cannulae.

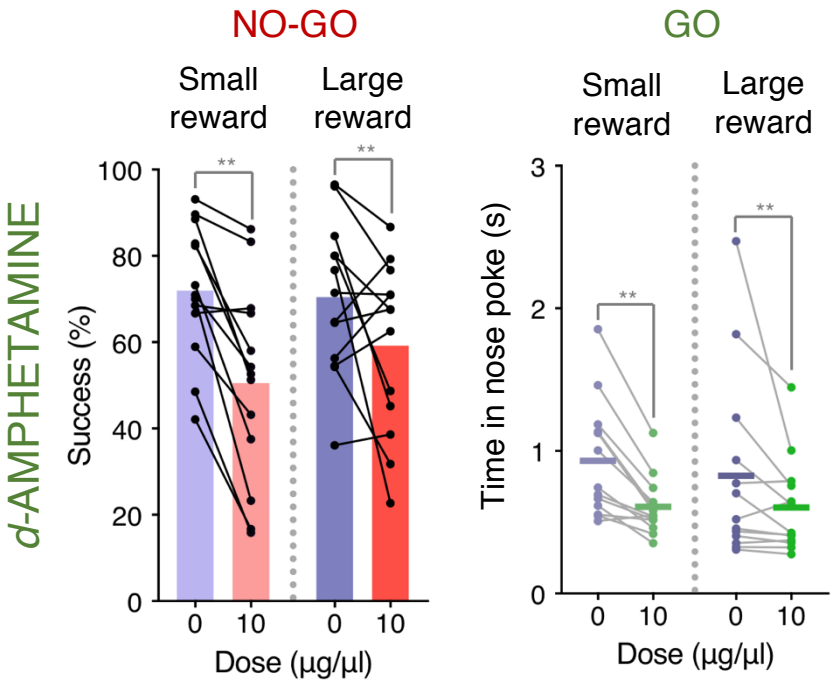


Fig. 14. Effects of local *d*-amphetamine infusion on (a) average and individual No-Go performance and (b) average and individual successful Go trial reaction times.

Systemic 5-HT$_{2C}$ receptor manipulation - replication

3.9. Systemic effects on Go and No-Go trial performance are not due to differences in task parameters or training

The systemic data described in this chapter were collected in a version of the task where animals had to withhold action for 1.7 – 1.9s, rather than the standard 1.5 – 1.7s used throughout the rest of this thesis and in the local infusion experiment. To ensure that the effects found were not attributable to any such differences in task parameters, or due to training experience, we re-ran the systemic administration of 5-HT2c ligand in the cohort of animals that received the local infusions using the standard 1.5 – 1.7s holding time in No-Go trials. We performed this replication twice: for the first time, before animals had undergone surgery for the local infusion

experiment, and for the second time post-infusions. In both cases, we applied only the 0.1 mg/kg dose, along with additional separate vehicle sessions.

In Go trials (Figure 15a and 15c), we replicated the effect of an improvement in success rate, particularly in small reward trials [significant reward*drug interaction: $F_{1,13}$ = 13.567, η_p^2 = .511, p = .003; small reward vehicle vs. drug comparison, p = .001; large reward comparison, p > .8]. However, there were no effects of training experience [no significant main effect of training or interactions with reward or drug: all Fs < 1.5, ps > .2]. In No-Go trials (Figure 15b and 15d), we also replicated the reduction in success rate in small reward trials [significant reward*drug interaction: $F_{1,13}$ = 7.145, η_p^2 = .355, p = .019; small reward vehicle vs. drug comparison: p = .088 Sidak-corrected; large reward comparison: p > .5], and again training experience had no influence [no main effect or interactions with training: all Fs < 0.4, ps > .5].

Therefore, the lack of effects seen after the NAc core infusions of the 5-HT2c ligand were not due changes in specific task parameters or training experience.

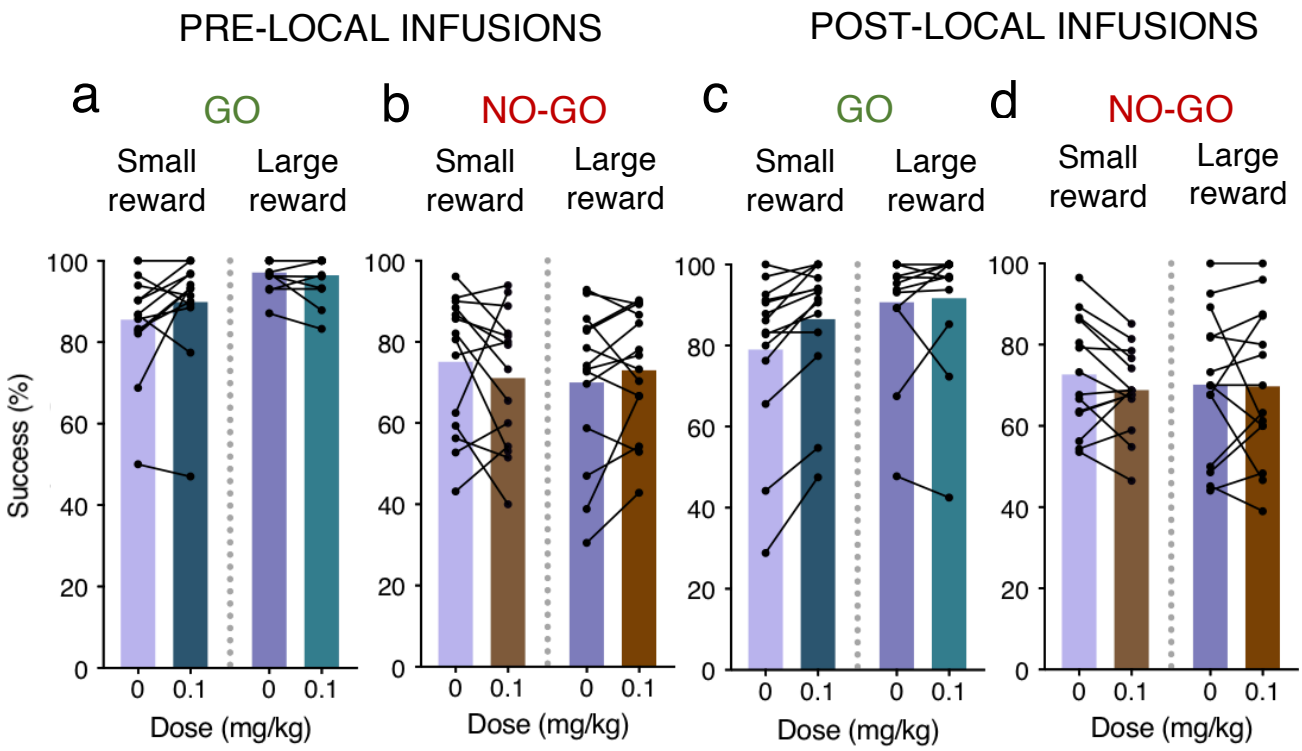


Fig. 15. Effect of replicated systemic 5-HT_{2C} manipulation on success rate in the Go/No-Go task, separated by small reward (left of dotted line) and large reward (right of dotted line) trials, and by replication pre- and post-local infusion. Bars indicate overall mean, and dots are averages for individual animals. (a) The 5-HT_{2C} ligand increased success rate in small reward Go trials. (b) The drug decreased success rate in small reward No-Go trials. (c) Same as in (a) but after animals had undergone the local infusion studies. (d) Same as in (b) but after animals had undergone the local infusion studies.

4. Discussion

As the supposed 'opponent' to dopamine, serotonin has long been associated with concepts of aversion and behavioural inhibition (Soubrie, 1986), with an increase in serotonin levels enabling restraint for reward and a reduction resulting in the converse (Fonseca et al., 2015; Harrison et al., 1997). However, the precise role of serotonin in withholding actions for future reward, the role of the 5-HT2c receptor transmission, and the locus of this effect is not well understood. This is particularly as transmission at this receptor has also been shown to play a role in modulating goal-directed actions (Bailey et al., 2016).

Here we applied a 5-HT_{2C} ligand, SB 242084, both systemically and locally into the NAc core, as animals performed the Go/No-Go task. We found that success rate in No-Go trials was reduced by systemic administration of the drug. Unlike dopamine manipulations, this was caused by a potentiation of premature No-Go responses that occurred later in the No-Go holding period and only on small reward trials. Conversely, the drug improved Go trial performance through a reduction in Response Omission errors, and also speeded progress throughout Go trials. We also discovered that these behavioural effects are not localisable to 5-HT_{2C} receptors in the NAc core. A summary of these findings is contained in Table 2, and they are discussed in more detail subsequently.

	Systemic 5-HT_{2C} ligand		Local 5-HT_{2C} ligand	
MEASURES OF PERFORMANCE				
Go errors	Wrong lever	Response omission	Wrong lever	Response omission
	—	↓	—	—
No-Go errors	↑		—	
Aborted trials	—		—	
MEASURES OF LATENCY				
Time in poke – successful Go	—		—	
Travel time	↓		—	
Lever response latency	↓		—	
Food magazine latency	↓		—	
Re-engagement latency	—		—	

Table 2. Summary of effects of systemic or local application of the 5-HT_{2C} ligand on measures of performance and measures of latency in the Go/No-Go task. Large arrows indicate a strong effect, small arrows indicate smaller or more specific effects. For measures of performance, green arrows indicate an improvement and red arrows indicate a deterioration.

4.1. Pharmacological effects of the 5-HT$_{2C}$ receptor ligand, SB 242084

In order to better understand the behavioural effects presented in this chapter, it is worth briefly expanding upon the pharmacological mechanisms of the compound used here. SB 242084 was selected as the 5-HT$_{2C}$ receptor ligand for use in these studies due to its very high affinity (pKi 9.0) and for its high selectivity over the other 5-HT$_2$ receptor subtypes (Kennett et al., 1997). In addition, it is the pharmacological agent of choice in many of the studies of impulse control and goal-directed behaviour that are pertinent to the questions being asked here.

However, despite being often referred to as an antagonist, it has been shown that that this ligand in fact acts as a partial agonist on intracellular signalling pathways such as phospholipase C (PLC) whilst acting as an inverse agonist on phospholipase A2 and the inhibitory G protein, Gα_i (De Deurwardere, Navailles, Berg, Clarke, & Spampinato, 2004). This complicates interpretations of its effects such that it cannot be thought to completely block serotonin transmission at 5-HT$_{2C}$ receptors, and may also account for the quadratic effect on Go trial success rate seen here due to the complex influence of different doses. On the other hand, its effects do often appear to act as the opposite of 5-HT$_C$ agonists, such that it can to some extent be thought to act as a 'functional' antagonist (Hayes, Clements, & Greenshaw, 2009), and this partial functional selectivity makes it a possible candidate for therapeutic use due to a lower likelihood of side effects that might result from complete receptor blockade (Bailey et al., 2016; Barker, Westphal, Schmidt, & Sanders-Bush, 1994; Conn, Lindsley, Meiler, & Niswender, 2014).

4.2. Ability to inhibit action for reward is mediated by the 5-HT$_{2C}$ receptor independent of cue responsivity

There is much evidence from both pharmacological (Harrison et al., 1997) and physiological (Fonseca et al., 2015; Miyazaki, Miyazaki, & Doya, 2012) approaches that central serotonin is a key modulator of the ability to wait for reward, or 'patience'. The serotonergic 5-HT$_{2C}$ receptor has been thought to be of particular importance to this behaviour, with its effective antagonism using the same drug and doses applied here having previously been shown to increase uncued premature responding in the 5-choice task (Winstanley, Theobald, et al., 2004). Here, we also found an increase in impulsivity, with animals less able to withhold action for reward in small, but not large, reward No-Go trials.

Whilst this reduction in No-Go trial success rate is in the same direction as the effect we found with stimulation of D1Rs in Chapters 4 and 5, a closer investigation of the *type* of errors animals made in fact suggests that a different mechanism may be at play. The D1R agonist rapidly drove early head exits, causing animals to leave prematurely early in the No-Go period. However, if they managed to overcome this, they were no more likely to make an error during the late period than controls. In contrast, the 5-HT$_{2C}$ ligand applied systemically increased the proportion of errors that occurred in the *late* period. Therefore, whilst the dopaminergic manipulation predominantly mediated action or inaction in response to the presentation of value-laden cues, the serotonergic manipulation instead affected instrumental drive to act (as evidenced by a decrease in No-Go success rate) but *without* a concomitant change

in cue-elicited drive. Further evidence for this contrasting effect can be found with analysis of time to respond in successful Go trials: whilst local infusion of the D1R agonist reduced this latency, again suggesting a cue-driven effect, the 5-HT_{2C} ligand here had no impact. This is also consistent with studies of cocaine-seeking behaviour, which have shown that application of the same ligand does not alter cued reinstatement of such seeking behaviours (Burmeister, Lungren, Kirschner, & Neisewander, 2004).

One caveat is that rate of aborted trials was surprisingly not altered, despite the fact that previous studies have demonstrated an increase in ability to wait for reward – with a behaviour almost identical to aborted trials here (remaining in a poke for an uncued period of variable length until a cue is delivered) – due to optogenetic stimulation of 5-HT neurons in the dorsal raphe nucleus (Fonseca et al., 2015). However, there the alteration in behaviour was found when animals had to withhold action in the order of seconds, whereas the pre-cue period here was never any longer than 0.7s. Therefore, it may be the case that we see little effect on aborted trials here because the ability to wait for reward is only mediated at longer latencies.

4.3. Manipulation of 5-HT_{2C} receptors increases execution and speed of goal-directed action as a function of opportunity cost

In Go trials, systemic administration of the 5-HT_{2C} ligand resulted in an increase in success rate due to fewer Response Omission errors. In addition, following action initiation, animals' progression throughout such trials was benefitted by the drug

such that they both responded more quickly – but without any detriment to performance, suggesting an increase in incentive motivation.

This effect was also mediated by the reward size on offer, such that small reward Go trials were most improved by the drug. Whilst this might partly be a ceiling effect, it is consistent with recent data from the Simpson lab who have used the same ligand to produce an enhancement in duration and vigour of effortful goal-directed behaviour (Bailey et al., 2016). Whilst several other studies have shown little effect on incentive motivation for food with administration of the 5-HT_{2C} ligand (Bezzina et al., 2015; Fletcher, Sinyard, & Higgins, 2010), Bailey et al. suggest that this is because the effects are specific to contexts of high effort requirements. Yet here, we find a substantial drug effect despite consistent effort requirements dependent on *reward size*, suggesting that mediation of goal-directed behaviour by 5-HT_{2C} receptors is, to some extent, dependent on a calculation of opportunity cost. This is counter to conclusions drawn by Bailey et al. (2018), who claimed no change in reactivity to rewards by the drug, but who did not manipulate the size of the reward on offer. This is also supported by the fact that the No-Go trial success rate effects were also reward-mediated, without suffering from possible ceiling limitations. In contrast, Go trial response *latencies* were similarly speeded regardless of whether the reward on offer was small or large, and this indeed may be due to the low effort requirements of the task.

4.4. Potential mechanisms of goal-directed enhancement of behaviour and reduction of inhibition by 5-HT_{2C} receptors

A potential mechanism to explain both the enhancement of goal-directed behaviour and reduction in ability to inhibit action by manipulation of the 5-HT_{2C} receptor as seen in the presented studies is via modulation of mesolimbic dopamine transmission. Indeed, it has been shown that systemic administration of this ligand increases midbrain dopamine neuron firing rate and dopamine release in the NAc core (De Deurwardere, Navailles, Berg, Clarke, & Spampinato, 2004b; Di Giovanni et al., 2000; Matteo et al., 1999a) which have in turn been associated with increased motivation in effort-based tasks and premature responding in impulse control paradigms. If that is indeed the case here, then this would be to some extent consistent with the results presented in Chapters 3 and 5, where phasic dopamine release coincided with correct action initiation, and activation of D1Rs drove action but worsened inhibition of action. However, this may not be the sole mechanism for the effects of SB 242084; as described above, the effects of dopamine agents on both Go and No-Go performance were distinct from those observed here. One possibility is that the drug effects may still be mediated by enhancement of dopamine transmission, but occurring in regions outside of the NAc. Bailey et al. (2018) recently showed that this drug increases extracellular dopamine levels in the dorsomedial striatum during goal-directed behaviour, but leaves both tonic and phasic dopamine levels in the NAc unaltered. However, an important caveat is that they did not record phasic dopamine levels as animals *worked* for reward, instead only measuring responses to reward-predictive cues. Therefore, it is unclear exactly what changes in dopamine might be observed in more dynamic situations.

4.5. The regulation of performance by 5-HT$_{2C}$ receptors in the Go/No-Go task is not mediated by the NAc core

Whilst ventral striatal dopamine levels may have contributed to these effects, the local infusion study presented here revealed that interruption of 5-HT$_{2C}$ receptor signalling within the NAc core itself does not. This is in contrast to previous findings that these receptors in the accumbens increases impulsivity in the 5-choice task (Robinson et al., 2008b), despite application of the same drug doses. Subsequent local infusion of *d*-amphetamine served to confirm that this lack of an effect was not due to a failure in administering the drug, as in all animals bar two, both No-Go success rate and Go head exit latencies were reduced by this manipulation, indicating that the cannulae were patent for the infusions of the 5-HT$_{2C}$ ligand. Additionally, cannulae location in the NAc core was similar to that used in Robinson et al. (2008). Therefore the lack of an effect on impulsive behaviour here is likely due to task differences, in particular suggesting that 5-HT$_{2C}$ receptors in the NAc core do not mediate inhibition of behaviour when *all* movement must be restrained. Additionally, the lack of effect here on goal-directed behaviour in Go trials has also been shown in a reversal learning study where animals had to press levers 3 times for reward (Boulougouris & Robbins, 2010).

However, whilst our findings suggest that NAc core 5-HT$_{2C}$ receptors do not mediate the systemic effects seen here, that is not to say that they definitively play no role in the behaviours measured. Previous studies have shown asymmetrical effects of activation or inhibition of ventral tegmental area neurons expressing the 5-HT$_{2C}$

receptor; whilst their inhibition did not alter behaviour, their activation was shown to reduce incentive motivation in terms of *ad libitum* food consumption and operant responding for reward (Valencia-Torres et al., 2017). As we did not stimulate 5-HT_{2C} receptors in the NAc core, we cannot conclude that they contribute nothing to performance on the Go/No-Go task, but rather that their activation is not necessary for successful performance.

4.6. Summary

By modulating the activity of the 5-HT_{2C} receptor systemically, we found that animals were more likely to execute inappropriate action but that, unlike effects found with D1R manipulation, this was not cue-potentiated. This coincides with much of the evidence tying serotonergic function to a notion of 'patience', particularly in uncued contexts (Fonseca et al., 2015b). In addition, manipulation of these receptors also increased degree of incentive motivation, as demonstrated by an increase in Go trial success rate and decrease in speed through which animals progressed through such trials. However, these effects were not replicated with local administration of the ligand into the NAc core. Therefore, normal function of 5-HT_{2C} receptors – outside of the accumbens – serves to inhibit goal-directed behaviour.

8 General discussion

Chapter abstract:
In this thesis, we have studied the contribution of the dopaminergic receptors and the 5-HT_{2C} serotonergic receptor to goal-directed action initiation, selection, and inhibition using a novel behavioural paradigm. Here, we summarise our findings and discuss interpretations of the behaviour as measured in this task, as well as possible circuit effects of these manipulations, limitations of these studies, and suggestions for future experiments to expand upon these results.

1. *Understanding dopamine and serotonin in reward and action*

The brain allows behaviour to be enacted on the basis of sensory input for the benefit of the organism, but exactly how motivation – both in terms of activation and direction of behaviour (Salamone & Correa, 2012) – arises is not yet fully understood. Additionally, there are questions surrounding how the brain multiplexes this drive for behaviour with information about *which* behaviours are most beneficial to perform. Two central monoamine systems, dopamine and serotonin, are thought to contribute to aspects of these different processes (Dayan & Huys, 2008; Robbins & Everitt, 1996; Salamone & Correa, 2012).

Dopamine historically has been linked to concepts of reward and reinforcement (Olds & Milner, 1954; Wise, 1978), but subsecond changes in the activity of midbrain dopamine neurons and subsequent mesolimbic dopamine release has more recently been hypothesised to signal the discrepancy between expected and received reward – a reward prediction error – with an impact on behaviour occurring due to processes outside of this signalling (Hart et al., 2014; Schultz et al., 1997). In contrast, pharmacological studies have suggested a more *active* role of the dopaminergic signal in motivating reward-seeking behaviours (Robbins & Everitt, 2007)

The aim of this thesis was to make headway in resolving this incongruity between these two literatures by using a task that incorporates *both* cue-reward and cue-action contingencies, in conjunction with methods for the recording of dopamine

release and manipulation of degree of activation of dopamine and serotonergic receptors. This unique combination of approaches allowed for a systematic understanding of the role of these neurotransmitters at different timescales to the same goal-directed behavioural activation and inhibition. A summary of the results garnered from the various pharmacological manipulations is presented in Table 1.

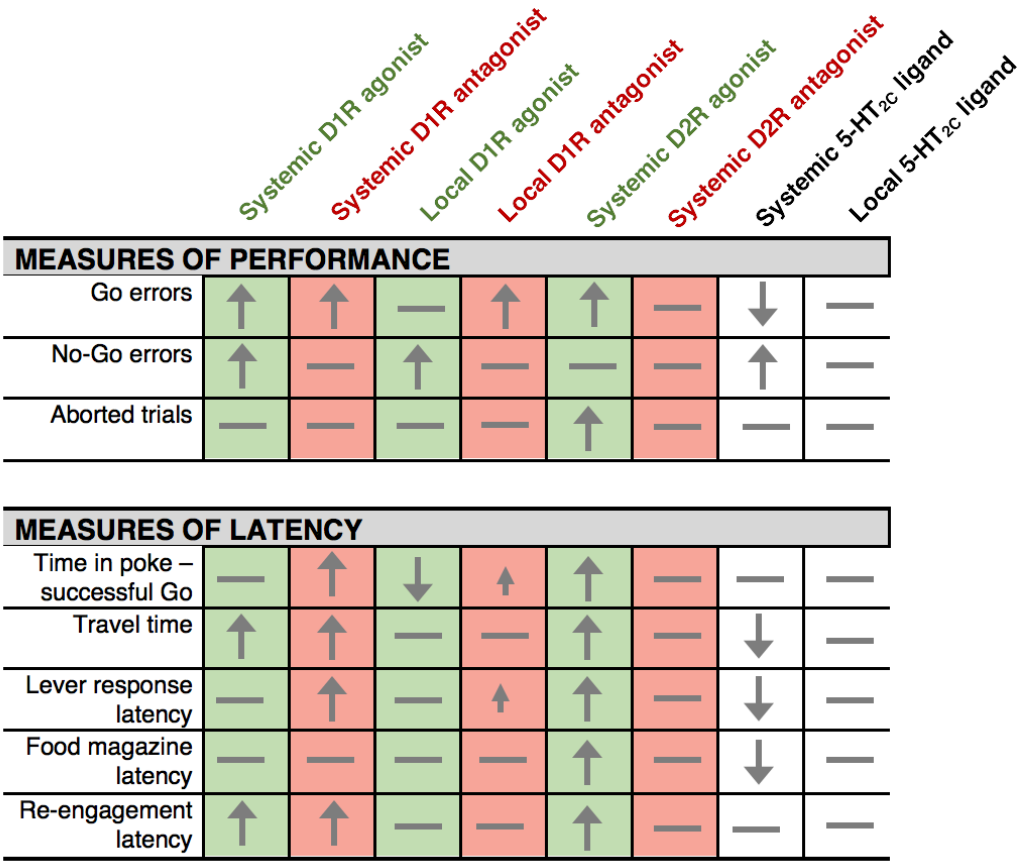


Table 1. A summary of all effects seen in the pharmacology experiments presented in this thesis, including both systemic and local manipulations of the dopaminergic and serotonergic systems.

In Chapter 3, by recording subsecond dopamine release in the NAc as animals performed the Go/No-Go task, we found that phasic efflux of dopamine levels reflected size of the reward on offer as indicated by well-learnt instructional cues,

consistent with the prediction error formulation of dopaminergic function. However, we also found that the dopamine signal was shaped by the action required to garner reward, and specifically was attenuated when inhibition of movement needed to be withheld to gain reward.

To understand whether this dopamine efflux was reporting or causally driving behavioural output, in Chapter 4 we systemically stimulated or blocked dopamine receptors that are thought to be most sensitive to phasic changes in dopamine release, D1-like receptors (D1Rs). We found that activation of these receptors promoted inappropriate action whilst also biasing action selection. In contrast, their blockade slowed goal-directed behaviour, but interestingly did not improve ability to *withhold* action, consistent with the voltammetric results.

In order to localise these effects, in Chapter 5 we infused the same pharmacological compounds into the NAc core – the region recorded from in Chapter 3. We found that artificial D1Rs stimulation in this region drives the initiation, but not ongoing performance, of cue-driven action, whilst blockade of these same receptors reduces cue responding, and neither manipulation influenced action selection. But, again, importantly we found no improvement in ability to withhold action for reward with NAc core D1R blockade, reinforcing our previous findings that phasic dopamine release and its subsequent impact on D1Rs is entirely absent when movement is restrained, even in the presence of a value-laden cue.

In the classical model of basal ganglia function, D2-like receptors (D2Rs) are thought to inhibit competing actions, but recent evidence suggests that their activation may instead work synergistically with D1Rs to drive behaviour. Additionally, they are also considered to be more sensitive to tonic changes in dopamine levels that have been hypothesised to encode reward rate, as reflected by alterations in action vigour. Here, we found confirmation that D2R stimulation slows movement towards a goal, but intriguingly has little effect on the ability to withhold action for reward, again suggesting that the influence of dopaminergic receptor activation on behaviour is dependent on movement.

Finally, as serotonin is thought to work in opponency to dopamine by inhibiting behaviour, we investigated whether this would also be the case when inhibition of all movement is required for reward. We found that by partially blocking the 5-HT_{2C} receptor, animals were both more impulsive but, in contrast to the dopaminergic manipulations, this was not a cue-driven effect. Additionally, the ligand also increased instrumental incentive motivation as indicated by an increase in Go trial performance.

2. *A deeper consideration of psychological processes in the Go/No-Go task*

In this thesis, a novel paradigm was developed to investigate the respective contributions of movement and reward prediction to the dopamine signal. Whilst the Go/No-Go task incorporates aspects of other tasks in the literature, it is also

unique in its approach in controlling behaviour, and so here we discuss the possible psychological processes involved to gain a deeper understanding of the results of these experiments.

2.1. Understanding behavioural inhibition in No-Go trials

Impulse control paradigms come in a variety of flavours that capture different aspects of inhibitory control including action postponement, action restraint, and action cancellation (Bari & Robbins, 2013). The No-Go trials used here combine elements of both action postponement – waiting until it is appropriate to execute action – and action restraint, most akin to the 5-choice task (Robbins, 2002) and to simpler Go/No-Go tasks (Fletcher, 1993). In those paradigms, animals must withhold responding until a cue indicates it is appropriate to do so, and any operant responses during the delay period are considered erroneous and thus impulsive.

Whilst this is similar to the No-Go adopted here, there are several key differences. First, animals had to withhold action for a short pre-cue period before the trial commenced, before *continuing* to withhold action upon presentation of the No-Go cue. This is in contrast to paradigms such as the 5-choice, where animals must indeed self-initiate a trial, for example by pushing the food magazine panel, but where successful performance is not dependent on continuation of this specific behaviour (Robbins, 2002). Therefore there is the possibility that No-Go trials differ from Go trials not just in their action requirement, but also in the behavioural 'state' of the animal; a No-Go cue instructs animals to remain in the same state, whilst a Go cue instructs animals to switch. Explicit physiological state has been shown to shape

dopaminergic reward prediction errors (Papageorgiou, Baudonnat, Cucca, & Walton, 2016), although dopaminergic control as a context-dependent behavioural switch from different motor patterns has also been posited (Vidal-Gadea & Pierce-Shimomura, 2012).

Second, withholding of responding in other impulse control paradigms refers to refraining from making a *specific* goal-directed action, usually poking into a port, whilst other actions remain unpunished. Therefore animals can still be moving whilst inhibiting the specific operant behaviour that is not permitted, and whilst that movement might not be goal-directed in the sense that responding in a port is, it may still serve a purpose of bringing animals psychologically closer to their goals (Salamone & Correa, 2012). Here, animals had to refrain from *all* movement to correctly perform in a No-Go trial, allowing for a precise disambiguation between goal-directed action and inaction. In addition, here animals must resist two potential drives to move: potential future reward in a different location (in the food magazine) and the presence of levers known to result in reward. In contrast, in other impulse control paradigms animals must resist poking or pressing a lever, but are not punished for wandering over to the food tray, an alternative goal-directed action. These key differences may explain why D1R antagonism, both systemically and locally, did not improve success rate in No-Go trials, as the ability to perform such a trial effectively is independent from the dopaminergic signal (as also demonstrated in Chapter 3). More broadly, this is the first conclusive evidence that an expression of impulsivity as driven by aberrant levels of dopamine receptor activity is dependent on movement.

A further difference, as discussed in more detail subsequently, is that here, animals withheld action during continuous presentation of an informative cue. Whilst this is more akin to traditional Go/No-Go tasks, where a cue continuously indicates that animals should refrain from an instrumental response (Harrison et al., 1999), it is quite different from tasks such as the 5-choice, where animals must respond only upon cue detection.

Finally, by virtue of the fact that animals had to withhold all movement, the duration of required inhibition in No-Go trials was, at the most, 2.4s (or 2.6s in the first systemic study of Chapter 7, in both cases including the pre-cue period and not accounting for the 0.1s buffer period). This is considerably less than in comparable tasks such as the 5-choice, where animals must usually wait between 5 to 9 seconds, or other Go/No-Go tasks or differential reinforcement of low rates (DRL) schedules, where the inhibitory period can be in the order of tens of seconds (Dalley et al., 2011). Whilst this meant that we were unable to investigate the effects of pharmacological manipulations on sustained behavioural inhibition beyond these several seconds, we were still able to differentiate between early and late responses with the knowledge of the animal's exact behaviour up until that point. Indeed, this allowed us to differentiate between D1R stimulation and 5-HT_{2C} receptor manipulation; whilst both decreased success rates in No-Go trials, the former resulted in an increase in inappropriate early responding and the later increase late responding.

2.2. *Cue-driven vs. internally-driven behaviour*

The Go/No-Go task allowed us to investigate the differential impact of dopaminergic and serotonergic receptor manipulation on cue-driven and internally-driven behaviour, both in the context of action initiation or inhibition. This is pertinent considering the importance of dopamine in learning cue values (Steinberg et al., 2013) but also in cue responding (Taha, Nicola, & Fields, 2007) and thereby in incentive motivation (Robbins & Everitt, 2007), whilst also considering less evidence that dopamine – particularly as part of the mesolimbic pathway – has an influence on general motor performance (Koch, Schmid, & Schnitzler, 2000). Further evidence of the importance of external cue presentation to the mediation of behaviour by the dopamine receptor subtypes has been shown in Pavlovian-to-instrumental transfer (PIT) tasks; blockade of either D1Rs or D2Rs reduces operant expression of this learning (Lex & Hauber, 2008). Similarly, the phasic dopaminergic response has been shown to be tightly coupled to both the reward-paired cue and reward-seeking lever pressing in PIT, again demonstrating the importance dopamine to *cue-driven* action (Wassum et al., 2012).

Here, we found that stimulation of accumbens D1Rs drove impulsive action in response to a cue without altering ongoing performance of internally-driven behaviour in Go trials or initiation latencies after success in No-Go trials, suggesting that the sustained dopamine release seen as animals travelled, lever pressed, and retrieved reward in Chapter 3 may be not be galvanising behaviour in such trials, but instead may be indicating proximity to reward (Howe et al., 2013). This was in contrast to the systemic manipulations of D1Rs, which *did* selectively affect travel

latencies. These effects were likely extra-accumbens; the dorsal striatum in particular has been linked to modulation of the vigour of movement (Alves et al., 2018).

2.3. Manipulation of opportunity cost by reward – and effort?

Here, again in contrast to typical impulse control paradigms, animals had to make or withhold action for either a small or large reward. As the reward size on offer was conveyed at trial initiation, this allowed us to understand whether any detrimental or enhancing effects of the pharmacological agents were mediated as a function of opportunity cost. Indeed, we found that the degree of several key effects of the D1R manipulations on success rate, both in terms of receptor stimulation and blockade, varied as a function of reward size on offer. This suggests that D1R activation in particular may play an important role in encoding of future value, but also in setting a threshold for when action is worth executing for reward.

Opportunity cost can also be altered by the effort required to garner reward and manipulation of both dopaminergic receptor subtypes has been associated with guidance of effort-based choice (Bardgett et al., 2009). Here, effort was not explicitly manipulated in the Go/No-Go task. However, whether Go trials were more effortful than No-Go trials, or vice versa, is worth a brief discussion. At baseline, across all cohorts of animals, success rate in Go trials was modulated by reward size on offer, whilst success rate in No-Go trials was not. However, in most cases, this was borne out as success rate being particularly high in Go Large trials, and animals were generally able to perform No-Go trials at similar levels to Go Small trials, suggesting comparable utility across these trial types. Similarly, time from trial initiation to

reward collection was very similar across Go and No-Go trials, meaning that, unlike many effort-based tasks, reward rate was held approximately constant.

In addition, the fact that animals often remained in the nose poke for longer than necessary in No-Go trials suggests that, whilst inhibition of action was to some degree effortful, it was not unduly so. Others have also found that antagonists of either dopamine receptor subtype affect operant responding for reward – but not waiting for reward, again suggesting that waiting (at least in terms of a calculation of utility by dopamine) cannot be seen as 'effortful' in the same way that physical exertion is effortful (Wakabayashi, Fields, & Nicola, 2004). Furthermore, amphetamine has been shown to boost choice of effortful options (Bardgett et al., 2009), and yet here reduced No-Go performance, again suggesting that there is not a simple link between effort and withholding action for reward.

3. Circuit effects of pharmacological manipulations

The focus of this thesis has been interpreting the effects of pharmacological stimulation or blockade of ventral striatal dopamine receptors. However, the downstream consequences of such receptor activation are key in ultimately mediating behavioural output, and while the interpretations of the pharmacological effects here have focused on modulations of the targeted receptors, it may also be that these in part relate to downstream changes in striatal dopamine release.

Specifically, activation of these receptors reduces levels of dopamine in this region as measured by microdialysis (Zetterstrom & Ungerstedt, 1984), although studies *in vivo* have suggested that this may be predominantly mediated by D2Rs (Boyar & Altar, 1987; See et al., 1991). Studies using dopamine receptor antagonists are more mixed, with some showing little effect on striatal dopamine levels (Zetterström et al., 1986) and others showing an increase in dopamine and its metabolites, but more so with a D2R than a D1R antagonist (See et al., 1991). These effects may also be dependent on how long after administration the measurements are taken (Garris et al., 2003)

However, many of these previous studies have used microdialysis or other measures of slow fluctuations of dopamine release. Additionally, the D1R agonist applied in all cases was a partial rather than full agonist. To understand whether the D1R agonist we specifically applied in Chapters 4 and 5, SKF 81297, would have an effect on size of striatal phasic dopamine release, we conducted a pilot study (n = 2) in DAT::Ires-Cre Sprague Dawley rats using optical stimulation of VTA dopamine neurons, transfected with channel rhodopsin2, to evoke transients in the NAc core as measured using fast-scan cyclic voltammetry (stimulation parameters: 50Hz, 20 pulses at 5mW with a 5ms pulse). We then systemically applied SKF 81297 at 1.0 mg/kg (the high dose used in Chapter 4), and found that the peak amplitude of accumbens transients was reduced shortly after injection (Figure 1a and 1b). Whilst highly preliminary, this suggests that dopamine release at short timescales may also be modulated by the application of D1R receptor pharmacological compounds, complicating the simplistic notion that the D1R agonist may partially mimic stimulation following large phasic releases of dopamine.

The key possible consequence of this is, when using a compound that is receptor subtype specific, activity at the other, unaltered, receptor subtype will be influenced in the converse direction. Therefore here, the D1R agonist, by reducing the amplitude of phasic dopamine transients, may lead to a reduction in activation of D2Rs – which are known to be sensitive to pauses in dopamine neuron firing (Dreyer et al., 2010). It is not possible to determine whether any such effects drove the results found in this thesis, or indeed, in other studies using similar approaches, but this should be considered for future experiments. In addition, dopamine is not simply a neurotransmitter but, in particular, a neuro*modulator*; therefore its effects, and thus the effects of activation of MSN subpopulations, might be dependent on circuit state and other neuromodulatory influences – including those of other neuromodulators – that were not identified in the experiments here (Avery & Krichmar, 2017).

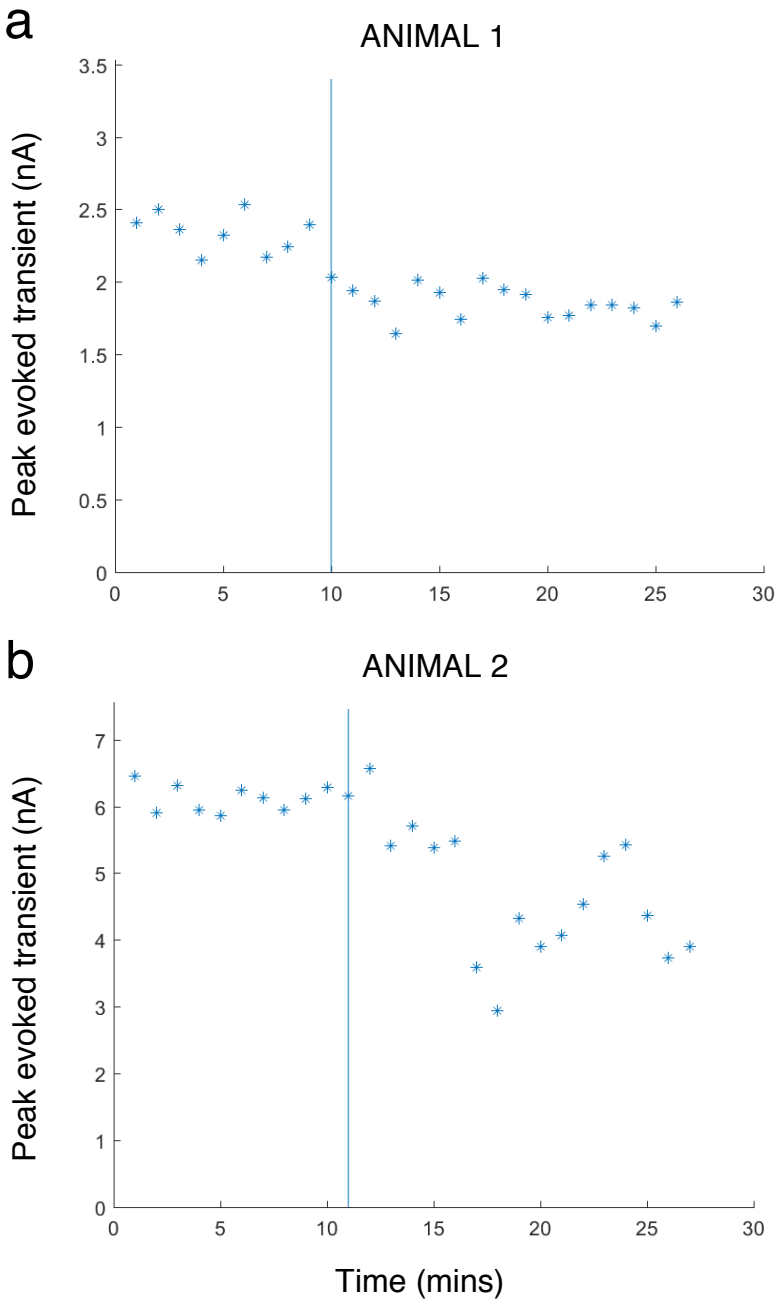


Fig. 1. Effect of systemic injection of a D1R agonist, SKF 81297 (1.0 mg/kg) on peak optically evoked dopamine transients as recorded by FCV in the NAc core. Blue lines indicate time of injection. Adapted from Werlen, unpublished).

4. Tonic and phasic dopamine: a valid distinction?

Fluctuations of dopamine levels are canonically thought to have two separable modes: large amplitude subsecond phasic release, and steady-state 'background' tonic release occurring over a timescale of minutes (Grace, 1991). Thus it has been

proposed that phasic transmission may convey reward learning information, whilst tonic changes may instead regulate overall levels of motivation or arousal (Schultz, 1998) or, more specifically, enables animals to exploit the reward-related information they have learnt (Beeler, Daw, Frazier, & Zhuang, 2010). Because of this, reported differences in the affinity of D1Rs and and D2Rs to these modes of dopamine (Dreyer et al., 2010) were utilised in the experiments within this thesis to tackle their differential behavioural contributions.

Mechanistically, the two modes of firing and release are physiologically separable (Floresco, West, Ash, Moorel, & Grace, 2003; Grace, 1991). However, whether this results in a meaningful distinction at a *behavioural* level has come under question. As aforementioned, reward learning and motivation have been attributed to the phasic and tonic signals, and Schultz (2007) has gone further to claim that this is because the behaviours that are mediated by these signals themselves occur at these different timescales. Yet direct evidence of a covariance between tonic dopamine firing rates and motivational state is lacking (Cohen, Amoroso, & Uchida, 2015), whilst we (in Chapter 3) and others have observed modulation of phasic signalling by motivated behaviour (Phillips et al., 2003; Wassum et al., 2012), suggesting less of a clear-cut distinction.

A recent review agrees that changes in dopamine levels occur at the timecourses of the relevant behaviours, but suggests that tonic/phasic distinction has arisen due to the temporal resolution of the methods used, rather than being a meaningful separation of the timecourses of the dopamine signal (Berke, 2018). The author also

suggests an alternative mechanism by which both prediction error and motivational signals can be conveyed, with cholinergic interneurons acting to switch the downstream interpretation of the signal from one to the other by pausing their activity to act as temporal windows for striatal plasticity (Morris, Arkadir, Nevet, Vaadia, & Bergman, 2004). Therefore, whilst the tonic/phasic distinction has been used throughout this thesis due to its establishment in the literature, it is recognised here that such a distinction may be artificial. On the other hand, regardless of whether such a distinction is truly physiological, the dopamine receptor subtypes have repeatedly been demonstrated to have different baseline affinities, and therefore the *downstream* consequences of dopamine efflux may still be separable in terms of slow and fast changes in dopamine levels.

5. *Dopamine in decision-making contexts: foraging for reward*

All of the data presented in this thesis were deliberately collected from well-learned tasks, but considering the role of prediction errors in updating value estimates in learning (Bayer & Glimcher, 2005), the question arises of the function of the dopaminergic signal in more dynamic situations. In addition, the Go/No-Go task is deterministic: the cues instruct animals as to what they should do to garner reward, but there is no explicit decision-making aspect to the task.

Dopamine has broadly been linked to decisions about effort (Bardgett et al., 2009), risk (Simon et al., 2011) and delay (Floresco, Tse, & Ghods-Sharifi, 2008). However,

mesolimbic dopamine does not seem to map directly onto decision-making when having to decide between *simultaneously* presented options (Hollon et al., 2014; Roesch, Calu, & Schoenbaum, 2007). Instead, consistent both with ideas of dopamine ramping throughout trial progression (Hamid et al., 2016; Niv, 2013) and that slower changes in dopamine levels encode a long-term average reward rate (Niv et al., 2007), there is the possibility that midbrain dopamine may instead be integral to sequential, or foraging decisions.

An example of sequential decision-making is 'patch-leaving', a 'patch' being a localised concentration of resources that depletes as an organism harvests reward within it (Stephens, 2008). Importantly, the reward rate in a patch depletes until it is economically advantageous to move into a new, replenished patch at a cost of time or energy. Therefore it may be the case that mesolimbic dopamine is key to the decision of whether to move on to a new patch, or to continue to exploit the current location.

We have begun to explore this possibility by developing a novel foraging paradigm (Figure 2a), where animals can harvest reward from a depleting patch, as controlled by increasing the length of the delay between each consecutive reward delivery. Here, the patch is a nose poke in which animals must sustain a poke for increasing durations of time. At any point (but particularly when the animal decides that the rate of reward in the current patch is no longer worth the cost of continuing to poke) they can decide to open up a new, replenished patch – a poke on the other side of the panel – but have to pay the cost of travel by poking into a centre port for 1, 2, or 4s.

We have established that animals will harvest a greater number of rewards in the current patch when travel time is longer (Figure 2b) and also leave sooner when the incrementing delay ratio is lower and thus environmental richness is higher (Figure 2c), as predicted by economic foraging theory. Our next step is to combine this behaviour with fibre photometry, a calcium imaging method that allows for the recording of defined populations of neurons in freely-moving animals. By recording the calcium levels as a proxy of the activity of midbrain dopamine neurons, both at cell bodies and at their terminals, we hope to elucidate the role of these neurons in sequential decision-making.

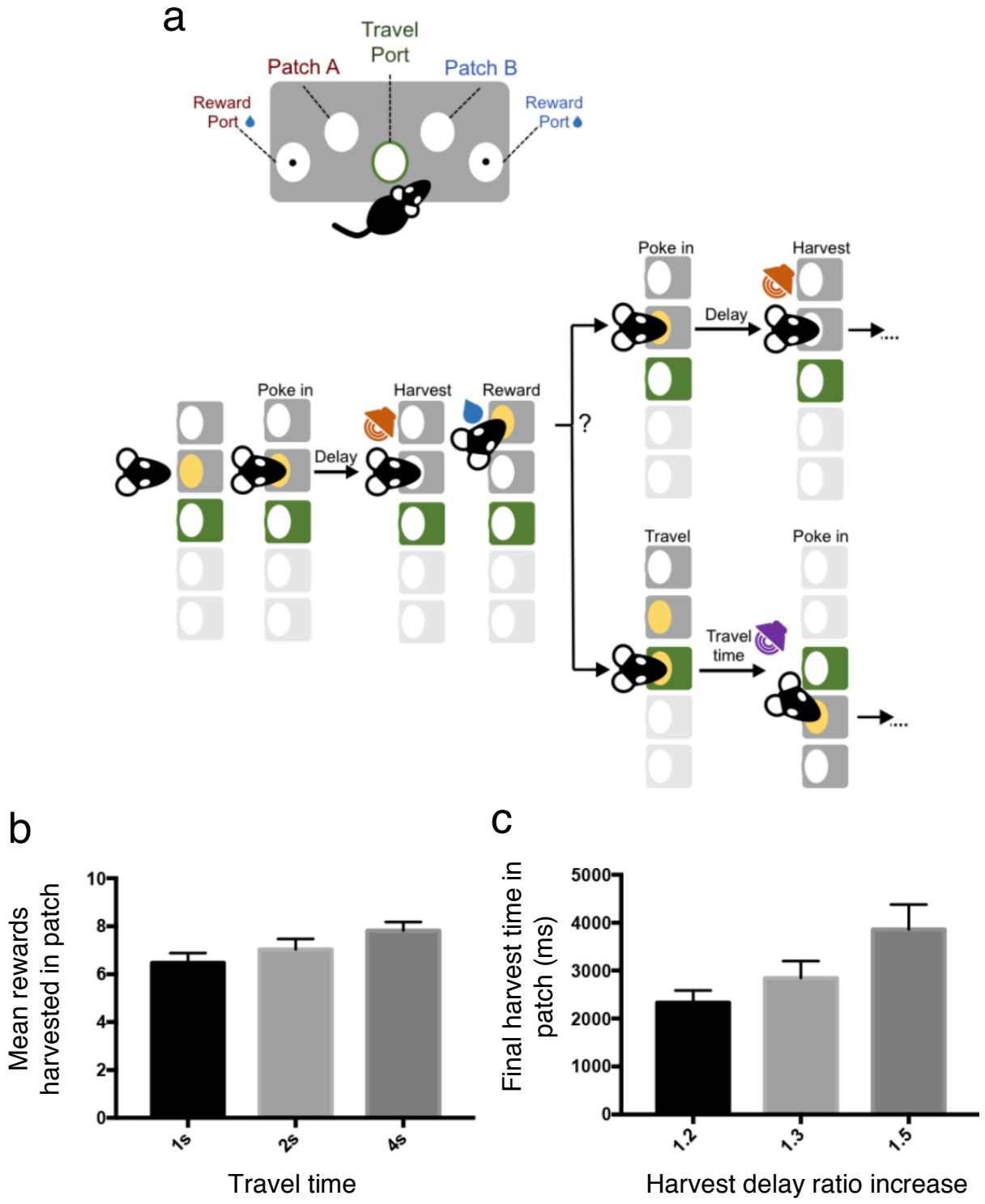


Fig. 2. Pilot design and behavioural data from the patch-leaving foraging task. (a) Task schematic. Mice have to choose between harvesting depleting water rewards in one patch, or travelling to open a new, replenished patch. (b) On average, mice harvest more rewards in the current patch when the travel time is longer. (c) On average, mice remained in the current patch for longer periods of time when the delay ratio was steeper due to poorer average environmental richness.

6. Limitations, future directions, and conclusions

The studies in this thesis sought first to understand the contribution of the goal-directed movement to the mesolimbic dopamine signal, and to elucidate the behavioural consequences of said signal when conveyed via the distinct dopaminergic receptor subtypes. In the final experimental chapter, the role of the serotonergic 5-HT_{2C} receptor on action and inaction was also explored.

Across the literature it is readily apparent that temporal resolution is important when understanding dopamine's effects on behaviour. Regardless of the caveats re: phasic/tonic dopamine as discussed in section 4, it is still the case that pharmacological manipulations as employed here act on a timescale that is biologically implausible, with dopamine or serotonergic receptors being tonically stimulated or blocked. Whilst these and other similar studies can still contribute to an understanding of the role of these receptors in behaviour, this limits the extent to which such manipulations can be interpreted as mimicking the signalling seen using methods with a higher temporal resolution, such as in Chapter 3. Indeed, exactly when the phasic signal is conveyed may be integral to how it is used to influence behaviour and therefore approaches such as optogenetics, that can excite or inhibit the activity of genetically-identified neurons at shorter timescales, may go further in elucidating the functional role of the signal we observed. Similarly, methods such as fibre photometry, when used in conjunction with strategies that allow for the

recording of populations of neurons that express specific receptor subtypes, may also further understanding of the downstream effects of this activity.

Spatial resolution is also integral to a fuller understanding of the dopaminergic signal. The systemic results presented in Chapters 4 and 6 therefore cannot be said to speak to mesolimbic dopamine only, but may have taken effect at dopamine receptors elsewhere in the brain. Indeed, considering the effects seen and doses of the compound used, it is likely that the locus of the D2R agonist effects was at dorsal striatal autoreceptors. In Chapters 5 and 7, we therefore targeted our manipulations to the NAc core – a site chosen to parallel that used in Chapter 3, but also due to being the reported locus of prediction error encoding. We were unable to infuse the D2R ligands due to a limitation in the number of effective infusions possible in a single cohort, but future studies would benefit from exploring the role of this receptor subtype in this region in the type of behaviour described here, as well as to adjacent regions such as the NAc shell or dorsomedial striatum, where their contributions to behaviour may be differentiable from those shown here (Di Chiara et al., 2004; Marcott et al., 2018; Shin, Kim, & Jung, 2018).

In conclusion, we have shown that phasic dopamine release is mediated by goal-directed movement, and that D1Rs in the NAc core are critical in driving the initiation of this behaviour, whilst D1Rs elsewhere contribute to the ongoing performance of action, including action selection, for reward. However, the activation of these receptors for goal-directed behaviour is critical only when such behaviour involves *movement*. Therefore the expression of incentive motivation by dopamine is

separable by action initiation and sustainment of action, and is dependent on movement.

In contrast, D2R activation was not necessary for successful performance in the task, suggesting that D1Rs play a unique role in the activation of behaviour for reward. Finally, effective blockade of serotonergic 5-HT_{2C} receptors also drove the expression of incentive motivated behaviours, suggesting a role for this receptor subtype in behaviour for reward but less dependent on cue-responding. Together, these studies demonstrate for the first time that the expression of goal-directed behaviour via dopaminergic mediation is contingent specifically on movement, bringing the field one step closer to a holistic understanding of the role of two central monoamines, dopamine and serotonin, in reward and motivated action.

References

Agnoli, L., Mainolfi, P., Invernizzi, R. W., & Carli, M. (2013). Dopamine D1-like and D2-like receptors in the dorsal striatum control different aspects of attentional performance in the five-choice serial reaction time task under a condition of increased activity of corticostriatal inputs. *Neuropsychopharmacology*, *38*(5), 701–714.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The Functional Anatomy of Basal Ganglia Disorders. *Trends in Neurosciences*, *12*, 366–375.

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

Alex, K. D., Yavanian, G. J., McFarlane, H. G., Pluto, C. P., & Pehek, E. A. (2005). Modulation of dopamine release by striatal 5-HT2C receptors. *Synapse*, *55*(4), 242–251.

Alleweireldt, A. T., Weber, S. M., Kirschner, K. F., Bullock, B. L., & Neisewander, J. L. (2002). Blockade or stimulation of D1 dopamine receptors attenuates cue reinstatement of extinguished cocaine-seeking behavior in rats. *Psychopharmacology*, *159*(3), 284–293.

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

Anastasio, N. C., Lanfranco, M. F., Bubar, M. J., Seitz, P. K., Stutz, S. J., McGinnis, A. G., … Cunningham, K. A. (2010). Serotonin 5-HT2C receptor protein expression is enriched in synaptosomal and post-synaptic compartments of rat cortex. *Journal of Neurochemistry*, *113*(6), 1504–1515.

Anden, N. E., Carlsoon, A., Dahlstroem, A., Fuxe, K., Hillarp, N. A., & Larsson, K. (1964). Demonstration and mapping out of nigro-neostriatal dopamine neurons. *Life Sciences*, *3*, 523–530.

Andersen, P. H., Gingrich, J. A., Bates, M. D., Dearry, A., Falardeau, P., Senogles, S. E., & Caron, M. G. (1990). Dopamine receptor subtypes: beyond the D1/D2 classification. *Trends in Pharmacological Sciences*, *11*, 231–236.

Anzalone, A., Lizardi-Ortiz, J. E., Ramos, M., De Mei, C., Hopf, F. W., Iaccarino, C., … Borrelli, E. (2012). Dual Control of Dopamine Synthesis and Release by Presynaptic and Postsynaptic Dopamine D2 Receptors. *Journal of Neuroscience*, *32*(26), 9023–9034.

Arnold, M. M., Burgeno, L. M., & Phillips, P. E. M. (2015). Fast-scan cyclic voltammetry in behaving animals. In *Basic Electrophysiological Methods*. Oxford University Press.

Avery, M. C., & Krichmar, J. L. (2017). Neuromodulatory Systems and Their Interactions: A Review of Models, Theories, and Experiments. *Frontiers in Neural Circuits*, *11*, 108.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Bailey, M. R., Goldman, O., Bello, E. P., Chohan, M. O., Jeong, N., Winiger, V., … Simpson, E. H. (2018). An Interaction between Serotonin Receptor Signaling

and Dopamine Enhances Goal-Directed Vigor and Persistence in Mice. *Journal of Neuroscience*, 2088–17.
Bailey, M. R., Williamson, C., Mezias, C., Winiger, V., Silver, R., Balsam, P. D., & Simpson, E. H. (2016). The effects of pharmacological modulation of the serotonin 2C receptor on goal-directed behavior in mice. *Psychopharmacology*, *233*(4), 615–624.
Bardgett, M. E., Depenbrock, M., Downs, N., Points, M., & Green, L. (2009). Dopamine modulates effort-based decision making in rats. *Behavioral Neuroscience*, *123*(2), 242–251.
Bari, A., & Robbins, T. W. (2013). Noradrenergic versus dopaminergic modulation of impulsivity, attention and monitoring behaviour in rats performing the stop-signal task: Possible relevance to ADHD. *Psychopharmacology*, *230*(1), 89–111.
Bari, A., & Robbins, T. W. (2013). Progress in Neurobiology Inhibition and impulsivity : Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44–79.
Barker, E. L., Westphal, R. S., Schmidt, D., & Sanders-Bush, E. (1994). Constitutively active 5-hydroxytryptamine2C receptors reveal novel inverse agonist activity of receptor ligands. *The Journal of Biological Chemsitry*, *269*, 11687–11690.
Barter, J. W., Li, S., Lu, D., Bartholomew, R. A., Rossi, M. A., Shoemaker, C. T., … Yin, H. H. (2015). Beyond reward prediction errors : the role of dopamine in movement kinematics, *9*(39), 1–22.
Bartlett, S. E., Enquist, J., Hopf, F. W., Lee, J. H., Gladher, F., Kharazia, V., … Whistler, J. L. (2005). Dopamine responsiveness is regulated by targeted sorting of D2 receptors. *Proceedings of the National Academy of Sciences*, *102*(32), 11521–11526.
Bassareo, V., & Di Chiara, G. (1999). Modulation of feeding-induced activation of mesolimbic dopamine transmission by appetitive stimuli and its relation to motivational state. *European Journal of Neuroscience*, *11*(12), 4389–4397.
Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.
Beaulieu, J.-M., & Gainetdinov, R. R. (2011). The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological Reviews*, *63*(1), 182–217.
Beckstead, M. J., Grandy, D. K., Wickman, K., & Williams, J. T. (2004). Vesicular dopamine release elicits an inhibitory postsynaptic current in midbrain dopamine neurons. *Neuron*, *42*(6), 939–946.
Beeler, J. A., Daw, N., Frazier, C. R. M., & Zhuang, X. (2010). Tonic Dopamine Modulates Exploitation of Reward Learning. *Frontiers in Behavioral Neuroscience*, *4*(170), 1–14.
Beeler, J. A., Frazier, C. R. M., Zhuang, X., & Salamone, J. D. (2012). Putting desire on a budget : dopamine and energy expenditure , reconciling reward and resources, *6*(July), 1–22.
Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., & Dayan, P. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology*, *38*(8), 1495–1503.
Bello, E. P., Mateo, Y., Gelman, D. M., Noaín, D., Shin, J. H., Low, M. J., … Rubinstein, M. (2011). Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2autoreceptors. *Nature Neuroscience*, *14*(8),

1033–1038.
Benoit-Marand, M., Borrelli, E., & Gonon, F. (2001). Inhibition of dopamine release via presynaptic D2 receptors: time course and functional characteristics in vivo. *The Journal of Neuroscience*, *21*(23), 9134–41.
Bentivoglio, M., & Morelli, M. (2005). The organisation and circuits of mesencephalic dopaminergic neurons and the distribution of dopamine receptors in the brain. In *Handbook of Chemical Neuroanatomy* (pp. 1–107).
Berg, K. A., Harvey, J. A., Spampinato, U., & Clarke, W. P. (2005). Physiological relevance of constitutive activity of 5-HT2A and 5-HT2C receptors. *Trends in Pharmacological Sciences*, *26*(12), 625–630.
Berger, M., Gray, J. A., & Roth, B. L. (2009). The Expanded Biology of Serotonin. *Annual Review of Medicine*, *60*(1), 355–366.
Berke, J. D. (2018). What does dopamine mean? *Nature Neuroscience*, *21*, 787-793.
Berridge, K. C. (1996). Food reward – brain structures of wanting and liking. *Neuroscience and Biobehavioral Reviews*, *20*(1), 1–25.
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. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research*
Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: "liking", "wanting", and learning. *Current Opinion in Pharmacology*.
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.
Bertran-Gonzalez, J., Bosch, C., Maroteaux, M., Matamales, M., Herve, D., Valjent, E., & Girault, J.-A. (2008). Opposing Patterns of Signaling Activation in Dopamine D1 and D2 Receptor-Expressing Striatal Neurons in Response to Cocaine and Haloperidol. *Journal of Neuroscience*, *28*(22), 5671–5685.
Bezzina, G., Body, S., Cheung, T. H. C., Hampson, C. L., Bradshaw, C. M., Glennon, J. C., & Szabadi, E. (2015). Evidence for a role of 5-HT 2C< receptors in the motor aspects of performance, but not the efficacy of food reinforcers, in a progressive ratio schedule. *Psychopharmacology*, *232*(4), 699–711.
Bischoff, S., Heinrich, M., Sonntag, J. M., & Krauss, J. (1986). The D-1 dopamine receptor antagonist SCH 23390 also interacts potently with brain serotonin (5-HT 2) receptors. *European Journal Pharmacology*, *129*, 367–370.
Bizot, J. C., Le Bihan, C., Puech, A. J., Hamon, M., & Thiébot, M. H. (1999). Serotonin and tolerance to delay of reward in rats. *Psychopharmacology*, *146*(4), 400–412.
Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an update. *Trends in Neurosciences*, *30*(5), 194–202.
Blanchard, D. C., Blanchard, R. J., & Briebel, G. (2005). Defensive responses to predator thread in the rat and mouse. *Current Protocols in Neuroscience*, *30*(1), 8.19.1-8.19.20.
Blokland, A., Sik, A., & Lieben, C. (2005). Evaluation of DOI, 8-OH-DPAT, eticlopride and amphetamine on impulsive responding in a reaction time task in rats. *Behavioural Pharmacology*, *16*(2), 93–100.
Bogacz, R. (2017). Theory of reinforcement learning and motivation in the basal

ganglia, *bioRxiv*, 174524.
Bolam, J. P., Hanley, J. J., Booth, P. A. C., & Bevan, M. D. (2000). Synaptic organisation of the basal ganglia. *Journal of Anatomy*, *196*(4), 527–542.
Bolonna, A. A., & Kerwin, R. W. (2005). Partial agonism and schizophrenia. *British Journal of Psychiatry*, *186*, 7–10.
Boulougouris, V., & Robbins, T. W. (2010). Enhancement of Spatial Reversal Learning by 5-HT2C Receptor Antagonism Is Neuroanatomically Specific. *Journal of Neuroscience*, *30*(3), 930–938.
Boureau, Y. L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*(1), 74–97.
Bourne, J. a. (2001). SCH 23390: the first selective dopamine D1-like receptor antagonist. *CNS Drug Reviews*, *7*(4), 399–414.
Boyar, W. C., & Altar, C. A. (1987). Modulation of In Vivo Dopamine Release by D2 but Not D1 Receptor Agonists and Antagonists. *Journal of Neurochemistry*, *48*(3), 824–831.
Brady, S. T., Siegel, G. J., Albers, R. W., & Price, D. L. (2012). Basic Neurochemistry: Principles of Molecular, Cellular, and Medical Neurobiology. Academic Press.
Bromberg-Martin, E. S., Hikosaka, O., & Nakamura, K. (2010). Coding of Task Reward Value in the Dorsal Raphe Nucleus. *Journal of Neuroscience*, *30*(18), 6262–6272.
Brown, H. D., Mccutcheon, J. E., Cone, J. J., Ragozzino, M. E., & Roitman, M. F. (2011). Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *European Journal of Neuroscience*, *34*(12), 1997–2006.
Brown, P., & Molliver, M. E. (2000). Dual serotonin (5-HT) projections to the nucleus accumbens core and shell: relation of the 5-HT transporter to amphetamine-induced neurotoxicity. *The Journal of Neuroscience*, *20*(5), 1952–1963.
Brown, V. J., & Bowman, E. M. (1995). Discriminative Cues Indicating Reward Magnitude Continue to Determine Reaction Time of Rats Following Lesions of the Nucleus Accumbens. *European Journal of Neuroscience*, *7*(12), 2479–2485.
Browne, C. J., & Fletcher, P. J. (2016). Decreased Incentive Motivation Following Knockout or Acute Blockade of the Serotonin Transporter: Role of the 5-HT2C Receptor. *Neuropsychopharmacology*, *41*(10), 2566–2576.
Bubar, M. J., Stutz, S. J., & Cunningham, K. A. (2011). 5-HT2c Receptors localize to dopamine and gaba neurons in the rat mesoaccumbens pathway. *PLoS ONE*, *6*(6).
Bunner, K. D., & Rebec, G. V. (2016). Voltammetry in behaving animals. In: *Receptor and Ion Channel Detection in the Brain, vol 110*. Spring Science + Business Media: New York.
Burmeister, J. J., Lungren, E. M., Kirschner, K. F., & Neisewander, J. L. (2004). Differential roles of 5-HT receptor subtypes in cue and cocaine reinstatement of cocaine-seeking behavior in rats. *Neuropsychopharmacology*, *29*(4), 660–668.
Cachope, R., Mateo, Y., Mathur, B. N., Irving, J., Wang, H. L., Morales, M., … Cheer, J. F. (2012). Selective activation of cholinergic interneurons enhances accumbal

phasic dopamine release: Setting the tone for reward processing. *Cell Reports*, *2*(1), 33–41.
Caille, I., Dumartin, B., & Bloch, B. (1996). Ultrastructural localization of D 1 dopamine receptor immunoreactivity in rat striatonigral neurons and its relation with dopaminergic innervation. *Brain Research*, *730*, 17–31.
Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., & Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: A critical reappraisal. *Nature Neuroscience*, *17*(8), 1022–1030.
Callier, S., Snapyan, M., Le Crom, S., Prou, D., Vincent, J. D., & Vernier, P. (2003). Evolution and cell biology of dopamine receptors in vertebrates. *Biology of the Cell*, *95*(7), 489–502.
Cardinal, R. N., Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., & Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science (New York, N.Y.)*, *292*(5526), 2499–2501.
Carlsson, A., Lindqvist, M., Magnusson, T., & Waldeck, B. (1958). On the presence of 3-hydroxytyramine in brain. *Science*, *127*(3296), 471.
Carr, D. B., Cooper, D. C., Ulrich, S. L., Spruston, N., & Surmeier, D. J. (2002). Serotonin receptor activation inhibits sodium current and dendritic excitability in prefrontal cortex via a protein kinase C-dependent mechanism. *J Neurosci*, *22*(16), 6846–6855.
Cazorla, M., Carvalho, F. D. De, Chohan, M. O., Shegda, M., Chuhma, N., Rayport, S., … Kellendonk, C. (2014). Dopamine D2 Receptors Regulate the Anatomical and Functional Balance of Basal Ganglia Circuitry. *Neuron*, *81*(1), 153–164.
Celada, P., Casanovas, J. M., Paez, X., & Artigas, F. (2002). Control of serotonergic neurons in the dorsal raphe nucleus by the lateral hypothalamus, *932*, 79–90.
Cepeda, C., Buchwald, N. A., & Levine, M. S. (1993). Neuromodulatory actions of dopamine in the neostriatum are dependent upon the excitatory amino acid receptor subtypes activated. *Proceedings of the National Academy of Sciences*, *90*(20), 9576–9580.
Charnay, Y., & Leger, L. (2010). Pharmacological aspects: brain serotonergic circuitries. *Dialogues in Clinical Neuroscience*, *12*(4), 471–487.
Chaudhuri, K. R., & Schapira, A. H. V. (2009). Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *The Lancet*, *8*(5), 464–474.
Chefer, V. I., Thompson, A. C., Zapata, A., & Shippenberg, T. S. (2010). Overview of brain microdialysis. In *Current protocols in neuroscience* (pp. 1–35).
Christakou, A., Robbins, T. W., & Everitt, B. J. (2004). Prefrontal Cortical-Ventral Striatal Interactions Involved in Affective Modulation of Attentional Performance: Implications for Corticostriatal Circuit Function. *Journal of Neuroscience*, *24*(4), 773–780.
Chudasama, Y., & Robbins, T. W. (2004). Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology*, *29*(9), 1628–1636.
Clark, J. J., Sandberg, S. G., Wanat, M. J., Gan, J. O., Horne, E. a, Hart, A. S., … Phillips, P. E. M. (2010). Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. *Nature Methods*, *7*(2), 126–129.
Clark, L., Roiser, J. P., Cools, R., Rubinsztein, D. C., Sahakian, B. J., & Robbins, T. W.

(2005). Stop signal response inhibition is not modulated by tryptophan depletion or the serotonin transporter polymorphism in healthy volunteers: Implications for the 5-HT theory of impulsivity. *Psychopharmacology*, *182*(4), 570–578.
Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, *33*, 107–112.
Cohen, J. Y., Amoroso, M. W., & Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *eLife*, *2015*(4), 1–25.
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.
Conn, J. P., Lindsley, C. W., Meiler, J., & Niswender, C. M. (2014). Opportunities and challenges in the discovery of allosteric modulators of GPCRs for treating CNS disorders. *Nature Reviews Drug Discovery*, *13*, 692–708.
Cools, R., Nakamura, K., & Daw, N. D. (2011). Serotonin and Dopamine : Unifying Affective , Activational , and Decision Functions. *Neuropsychopharmacology*, *36*, 98–113.
Correia, P. A., Lottem, E., Banerjee, D., Machado, A. S., Carey, M. R., & Mainen, Z. F. (2017). Transient inhibition and long-term facilitation of locomotion by phasic optogenetic activation of serotonin neurons. *eLife*, *6*(e20975), 1–27.
Crockett, M. J., Clark, L., Apergis-Schoute, A. M., Morein-Zamir, S., & Robbins, T. W. (2012). Serotonin modulates the effects of pavlovian aversive predictions on response vigor. *Neuropsychopharmacology*, *37*(10), 2244–2252.
Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the Role of Serotonin in Behavioral Inhibition and Aversion: Acute Tryptophan Depletion Abolishes Punishment-Induced Inhibition in Humans. *Journal of Neuroscience*, *29*(38), 11993–11999.
Cubeddu, L. X., Hoffman, I. S., & Talmaciu, R. K. (1990). Is the Release of Dopamine from Medial Prefrontal Cortex Modulated by Presynaptic Receptors? Comparison with Nigrostriatal and Mesolimbic Terminals. *Annals of the New York Academy of Sciences*, *604*(1), 452–461.
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., & Fuxe, K. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system. *Acta Phsyiologica Scandinavica*, *62*, 1–55.
Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2011). Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron*, *69*(4), 680–694.
Dalley, J. W., & Roiser, J. P. (2012). Dopamine, Serotonin and Impulsivity. *Neuroscience*, *215*, 42–58.
Davies, J., & Tongroach, P. (1978). Neuropharmacological studies on the nigro-striatal and raphe-striatal system in the rat. *European Journal of Pharmacology*, *51*(2), 91–100.
Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, *15*, 603–616.
Day, J. J., Jones, J. L., Wightman, R. M., & Carelli, R. M. (2010). Phasic nucleus

accumbens dopamine release encodes effort- and delay-related costs. *Biological Psychiatry*, *68*(3), 306–309.
Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature Neuroscience*, *10*(8), 1020–1028.
Dayan, P., & Huys, Q. J. M. (2008). Serotonin, inhibition, and negative mood. *PLoS Computational Biology*, *4*(2).
Dayan, P., & Huys, Q. J. M. (2009). Serotonin in Affective Control. *Annual Review of Neuroscience*, *32*(1), 95–126.
De Deurwardere, P., Navailles, S., Berg, K. A., Clarke, W. P., & Spampinato, U. (2004a). Constitutive Activity of the Serotonin 2C Receptor Inhibits In Vivo Dopamine Release in the Rat Striatum and Nucleus Accumbens. *Journal of Neuroscience*, *24*(13), 3235–3241.
De Deurwardere, P., Navailles, S., Berg, K. A., Clarke, W. P., & Spampinato, U. (2004b). Constitutive Activity of the Serotonin 2C Receptor Inhibits In Vivo Dopamine Release in the Rat Striatum and Nucleus Accumbens. *Journal of Neuroscience*, *24*(13), 3235–3241.
Deakin, J. F. W., & Graeff, F. G. (1991). 5-HT and mechanisms of defence. *Journal of Psychopharmacology*, *5*(4), 339–341.
De Mei, C., Ramos, M., Iitaka, C., & Borrelli, E. (2009). Getting specialized: presynaptic and postsynaptic dopamine D2 receptors. *Current Opinion in Pharmacology*, *9*(1), 53–58.
Delle Donne, K. T., Sesack, S. R., & Pickel, V. M. (1997). Ultrastructural immunocytochemical localization of the dopamine D2 receptor within GABAergic neurons of the rat striatum. *Brain Research*, *746*, 239–255.
Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, *179*(3), 587–596.
Deutch, A. Y., & Cameron, D. S. (1992). Pharmacalogical characterization of dopamine systems in the nucleus accumbens core and shell. *Neuroscience*, *46*(1), 49–56.
Di Chiara, G., Bassareo, V., Fenu, S., De Luca, M. A., Spina, L., Cadoni, C., … Lecca, D. (2004). Dopamine and drug addiction: The nucleus accumbens shell connection. *Neuropharmacology*, *47*, 227–241.
Di Chiara, G., & Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences*, *85*(14), 5274–5278.
Di Giovanni, G., Di Matteo, V., Di Mascio, M., & Esposito, E. (2000). Preferential modulation of mesolimbic vs. nigrostriatal dopmanergic function by serotonin 2C/2B receptor agonists: a combined in vivo electrophysiology and microdialysis study. *Synapse*, *35*, 53–61.
Di Giovanni, G., Esposito, E., & Di Matteo, V. (2010). Role of serotonin in central dopamine dysfunction. *CNS Neuroscience and Therapeutics*, *16*(3), 179–194.
Di Matteo, V., Di Giovanni, G., & Esposito, E. (2000). SB 242084: A selective 5-HT(2C) receptor antagonist. *CNS Drug Reviews*, *6*(3), 195–205.

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

Dreher, J. K., & Jackson, D. M. (1989). Role of D1 and D2 dopamine receptors in mediating locomotor activity elicited from the nucleus accumbens of rats. *Brain Res.*, *487*(1989), 267–277.

Dreyer, J. K., Herrik, K. F., Berg, R. W., & Hounsgaard, J. D. (2010). Influence of phasic and tonic dopamine release on receptor activation. *The Journal of Neuroscience*, *30*(42), 14273–14283.

du Hoffmann, J., & Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the Nucleus Accumbens. *Journal of Neuroscience*, *34*(43), 14349–14364.

Durana, J. H., & Barnes, P. A. (1993). A neurodevelopmental view of impulsivty and its relationship to the superfactors of personality. In W. G. McCown, J. L. Johnson, & M. B. Shure (Eds.), *The Impulsive Client; Theory, Research and Treatment*. Washington D. C.: American Psychological Association.

Durcan, M. J., Rigdon, G. C., Norman, M. H., & Morgan, P. F. (1995). Is clozapine selective for the dopamine D4 receptor? *Life Sciences*, *57*(18), 275–283.

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., Lehmann, O., Theobald, D. E. H., Pena, Y., Zakaria, R., Ghosh, R., … Robbins, T. W. (2009). Serotonin depletion impairs waiting but not stop-signal reaction time in rats: Implications for theories of the role of 5-HT in behavioral inhibition. *Neuropsychopharmacology*, *34*(5), 1311–1321.

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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(20), 7349–7356.

Eberle-Wang, K., Mikeladze, Z., Uryu, K., & Chesselet, M.-F. (1997). Pattern of expression of the serotonin 2C receptor messenger RNA in the basal ganglia of adult rats. *The Journal of Comparative Neurology*, *384*(2), 233–247.

Eilam, D., & Szechtman, H. (1989). Biphasic effect of D-2 agonist quinpirole on locomotion and movements. *European Journal of Pharmacology*, *161*(2–3), 151–157.

Elverfors, A., & Nissbrandt, H. (1992). Effects of d-amphetamine on dopaminergic neurotransmission; a comparison between the substantia nigra and the striatum. *Neuropharmacology*, *31*(7), 661–670.

Escobar, A. P., Cornejo, F. A., Olivares-Costa, M., González, M., Fuentealba, J. A., Gysling, K., … Andrés, M. E. (2015). Reduced dopamine and glutamate neurotransmission in the nucleus accumbens of quinpirole-sensitized rats hints at inhibitory D2 autoreceptor function. *Journal of Neurochemistry*, *134*(6), 1081–1090.

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

246.
Fernando, A. B. P., Economidou, D., Theobald, D. E., Zou, M. F., Newman, A. H., Spoelder, M., … Dalley, J. W. (2012). Modulation of high impulsivity and attentional performance in rats by selective direct and indirect dopaminergic and noradrenergic receptor agonists. *Psychopharmacology*, *219*(2), 341–352.
Fiorillo, C. D., Song, M. R., & Yun, S. R. (2013). Multiphasic Temporal Dynamics in Responses of Midbrain Dopamine Neurons to Appetitive and Aversive Stimuli. *Journal of Neuroscience*, *33*(11), 4710–4725.
Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(March), 1898–1902.
Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., … Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, *469*(7328), 53–57.
Fletcher, P. J. (1993). A comparison of the effects of dorsal or median raphe injections of 8-OH-DPAT in three operant tasks measuring response inhibition. *Behavioural Brain Research*, *54*(2), 187–197.
Fletcher, P. J., Chambers, J. W., Rizos, Z., & Chintoh, A. F. (2009). Effects of 5-HT depletion in the frontal cortex or nucleus accumbens on response inhibition measured in the 5-choice serial reaction time test and on a DRL schedule. *Behavioural Brain Research*, *201*(1), 88–98.
Fletcher, P. J., Sinyard, J., & Higgins, G. A. (2010). Genetic and pharmacological evidence that 5-HT2C receptor activation, but not inhibition, affects motivation to feed under a progressive ratio schedule of reinforcement. *Pharmacology Biochemistry and Behavior*, *97*(1), 170–178.
Fletcher, P. J., Tampakeras, M., Sinyard, J., & Higgins, G. a. (2007). Opposing effects of 5-HT2A and 5-HT2C receptor antagonists in the rat and mouse on premature responding in the five-choice serial reaction time test. *Psychopharmacology*, *195*(2), 223–234.
Floresco, S. B., St. Onge, J. R., Ghods-Sharifi, S., & Winstanley, C. A. (2008). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cognitive, Affective and Behavioral Neuroscience*, *8*(4), 375–389.
Floresco, S. B., Tse, M. T. L., & Ghods-Sharifi, S. (2008). Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. *Neuropsychopharmacology*, *33*(8), 1966–1979.
Floresco, S. B., West, A. R., Ash, B., Moorel, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience*, *6*(9), 968–973.
Floresco SB. (2006). Dopaminergic regulation of limbic-striatal interplay. *J Psychiatry Neurosci*, *32*(6), 400–411.
Fonseca, M. S., Murakami, M., & Mainen, Z. F. (2015a). Activation of Dorsal Raphe Serotonergic Neurons Promotes Waiting but Is Not Reinforcing. *Current Biology*, *25*(3), 306–315.
Fonseca, M. S., Murakami, M., & Mainen, Z. F. (2015b). Activation of Dorsal Raphe Serotonergic Neurons Promotes Waiting but Is Not Reinforcing. *Current Biology*, *25*, 306–315.
Ford, C. P. (2014). The role of D2-autoreceptors in regulating dopamine neuron

activity and transmission. *Neuroscience*, *282*, 13–22.
Fox, M. E., & Wightman, R. M. (2017). Contrasting Regulation of Catecholamine Neurotransmission in the Behaving Brain : Pharmacological Insights from an Electrochemical Perspective. *Pharmacalogical Reviews*, (January), 12–32.
Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*(1), 51–72.
Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience*, *120*(3), 497–517.
Frank, M. J., Santamaria, A., O'Reilly, R. C., & Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*, *32*(7), 1583–1599.
Franklin, K. B. J., & Paxinos, G. (2007). *The Mouse Brain in Stereotaxic Coordinates, 3rd Edition*. San Diego, Elsevier Academic Press.
Frazer, A., & Hensler, J. G. (1999). Serotonin involvement in physiological function and behavior. In G. J. Siegel, B. W. Agranoff, R. W. Albers, S. K. Fisher, & M. D. Uhler (Eds.), *Basic Neurochemistry: Molecular, Cellular, and Medical Aspects* (6th ed.). Philadelphia: Lippincott-Raven.
Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.
Fuxe, K., Dahlström, A. B., Jonsson, G., Marcellino, D., Guescini, M., Dam, M., … Agnati, L. (2010). The discovery of central monoamine neurons gave volume transmission to the wired brain. *Progress in Neurobiology*, *90*(2), 82–100.
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.
Gangarossa, G., Espallergues, J., Mailly, P., De Bundel, D., de Kerchove d'Exaerde, A., Hervé, D., … Krieger, P. (2013). Spatial distribution of D1R- and D2R-expressing medium-sized spiny neurons differs along the rostro-caudal axis of the mouse dorsal striatum. *Frontiers in Neural Circuits*, *7*(July), 1–16.
Gantz, S. C., Robinson, B. G., Buck, D. C., Bunzow, J. R., Neve, R. L., Williams, J. T., & Neve, K. A. (2015). Distinct regulation of dopamine D2S and D2L autoreceptor signaling by calcium. *eLife*, *4*, 1–19.
Garris, P. A., Budygin, E. A., Phillips, P. E. M., Venton, B. J., Robinson, D. L., Bergstrom, B. P., … Wightman, R. M. (2003). A role for presynaptic mechanisms in the actions of nomifensine and haloperidol. *Neuroscience*, *118*(3), 819–829.
Gaspar, P., Bloch, B., & Le Moine, C. (1995). D1 and D2 receptor gene expression in the rat frontal cortex: Cellular localization in different classes of efferent neurons. *European Journal of Neuroscience*, *7*(5), 1050–1063.
Gerfen, C. R. (1992). The neostriatal mosaic: multiple levels of compartmental organization. *Trends in Neurosciences*, *15*(4), 133–139.
Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of Striatal Projection Systems by Dopamine. *Annual Review of Neuroscience*, *34*(1), 441–466.
Gershman, S. J. (2014). Dopamine ramps are a consequence of reward prediction errors. *Neural Computation*, *26*(3), 467–471.
Geyer, M. A., Russo, P. V., Segal, D. S., & Kuczenski, R. (1987). Effects of

apomorphine and amphetamine on patterns of locomotor and investigatory behavior in rats. *Pharmacology, Biochemistry and Behavior*, *28*(3), 393–399.
Ghods-Sharifi, S., & Floresco, S. B. (2010). Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behavioral Neuroscience*, *124*(2), 179–191.
Giros, B., Sokoloff, P., Martres, M.-P., Riou, J.-F., Emorine, L. J., & Schwartz, J.-C. (1989). Alternative splicing directs the expression of two D2 dopamine receptor isoforms. *Nature*, *342*(21), 923–926.
Glickstein, S. B., DeSteno, D. A., Hof, P. R., & Schmauss, C. (2005). Mice lacking dopamine D2 and D3 receptors exhibit differential activation of prefrontal cortical neurons during tasks requiring attention. *Cerebral Cortex*, *15*(7), 1016–1024.
Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement 3), 15647–15654.
Glowinski, J., Chéramy, A., Romo, R., & Barbeito, L. (1988). Presynaptic regulation of dopaminergic transmission in the striatum. *Cellular and Molecular Neurobiology*, *8*(1), 7–17.
Gobert, A., Rivet, J. M., Lejeune, F., Newman-Tancredi, A., Adhumeau-Auclair, A., Nicolas, J. P., … Millan, M. J. (2000). Serotonin 2C receptors tonically suppress the activity of mesocortical dopaminergic and adrenergic, but not serotonergic, pathways: A combined dialysis and electrophysiological analysis in the rat. *Synapse*, *36*(3), 205–221.
Goldman-Rakic, P. S. (1998). Th ecortical dopamine system: role in memory and cognition. *Advances in Pharmacology*, *42*, 707–711.
Goto, Y., & Grace, A. a. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, *8*(6), 805–812.
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.
Granon, S., Passetti, F., Thomas, K. L., Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2000). Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *The Journal of Neuroscience*, *20*(3), 1208–1215.
Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the Behavioral Sciences*.
Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
Griffon, N., Pilon, C., Sautel, F., Schwartz, J.-C., & Sokoloff, P. (1996). Antipsychotics with inverse agonist activity at the dopamine D3 receptor. *Journal of Neural Transmission*, *103*(10), 1163–1175.
Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Action controls dopaminergic enhancement of reward representations. *Proceedings of the National Academy of Sciences, 109*(19), 7511-7516.
Hall, H., Kohler, C., & Gawell, L. (1985). Some in vitro receptor binding properties of |3H|eticlopride, a novel substituted benzamide, selective for dopamine-D2 receptors in the rat brain. *European Journal of Pharmacology*, *111*(2), 191–199.

Hallett, P. J., Spoelgen, R., Hyman, B. T., Standaert, D. G., & Dunah, A. W. (2006). Dopamine D1 Activation Potentiates Striatal NMDA Receptors by Tyrosine Phosphorylation-Dependent Subunit Trafficking. *Journal of Neuroscience*, *26*(17), 4690–4700.

Haluk, D. M., & Floresco, S. B. (2009). Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology*, *34*(8), 2041–2052.

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Weele, C. M. Vander, … Berke, J. D. (2016). Mesolimbic dopamine signals the value of work, *Nature Neuroscience, 19*(1), 117-125.

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*, 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. *The Journal of Neuroscience*, *34*(3), 698–704.

Hassani, O.-K., Francois, C., Yelnik, J., & Feger, J. (1997). Evidence for a dopaminergic innervation of the subthakmic nucleus in the rat. *Brain Research*, *993*(96), 88–94.

Haverkamp, N., & Beauducel, A. (2017). Violation of the sphericity assumption and its effect on type-I error rates in repeated measures ANOVA and multi-level linear models (MLM). *Frontiers in Psychology*, *8*, 1841.

Hayes, D. J., Clements, R., & Greenshaw, A. J. (2009). Effects of systemic and intra-nucleus accumbens 5-HT2C receptor compounds on ventral tegmental area self-stimulation thresholds in rats. *Psychopharmacology*, *203*(3), 579–88.

Heien, M. L. A. V, Johnson, M. A., & Wightman, R. M. (2004). Resolving neurotransmitters detected by fast-scan cyclic voltammetry. *Analytical Chemistry*, *76*(19), 5697–5704.

Heien, M. L. A. V, Phillips, P. E. M., Stuber, G. D., Seipel, A. T., & Wightman, R. M. (2003). Overoxidation of carbon-fiber microelectrodes enhances dopamine adsorption and increases sensitivity. *The Analyst*, *128*(12), 1413–1419.

Herve, D., Simon, H., Blanc, G., Lisoprawski, A., Le Moal, M., Glowinski, J., & Tassin, J. P. (1979). Increased utilization of dopamine in the nucleus accumbens but not in the cerebral cortex after dorsal raphe lesion in the rat. *Neuroscience Letters*, *15*, 127–133.

Higgins, G. A., Enderlin, M., Haman, M., & Fletcher, P. J. (2003). The 5-HT2A receptor antagonist M100,907 attenuates motor and "impulsive-type" behaviours produced by NMDA receptor antagonism. *Psychopharmacology*, *170*(3), 309–319.

Hikida, T., Kimura, K., Wada, N., Funabiki, K., & Nakanishi Shigetada, S. (2010). Distinct Roles of Synaptic Transmission in Direct and Indirect Striatal Pathways to Reward and Aversive Behavior. *Neuron*, *66*(6), 896–907.

Hiranita, T., & Collins, G. T. (2015). Differential roles for dopamine D1-like and D2-like receptors in mediating the reinforcing effects of cocaine: Convergent evidence from pharmacological and genetic studies. *Journal of Alcoholism and Drug Dependence*, *3*(4), 1922–2013.

Hollon, N. G., Arnold, M. M., Gan, J. O., Walton, M. E., & Phillips, P. E. M. (2014). Dopamine-associated cached values are not sufficient as the basis for action selection. *Proceedings of the National Academy of Sciences*, *111*(51), 18357–18362.

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

Hosking, J. G., Floresco, S. B., & Winstanley, C. a. (2014). Dopamine Antagonism Decreases Willingness to Expend Physical, But Not Cognitive, Effort: A Comparison of Two Rodent Cost/Benefit Decision-Making Tasks. *Neuropsychopharmacology*, *40*(4), 1005–1015.

Hotte, M., Naudon, L., & Jay, T. M. (2005). Modulation of recognition and temporal order memory retrieval by dopamine D1 receptor in rats. *Neurobiology of Learning and Memory*, *84*(2), 85–92.

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*, 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*, *500*(7464), 575–579.

Hoyer, D., Hannon, J. P., & Martin, G. R. (2002). Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacology Biochemistry and Behavior*, *71*(4), 533–554.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational and Behavioral Statistics*, *1*(1), 69–82.

Ikemoto, S. (2002). Ventral striatal anatomy of locomotor activity induced by cocaine,-amphetamine, dopamine and D1/D2 agonists1. *Neuroscience*, *113*(4), 939–955.

Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Research Reviews*, *56*(1), 27–78.

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior. *Brain Research Reviews*, *31*, 6–41.

Iversen, S. D., & Iversen, L. L. (2007). Dopamine: 50 years in perspective. *Trends in Neurosciences*, *30*(5), 188–193.

Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, *466*(7305), 457–462.

Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: New anatommical and computational perspective. *Neural Networks*, *15*, 535–547.

Jupp, B., & Dalley, J. W. (2014). Convergent pharmacological mechanisms in impulsivity and addiction: insights from rodent models. *British Journal of Pharmacology*, *171*(20), 4729–4766.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of Neural Science*

(4th Editio). McGraw-Hill Medical.
Kato, A., & Morita, K. (2016). Forgetting in reinforcement learning links sustained dopamine signals to motivation. *PLoS Computational Biology*, *12*(10): e1005145.
Kebabian, J. W., & Calne, D. B. (1979). Multiple receptors for dopamine. *Nature*, *277*, 93–96.
Keeler, J. F., Pretsell, D. O., & Robbins, T. W. (2014). Functional implications of dopamine D1 vs D2 receptors: A "Prepare and Select" model of the striatal direct vs. indirect pathways. *Neuroscience*, *12*(282), 156–175.
Kemp, J. M., & Powell, T. P. (1971). The structure of the caudate nucleus of the cat: Light and electron microscopy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *262*, 255–265.
Kennett, G. a., Wood, M. D., Bright, F., Trail, B., Riley, G., Holland, V., … Blackburn, T. P. (1997). SE 242084, a selective and brain penetrant 5-HT(2C) receptor antagonist. *Neuropharmacology*, *36*(4–5), 609–620.
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.
Koch, M., Schmid, A., & Schnitzler, H. U. (2000). Role of nucleus accumbens dopamine D1and D2receptors in instrumental and pavlovian paradigms of conditioned reward. *Psychopharmacology*, *152*(1), 67–73.
Koffarnus, M. N., Newman, A. H., Grundt, P., Rice, K. C., & Woods, J. H. (2011). Effects of selective dopaminergic compounds on a delay-discounting task. *Behavioural Pharmacology*, *22*(4), 300–311.
Koob, G. F., Riley, S. J., Smith, S. C., & Robbins, T. W. (1978). Effects of 6-hydroxydopamine lesions of the nucleus accumbens septi and olfactory tubercle on feeding, locomotor activity, and amphetamine anorexia in the rat. *Journal of Comparative and Physiological Psychology*, *92*(5), 917–927.
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*, *466*(7306), 622–626.
Kravitz, A. V., & Kreitzer, A. C. (2012). Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology*, *27*(3), 167–177.
Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818.
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*, *18*(9), 1230–1233.
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*, *111*(6), 2343–2348.
Lammel, S., Hetzel, A., Häckel, O., Jones, I., Liss, B., & Roeper, J. (2008). Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System. *Neuron*, *57*(5), 760–773.
Lau, B., Monteiro, T., & Paton, J. J. (2017). The many worlds hypothesis of dopamine prediction error: implications of a parallel circuit architecture in the basal

ganglia. *Current Opinion in Neurobiology*, *46*, 241–247.
Leff, S. E., & Creese, I. (1985). Interactions of dopaminergic agonists and antagonists with dopaminergic D3 binding sites in rat striatum. Evidence that [3H]dopamine can label a high affinity agonist-binding state of the D1 dopamine receptor. *Molecular Pharmacology*, *27*(2), 184–192.
Levant, B., Grigoriadis, D. E., & DeSouza, E. B. (1992). Characterization of [3H]quinpirole binding to D2-like dopamine receptors in rat brain. *The Journal of Pharmacology and Experimental Therapeutics*, *262*(3), 929–35.
Leventhal, D. K., Stoetzner, C., Abraham, R., Pettibone, J., DeMarco, K., & Berke, J. D. (2014). Dissociable effects of dopamine on learning and performance within sensorimotor striatum. *Basal Ganglia*, *4*(2), 43–54.
Lex, A., & Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learning and Memory*, *15*(7), 483–491.
Lex, B., & Hauber, W. (2010). The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning. *Neurobiology of Learning and Memory*, *93*(2), 283–290.
Li, Y., Zhong, W., Wang, D., Feng, Q., Liu, Z., Zhou, J., … Luo, M. (2016). Serotonin neurons in the dorsal raphe nucleus encode reward signals. *Nature Communications*, *7*, 10503.
Lindvall, O., & Björklund, A. (1979). Dopaminergic innervation of the globus pallidus by collaterals from the nigrostriatal pathway. *Brain Research*, *172*(1), 169–173.
Liu, S., Bubar, M. J., Lanfranco, M. F., Hillman, G. R., & Cunningham, K. A. (2007). Serotonin 2C receptor localization in GABA neurons of the rat medial prefrontal cortex: Implications for understanding the neurobiology of addiction. *Neuroscience*, *146*(4), 1677–1688.
Lloyd, K., & Dayan, P. (2015). Tamping Ramping: Algorithmic, Implementational, and Computational Explanations of Phasic Dopamine Signals in the Accumbens. *PLoS Computational Biology*, *11*(12), 1–34.
Lodge, D. J., & Grace, A. A. (2006). The hippocampus modulates dopamine neuron responsivity by regulating the intensity of phasic neuron activation. *Neuropsychopharmacology*, *31*(7), 1356–1361.
Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the Ability to Inhibit Simple and Choice Reaction-Time Responses - a Model and a Method. *Journal of Experimental Psychology-Human Perception and Performance*, *10*(2), 276–291.
Lomanowska, A., Gormley, S., & Szechtman, H. (2004). Presynaptic stimulation and development of locomotor sensitization to the dopamine agonist quinpirole. *Pharmacology Biochemistry and Behavior*, *77*(3), 617–622.
Maier, S. F., & Watkins, L. R. (2005). Stressor controllability and learned helplessness: The roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neuroscience and Biobehavioral Reviews*, *29*(4–5), 829–841.
Marcott, P. F., Gong, S., Donthamsetti, P., Grinnell, S. G., Nelson, M. N., Newman, A. H., … Ford, C. P. (2018). Regional Heterogeneity of D2-Receptor Signaling in the Dorsal Striatum and Nucleus Accumbens. *Neuron*, 575–587.
Marcott, P. F., Mamaligas, A. a., & Ford, C. P. (2014a). Phasic dopamine release

drives rapid activation of striatal D2-receptors. *Neuron*, *84*(1), 164–176.
Marcott, P. F., Mamaligas, A. A., & Ford, C. P. (2014b). Article Phasic Dopamine Release Drives Rapid Activation of Striatal D2-Receptors.
Marin, O., Anderson, S. a, & Rubenstein, J. L. (2000). Origin and molecular specification of striatal interneurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(16), 6063–6076.
Martelle, J. L., & Nader, M. A. (2008). A review of the discovery, pharmacological characterization, and behavioral effects of the dopamine D2-like receptor antagonist eticlopride. *CNS Neuroscience and Therapeutics*, *14*(3), 248–262.
Martin, J. R., Ballard, T. M., & Higgins, G. a. (2002). Influence of the 5-HT2C receptor antagonist, SB-242084, in tests of anxiety. *Pharmacology Biochemistry and Behavior*, *71*(4), 615–625.
Matias, S., Lottem, E., Dugué, G. P., & Mainen, Z. F. (2017). Activity patterns of serotonin neurons underlying cognitive flexibility. *eLife*, *6*, 1–24.
Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, *447*(7148), 1111–1115.
Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841.
Matteo, V. Di, Giovanni, G. Di, Mascio, M. Di, & Esposito, E. (1999). Dopaminergic Transmission in the Mesolimbic System, *38*, 1195–1205.
Mauchley, J. W. (1940). Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to. *The Annals of Mathematical Statistics*, *11*(2), 204–209.
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.
Meck, W. H., & Benson, A. M. (2002). Dissecting the brain's internal clock: how frontal-striatal circuitry keeps time and shifts attention. *Brain and Cognition*, *48*(1), 195–211.
Menegas, W., Babayan, B. M., Uchida, N., & Watabe-uchida, M. (2017). Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice, 1–26.
Meredith, G. E., Agolia, R., Arts, M. P. M., Groenewegen, H. J., & Zahm, D. S. (1992). Morphological differences between projections neurons of the core and shell in the nucleus accumbens of the rat. *Neuroscience*, *50*(1), 149–162.
Michael, D., Travis, E. R., & Wightman, R. M. (1998). Color images for fast-scan CV measurements in biological systems. *Analytical Chemistry*, *70*(17), 586A–592A.
Miguelez, C., Morera-Herreras, T., Torrecilla, M., Ruiz-Ortega, J. A., & Ugedo, L. (2014). Interaction between the 5-HT system and the basal ganglia: functional implication and therapeutic perspective in Parkinson's disease. *Frontiers in Neural Circuits*, *8*(March), 1–9.
Millan, M., Dekeyne, a., & Gobert, a. (1998). Serotonin 2C receptors tonically inhibit dopamine and noradrenaline, but not 5-HT, release in the frontal cortex in vivo. *Neuropharmacology*, *37*, 953–5.
Missale, C., Nash, S. R., Robinson, S. W., Jaber, M., & Caron, M. G. (1998).

Dopamine receptors: from structure to function. *Physiological Reviews*, *78*(1), 189–225.
Miyazaki, K., Miyazaki, K. W., & Doya, K. (2012). The role of serotonin in the regulation of patience and impulsivity. *Molecular Neurobiology*, *45*(2), 213–224.
Mobini, S., Chiang, T. J., Ho, M. Y., Bradshaw, C. M., & Szabadi, E. (2000). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology*, *152*(4), 390–397.
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.
Montagu, K. A. (1957). Catechol compounds in rat tissues and in brains of different animals. *Nature*, *180*, 244–245.
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.
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.
Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience*, *9*(8), 1057–1063.
Murphy, E. R., Robinson, E. S. J., Theobald, D. E. H., Dalley, J. W., & Robbins, T. W. (2008). Contrasting effects of selective lesions of nucleus accumbens core or shell on inhibitory control and amphetamine-induced impulsive behaviour. *European Journal of Neuroscience*, *28*(2), 353–363.
Nakamura, K. (2013). The role of the dorsal raphé nucleus in reward-seeking behavior. *Frontiers in Integrative Neuroscience*, *7*(August), 1–18.
Nakamura, K., Matsumoto, M., & Hikosaka, O. (2008). Reward-Dependent Modulation of Neuronal Activity in the Primate Dorsal Raphe Nucleus. *Journal of Neuroscience*, *28*(20), 5331–5343.
Navailles, S., & De Deurwaerdère, P. (2011). Presynaptic control of serotonin on striatal dopamine function. *Psychopharmacology*, *213*(2–3), 213–242.
Navailles, S., De Deurwaerdère, P., Porras, G., & Spampinato, U. (2004). In Vivo Evidence that 5-HT2C Receptor Antagonist but not Agonist Modulates Cocaine-Induced Dopamine Outflow in the Rat Nucleus Accumbens and Striatum. *Neuropsychopharmacology*, *29*(2), 319–326.
Neumeyer, J. L., Kula, N. S., Bergman, J., & Baldessarini, R. J. (2003). Receptor affinities of dopamine D1 receptor-selective novel phenylbenzazepines. *European Journal of Pharmacology*, *474*(2–3), 137–140.
Nichols, D. E., & Nichols, C. D. (2008). Serotonin Receptors. *Chemical Reviews*, *108*(5), 1614–1641.
Nicola, S. M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology*, *191*(3), 521-550.
Nicola, S. M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *The Journal of Neuroscience : The Official*

Journal of the Society for Neuroscience, *30*(49), 16585–16600.
Nicola, S. M., Taha, S. A., Kim, S. W., & Fields, H. L. (2005). Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. *Neuroscience*, *135*(4), 1025–1033.
Niv, Y. (2013). Dopamine ramps up. *Nature*, *500*, 533–534.
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.
Niv, Y., Duff, M. O., & Dayan, P. D. (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions*, *1*, 1–9.
Nowend, K. L., Arizzi, M., Carlson, B. B., & Salamone, J. D. (2001). D1 or D2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacology Biochemistry and Behavior*, *69*(3–4), 373–382.
Nutt, D., Stahl, S., Blier, P., Drago, F., Zohar, J., & Wilson, S. (2017). Inverse agonists – What do they mean for psychiatry? *European Neuropsychopharmacology*, *27*(1), 87–90.
O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Role of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, *304*(5669), 452–454.
O'Donnell, P., & Grace, A. a. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *The Journal of Neuroscience*, *15*(5), 3622–3639.
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.
Olpe, H. R., & Koella, W. P. (1977). The response of striatal cells upon stimulation of the dorsal and median raphe nuclei. *Brain Research*, *122*(2), 357–360.
Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., … Dudman, J. T. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell*, *162*(6), 1418–1430.
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.
Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., … Witten, I. B. (2016a). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience, 19*(6), 845-854.
Passetti, F., Levita, L., & Robbins, T. W. (2003). Sulpiride alleviates the attentional impairments of rats with medial prefrontal cortex lesions. *Behavioural Brain Research*, *138*, 59–69.
Paterson, N. E., Wetzler, C., Hackett, A., & Hanania, T. (2012). Impulsive action and impulsive choice are mediated by distinct neuropharmacological substrates in rat. *International Journal of Neuropsychopharmacology*, *15*(10), 1473–1487.
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.

Paxinos, G., & Watson, C. (2009). The rat brain in stereotaxic coordinates. Academic Press.
Pazos, A., & Palacios, J. M. (1985). Quantitative autoradiographic mapping of serotonin receptors in the rat brain. I. Serotonin-1 receptors. *Brain Research*, *346*(2), 205–230.
Perreault, M. L., Hasbi, A., O'dowd, B. F., & George, S. R. (2014). Heteromeric dopamine receptor signaling complexes: Emerging neurobiology and disease relevance. *Neuropsychopharmacology*, *39*(1), 156–168.
Perreault, M. L., Hasbi, A., O'Dowd, B. F., & George, S. R. (2011). The Dopamine D1–D2 Receptor Heteromer in Striatal Medium Spiny Neurons: Evidence for a Third Distinct Neuronal Pathway in Basal Ganglia. *Frontiers in Neuroanatomy*, *5*(31), 1–8.
Peyron, C., Luppi, P.-H., Kitahama, K., Fort, P., Hermann, D. M., & Jouvet, M. (1995). Origin of the dopaminergic innervation of the rat dorsal raphe nucleus. *Neuroreport*, *6*(18), 2527–25.
Pezze, M. A., Dalley, J. W., & Robbins, T. W. (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.
Phillips, A. G., Atkinson, L. J., Blackburn, J. R., & Blaha, C. D. (1993). Increased extracellular dopamine in the nucleus accumbens of the rat elicited by a conditional stimulus for food: an electrochemical study. *Journal of Physiology and Pharmacology*, *71*(5–6), 387–393.
Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V, Wightman, R. M., & Carelli, R. M. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature*, *422*(6932), 614–618.
Phillips, P. E. M., Walton, M. E., & Jhou, T. C. (2007). Calculating utility: Preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology*, *191*(3), 483-495.
Phillips, P. E. M., & Wightman, R. M. (2004). Extrasynaptic dopamine and phasic neuronal activity. *Nature Neuroscience*, *7*(3), 199.
Phillipson, O. T., & Griffiths, A. C. (1985). The topographic order of inputs to nucleus accumbens in the rat. *Neuroscience*, *16*(2), 275–296.
Pleuvry, B. J. (2004). Receptors , agonists and antagonists. *Anaesthesia & Intensive Care Medicine*, *5*(10), 350–352.
Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., … Lang, A. E. (2017). Parkinson disease. *Nature Reviews Disease Primers*, *3*, 1–21.
Pompeiano, M., Palacios, J. M., & Mengod, G. (1994). Distribution of the serotonin 5-HT2 receptor family mRNAs: comparison between 5-HT2A and 5-HT2C receptors. *Molecular Brain Research*, *23*(1–2), 163–178.
Pontieri, F. E., Tanda, G., Orzi, F., & Di Chiara, G. (1996). Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. *Nature*, *382*, 255–257.
Popescu, A. T., Zhou, M. R., & Poo, M. (2016). Phasic dopamine release in the medial prefrontal cortex enhances stimulus discrimination. *Proceedings of the National Academy of Sciences*, *113*(22): E3169-76.
Pothuizen, H. H. J., Jongen-Rêlo, A. L., Feldon, J., & Yee, B. K. (2005). Double

dissociation of the effects of selective nucleus accumbens core and shell lesions on impulsive-choice behaviour and salience learning in rats. *European Journal of Neuroscience*, *22*(10), 2605–2616.
Pozzi, L., Acconcia, S., Ceglia, I., Invernizzi, R. W., & Samanin, R. (2002). Stimulation of 5-hydroxytryptamine (5-HT2C) receptors in the ventrotegmental area inhibits stress-induced but not basal dopamine release in the rat prefrontal cortex. *Journal of Neurochemistry*, *82*(1), 93–100.
Prisco, S., & Esposito, E. (1995). Differential-Effects of Acute and Chronic Fluoxetine Administration on the Spontaneous Activity of Dopaminergic-Neurons in the Ventral Tegmental Area. *British Journal of Pharmacology*, *116*(2), 1923–1931.
Puig, M. V., & Gulledge, A. T. (2011). Serotonin and prefrontal cortex function: Neurons, networks, and circuits. *Molecular Neurobiology*, *44*(3), 449–464.
Puig, M. V., Rose, J., Schmidt, R., & Freund, N. (2014). Dopamine modulation of learning and memory in the prefrontal cortex: insights from studies in primates, rodents, and birds. *Frontiers in Neural Circuits*, *8*(93), 1–15.
Puryear, C. B., Kim, J. M., & Mizumori, S. J. Y. (2010). Conjunctive encoding of movement and reward by ventral tegmental area neurons in the freely navigating rodent. *Behavioural Neuroscience*, *142*(2), 234–247.
Rashid, A. J., So, C. H., Kong, M. M. C., Furtak, T., El-ghundi, M., Cheng, R., … George, S. R. (2007). D1 – D2 dopamine receptor heterooligomers with unique pharmacology are coupled to rapid activation of G q / 11 in the striatum. *Proceedings of the National Academy of Sciences*, *104*(2), 654–659.
Ratcliff, R. (1979). Group Reaction-Time Distributions and an Analysis of Distribution Statistics. *Psychological Bulletin*, *86*(3), 446–461.
Richfield, E. K., Penney, J. B., & Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience*, *30*(3), 767–77.
Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In *Mechanisms in Perception and Action* (pp. 494–519).
Rivet, J.-M., Audinot, V., Gobert, A., Peglion, J.-L., & Millan, M. J. (1994). Modulation of mesolimbic dopamine release by the selective dopamine D3 receptor antagonist, (+)-S 14297. *European Journal of Pharmacology*, *265*(3), 175–177.
Robbins, T. W. (2002). The 5-choice serial reaction time task: Behavioural pharmacology and functional neurochemistry. *Psychopharmacology*, *163*(3–4), 362–380.
Robbins, T. W., Cador, M., Taylor, J. R., & Everitt, B. J. (1989). Limbic-striatal interactions in reward-related processes. *Neuroscience and Biobehavioral Reviews*, *13*(2–3), 155–162.
Robbins, T. W., & Everitt, B. J. (1992). Functions of dopamine in the dorsal and ventral striatum. *Seminars in the Neurosciences*, *4*(2), 119–127.
Robbins, T. W., & Everitt, B. J. (1996). Neurobehavioural mechanisms of reward and motivation. *Current Opinion in Neurobiology*, *6*(2), 228–236.
Robbins, T. W., & Everitt, B. J. (2007). A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). *Psychopharmacology*, *191*(3), 433–437.

Robinson, E. S. J., Dalley, J. W., Theobald, D. E. H., Glennon, J. C., Pezze, M. A., Murphy, E. R., & Robbins, T. W. (2008a). Opposing Roles for 5-HT2A and 5-HT2C Receptors in the Nucleus Accumbens on Inhibitory Response Control in the 5-Choice Serial Reaction Time Task. *Neuropsychopharmacology*, *33*(10), 2398–2406.

Robinson, E. S. J., Eagle, D. M., Economidou, D., Theobald, D. E. H., Mar, A. C., Murphy, E. R., … Dalley, J. W. (2009). Behavioural characterisation of high impulsivity on the 5-choice serial reaction time task: Specific deficits in "waiting" versus "stopping." *Behavioural Brain Research*, *196*(2), 310–316.

Robinson, T. E., Jurson, P. A., Bennett, J. A., & Bentgen, K. M. (1988). Persistent sensitization of dopamine neurotransmission in ventral striatum (nucleus accumbens) produced by prior experience with (+)-amphetamine: a microdialysis study in freely moving rats. *Brain Research*, *462*(2), 211–222.

Rodeberg, N. T., Sandberg, S. G., Johnson, J. A., Phillips, P. E. M., & Wightman, R. M. (2017). Hitchhiker's Guide to Voltammetry: Acute and Chronic Electrodes for in vivo Fast-Scan Cyclic Voltammetry. *ACS Chemical Neuroscience, 15*(8), 221-234.

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.

Rogers, R. D., Wong, A., McKinnon, C., & Winstanley, C. A. (2013). Systemic administration of 8-OH-DPAT and eticlopride, but not SCH23390, alters loss-chasing behavior in the rat. *Neuropsychopharmacology*, *38*(6), 1094–1104.

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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(6), 1265–1271.

Rouder, J. N., & Speckman, P. L. (2004). An evlatuation of the {V}incentizing method of forming group-level response time distributions. *Psychonomic Bulletin and Review*, *11*(3), 419–427.

Rueter, L. E., Tecott, L. H., & Blier, P. (2000). In vivo electrophysiological examination of 5-HT2 responses in 5-HT2C receptor mutant mice. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *361*(5), 484–491.

Sadacca, B. F., Jones, J. L., & Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *eLife*, *5*(e13665), 1–13.

Salamone, J. D. (1988). Dopaminergic involvement in activational aspects of motivation: Effects of haloperidol on schedule-induced activity, feeding, and foraging in rats. *Psychobiology*, *16*(3), 196–206.

Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, *137*(1–2), 3–25.

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., Steinpreis, R., McCullough, L., Smith, P., Grebel, D., & Mahan, K. (1991). Haloperidol and nucleus accumbens dopamine depletion suppresses lever pressing for food but increase free food consumption in a novel food-choice procedure. *Psychopharmacology*, *104*, 515–521.

Salgado, S., & Kaplitt, M. G. (2015). The nucleus accumbens: A comprehensive review. *Stereotactic and Functional Neurosurgery*, *93*(2), 75–93.

Santana, N., & Artigas, F. (2017). Expression of Serotonin 2C Receptors in Pyramidal and GABAergic Neurons of Rat Prefrontal Cortex: A Comparison with Striatum. *Cerebral Cortex*, *27*(6), 3125–3139.

Santana, N., Mengod, G., & Artigas, F. (2009). Quantitative analysis of the expression of dopamine D1 and D2 receptors in pyramidal and GABAergic neurons of the rat prefrontal cortex. *Cerebral Cortex*, *19*(4), 849–860.

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

Saunders, B. T., Richard, J. M., Margolis, E. B., & Janak, P. H. (2017). Instantiation of incentive value and movement invigoration by distinct midbrain dopamine circuits. *bioRxiv*, 186502.

Sawaguchi, T., & Goldman-Rakic, P. S. (1994). The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *Journal of Neurophysiology*, *71*(2), 515–528.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is It Really Robust?: Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology*, *6*(4), 147–151.

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

Schultz, W. (2007a). Behavioral dopamine signals. *Trends in Neurosciences*, *30*(5), 203–210.

Schultz, W. (2007b). Multiple dopamine functions at different time courses. *Annual Review of Neuroscience*, *30*(May), 259–88.

Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews. Neuroscience*, *17*(3), 183–95.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *13*(3), 900–13.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*, 1593–1599.

Schultz, W., Ruffieux, A., & Aebischer, P. (1983). The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Experimental Brain Research*, *51*, 377–387.

Seamans, J. K., & Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*, *74*(1), 1–57.

See, R. E., Sorg, B. A., Chapman, M. A., & Kalivas, P. W. (1991).In vivo assessment of release and metabolism of dopamine in the ventrolateral striatum of awake rats

following administration of dopamine D1 and D2 receptor agonists and antagonists. *Neuropharmacology, 30*(12A), 1269-1274.

Sesack, S. R., Carr, D. B., Omelchenko, N., & Pinto, A. (2003). Anatomical Substrates for Glutamate-Dopamine Interactions: Evidence for Specificity of Connections and Extrasynaptic Actions. *Annals of the New York Academy of Sciences*, *1003*, 36–52.

Shi, W. X., Smith, P. L., Pun, C. L., Millet, B., & Bunney, B. S. (1997). D1-D2 interaction in feedback control of midbrain dopamine neurons. *The Journal of Neuroscience*, *17*(20), 7988–94. Retrieved from

Shin, J. H., Kim, D., & Jung, M. W. (2018). Differential coding of reward and movement information in the dorsomedial striatal direct and indirect pathways. *Nature Communications*, *9*(1), 404.

Simon, N. W., Montgomery, K. S., Beas, B. S., Mitchell, M. R., LaSarge, C. L., Mendez, I. A., … Setlow, B. (2011). Dopaminergic Modulation of Risky Decision-Making. *Journal of Neuroscience*, *31*(48), 17460–17470.

Simpson, E. H., Kellendonk, C., Ward, R. D., Richards, V., Lipatova, O., Fairhurst, S., … Balsam, P. D. (2011). Pharmacologic rescue of motivational deficit in an animal model of the negative symptoms of schizophrenia. *Biological Psychiatry*, *69*(10), 928–935.

Sinha, N., Manohar, S., & Husain, M. (2013). Impulsivity and apathy in Parkinson's disease. *Journal of Neuropsychology*, *7*(2), 255–283.

Smith, Y., Bennett, B. D., Bolam, J. P., Parent, A., & Sadikot, A. F. (1994). Synaptic relationships between dopaminergic afferents and cortical or thalamic input in the sensorimotor territory of the striatum in monkey. *Journal of Comparative Neurology*, *344*(1), 1–19.

Soares-Cunha, C., Coimbra, B., David-Pereira, A., Borges, S., Pinto, L., Costa, P., … Rodrigues, A. J. (2016). Activation of D2 dopamine receptor-expressing neurons in the nucleus accumbens increases motivation. *Nature Communications*, *7*(May), 11829.

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

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

Soares, S., Atallah, B. V., & Paton, J. J. (2016). Midbrain dopamine neurons control judgment of time. *Science*, *354*(6317), 1273–1277.

Sokoloff, P., Giros, B., Martres, M.-P., Bouthenet, M.-L., & Schwartz, J.-C. (1990). Molecular cloning and characterization of a novel dopamine receptor (D3) as a target for neuroleptics. *Nature*, *347*, 146–151.

Solanto, M. V, Abikoff, H., Sonuga-Barke, E., Schachar, R., Logan, G. D., Wigal, T., … Turkel, E. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: a supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology*, *29*(3), 215–228.

Soubrie, P. (1986). Reconciling the role of central serotonin neurons in human and

animal behavior. *The Behavioural and Brain Sciences*, *9*, 319–364.
Stamford, J. A., Kruk, Z. L., & Millar, J. (1991). Differential effects of dopamine agonists upon stimulated limbic and striatal dopamine release : in vivo voltammetric data, 45–50.
St Onge, J. R., & Floresco, S. B. (2009). Dopaminergic Modulation of Risk-Based Decision Making. *Neuropsychopharmacology*, *34*(3), 681–697.
Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, *20*(4), 581-589.
Stefanik, M. T., Kupchik, Y. M., Brown, R. M., & Kalivas, P. W. (2013). Optogenetic Evidence That Pallidal Projections, Not Nigral Projections, from the Nucleus Accumbens Core Are Necessary for Reinstating Cocaine Seeking. *Journal of Neuroscience*, *33*(34), 13654–13662.
Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973.
Stephens, D. W. (2008). Decision ecology: foraging and the ecology of animal decision making. *Cognitive, Affective & Behavioral Neuroscience*, *8*(4), 475–484.
Stopper, C. M., Khayambashi, S., & Floresco, S. B. (2013). Receptor-specific modulation of risk-based decision making by nucleus accumbens dopamine. *Neuropsychopharmacology*, *38*(5), 715–728.
Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R., & Floresco, S. B. (2014). Overriding Phasic Dopamine Signals Redirects Action Selection during Risk/Reward Decision Making. *Neuron*, *84*(1), 177–189.
Strange, P. G. (2008). Agonist binding, agonist affinity and agonist efficacy at G protein-coupled receptors. *British Journal of Pharmacology*, *153*(7), 1353–1363.
Suaud-Chagny, M. F., Ponec, J., & Gonon, F. (1991). Presynaptic autoinhibition of the electrically evoked dopamine release studied in the rat olfactory tubercle by in vivo electrochemistry. *Neuroscience*, *45*(3), 641–652.
Sunahara, R. K., Guan, H.-C., O'Dowd, B. F., Seeman, P., Laurier, L. G., Ng, G., … Niznik, H. B. (1991). Cloning of the gene for a human dopamine D5 receptor with higher affinity for dopamine than D1. *Nature*, *350*, 614–619.
Surmeier, D. J., Ding, J., Day, M., Wang, Z., & Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in Neurosciences*, *30*(5), 228-235.
Sutton, R. S., & Barto, A. G. (1998). Reinforcement Learning: An Introduction. MIT Press: Cambridge, Massachusetts.
Swanson, L., & Cowan, W. (1975). A note on the connections and development of the nucleus accumbens. *Brain Research*, *92*, 324–330.
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), 4–9.
Taha, S. A., Nicola, S. M., & Fields, H. L. (2007). Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *Journal of Physiology*, *584*(3), 801–818.
Tai, L.-H., Lee, A. M., Benavidez, N., Bonci, A., & Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in

action value. *Nature Neuroscience*, *15*, 1281-1289.
Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., & Schoenbaum, G. (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. *Neuron*, *95*(6), 1395–1405.e3.
Takahashi, Y. K., Langdon, A. J., Niv, Y., Schoenbaum, G., Takahashi, Y. K., Langdon, A. J., … Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, *91*(1), 182-193.
Takase, L. F., Nogueira, M. I., Baratta, M., Bland, S. T., Watkins, L. R., Maier, S. F., … Jacobs, B. L. (2004). Inescapable shock activates serotonergic neurons in all raphe nuclei of rat. *Behavioural Brain Research*, *153*(1), 233–239.
Tanaka, T., Vincent, S. R., Nomikos, G. G., & Fibiger, H. C. (1992). Effect of Quinine on Autoreceptor-Regulated Dopamine Release in the Rat Striatum. *Journal of Neurochemistry*, *59*(5), 1640–1645.
Tang, L., Todd, D., & O'Malley, L. (1994). Dopamine D2 and D3 receptors inhibit dopamine release. *The Journal of Pharmacology and Experimental Therapeutics*, *270*(2), 475–479.
Taylor, J. R., & Robbins, T. W. (1984). Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. *Psychopharmacology*, *84*(3), 405–412.
Tecuapetla, F., Jin, X., Lima, S. Q., Costa, R. M., Tecuapetla, F., Jin, X., … Costa, R. M. (2016). Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. *Cell, 166*(3), 703–715.
Threlfell, S., Lalic, T., Platt, N. J., Jennings, K. A., Deisseroth, K., & Cragg, S. J. (2012). Striatal dopamine release is triggered by synchronised activity in cholinergic interneurons. *Neuron*, *75*(1), 58-64.
Tobler, P. N., Dickinson, A., & Schultz, W. (2003). Coding of Predicted Reward Omission by Dopamine Neurons in a Conditioned Inhibition Paradigm. *The Journal of Neuroscience*, *23*(32), 10402–10410.
Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human Neural Learning Depends on Reward Prediction Errors in the Blocking Paradigm. *Journal of Neurophysiology*, *95*(1), 301–310.
Tomasella, E., Bechelli, L., Ogando, M. B., Mininni, C., Di Guilmi, M. N., De Fino, F., … Gelman, D. M. (2018). Deletion of dopamine D $_2$ receptors from parvalbumin interneurons in mouse causes schizophrenia-like phenotypes. *Proceedings of the National Academy of Sciences*, *115*(13), 201719897.
Tops, M., Russo, S., Boksem, M. A. S., & Tucker, D. M. (2009). Serotonin: Modulator of a drive to withdraw. *Brain and Cognition*, *71*(3), 427–436.
Tritsch, N. X., & Sabatini, B. L. (2012). Dopaminergic Modulation of Synaptic Transmission in Cortex and Striatum. *Neuron*, *76*(1), 33–50.
Truong, J. G., Newman, A. H., Hanson, G. R., & Fleckenstein, A. E. (2004). Dopamine D2 receptor activation increases vesicular dopamine uptake and redistributes vesicular monoamine transporter-2 protein. *European Journal of Pharmacology*, *504*(1–2), 27–32.
Usiello, A., Baik, J.-H., Rouge-Pomt, F., Picetti, R., Dierich, A., LeMeur, M., … Borrelli, E. (2000). Distinct functions of the two isoforms of dopamine D2 receptors. *Nature*, *408*(6809), 199–203.

Valencia-Torres, L., Olarte-Sánchez, C. M., Lyons, D. J., Georgescu, T., Greenwald-Yarnell, M., Myers, M. G., … Heisler, L. K. (2017). Activation of Ventral Tegmental Area 5-HT2C Receptors Reduces Incentive Motivation. *Neuropsychopharmacology*, *42*(7), 1511–1521.

Van Bockstaele, E. J., Cestari, D. M., & Pickel, V. M. (1994). Synaptic structure and connectivity of serotonin terminals in the ventral tegmental area: potential sites for modulation of mesolimbic dopamine neurons. *Brain Research*, *647*(2), 307–322.

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 Gaalen, M. M., van Koten, R., Schoffelmeer, a. N. M., & Vanderschuren, L. J. M. J. (2006). Critical Involvement of Dopaminergic Neurotransmission in Impulsive Decision Making. *Biological Psychiatry*, *60*(1), 66–73.

Van Tol, H. H. M., Bunzow, J. R., Guan, H.-C., Sunahara, R. K., Seeman, P., Niznik, H. B., & Civelli, O. (1991). Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature*, *350*(5), 610–614.

Vicente, A. M., Galvao-Ferreira, P., Tecuapetla, F., & Costa, R. M. (2016). Direct and indirect dorsolateral striatum pathways reinforce different action strategies. *Current Biology*, *26*(7), R267–R269.

Vidal-Gadea, A. G., & Pierce-Shimomura, J. T. (2012). Conserved role of dopamine in the modulation of behavior. *Communicative & Integrative Biology*, *5*(5), 440–447.

Vogt Weisenhorn, D. M., Giesert, F., & Wurst, W. (2016). Diversity matters – heterogeneity of dopaminergic neurons in the ventral mesencephalon and its relation to Parkinson's Disease. *Journal of Neurochemistry*, *139*, 8–26.

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.

Wade, T. R., De Wit, H., & Richards, J. B. (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology*, *150*(1), 90–101.

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

Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. *Behavioural Brain Research*, *154*(1), 19–30.

Walters, J. R., Bergstrom, D. A., Carlson, J. H., Chase, T. N., & Braun, a R. (1987). D1 dopamine receptor activation required for postsynaptic expression of D2 agonist effects. *Science*, *236*(4802), 719–722.

Walton, M. E., Bannerman, D. M., & Rushworth, M. F. S. (2002). The role of rat medial frontal cortex in effort-based decision making. *The Journal of Neuroscience*, *22*(24), 10996–11003.

Walton, M. E., Groves, J., Jennings, K. A., Croxson, P. L., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2009). Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-

based decision making. *European Journal of Neuroscience*, *29*(8), 1678–1691.
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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(36), 12020–12027.
Wang, M., Roman, G. T., Schultz, K., Jennings, C., & Kennedy, R. T. (2008). Improved temporal resolution for in vivo microdialysis by using segmented flow. *Analytical Chemistry*, *80*(14), 5607–5615.
Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H. H., Ding, J., … Surmeier, D. J. (2006). Dopaminergic Control of Corticostriatal Long-Term Synaptic Depression in Medium Spiny Neurons Is Mediated by Cholinergic Interneurons. *Neuron*, *50*(3), 443–452.
Waselus, M., Galvez, J. P., Valentino, R. J., & Van Bockstaele, E. J. (2006). Differential projections of dorsal raphe nucleus neurons to the lateral septum and striatum. *Journal of Chemical Neuroanatomy*, *31*(4), 233–242.
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., Zhu, L., Ogawa, S. K., Vamanrao, A., & Uchida, N. (2012). Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. *Neuron, 74*(5), 858-873.
Watanabe, M., Cromwell, H. C., Tremblay, L., Hollerman, J. R., Hikosaka, K., & Schultz, W. (2001). Behavioral reactions reflecting differential reward expectations in monkeys. *Experimental Brain Research*, *140*(4), 511–518.
Weiner, R. I., & Ganong, W. F. (1978). Role of brain monoamines and histamine in regulation of anterior pituitary secretion. *Physiological Reviews*, *58*(4), 905–976.
Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*, 475–482.
White, I. M., & Rebec, G. V. (1994). Performance on a lever-release, conditioned avoidance response task involves both dopamine D1 and D2 receptors. *European Journal of Pharmacology*, *253*(1–2), 167–169.
Wickens, J. R., Horvitz, J. C., Costa, R. M., & Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(31), 8181–8183.
Winstanley, C. A., Cocker, P. J., & Rogers, R. D. (2011). Dopamine modulates reward expectancy during performance of a slot machine task in rats: evidence for a "near-miss" effect. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *36*(5), 913–925.
Winstanley, C. A., Dalley, J. W., Theobald, D. E. H., & Robbins, T. W. (2004). Fractioning impulsivity: Contrasting effects of central 5-HT depletion on different measures of impulsive behaviour. *Neuropsychopharmacology*, *29*(7), 1331–1343.
Winstanley, C. A., Theobald, D. E. H., Dalley, J. W., Glennon, J. C., & Robbins, T. W. (2004). 5-HT2A and 5-HT2C receptor antagonists have opposing effects on a measure of impulsivity: interactions with global 5-HT depletion. *Psychopharmacology*, *176*(3–4), 376–385.
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. (1978a). Catecholamine theories of reward: A critical review. *Brain Research*, *152*(2), 215–247.
Wise, R. A. (1980). The dopamine synapse and the notion of "pleasure centers" in the brain. *Trends in Neurosciences*, *3*(C), 91–95.
Wise, R. A. (1982). Neuroleptics and operant behavior: The anhedonia hypothesis. *Behavioral and Brain Sciences*, *5*(1), 39–53.
Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494.
Witten, I. B., Lin, S.-C., Brodsky, M., Prakash, R., Diester, I., Anikeeva, P., … Deisseroth, K. (2010). Cholinergic interneuron control local circuit activity and cocaine conditioning. *Science*, *330*(6011), 1677–1681.
Wolf, M. E., & Roth, R. H. (1990). Autoreceptor regulation of dopamine synthesis. *Annals of the New York Academy of Sciences*, *604*(1), 323–342.
Wood, J., Simon, N. W., Koerner, F. S., Kass, R. E., & Moghaddam, B. (2017). Networks of VTA Neurons Encode Real-Time Information about Uncertain Numbers of Actions Executed to Earn a Reward. *Frontiers in Behavioral Neuroscience*, *11*(140).
Wright, D. E., Seroogy, K. B., Lundgren, K. H., Davis, B. M., & Jennes, L. (1995). Comparative localization of serotonin 1A, 1C, and 2 receptor subtype mRNAs in rat brain. *The Journal of Comparative Neurology*, *351*(3), 357–373.
Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: Role in drug addiction. *Neuroscience*, *301*, 529–541.
Yapo, C., Nair, A. G., Clement, L., Castro, L. R., Hellgren Kotaleski, J., & Vincent, P. (2017). Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. *Journal of Physiology*, *595*(24), 7451–7475.
Yilmaz, M., & Meister, M. (2013). Rapid innate defensive responses of mice to looming visual stimuli. *Current Biology, 23*(20), 2011-2015.
Yohn, S. E., Santerre, J. L., Nunes, E. J., Kozak, R., Podurgiel, S. J., Correa, M., & Salamone, J. D. (2015). The role of dopamine D1 receptor transmission in effort-related choice behavior: Effects of D1 agonists. *Pharmacology Biochemistry and Behavior*, *135*, 217-226.
Yokel, R. A., & Wise, R. A. (1975). Increased lever pressing for amphetamine after pimozide in rats: implications for a dopamine theory of reward. *Science*, *187*(4176), 547–9.
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.
Zahm, D. S., & Brog, J. S. (1992). On the significance of subterritories in the "accumbens" part of the rat ventral striatum. *Neuroscience*, *50*(4), 751–767.

Zahm, D. S., & Heimer, L. (1990). TwoTranspallidalPathwaysOriinatingin the Rat Nucleus Accumbens. *Journal of Comparative Neurology*, *302*, 437–446.
Zenon, A., Devesse, S., & Olivier, E. (2016). Dopamine Manipulation Affects Response Vigor Independently of Opportunity Cost TL - 36. *Journal of Neuroscience*, *36 VN-r*(37), 9516–9525.
Zetterström, T., Sharp, T., & Ungerstedt, U. (1986). Effect of dopamine D-1 and D-2 receptor selective drugs on dopamine release and metabolism in rat striatum in vivo. *Naunyn-Schmiedeberg's Arch. Pharmacol.*, *334*, 117–124.
Zetterstrom, T., & Ungerstedt, U. (1984). European Journal of Pharmacology, 97 (1984) 29-36, *97*,