# Elucidating the neural correlates of cost-benefit decisions in a rat cortico-basal ganglia network




Dr Oliver Härmson

Somerville College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity 2022

# Acknowledgements

While the months spent in windowless rooms listening to monotonous reward cues and mopping up rodent droppings while my rats were making value-guided decisions have certainly left their mark on me, I look back at years in this D.Phil. program with fond memories. While I can recollect most of the experimental procedures only in monochrome tones with blur around the edges, my experience was also punctuated by moments of sheer bliss. I can vividly recall the first time that one of my pilot rats (oh011) finally seemed to realise the underlying structure of the apparatus and paradigm that I had spent years setting up. I also remember the feeling of immeasurable happiness as I successfully concluded my first-ever implantation of a powerful multi-site microelectrode device.

None of this would have been possible without the people around me who persistently inspired and supported me. First and foremost, I would like to thank (in no particular order) my magnificent supervisors Mark Walton and Andrew Sharott. Both have unfalteringly answered my very naive questions and have offered moral and material support along the entire way. Andrew is a friend and a mentor who has been there for me with encouragement and intellectual rigour every day of the D.Phil. Further, he has helped me to a soft landing in the tricky landscape of electrophysiology. Mark has complemented this with particular detail and consideration for the behavioural side of the experiments, as well as inspiring me with his beautiful style of scientific writing.

In fact, me and Mark go back all the way to 2015 when he kindly hosted me in his lab for a year. For many years he gently withstood my limping efforts to write up a paper from our first experiment together, teaching me a wealth regarding scientific enquiry, data analysis and publishing. For those same reasons I also need to thank my mentor Masud Husain, who hosted me in 2015 alongside Mark and who was instrumental to me finding my match with neuroscience. Laura Grima, with whom we worked closely in 2015/16 and subsequently co-authored 2 papers has been both a friend and a mentor and I miss all of our banter both at Oxford and abroad dearly. Many of my ideas since 2015 time have also been shaped by another brilliant scientist who was closely affiliated with both Mark and Masud - Sanjay Manohar. Sanjay, thank you for inspiring and mentoring me, and helping me find neat ways of summarising complex data in presentations ("give me the one-liner!").

My success has also heavily built on the collective intellect of Andrew's and Mark's labs. Thomas Akam has helped me reach several Eureka moments relating to the interpretation of my data. His analytically apt mind is the cornerstone of many of the statistical approaches that I have subsequently developed in my work to date. Colin McNamara and Max Rothwell spent many years carefully designing and testing the multisite micro-electrode drive that I subsequently repurposed and used. Both of them also laid the foundation to my transfer to the Python programming language. Moreover, Max was a close companion throughout my first year of D.Phil. and it is thanks to him that I acquired many of the workshop and surgical skills that my later years heavily

depended on. Importantly, his holistic approach to what constitutes a healthy work-life balance has been a source of inspiration for me throughout my D.Phil. years. Like Max, Isaac Grennan has been an invaluable companion throughout the years, but has also helped me both collect some of the data for this work as well as refine my statistical approaches. And as if this were not enough, Isaac has also helped me reach new heights on the bouldering wall during our climbing sessions together. Brook Perry also helped me collect some of the data reported in the experiments here, but has also been very resourceful throughout my journey in data analysis and more generally great company to be around. Robert Toth was instrumental to me being able to set up the controller and data acquisition systems for both versions of our behavioural paradigm. However, his lessons on electronics and our late evening exchanges in the MRC BNDU electronics workshop and dimly lit Oxford streets will far outlive the months we spend developing these systems.

Life outside the lab (though brief) was also very transformative for my D.Phil. experience in Oxford. A particularly vivid series of memories was created at the Walton lab outings across various locations in England and Wales. Thank you, Thomas Akam, Laura Grima, Marta Blanco Pozo, Marios Panayi, Lauren Burgeno, Tom JP, TB2, Veronika Samborska, Mason Silveira, Freya Marijatta and George Jenkins for setting up the stage for some truly incredible lab experiences. In addition to those named above from Andrew's lab and the MCR BNDU at large, I would like to thank Yangfan Peng, Julia Sala-Bayo and Joram van Rheede for their sobering presence and companionship. In 2018/19 I also spent a considerable amount of time grinding the gears of rowing machines and paddling in boats with a great group of people from Somerville college, including Christopher Whiteman, Joel Summerfield, David Ahern, Viktor Birkefeldt, Toby Sanderson, Alex Nash and Edwin Silverthorne. Their humorous banter, punctuated by shouts like "shall we do some rowing?" before setting off to paddle have rewarded me with generous amounts of laughter and good spirits.

Finally, my life throughout most of the D.Phil. was transformed by one person in particular - Muniza Ahmed Siddiqui. Despite having to spend an unjust amount of time listening to my musings about rats and their brains, she nevertheless stayed at my side, offering me moral and emotional support, inspiration and love from two different continents. Marrying her and concluding this D.Phil. - both of which have happened simultaneously - have been some of the most exhilarating and rewarding experiences of my life. I am also sincerely grateful for the love and nurture that my mother Margit, my grandmother Malle, my sister Emma and step-father Raimo have offered me throughout the years.

This work was funded by The Clarendon Fund and Mary Somerville Clarendon Graduate Scholarship in conjunction with the Department of Experimental Psychology.

# Elucidating the neural correlates of cost-benefit decisions in a rat cortico-basal ganglia network

Dr Oliver Härmson
Somerville College
University of Oxford

A thesis submitted for the degree of
*Doctor of Philosophy*
Trinity 2022

## Abstract

Adaptive value-guided decision-making often requires weighing-up of costs and benefits of pursuing an available opportunity. Several brain areas, particularly in frontal-striatal circuits, have been reported to be important for this behaviour. However, little is currently known regarding how the decision variables are represented and evolve on the single cell and population levels across these circuits within the confines of a single behavioural paradigm. We developed a novel rodent cost-benefit "accept-reject" task, in which food-restricted rats choose whether or not to run to the end of a corridor to collect sucrose pellets based on the prospective reward (4 different levels, varied trial-by-trial, signalled by auditory cues) and effort cost (3 different levels, varied over blocks of trials, implemented with barriers placed in the corridor that needed to be scaled to reach the reward magazine). Behavioural data (n=12 rats, average 1119 trials per animal) demonstrate a positive effect of reward and a negative influence of effort on the likelihood of accepting an offer, without an interaction between the two. Outcome devaluation strongly reduces acceptance of any offer. A subset of rats (n=5) were also implanted with a bespoke driveable micro-electrode array targeting, anterior cingulate cortex (ACC), medial orbitofrontal cortex (mOFC), dorsomedial striatum (DMS), ventral pallidum (VP), and subthalamic area (STA) and electrophysiological recordings were collected as they performed the task. Analysis of spiking data (n=4 rats; average 119 cells per region per animal) suggests that individual neurons in the 3 targeted basal ganglia areas (DMS, VP and STA) encode reward and/or, to a lesser extent, effort with high fidelity during cue presentation, while neurons in the two cortical areas (ACC and mOFC) display weaker tuning to reward/effort across multiple task points. By contrast, decision is progressively signalled by mOFC, DMS and VP units after action initiation. Decoding analyses using all the units from individual areas in a given session (ensembles) revealed that reward is signalled with similar spatial and temporal characteristics to single units. By contrast, effort is represented by DMS ensembles in a time-distributed manner and is dependent on diverse underlying neuronal activity profiles. These results are consistent with a parallel and distributed system for processing cost-benefit decision variables in the frontal cortical-basal ganglia network.

# *Table of Contents*


*Table of Contents* ..... *6*

*1. Introduction* ..... *8*

**1 Overview of goal-directed behaviour** ..... **10**
1.1 What is goal-oriented behaviour ..... 10
1.1.1 Objects of goal-directed behaviour: rewards ..... 10
1.1.2 Objects of goal-directed behaviour: adversities ..... 13
1.1.3 Inference of net value ..... 17
1.2 The phases of goal-directed behaviour ..... 18
1.2 Derangements in mammalian goal-oriented behaviour ..... 22
1.3 Summary ..... 25

**2 Neural substrates of cost-benefit decision-making in mammals** ..... **26**
2.1 Lessons from neurology and psychiatry: caring for those that no longer care ..... 26
2.3 Anatomy and function of the frontal cortical basal ganglia circuit ..... 28
2.4 Cost-benefit decision-making and the brain ..... 43
2.4.1 From anatomy to function ..... 43
2.4.2 Indexing into nervous tissue activity ..... 44
2.4.3 Identifying value signals in the brain ..... 45
2.4.4 Cost-benefit decision-making in the frontal cortical basal ganglia circuit: divide and conquer? ..... 47
2.4.5 Cost-benefit decision-making in the frontal cortical basal ganglia circuit: everything, everywhere, all at once? ..... 52
2.5 Summary ..... 53

**3 Towards a brain network-level understanding of cost-benefit decision-making** ..... **54**

*2. Setting up the stage for the examination of cost-benefit decision-making in rats* ..... *56*
ABSTRACT ..... 56

**1. Introduction** ..... **57**

**2. Methods** ..... **59**
2.1. Animals ..... 59
2.2 Apparatus ..... 59
2.3 Behavioural task ..... 61
2.3.1 Testing ..... 62
2.3.2 Habituation and training ..... 65
2.4 Outcome devaluation ..... 69
2.5 Microwire electrode implant preparation ..... 70
2.6 Microwire electrode implant surgery ..... 74
2.7 Histology ..... 74
2.8 Data acquisition and storage ..... 74
2.9 Data pre-processing and spike sorting ..... 75
2.10 Data analyses ..... 76

**3. Results** ..... **78**
3.1 Paradigm v1 ..... 78
3.2 Paradigm v2 ..... 81
3.3 Paradigm v2 behaviour is resilient to electrophysiology-related modifications ..... 84
3.4 Electrophysiological signal quality ..... 85

**4. Discussion** ..... **87**

*3. Single-unit correlates of cost-benefit decision-making in the frontal cortical-basal ganglia network ..... 91*
ABSTRACT ..... 91
**1. Introduction ..... 92**
**2. Methods ..... 93**
2.1. Anatomical data analysis ..... 93
2.2 Neurophysiological data analysis ..... 94
**3. Results ..... 97**
3.1. Anatomy ..... 97
3.2. Overview of neuronal yield ..... 98
3.3. Firing rate relationships to task variables in identified individual units ..... 99
**4. Discussion ..... 109**
*4. Ensemble-level correlates of cost-benefit decision-making in the frontal cortical-basal ganglia network ..... 112*
ABSTRACT ..... 112
**1. Introduction ..... 114**
**2. Methods ..... 115**
2.1 Data pre-processing and selection ..... 115
2.2 Decoder models ..... 115
2.3 Statistical tests ..... 116
2.4 Data depiction ..... 116
**3. Results ..... 117**
3.1 Reward and effort are robustly coded on the population level ..... 117
3.2 The relationship between ensemble size and decoding accuracy ..... 119
3.3 Time-distributed multiplexing of reward and effort in DMS ensembles ..... 122
**4. Discussion ..... 123**
*5. General discussion ..... 125*
ABSTRACT ..... 125
**1. A novel cost-benefit decision-making task ..... 125**
**2. Neural signatures of cost-benefit decision-making in a frontal cortical-basal ganglia network ..... 128**
**3. Conclusions and future directions ..... 131**
*Bibliography ..... 133*

# 1. Introduction

What makes us get out of bed in the morning? This may seem like a trivial question *prima facie*. Perhaps it is the aromatic and creamy cup of coffee that we can make ourselves next, or the promise of an exciting day at work. But what sort of drive propels us toward these goals? And how do we figure out that getting out of bed is the way to achieve them? Would it not be just as sensible to stay in bed and wait for the coffee to come to us?

In the background, our brains are busily whirring away to define and measure our need for these goals, and in the process, drive the behaviours that seem most productive to achieve them. With billions of years of evolution engrained in its architecture, it has already defined for us the primary needs which help us stay alive and reproduce. And where there is a will, the brain is there to help us figure out *a way*.

The "way" has, for a long time, eluded neuroscientists and clinicians. We seem particularly skilled at figuring out appropriate means of achieving their goals, and in the process balancing the effort spent on acquisition (the costs) with the rate of the return (the benefits). We know that the coffee will probably not come to us itself and we should not wait until the end of our time until it does, but should rather put in the effort to make it. Further, if goals change, or if the previous means of achieving them ceases to yield, we flexibly change course. It would not be adaptive for us to get stuck in a loop of making coffee for ourselves the whole day. Focal brain lesions in the frontal cortical-basal ganglia circuit can, however, dramatically change the flexibility of our behaviour, sometimes making it hard for us to get out of bed, at other times leaving us destined to unnecessary courses of action.

These observations have spurred a quest for the underlying brain mechanisms of cost-benefit decision-making. One fundamental assumption has driven this quest - that the brain must represent the variables needed to optimally carry out goal-driven behaviours. This thesis will not attempt to prove the validity of this assumption nor elucidate the neural mechanisms of goal-directed behaviour in their entirety. However, it will aim to shed light on the computation dynamics of some of the primary decision variables across a selection of brain areas that are crucial for cost-benefit decision-

making. To help with this, I will first provide an overview of the characteristics and determinants of goal-directed behaviour, crucial for understanding the contents of all chapters in this thesis. Second, I will outline some of the current views on the underlying brain mechanisms.

# 1 Overview of goal-directed behaviour

## 1.1 What is goal-oriented behaviour

Goal-oriented behaviour is a set of repertoires implemented by organisms to bring them closer to their goals. Goals revolve around the maximisation of desiderata, often referred to as rewards, the minimisation of costs, or the maintenance of a favourable *status quo*. Understanding these endpoints of behaviour is essential, which is why I will individually dissect each. To understand how an organism's perception of these endpoints impacts behaviour, I will later also discuss the various phases of goal-directed behaviour. I will focus mainly on rodent research as it is of specific relevance to this thesis.

### 1.1.1 Objects of goal-directed behaviour: rewards

Rewards are outcomes that increase the likelihood of organisms producing the behaviour that leads to them (Schultz 2006; Seymour et al. 2007). They are commonly categorised as *primary* - rewards which satisfy a physiologic need and do not need conditioning to acquire value, e.g. food - or *secondary* - outcomes which acquire value through pairing with primary rewards, e.g. money.

Ideas about what makes rewards rewarding have historically been varied. One intuitive set of ideas grounded motivation within a framework of homeostatic drive reduction (Berridge 2004). For example, one school of thought conceptualised motivation as a function of an organism's proximity to a setpoint (Cannon 1980). According to this framework, organisms monitor the difference between the level of a homeostatic variable and its set-point (or settling point, Wirtshafter and Davis 1977), and are driven

to take corrective action (such as eating) if the discrepancy is large enough. This framework predicts the presence of a setpoint, a deviance detector and a corrective mechanism in the brain. An alternative theory likened the drive to pursue reward to a hydraulic process, where the build-up of pressure exerted by a physiologic need, such as thirst to preserve circulating blood volume (Lorenz and Leyhausen 1973), would motivate organisms to take corrective action once the pressure reaches a certain threshold.

Parsimonious as the drive reduction theories were, they were not sufficient to answer to some puzzling experimental findings, such as why dogs that were parenterally fed the full daily calorie requirement persisted in eating their normal meals by mouth (Turner et al. 1975) and why rats preferred to drink milk rather than have it fed through an intra-gastric fistula (Miller and Kessen 1952). More importantly, the finding that hypothalamic electrical stimulation sites that evoked intra-cranial self-stimulation - presumably rewarding by virtue of its ability to reduce drive - also elicited eating and drinking - i.e. the very behaviours that reflected increased drive - posed a serious challenge to the framework (Valenstein et al. 1970).

However, the emphasis on drives as the origins of motivation gradually waned. Instead, the idea that rewards are rewarding by virtue of their incentive expectancies gained momentum (Bolles 1972). This line of thought also helped explain why organisms develop preferences for conditioned stimuli (CS), e.g. levers and house-lights, which in and of themselves do not replenish any physiologic needs: CS acquire some of the hedonic properties of the unconditioned stimuli (US). In an effort to re-integrate the theory with physiology, organisms' internal states were posited to modulate the incentive value of the US (Toates 1986).

It is helpful to not only think about what we would do for reward, but what reward can do for us. One noteworthy account in the field proposed that rewards elicit both "wanting" - an incentive motivational effect that makes organisms actively seek rewards - and "liking" - a hedonic effect resulting an interaction with the reward (Robinson and Berridge 1993; Berridge and Robinson 1998). In normal goal-directed behaviour,

"wanting" is congruent with "liking" and animals' drive to work reduces concomitantly with decreases in goal value (Seymour et al. 2007), however they can be dissociable in pathological brain states (Berridge 2004).

"Wanting" can be measured through the expression of actions directed toward the goal. Such *goal-oriented* behaviour relies on the integration between action-outcome association and the value of the outcome (Balleine and Dickinson 1998a). Were the value of the outcome decrease and organisms are relying on it to guide outcome-oriented actions, then in principle the rate and intensity of these actions should decrease. This forms the rationale for the test of outcome-specific devaluation, where organisms are typically allowed unconstrained access to the outcome, followed by measurement of their willingness to engage with instrumental action to extract it (Balleine and Dickinson 1998b). A related test of non-specific outcome devaluation, where animals are fed to satiety with food that is not instrumentally obtainable in the subsequent testing phase, will be discussed in chapter 2.

How do organisms judge the value of rewards? In a situation with no associated costs, the multi-attribute utility model framework characterises the evaluation as consisting of a summation of the value functions of each attribute of a reward (Tversky and Simonson 1993). For example, when evaluating the utility of an apple, we may consider the sum of value functions pertaining to sweetness, acidity, water content, etc. Despite the intuitive appeal of this model, it fails to account for context effects (e.g. attraction, compromise and similarity biases that violate the assumption of independence of irrelevant options; Louie et al. 2013; Tajima et al. 2019) and heuristic-based decision-making, e.g. the "priority heuristic" (Brandstätter et al. 2006; Hayden and Niv 2021a).

When it comes to the use of rewards as reinforcers in neuroscience, research in rodents has commonly used either liquid sucrose or solid sucrose/chow food rewards in operant tasks, often in combination with food restriction to enhance palatability. Other reward types include ethanol, drugs of addiction, intra-cranial self-stimulation and access to a conspecific of the opposite sex. Solid rewards usually come in the form of individual pellets of a few dozens of grams, helpful for the discretisation of reward sizes. Liquids,

on the other hand, are measured out by volume, allowing fine gradations of reward magnitude. There is good evidence that rodents keep track of the volume/count of these different reward types, however the exact level of discrimination between small differences in numbers of pellets is not clear (M. Walton, personal communication).

### 1.1.2 Objects of goal-directed behaviour: adversities

When Faust asked Mephistopheles "Well now, who are you then?", Mephistopheles replied: "Part of that force that always wills the evil and always produces the good" (Goethe 1862). The good and the bad often go hand in hand. Similarly, many scenarios feature a trade-off between rewards and adversities. Take, for example, the monetary cost associated with an expensive holiday, or the effort associated with building an athletic body.

Adversities come in two flavours: costs and punishments. Though these facets are addressed in mainly two separate streams of neuroscience literature, they overlap by virtue of organisms' propensity to avoid both. However, I acknowledge that the affective nature of, on one extreme, foot-shocks, and on the other, a large lever press requirement for reward, are very different, with objectively quantifiable reactions elicited in the former but not the latter (Jean-Richard-Dit-Bressel et al. 2018). To some degree this difference might be driven by helplessness in the context of many punishment-related experimental paradigms. Fundamentally, effort requirements that are demanding enough (e.g. hundreds of presses on a lever for a rodent) might have a considerable overlap with punishment, as both involve a degree of suffering.

Crucially for this dissertation, there is one key difference between the two concerning the way they have been studied. Whereas punishments are usually delivered as an endpoint and sometimes administered in a setting where an animal can work to avoid them (Seymour et al. 2007; Jean-Richard-Dit-Bressel et al. 2018), costs are mostly presented *en route* to reward. Cost-related brain signals may thus evolve with different temporal dynamics to punishments and, by virtue of their distance to rewards, may take

on a different shape. Costs may take various forms: physical effort, cognitive effort, time and risk. In the following, I will separately address each.

Manipulations of physical effort alter the balance between the allocation versus acquisition of energy that is thought to lie at the heart of reward-guided pursuit (Stephens and Krebs 1986). Converging evidence suggests that mammals ranging from primates (Walton et al. 2006; Varazzani et al. 2015; Klein-Flügge et al. 2016; Le Heron et al. 2018) to laboratory rodents (Salamone et al. 1991, 1993; Hillman and Bilkey 2010; Cowen et al. 2012; Holec et al. 2014) are able to discriminate between levels of experimentally defined effort when pursuing rewards. Physical effort is commonly manipulated either in the form of a force needs to be exerted (Ishiwari et al. 2004; Zénon et al. 2016a; Chong et al. 2017; Le Heron et al. 2018; Porter and Hillman 2019), a scalable barrier height (Salamone et al. 1994; Walton et al. 2002, 2003; Schweimer and Hauber 2006), a response rate (Tanno and Silberberg 2014) or amount (Hodos 1961; Salamone et al. 1991, 1993; Gan et al. 2010) requirement. For example, the progressive ratio task challenges subjects with an increasing number of required responses (e.g. lever presses in rodents, keyboard presses in humans) for reward until the primary outcome - a "breaking point" - is reached (fig. 1b). Though time - another cost - is frequently correlated with response difficulty especially in tasks that involve repeated responses, there is evidence to suggest that subjects are primarily concerned with evaluating effort rather than time cost (Wakabayashi et al. 2004; Mingote et al. 2005; Floresco et al. 2008b). It is currently unclear how much the rodent effort-based paradigms specifically tax calorific vs other reserves (Taylor et al. 1970; Walton and Bouret 2019)

Cognitive challenges, by contrast, are set up to tap into the limited reserves of attention and executive control. In the primate literature, reserves of cognitive effort have commonly been taxed using visuospatial tasks that combined with a variable amount of task-switching, e.g. needing to make assessments about numbers based on two alternating rules (Botvinick et al. 2009; McGuire and Botvinick 2010) or needing to find numbers in streams of distractors (Chong et al. 2017). Human cognitive effort tasks have inspired attempts to implement equivalent rodent paradigms, resulting in the rat cognitive effort task, where rats need to decide between either brief or long (high and

low effort, respectively) spatial cue durations to direct successful reward retrieval (Cocker et al. 2012; Hosking et al. 2014, 2015; Silveira et al. 2020, fig. 1C).

Time cost represents another commonly modelled form of negative motivator in neuroscience studies. This hinges on the ability of time to devalue rewards over time in all tested species (Green and Myerson 2004; Glimcher and Fehr 2014; Hayden 2016), such that immediate reinforcers are perceived to be more valuable than future ones. The procedure for testing this usually involves presenting the participants with a choice between a small immediate and a large delayed option (Berns et al. 2007, fig. 1D).

Time and physical effort are often inextricably linked with risk, echoed by the proverbial saying "a bird in the hand is worth two in the bush". Risk can result either from ambiguity - not knowing the likelihood of different outcomes, which is partially reducible by learning - or unchangeable mathematical properties of the outcome distribution (Knight 2002). The exact nature of risk perception has been conceptualised either as a sensation of outcome-probability pairings or a perception of the statistical moments of the outcome distribution (mean, variance, skewness)(Tobler et al. 2005; Dreher et al. 2006; Burke and Tobler 2011a, b; Glimcher and Fehr 2014) An exemplary format for paradigms of risky decision-making in rodents is one where rats are simultaneously presented with a choice between a small/certain lever which will always deliver reward (e.g. 1 pellet) or a large/risky lever associated with the probabilistic delivery of a large reward (e.g. 4 pellets, St. Onge and Floresco 2009, fig. 1E).

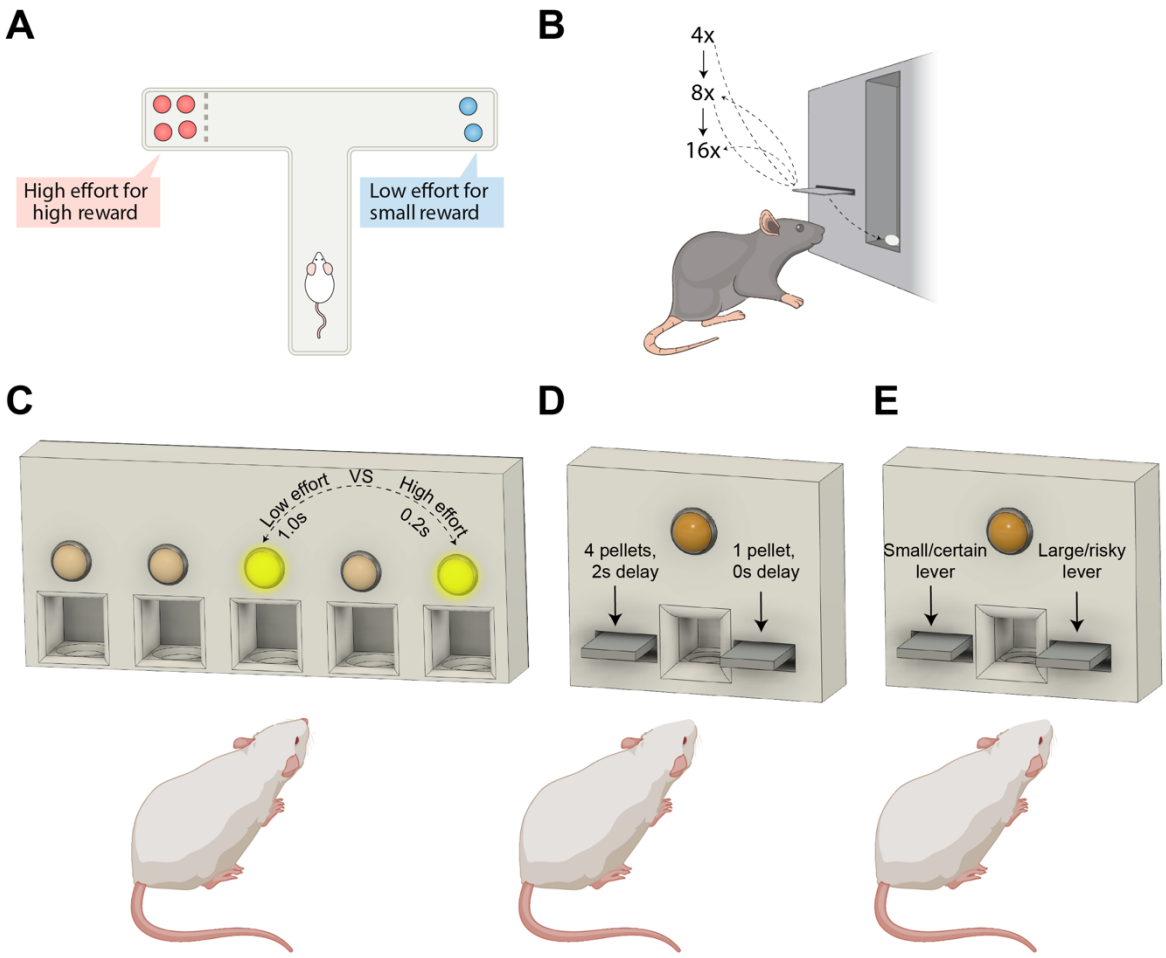



**Figure 1 | Overview of behavioural paradigms involving different forms of cost A,** an effort T-maze paradigm where an easily accessible small reward option is pitted against a large reward behind a barrier that needs to be scaled (Walton et al. 2003) **B,** the progressive ratio paradigm (Der-Avakian et al. 2016), where the lever press amount requirement for reward is steadily increased until a "breaking point" is reached. **C,** The rat cognitive effort task (Cocker et al. 2012), where subjects need to choose between monitoring a 0.2s (high effort) or 1s (high effort) visual stimulus and subsequently retrieve reward from the associated reward dispenser. **D,** a delay discounting paradigm (Berns et al. 2007) where subjects need to choose between an immediately available small vs a delayed large reward. **E,** a risk discounting paradigm (St. Onge and Floresco 2009) where rodents need to make a choice between pressing a lever associated with the deterministic delivery of a small reward vs a lever that delivers large reward on a probabilistic basis.

### 1.1.3 Inference of net value

We have thus far considered rewards and costs in separation. In many decision-making scenarios, organisms must consider both to make adaptive decisions. As effort is usually perceived *en route* to reward, it is conceptualised as something that discounts reward. Several different functions have been proposed to model the effect of physical, cognitive, time and risk discounting of reward. Notably, while delay and probability discount reward in a convex fashion (fig. 2A-B; Green and Myerson 2004), physical and cognitive effort discounting is concave (fig. 2C; Klein-Flügge et al. 2015, 2016; Białaszek et al. 2017). In other words, in contrast to delay and risk costs, small increases in physical and cognitive effort lead to less, whereas large increases lead to progressively more discounting.

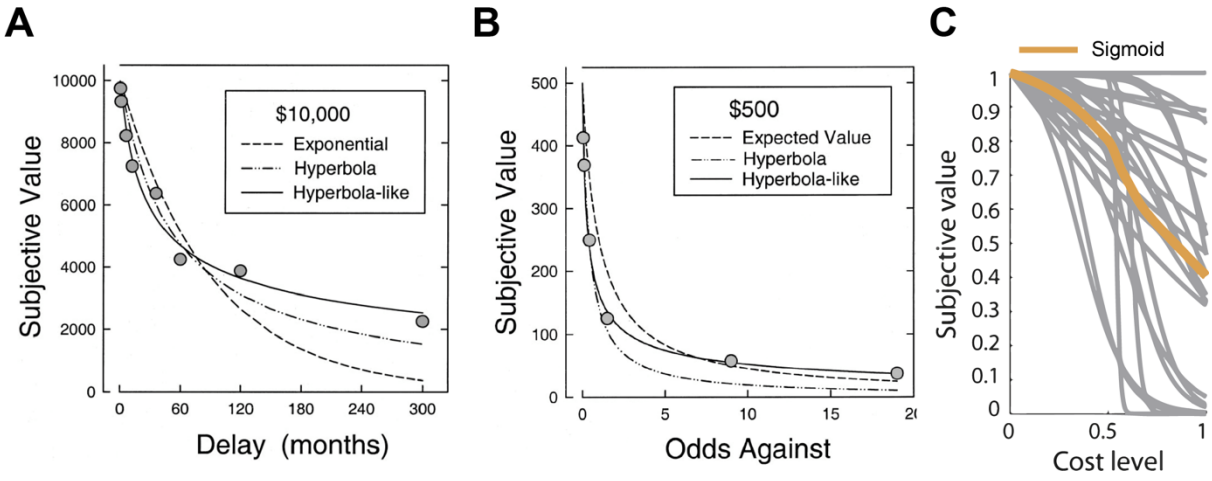



**Fig. 2 | Cost functions are specific to domain**. **A-B,** Different cost functions (lines) and empirical data (dots) are portrayed to exemplify depreciation of the subjective value of money as a function of time until its receipt (left panel) or as a function of the odds against its receipt (right panel). Adapted from Green and Myerson 2004. **C,** The depreciation of monetary offers as a function of physical effort. Individual subject data are depicted in grey lines; the best model of best fit (a sigmoid function) is superimposed as a yellow line. Adapted from Klein-Flügge et al. 2015.

## 1.2 The phases of goal-directed behaviour

The integration of rewards and costs is traditionally thought to be a hallmark of adaptive goal-directed behaviour. However, this is but one step in a sequence of processes leading from the perception of internal state to interaction with the outcome (Kalis et al. 2008; Sinha et al. 2013; Husain and Roiser 2018, fig. 3). Here, I will discuss each of those steps in separation, but acknowledge that they may in fact evolve in parallel.

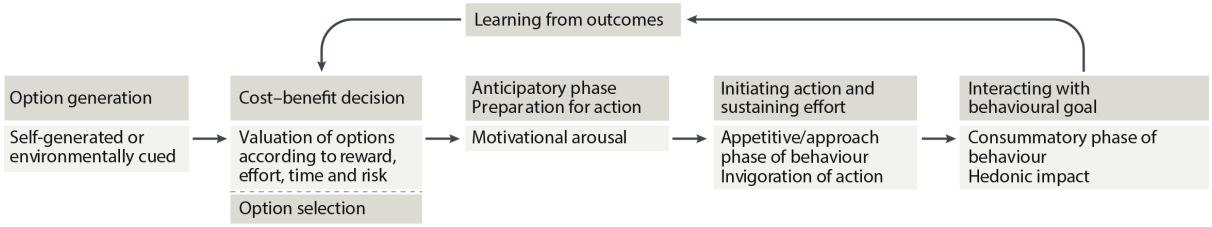



**Figure 3 | The stages of goal-directed action.** Dark boxes indicate the label of the behavioural stage, while light boxes provide mechanistic detail. Adapted from Husain and Roiser 2018.

#### *1.1.3.1 Option generation*

In many scenarios, especially outside over-constrained experimental scenarios, organisms need to first flesh out the alternative ways of reaching a goal, often referred to as option generation (as opposed to generating goals, fig 3; Kalis et al. 2008; Sinha et al. 2013; Husain and Roiser 2018). *Options* have actional and affective properties and the two are psychologically very closely linked (Sebanz et al. 2006). Together, they make up an action-value pair. Depending on the circumstances and the decision-maker's experience, alternatives may be generated anew internally, retrieved from memory, or may be suggested by the situation (Kalis et al. 2008). Option generation is typically part of an assessment of fluency in cognitive screening tests and indexes into executive control (Husain and Roiser 2018). The process of option generation and its neurobiological correlates have seldom been studied, but was recently examined in a report which instructed human participants to draw as many unique paths as possible between two dots (Ang et al. 2018).

#### *1.1.3.2 Cost-benefit decision*

After the options have been laid out, the organism needs to settle on a decision about which one(s) to pursue. Identifying the most valuable option is subject to cost discounting as discussed above, however deviations from normative decision-making may be observed if subjects are influenced by context and heuristics. Further, though having many options might help a rational decision-maker to the best outcome (Von Neumann and Morgenstern 2007), choosing between a large array of options may in itself carry a cognitive cost (Ortoleva 2013), biasing the organism to settle for a smaller (even if suboptimal) subset of the original experimenter-defined alternatives (Salgado 2011).

Value-guided decision-making is a process that has attracted intense modelling efforts. One commonly used mathematical description - logistic regression - bears resemblance to the multi-attribute utility model which is defined as:

$$U_i = \alpha + \sum_{q=1}^{Q} V_{iq}$$

where $V_{iq}$ denotes the value function of the $q$'th attribute of option $i$ (Glimcher and Fehr 2014). In a typical example, the ordinal quantities of currently offered reward and cost, defined by the experimenter, are regressed onto binary choice data to characterise their net effect on decision-making. Variations of such descriptions have also been developed to explain the effects of costs and benefits across time scales. For example, one report found that the inclusion of additional across-trial unrecoverable and recoverable fatigue terms added explanatory power to a null model without those terms in a task that required subjects to squeeze dynamometers with varying force for varied rewards (fig. 4; Müller et al. 2021). Similarly, recent reward experiences, captured as a reinforcement rate over time/trials or the reward outcomes of previous trials often exerts additional influence on current choice (Guitart-Masip et al. 2011; Hamid et al. 2016; Wittmann et al. 2016; Akam et al. 2021).

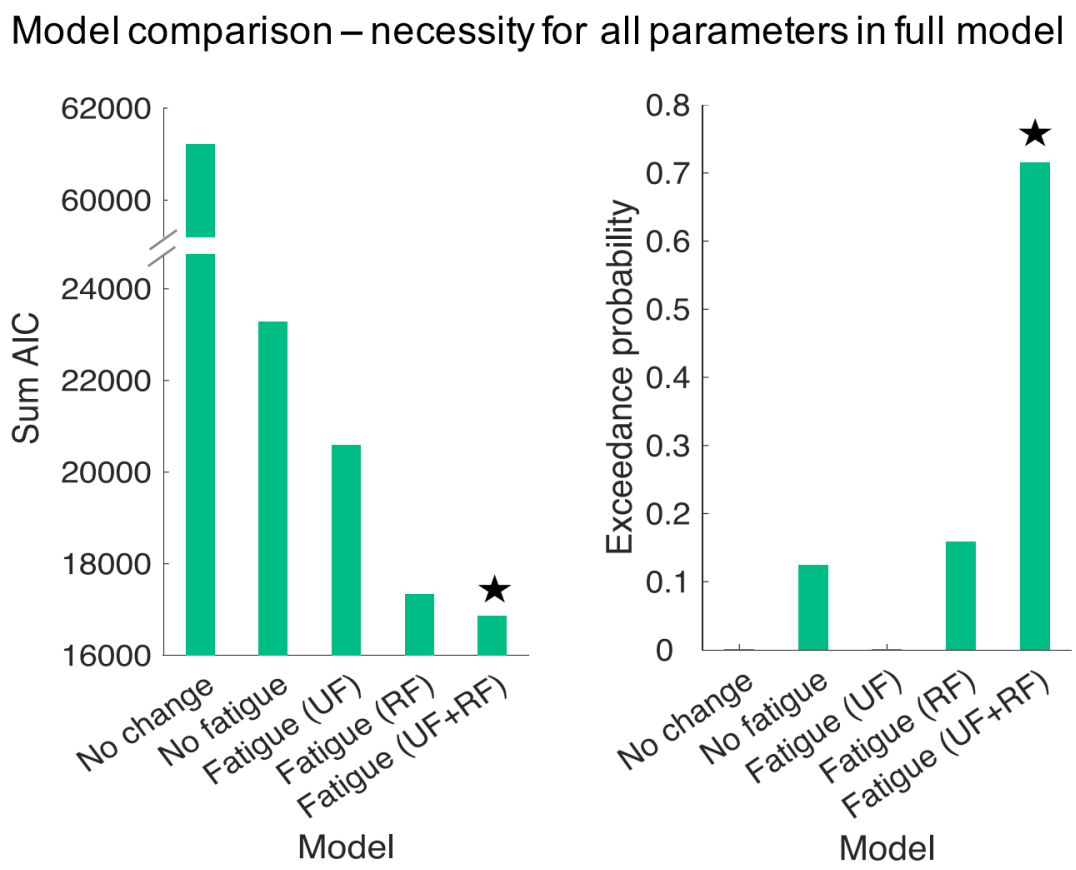



**Figure 4 | Model comparison indicates that the inclusion of fatigue terms improves fit on choice data.** AIC (left) and exceedance probabilities were compared between two null models which assumed static willingness to overcome effort for reward - the no change and no fatigue model - vs models that included a steadily ramping unrecoverable fatigue term, a recoverable fatigue term which accounted for the history of prior rest trials, and a combined fatigue model. Adapted from Müller et al. 2021.

While these models provide a good description of the determinants of trial-by-trial decisions, they do not capture the time-evolving dynamics of the decision process. The diffusion class of models offers a solution that is able to both capture empirical properties of response latencies and choice rates (Ratcliff and McKoon 2008). In this framework, decisions are captured as noisy accumulator processes that drift from a starting point toward one, two or more decision boundaries over time (fig. 5). The drift rate and decision bounds - two free parameters - can be fitted directly onto observed behavioural data (fig. 5B-C; Ratcliff and Tuerlinckx 2002). The model is also able to account for behavioural phenomena that are unrelated to the decision process but

nevertheless affect response latencies, such as stimulus representation, attention, memory retrieval and motoric responses (Glimcher and Fehr 2014).

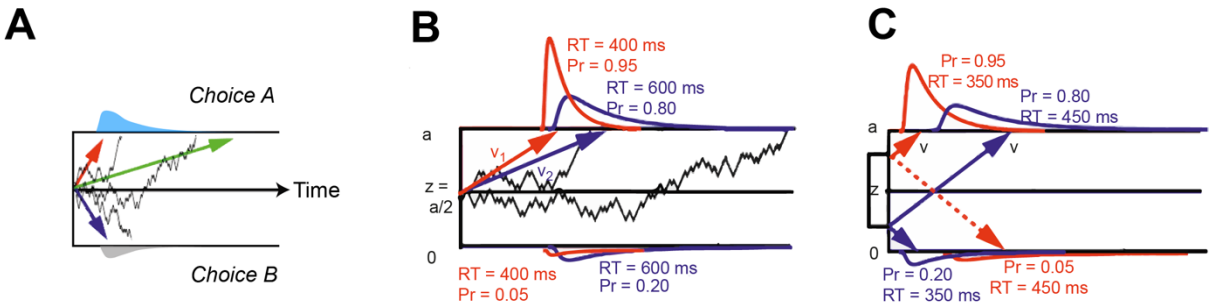



**Figure 5 | Illustration of diffusion models. A,** time-dynamics of a decision between choices A and B over time. Individual thin lines and superimposed coloured arrows indicate diffusion process instances with different drift rates. **B,** Example of the effects of variation in drift rate. The red curve illustrates a high drift rate which produces low-latency responses that favour choice A, whereas a slower drift rate (blue curve) produces slower responses. **C,** Example of variation in starting point. Red curves illustrate a process that starts closer the boundary of choice A, producing faster and more frequent choice of A over B, and vice versa for the blue curve. Adapted from Glimcher and Fehr 2014.

#### *1.1.3.3 Anticipatory phase*

After the decision has been made, individuals typically undergo a phase of preparation for goal-directed action. This preparation can manifest in physiological correlates, such as cardiovascular (Brehm and Self 1989) and pupillary (Muhammed et al. 2016) changes. Increased arousal confers the benefit of readiness for rapid goal-directed action (Ressler 2004).

#### *1.1.3.4 Initiating action and sustaining effort*

The initiation of action marks the onset of appetitive behaviour. Subsequently, the individual needs to be able to sustain reward-guided behaviour until the action requirement is satisfied. Interestingly, action vigour in physical paradigms often tends to be proportionate to the incentives on offer even if the subjects are not reinforced for this relationship (the Crespi effect; Crespi 1942; Berridge 2004). This may fit with a normative account of action vigour allocation in foraging scenarios, whereby increases

in reward expectation offset the balance of the calorific value of the reward vs energetic cost of movement, which subjects may be trying optimise (Niv et al. 2006; Yoon et al. 2018; Shadmehr et al. 2019).

#### *1.1.3.5 Interacting with the behavioural goal*

Successful reward pursuit leads to a consummatory phase where the subject interacts with the goal. This phase overlaps with the concept of "liking" described above. "Liking" can be indexed into on both objective and subjective fronts. Objectively, it can be measured as a set of hedonic reactions (e.g. lip licking vs headshakes as expressions of liking vs disgust, respectively; Treadway et al. 2015). Subjective accounts are commonly obtained through questionnaires in human studies (Rushworth et al. 2011). Consummatory reactions consist of stereotyped oropharyngeal movements (Berridge 2004).

#### *1.1.3.6 Learning from outcomes*

The loop of cost-benefit decision-making is closed by the learning phase. To keep goal-directed behaviour adaptive, both internal and external states need to be frequently re-evaluated. First, subjects need to weigh the outcome value of the behaviour in question against their changing internal signals such as hunger, thirst, social status etc, with resulting corrections in "liking" and "wanting" (Cabanac 1971; Berridge 2004; Treadway et al. 2015; Ottenheimer et al. 2020). Second, if the environmental context is unstable, as it is in naturalistic foraging settings (Sutherland et al. 1988; Stephens 2008) or in a probabilistic reward learning task (Samejima et al. 2005; Mohebi et al. 2019), internal representations of value in relation to external events need to be updated. The framework of reinforcement learning has been used as a biophysically plausible explanation for how outcome values are used for learning in the dopaminergic system (Schultz et al. 1997; Schultz 2016).

## 1.2 Derangements in mammalian goal-oriented behaviour

As is evident from the above discussion, adequate goal-directed behaviour relies on an intricate balance between several psychological processes and, by extension, wide-ranging brain circuitry (discussed below). It is thus easily exposed to derangements leading either to deficiency or excess. As a reflection of this, significant reduction in goal-directed (or motivated) behaviour - known as the clinical syndrome of *apathy* - is a common manifestation of a myriad of neurological and psychiatric diseases (Husain and Roiser 2018). The construct of apathy originates from clinical reports of individuals with ostensibly normal cognition but a striking lack of *spontaneous* behaviours, despite being able to react normally to external stimuli (Cairns et al. 1941; Németh et al. 1988).

One set of criteria defines apathy as a loss or reduction of motivation compared to an individual's prior baseline, with deficits observed in at least two out of three core domains: behavioural (internally or externally initiated), cognitive (loss of spontaneous ideation and curiosity) or emotional (blunted responsiveness to positively/negatively valenced stimuli or diminished spontaneous emotion; Robert et al. 2009). Broadly, four behavioural sub-domains are thought contribute to the apathetic phenotype, helping explain a variety of symptoms in patients (fig. 6; Pagonabarraga et al. 2015). Apathy can affect the stages of option generation (Ang et al. 2018), cost-benefit decision (Le Heron et al. 2018), motivational arousal (Muhammed et al. 2016), and invigoration of ongoing reward-guided action (Le Heron et al. 2018), whereas its impact on reward consumption and learning phases are less clear.

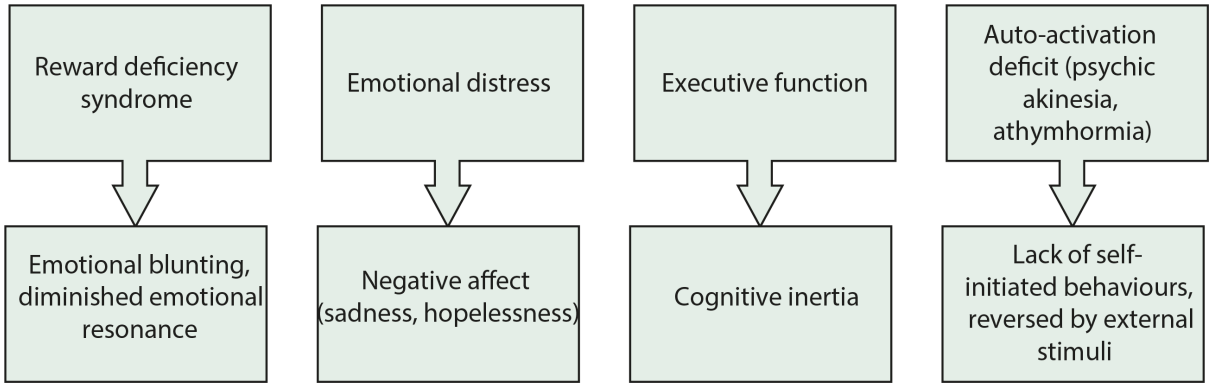



**Figure 6 | The 4 features of an apathetic phenotype and their resulting symptoms in human subjects.** Adapted from Pagonabarraga et al. 2015.

Apathy overlaps to some degree with *anhedonia* or lack of pleasure, defined from a clinical standpoint as a significant loss of pleasure in most daily activities (American Psychiatric Association 2013). Hedonic experiences, which diminish in anhedonia, can co-occur at several stages of reward pursuit, starting from initiation of motivated behaviour until well after the consummatory phase (fig. 7; Thomsen et al. 2015), and are vitally important to sustain goal-directed behaviour. Anhedonia is closely correlated with the clinical concept of *avolition*, characterised by an inability to initiate and sustain motivated behaviour activities (American Psychiatric Association 2013).

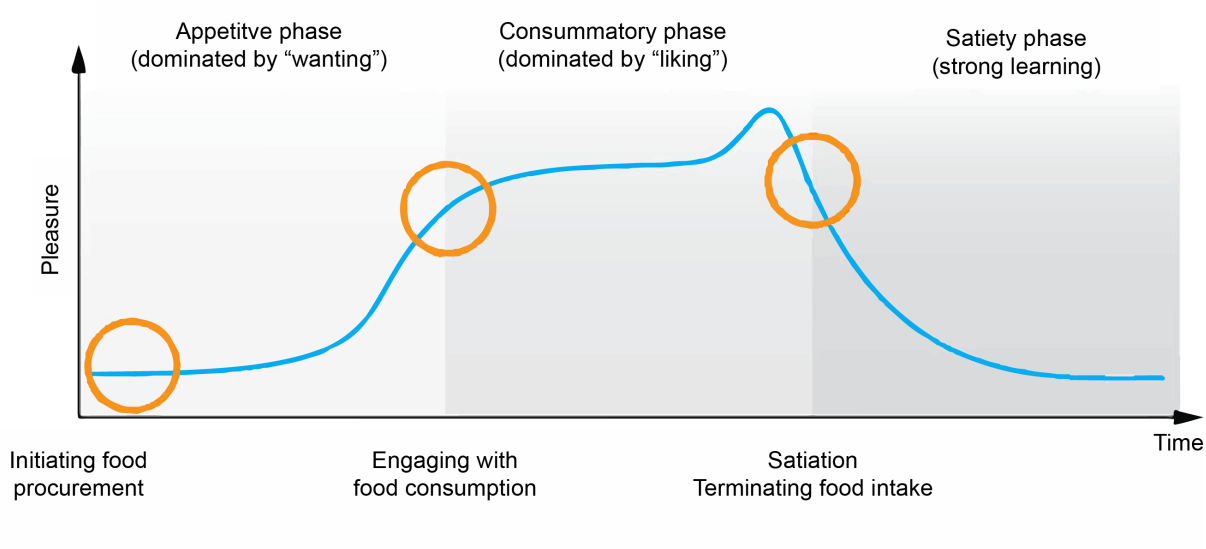



**Figure 7 | Please intensity during goal-directed behaviour.** Hedonic sensation (blue line) increases as a function of proximity to reward consumption. Adapted from Thomsen et al. 2015.

Apathy and anhedonia have typically been screened using interviews and self-report questionnaires (Fawcett et al. 1983; Snaith et al. 1995; Yazbek et al. 2014). In the Snaith-Hamilton pleasure scale, for example, patients need to indicate that they either (1) strongly disagree, (2) disagree (3) agree, or (4) strongly agree with statements like "I would enjoy reading a book, magazine or newspaper" or "I would enjoy looking smart when I have made an effort with my appearance". This poses a problem for translational research. Luckily, objective cross-species indices of apathy and anhedonia have been developed since the parcellation of motivation into the subdomains of "liking", "wanting" and "learning" (Robinson and Berridge 1993). For example, patterns of orofacial grimaces in response to pleasurable/aversive stimuli ("liking"), tendencies to work for reward or work to avoid punishment ("wanting") and learning mechanisms are

largely conserved between rodents and primates (Thomsen 2015; Thomsen et al. 2015; Treadway et al. 2015). As motivational disorders pose a substantial caregiver burden and do not yet have licensed treatments (Pagonabarraga et al. 2015) research into the neurobiological mechanisms of motivated behaviour, discussed in the latter half of the introduction, is especially timely and important.

### 1.3 Summary

In the first half of this chapter, I have provided an overview of the outcomes and conceptual stages of goal-directed behaviour. It can be viewed as a cycle consisting of the realisation of goals and the means to reach them, cost-benefit comparison, outcome anticipation, goal pursuit, consumption, and learning. One of the hallmarks of goal-directed behaviour is its flexibility, whereby changes in outcome value can readily update the overall likelihood and parameters of the behaviour. Many brain disorders are known to lead to dysregulation of goal-directed behaviour, with very disabling outcomes and little prospects of help from the healthcare sector. Understanding the neural mechanisms underlying these behaviours could provide insights for the development of treatments.

## 2 Neural substrates of cost-benefit decision-making in mammals

Since Edouard Brissaud's description of apathetic patients as "speak[ing] no more than they move", "withdrawn to themselves" and "indifferent to everything" over a century ago (Prange et al. 2018), the neural mechanisms of apathy have both fascinated and eluded neurologists. Research to date has revealed that there are several motivation "hotspots" in the brain which can lead to excessive or deficient behavioural drive when lesioned and which signal cost-benefit decision-making variables in health. However, these "hotspots" are also linked together, yielding a loop-like network of areas that co-ordinate to mediate motivation. The second half of the introduction focuses on the anatomy, physiology and function of this circuit.

### 2.1 Lessons from neurology and psychiatry: caring for those that no longer care

Documentation of an organic basis for apathy gathered pace at the end of the 20th century, owing to the spread of advanced imaging techniques such as magnetic resonance tomography. At that time, manifestations on the spectrum of apathy, ranging from akinetic mutism (absence of spontaneous speech and limb movements) to abulia (diminished self-initiated motor, cognitive and emotional behaviours), were reported to have developed in previously normally functioning people in the aftermath of insults to regions such as the caudate nucleus, globus pallidus, putamen, anterior cingulate cortex and the interconnecting white matter (Habib and Poncet 1988; Németh et al. 1988; Trillet et al. 1990; Bhatia and Marsden 1994; Benke et al. 2003). The common denominator to all of these regions is that they were positioned in an interconnected network of frontal cortex and basal ganglia (BG). More recent reports have further corroborated this by uncovering reward incentivisation deficits associated with parts of the ventromedial prefrontal cortex (vmPFC; Manohar and Husain 2016)[1] and globus pallidus (Adam et al. 2013) damage.

[1] But increased reward sensitivity was found in patients with selective damage to the subgenual ventromedial prefrontal cortex.

On the flipside, electrical stimulation of a partly overlapping network in human patients can lead to subjective increases incentive motivation or "wanting" to engage with various activities. In one case series of subjects institutionalized for severe depression, schizophrenia and epilepsy, one patient (B-19) was reported to develop "feelings of pleasure, alertness and warmth (goodwill); he had feelings of sexual arousal and described a compulsion to masturbate" when given the means to self-regulate stimulation of his nucleus accumbens and septal area (NAc; Heath 1972). Deep brain stimulation (DBS) therapy case reports have further provided accounts of e.g. a patient's sudden development of extreme affection for members of her care team following subthalamic nucleus (STN) DBS (Herzog et al. 2003), new-found motivation to engage with hobbies after NAc stimulation (Schlaepfer et al. 2008) and the onset of an anticipation of challenge that the patient was determined to overcome, associated with anterior cingulate cortex (ACC) stimulation (Parvizi et al. 2013).

Incentivisation deficits in humans have largely been mirrored by lesion phenotypes of homologous regions in rodents. For example, quinolinic acid lesions of the ACC, but not the pre-limbic (PrL) or infralimbic (IL) cortex, have consistently led to reductions in rats' choice of a high effort-high reward option over a low effort-low reward option on a T-maze paradigm (described above; Walton et al. 2002, 2003; Schweimer and Hauber 2005). Similar preference changes for low vs high effort options have been documented in association with dopamine depletion in the interconnected NAc (Salamone et al. 1991, 1994) and reductions in instrumental vigour have been demonstrated following lesions in the dorsal medial striatum (DMS; Wang et al. 2013) - another afferent region of the ACC (Mailly et al. 2013).

But lesions do not invariably reduce motivation. On the contrary, lesions of the medial orbito-frontal cortex (mOFC) lead to increased motivation to work for rewards as indexed by the progressive ratio schedule of reinforcement (Münster and Hauber 2018), mirroring clinical findings (Manohar and Husain 2016). This perplexing finding can be reconciled with the view of some nodes of the frontal cortical BG network implicated in dysfunctional motivation acting to *modulate*, rather than directly cause, incentivisation for rewards (Wilson et al. 2014; Manohar and Husain 2016). This raises the question of

what kind of signals and in which brain areas the mOFC is in the position to regulate and how the frontal cortical BG circuit functions at large. To attempt to address this, I will next dissect the anatomy and functional correlates of the frontal cortical BG network, with particular emphasis on a selection of areas with clear clinical and pre-clinical associations with dysfunctional motivation (fig. 8). These areas will also be the focus of study in chapters 2-4.

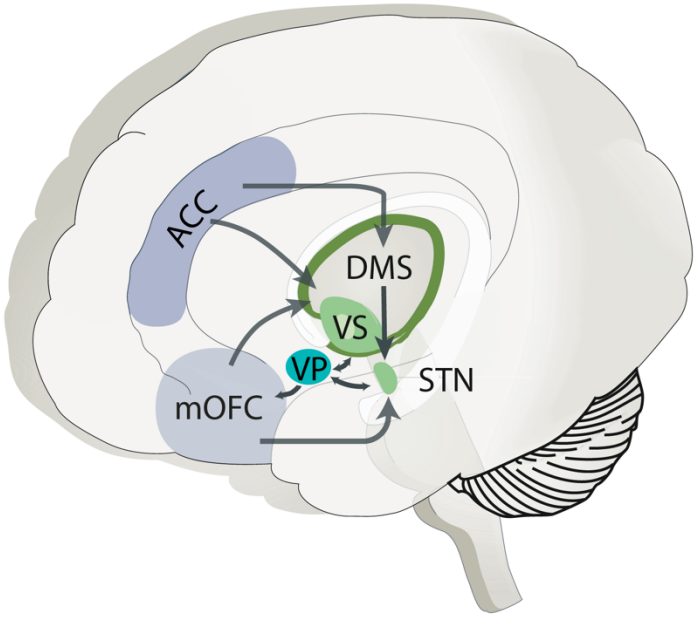



**Figure 8 | Brain regions implicated in dysfunctional motivation.** Frontal cortical regions form a feedback loop through the BG and thalamus. Adapted from Husain and Roiser 2018.

## 2.3 Anatomy and function of the frontal cortical basal ganglia circuit

### 2.3.1 PFC

#### 2.3.1.1 Evidence for discrete subdivisions

Though the parcellation and definition of PFC is a topic of debate (Laubach et al. 2018), divisions can be laid out based on cytoarchitecture, connectivity and gene expression profiles. Cytoarchitectonics leverages the visually appreciable patterns of neuronal spatial arrangement to create divisions in the PFC. One particularly fine-grained report divided the PFC into frontal association area 1 and 2 (Fr1, Fr2, respectively), dorsal or

ventral anterior cingulate cortex (ACd and ACv, respectively), dorsal pre-limbic cortex (PLd), ventral pre-limbic cortex (PLv), infra-limbic cortex (IL), mOFC (also abbreviated as MO), ventral orbitofrontal cortex (VO), posterior ventrolateral orbitofrontal cortex (VLOp), lateral orbitofrontal cortex (LO), dorsal anterior insula 1 and 2 (AId1 and AId2, respectively; fig. 9; Van De Werd et al. 2010; , but see also a review of other parcellations in Le Merre et al. 2021).

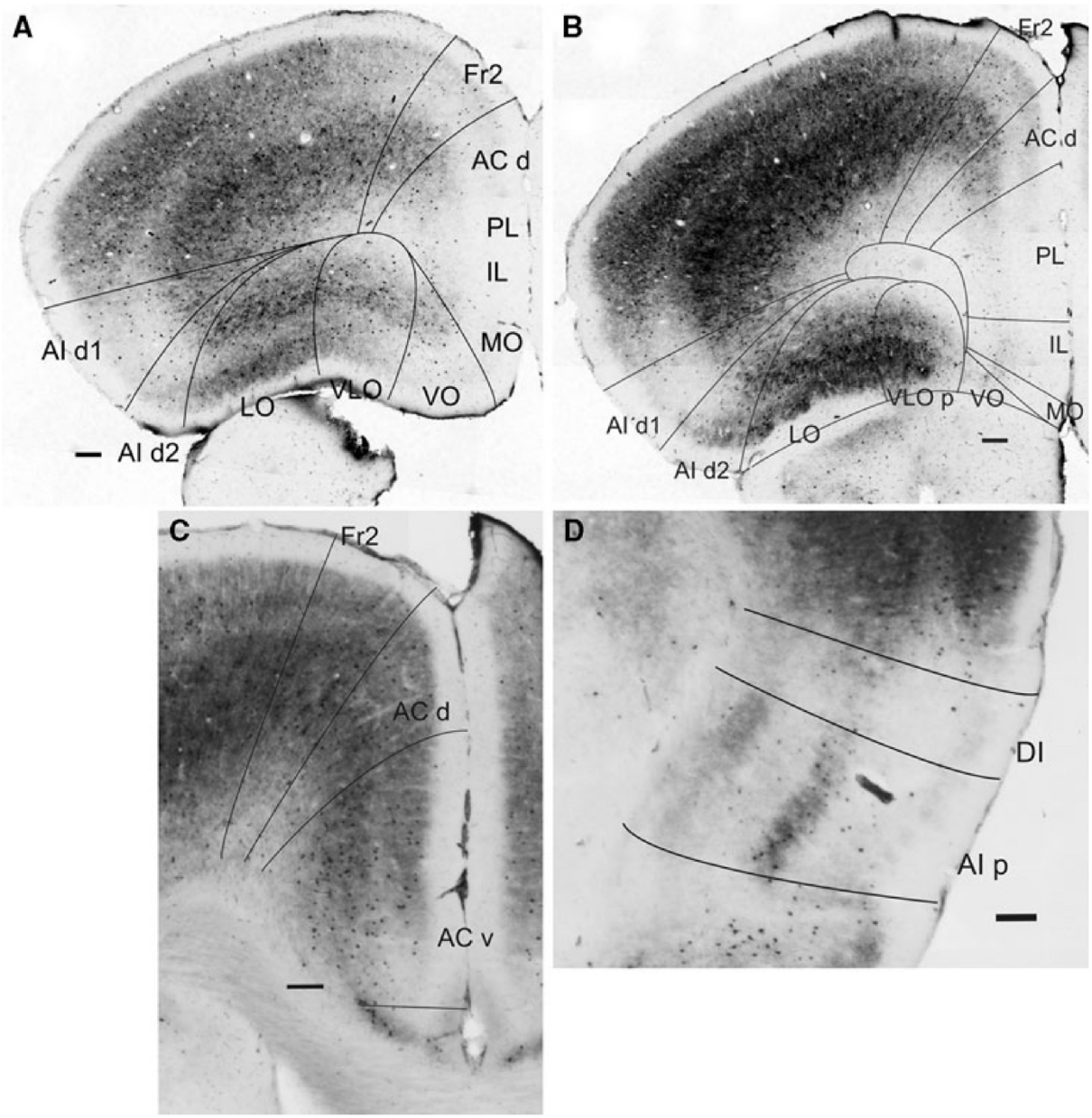



**Figure 9 | An example of a cytoarchitectonic parcellation of the PFC based on parvalbumin staining.** (A-D) parvalbumin-stained sections at differing AP locations. Adapted from Van De Werd et al. 2010.

Cortico-cortical connectivity mapping using anterograde viral tracing, by contrast, has suggested a slightly different picture, revealing the mouse PFC to be composed of the pre-limbic area (PrL), infralimbic area (IL), dorsal and ventral anterior cingulate areas (ACAd and ACAv, respectively), frontal polar cortex (FPC), and orbital areas (lOFC, vlOFC, mOFC for lateral, ventrolateral and medial subdivisions, respectively). Notably, insular areas were not categorised as PFC based on divergent connectivity profiles - a departure from the cytoarchitectonic view (Harris et al. 2019). One meta-analysis that additionally considered subcortical connectivity categorised the mouse PFC into three major subdivisions: the *dorsomedial PFC* (dmPFC), characterised by reciprocal connectivity with sensorimotor cortical and thalamic regions and rostro-dorsal striatum; *ventromedial PFC* (vmPFC), defined by its limbic ventral hippocampal formation inputs and efferents to the ventral tegmental area (VTA), basolateral amygdalar nucleus (BLA), rostral ventromedial striatum, NAc, dorsal raphe nucleus (DRN) and lateral habenula (LHb); and *ventrolateral PFC* (vlPFC), distinguished by its combination of sensory and limbic (but not motor) input (Le Merre et al. 2021; fig 10).

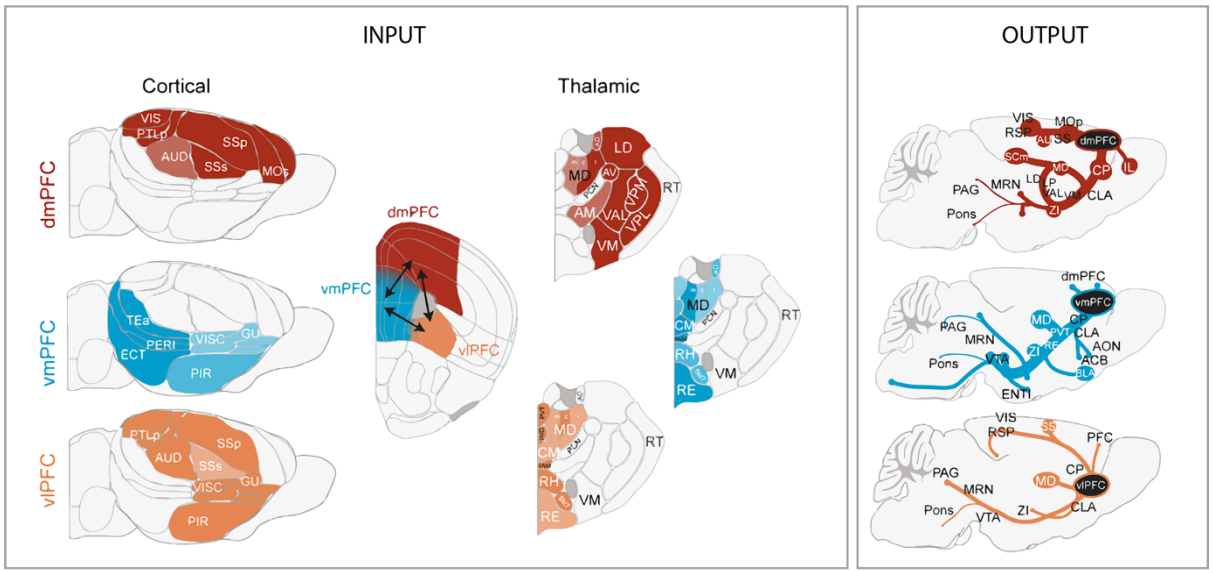



**Figure 10 | PFC parcellation based on cortico-cortical and cortico-subcortical connectivity profiles.** Adapted from Le Merre et al. 2021.

Quite a different rendition has been put forward by a study of the spatial patterns of gene expression (termed spatial transcriptomics; Ortiz et al. 2020). The method used in

this study relies on dimensionality reduction and T-distributed stochastic neighbour embedding approaches to categorise areas based on molecular similarity. By contrast to cytoarchitectonic and cortico-cortical connectivity patterns, genetic profiles were discovered to be shared by several previously conceived PFC sub-territories, putting the findings somewhat at odds with previously accepted PFC boundaries.

These distinct molecular and connectomic patterns should, at least in theory, impart functional segregation within the PFC. Converging evidence in rats does indeed suggest specialisation along the mediolateral (ML), dorsoventral (DV) and anteroposterior (AP) axes. AP dissociations in Pavlovian outcome devaluation and reversal learning have been observed in the lOFC (Panayi and Killcross 2018). Along the DV axis, dissociations have been demonstrated between the OFC and ACC. For example, the latter has been deemed to be necessary to titrate the amount of effort spent in pursuit of reward, whereas the more dorsal PrL and IL are not (Walton et al. 2003; Rudebeck et al. 2006). This is further supported by available evidence of the necessity of lOFC in pavlovian-instrumental transfer - a measure of stimulus-outcome learning - but not instrumental outcome devaluation - reflective of action-outcome learning - in rats (Ostlund and Balleine 2007). Contrasting the OFC along the ML axis, however, has revealed that medial territories are associated with higher necessity for instrumental action-outcome learning (Gardner et al. 2017; Panayi and Killcross 2018).

However, caution must be exercised when it comes to rigorous delineation of the roles of PFC regions in the reviewed behaviours. Clear-cut ML, AP and DV differences are subject to the utilised nomenclature and studied behaviour (reviewed in Wilson et al. 2010; Rushworth et al. 2012; Barreiros et al. 2021). For example, the term PFC has been used rather loosely to describe ACC, PrL and IL, whereas OFC has almost invariably been used to refer to lateral and ventrolateral OFC (Laubach et al. 2018; Le Merre et al. 2021). Further, adaptive decision-making between two concurrently available cost-free options might only require a value representation (possibly implemented in the mOFC) and value comparison (possibly in the ACC), whereas optimal foraging behaviour relies on comparisons the foreground and background reward acquisition rate (Stephens and

Krebs 1986), possibly requiring distinct computations in the PFC circuit (Kolling et al. 2012; Rushworth et al. 2012)

#### 2.3.1.2 Cellular anatomy and physiology

The mouse PFC consists of approximately 82% excitatory, 17% inhibitory and 1% modulatory neurons as defined by NRN1 (excitatory), GAD67 (inhibitory) and dopaminergic/serotonergic/cholinergic marker expression, respectively. Specifically, frontal cortical regions have more prevalent somatostatin-expressing interneurons, compared to the more predominant parvalbumin-expressing interneuron type found in other cortical areas (Erö et al. 2018). Similarly to the rest of the neocortex, these cells can be categorised into 6 cytoarchitectonic layers (L), with the exception of an absent L4 in the mouse PFC (Lein et al. 2007; Tasic et al. 2018).

Though the layering seems a useful means of ordering, an alternative picture has been revealed by considering axonal projection patterns. Based on these grounds, excitatory cells have thus been categorised into three main groups: *intratelencephalic* neurons, characterised by sparse firing, distribution through L2-6 and efferents only within neocortex, striatum, amygdala and claustrum (Harris and Shepherd 2015), *pyramidal tract* neurons, distinguished by their specific distribution in L5B, their (mostly) sub-cerebral projections to areas such as the brainstem, spinal cord, striatum and thalamus and their propensity to either fire in bursts or densely (Molnár and Cheung 2006), and *corticothalamic* neurons, defined by their predominant projection to ipsilateral thalamus, distribution in L6 and very sparse firing (Harris and Shepherd 2015).

Inhibitory neurons, on the other hand, do not easily lend themselves to categorisation based on efferent projections. Instead, one common classification scheme parses them based on three genetic markers: 5-hydroxytryptamine receptor 3A (Htr3a), parvalbumin (PV) and somatostatin (Sst). These cells are receive feedforward input from local excitatory cells, display inter-regulation (e.g. vasoactive intestinal peptide cells - a subclass of Htr3a cells - inhibit Sst and PV cells), send feedback signals to excitatory cells and display either dense firing with little relation to behaviour (Basket cell subclass

of PV neurons) or behavioural tuning (some subclasses of Htr3a and ST cells; Harris and Shepherd 2015).

### 2.3.2 Basal ganglia

#### 2.3.2.1 Regional subdivision

The BG are a set of interconnected subcortical nuclei encompassing the telencephalon, diencephalon and midbrain (Albin et al. 1989; DeLong 1990). The basal ganglia can be categorised as members of the input nuclei (striatum and STN), intrinsic nuclei (globus pallidus externa, GPe; substantia nigra pars compacta; SNc) or output nuclei (globus pallidus interna, GPi; substantia nigra pars reticulata, SNr; fig. 11; (Shipp 2017).

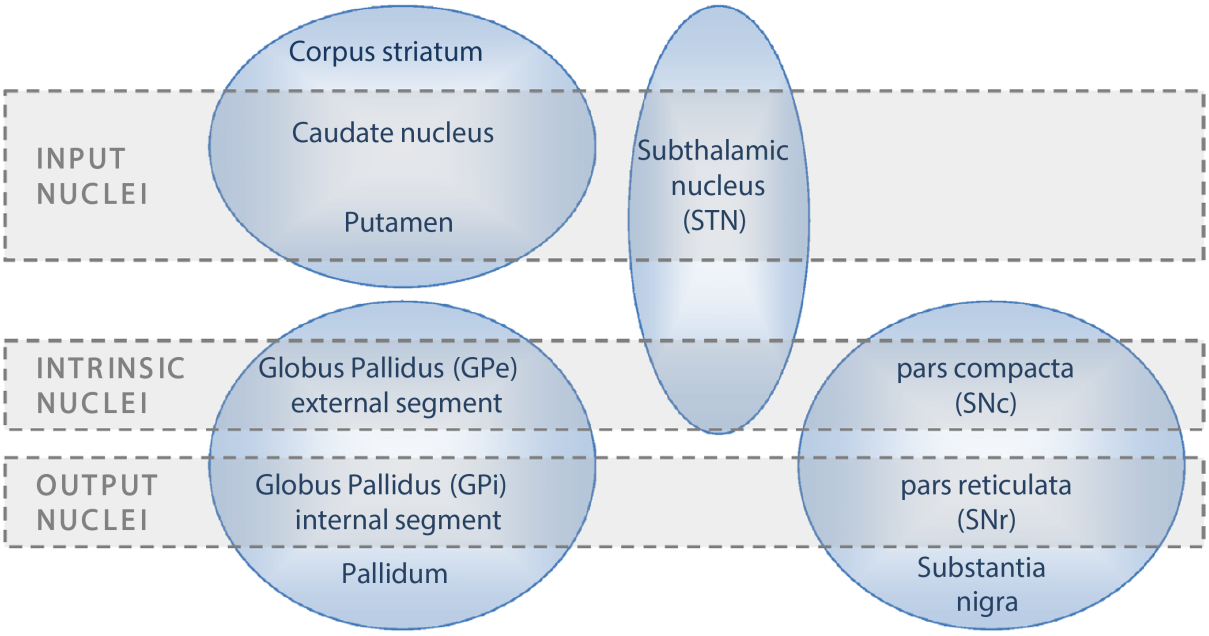



**Figure 11 | Anatomical and functional overview of basal ganglia nuclei.** The anatomical identity of each BG nucleus is listed as a blue oval, whereas functional identities are superimposed with dashed rectangles. Adapted from Shipp 2017.

#### 2.3.2.2 Cellular anatomy and physiology

#### 2.3.2.2.1 Corpus striatum

If the BG were a railroad network, the STR would be its busiest station. The STR is the convergence point between glutamatergic corticostriatal inputs, glutamatergic thalamic afferents and dopaminergic modulation (discussed below). At the heart of all of this sit two types of cells - the medium spiny neurons (MSNs), which make up > 95% of the population, and aspiny interneurons (Kemp and Powell 1971).

Collectively, MSNs are a population of cells with a hyperpolarised membrane potential (Kita et al. 1984). These cells feature several varieties of potassium conductance resulting from e.g. inwardly rectifying potassium channels (Kirs), fast- or slow-inactivating A-type potassium currents and persistent potassium channels, which together contribute to the MSNs characteristic slow depolarisation and delayed action potential onset. MSNs characteristically also switch between hyperpolarised (-90 to -70 mV) "down" and relatively depolarised (-60 to -40 mV) "up" states, associated respectively with a reduced and increased probability of action potential generation. Transitions between the two are determined by intrinsic membrane properties and coherent excitatory drive from the cortex and thalamus (Kreitzer 2009).

MSN properties are further shaped by their projection properties. Striatonigral pathway neurons, which project to SNr/GPi, express D1 receptors (D1R) which increase MSN firing by making it more likely that an action potential is generated while the neurons are in their "up" state. Striatonigral neurons also co-express the muscarinic M4 receptors, dynorphin and substance P (Hernández-López et al. 1997; Kreitzer 2009). By contrast, striatopallidal neurons, defined by their projection to the GPe, express D2 receptors (D2R) which inhibit MSN firing by decreasing the probability of depolarisation while the neurons are in their "Up" state (-60 to -40 mV, compared to the more hyperpolarized "Down" state, -90 to -70 mV). Striatopallidal neurons co-express adenosine A2A receptors and enkephalin (Hernádez-López et al. 2000; Kreitzer 2009). Regulation of these opponent pathways is achieved through dopaminergic afferents in the ventral tegmental area (VTA) and SNc (Fallon and Moore 1978).

Striatal aspiny interneurons, on the other hand, serve to modulate MSN activity (Tepper and Bolam 2004). They are classified as either large cholinergic or medium-sized GABAergic cells (Kreitzer 2009). Large cholinergic cells are characterised by tonic spontaneous firing patterns, which potentially set these cells up to provide constant cholinergic input to MSNs (Wilson et al. 1990). Medium-sized GABAergic cells are classified into (i) parvalbumin-positive, (ii), somatostatin-, neuropeptide Y-, and nitric oxide synthase-positive, and (iii) calretinin-positive cells. Parvalbumin-positive cells characteristically display rapid and sustained firing bursts in response to current injection, are highly excitable and can modulate up to hundreds of MSNs via synapses on their proximal dendritic regions (Cowan et al. 1990; Kreitzer 2009). They are also known as fast-spiking interneurons (FSIs). The other interneuron subclasses, by contrast, display low firing rates and are also known as low-threshold spiking interneurons (Tepper and Bolam 2004; Kreitzer 2009).

Macroscopically, the STR can be considered one single nucleus, though it is perforated by internal capsule (Shipp 2017). Its name refers to the interleaved strands of white matter from the internal capsule and grey matter of the STR in coronal sections. The striation is particularly pronounced in primates, where one can visually appreciate the medially placed caudate nucleus and laterally oriented putamen. By contrast, this division is not appreciable in the rodent and the white matter strands instead perforate the STR throughout. Collectively, the caudate nucleus and putamen (in the primate) are referred to as the dorsal striatum, by contrast to the ventral striatum, comprised of the NAc and olfactory tubercle. The striatum also features another prominent macroscopic division into patch (also known as striosome) and matrix compartments. Patches represent ~10% of the striatal volume, contain MSNs that primarily receive afferents from limbic and frontal regions and connect reciprocally to SNc. Patches are characterised by poor cholinergic marker staining. The matrix compartment, on the other hand, stains intensely for cholinergic markers, suggesting prominent cholinergic modulation. Matrix MSNs are associated with more prominent sensorimotor connectivity - they receive inputs primarily from the cortex and thalamus and connect to the SNr/GP (Kreitzer 2009).

#### 2.3.2.2.2 Globus pallidus

Primate globus pallidus consists of an external (GPe) and internal (GPi) segment. These have been linked to the rodent globus pallidus and entopeduncular nucleus, respectively. One area of particular relevance for disorders of motivation is the ventral pallidum (VP) - the ventral segment of the rodent globus pallidus - which is necessary for appropriate incentive motivation as well as hedonic reactions to reward, and which has been proposed to be the "limbic final common pathway" for mesocorticolimbic reward processing (Smith et al. 2009; Root et al. 2015). VP is comprised of GABAergic, glutamatergic, and cholinergic neurons. Physiologically these neurons have been categorised into type I (46%) - cells that are highly reactive to reward-predictive cues - vs type II and type III (14% and 16%, respectively), which are instead modulated by other reward-guided behaviours such as pursuit and consumption (reviewed in Root et al. 2015). VP cells interconnect with a wide network of cortical and subcortical areas. Notably, they receive a relatively large and topographically organised GABAergic input from NAc, a dopaminergic input from the VTA, glutamatergic inputs from the STN, PrL and amygdala; and serotonergic input. VP sends efferent projections to the dopaminergic mesencephalon, the medio-dorsal thalamus and lateral hypothalamus, lateral habenula, STN, entopeduncular nucleus, NAc, amygdala and PFC (Root et al. 2015).

#### 2.3.2.2.3 Subthalamic nucleus

The STN is a biconvex (in primates) or elongated (in rodents) nucleus that is juxtaposed between white matter tracts (internal capsule, cerebral peduncle), midbrain nuclei (red nucleus, substantia nigra), entopeduncular nucleus and other nuclei in the subthalamic area (e.g. zona incerta; (Hamani et al. 2004; Paxinos and Watson 2006; Marani et al. 2008). STN cells are all glutamatergic with an average firing rate of 13-18Hz in normal rats and 18-25Hz in non-human primates. Irregular (55-65%), regular (15-25%) and bursting (15-50%) firing profiles have been observed in non-human primates. STN has motor, associative and limbic subdivisions based on its afferent cortical innervation patterns. The motor STN, situated dorsally, receives afferents from the primary motor cortex, supplementary motor area (SMA), pre-SMA and the pre-motor cortex in primates. The limbic STN, situated ventrally, receives afferents from the PrL and mOFC.

The STN additionally receives afferent projections from the globus pallidus, thalamus, SNc, pedunculopontine nucleus (PPN), and the dorsal raphe nucleus (DRN). The STN sends efferent projections to both segments of the globus pallidus, substantia nigra, striatum, PPN and VTA (Parent and Hazrati 1995b; Hamani et al. 2004).

### 2.3.3 Frontal cortex and basal ganglia form interconnected loops

Early conceptualisations of the network architecture of the frontal cortex and BG were driven by efforts to explain clinical phenomenology. Key was the parcellation of cortico-BG pathways into those that effect either positive or negative feedback on the (motor) cortex (fig. 8). Though both have their origins in the striatum - a BG input nucleus - the former, termed the *direct pathway*, was proposed to start from dopamine 1 receptor (D1R)-expressing medium spiny neurons that project straight to the output nuclei - substantia nigra pars reticulata (SNr)/globus pallidus interna (GPi) - before proceeding to form another synapse in the motor thalamus and the cortex. The latter - termed the *indirect pathway* - was instead found to be initiated by D2R-expressing MSNs that loop to the external segment of the globus pallidus (GPe) and STN, giving rise to synapses in each, before returning to GPi/SNr. This circuit architecture provided an intuitive scaffold for explaining the basis of hypo- and hyperkinetic disorders, thought to result from disruptions in either the direct or indirect pathway, respectively (Albin et al. 1989; DeLong 1990).

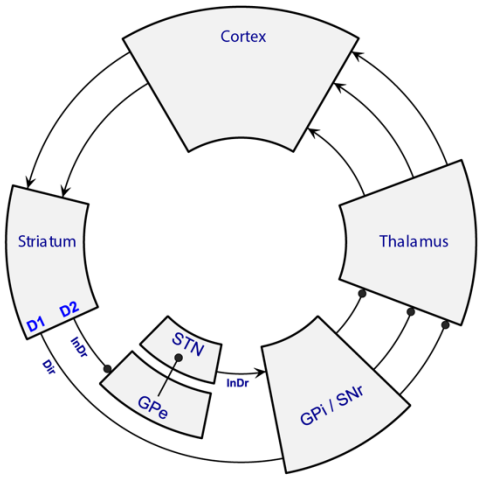



**Figure 8 | A rendition of the classic model of direct/indirect pathway connectivity in cortico-basal ganglia loops.** Lines with arrowheads or dots denote excitatory or inhibitory

connections, respectively. D1 - dopamine 1 receptors; D2 - dopamine 2 receptors. Dir - direct pathway axon; InDr - indirect pathway axon. Adapted from Shipp 2017.

Three additions have been made to this initial circuit model over time based on anatomical findings (fig. 9). First is the addition of a second short limb between the GPe and GPi/SNr to the indirect pathway, formed GPe-STN axon collaterals which can disinhibit the BG output nuclei (Smith et al. 1998; Sato et al. 2000a). Second and third are additional excitatory projections from (i) the STN to the GPe, formed by axon collaterals terminating in GPi/SNr (Sato et al. 2000b) and (ii) the cortex to the STN, termed the *hyper-direct pathway* (Parent and Hazrati 1995b), proposed to be important for prompt pruning of volitional action programmes to ensure the right one is selected (Nambu et al. 2002). Further regulatory loops will be discussed later.

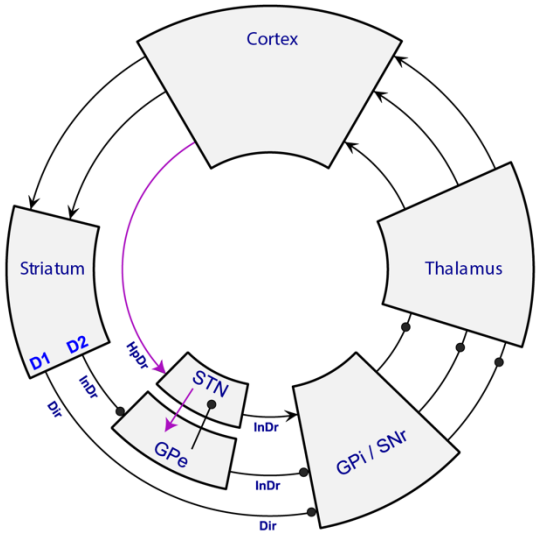



**Figure 9 | A revised model of cortico-basal ganglia connectivity.** Additions to the previous model are marked with purple lines. Adapted from Shipp 2017.

These cortico-basal ganglia loops are not, however, only concerned with motor function. In fact, anatomical work in primates indicate that inputs to these loops originate from most of the cortical sheet with preserved topography (Shipp 2017). These have been categorised into three domains: *sensorimotor, associative* and *limbic*, reflecting their distinct cortical origins, striatal target areas (Alexander 1986; Parent and Hazrati 1995a; Postuma and Dagher 2006). The sensorimotor projection originates in

the primary motor and somatosensory cortical areas and projects to caudal portions of dorsolateral striatum. The associative projection arises in the associative areas of the frontal, temporal and parietal cortices and terminates in most of the caudate and rostral parts of the putamen. The limbic input projects from the medial wall of the prefrontal cortex, the amygdala and hippocampus to the ventral and medial striatum (Parent and Hazrati 1995a; Haber 2016; Haber et al. 2020).

There is, however, notable overlap between the projections from different cortical areas, suggesting functional integration of diverse cortical function within the striatum (Parent and Hazrati 1995a; Mailly et al. 2013; Shipp 2017; Haber et al. 2020). For example, projections from the frontal cortex converge extensively in the rostral striatum of macaque monkeys (fig. 10A; Averbeck et al. 2014) and rats (fig. 10B; Mailly et al. 2013).

Convergence on the microscopic scale is mirrored in striatal microarchitecture. Cortico-striatal terminals mainly target MSNs and FSIs However, each cortico-striatal axon mainly targets only up to 1% of the MSNs target area (Zheng and Wilson 2002). FSIs have therefore been investigated as possible substrate of adjudication between various cortical inputs. Indeed, convergence of input from primary motor and sensory cortices has been demonstrated on these FSIs, highlighting this cell type to be a hotspot of integration antecedent to MSN modulation (Ramanathan et al. 2002).

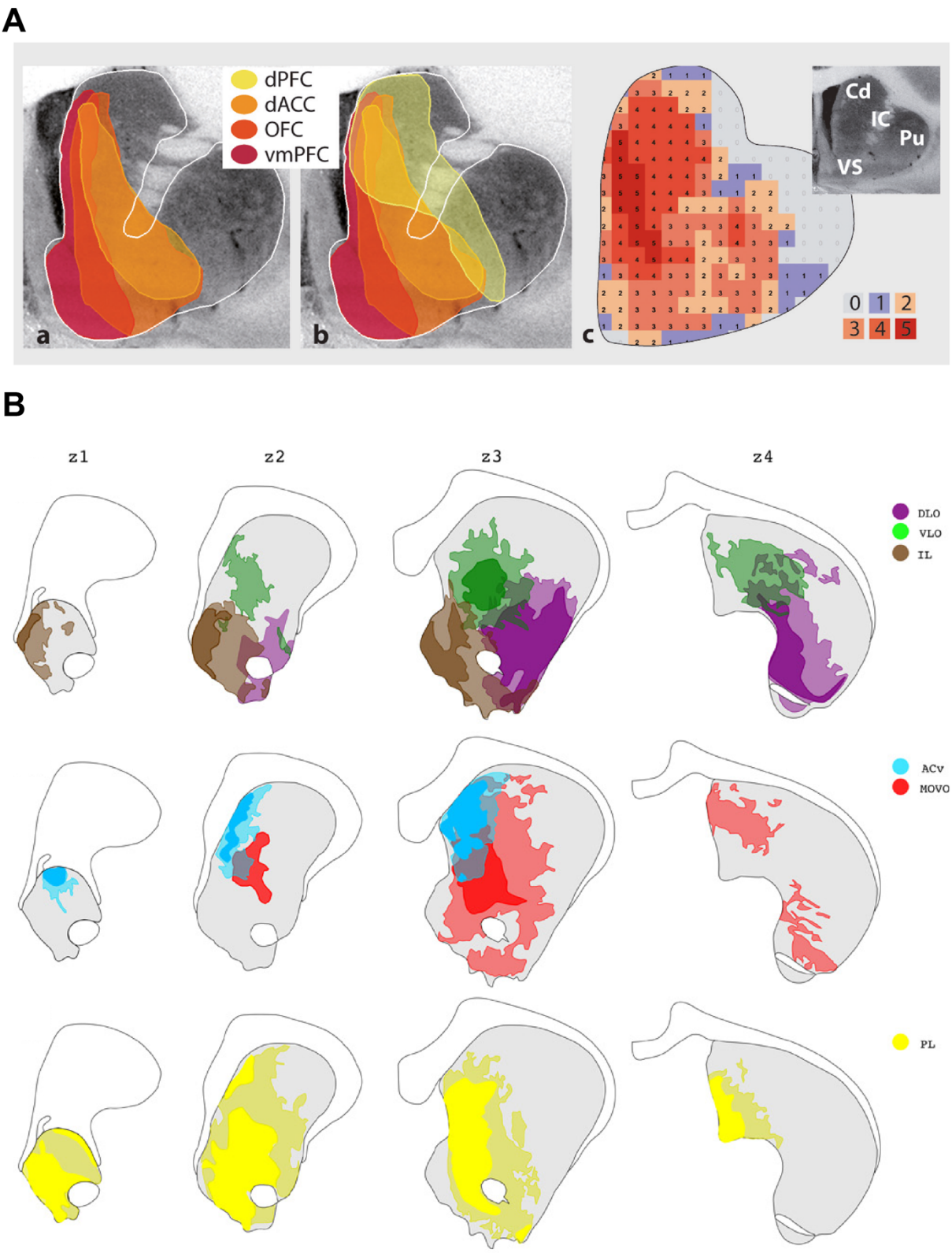



**Figure 10 | Illustration of the convergence of frontal projection territories in the striatum of the macaque monkey and the rat. Aa-Ab,** overlap of frontal cortical projection territories in the macaque monkey**. Ac,** probability map of multi-area converge in the striatum

of the macaque monkey. Numbers and the associated heatmap, detailed in the bottom right legend, indicate the number of prefrontal cortical areas that projected to each voxel on the mapped striatal territory. dPFC - dorsal prefrontal cortex; dACC - dorsal ACC; OFC - orbitofrontal cortex. Adapted from Averbeck et al. 2014. **B,** Terminal fields of anterogradely labelled frontal cortical areas in the striatum of the rat along different anteroposterior co-ordinates (ordered as columns). DLO - dorsolateral OFC, VLO - ventrolateral OFC; IL - infra-limbic cortex; ACv - ventral ACC; MOVO - medial and ventral OFC; PL - pre-limbic cortex.

### 2.3.2 Neurotransmitter modulation

Aside from the fast and focused neurotransmission allowed by glutamate and GABA, the central nervous system also relies on the broadcast-type relatively slower timescale signals transmitted by monoamine neuromodulators such as dopamine, serotonin, noradrenaline, adrenaline and histamine. Here, only dopaminergic and serotonergic and noradrenergic systems, which influence function in the targeted regions in chapters 2-3, will briefly be considered.

#### 2.3.2.1 Dopamine

Dopamine (DA) is involved in cognition, motivation and movement (Cools et al. 2011). Importantly for the context of motivational disorders, DA is crucial for regulating appropriate reward-seeking behaviour (Salamone 1996; Berridge and Robinson 1998). It is produced in nuclei A8-A16 of the mammalian brain. Among those are A9 and A10, better known as SNc and VTA, located in the midbrain. These cells are defined as dopaminergic based on their unique ability to synthesize from L-tyrosine with the help of enzymes tyrosine hydroxylase and dopa decarboxylase. SNc and VTA neurons typically exhibit tonic 1-8Hz spontaneous firing *in vivo*. However, behavioural events can also lead to the generation of bursts of action potentials, leading to phasic increases in DA concentration (Schultz 2007). Dopaminergic projections are grouped into nigrostriatal, mesolimbic, mesocortical and tuberoinfundibular. The originator cells of these projections are interdigitated in the mesencephalon (fig. 15). Further, dopaminergic axon collaterals also reach targets such as VP, STN and globus pallidus/entopeduncular nucleus *en route* to the striatum (Björklund and Dunnett 2007).

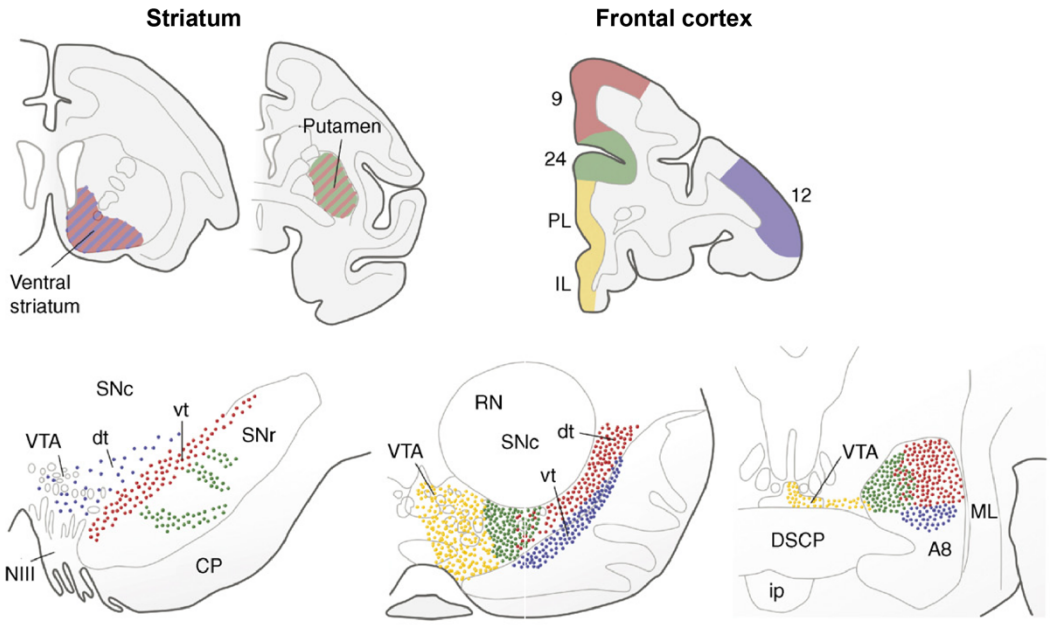



**Figure 15 | Projection-based topographic patterning of midbrain dopamine cells.** While vmPFC projections are mostly restricted to the VTA, projections to other PFC and striatal regions are distributed throughout the VTA and SNc. Adapted from Björklund and Dunnett 2007.

### 2.3.2.2 Serotonin

Similarly to DA, serotonin (5-HT) is been implicated in affective and activational processes but, contrary to DA, has been proposed to mediate aversive processing and behavioural inhibition (Soubrié 1986; Cools et al. 2011). Indeed, depletion of central 5-HT or administration of 5-HT$_{2C}$ receptor antagonists speeds up reward-oriented responses and increases willingness to work (Higgins and Fletcher 2003; Fletcher et al. 2007; Simpson et al. 2011; Bailey et al. 2016b, a; Härmson et al. 2022). 5-HT neurons are mostly located in the brainstem in either the DRN or median raphe nucleus (MRN; Judge and Gartside 2006). Ascending serotonergic projections innervate cortical, limbic, midbrain and hindbrain regions and serotoninergic receptors are found throughout the brain (Berger et al. 2009).

### 2.3.2.3 Noradrenaline

In contrast to DA and 5-HT, noradrenaline (NA) has been proposed to carry out quite a different function - implementing changes in brain network states in response to

shifting behavioural contexts (Bouret and Sara 2005), facilitating arousal (Sara and Bouret 2012) and indexing effort cost (Bouret et al. 2012; Varazzani et al. 2015) . NA is produced in noradrenergic cells in the locus coeruleus (LC) and is broadcast to brainstem neuromodulatory nuclei as well as all cortical regions, thalamic nuclei, the septal region, hippocampus and amygdala, though there is only a sparse projection to striatum (Loughlin et al. 1986).

## 2.4 Cost-benefit decision-making and the brain

Having reviewed the anatomy, physiology and neuromodulatory properties of the frontal cortical basal ganglia circuit, attention can now be turned to its function with regards to value-guided decision-making. In the following, I will examine some of the typical ways of indexing into neural function and highlight how the anatomical connections of the hitherto highlighted regions shapes their role in cost-benefit decision-making.

### 2.4.1 From anatomy to function

Anatomical connectivity might provide a useful means to dissociate regional function in the cortico-basal ganglia circuit. In this vein, the efferent connections of the rodent **mOFC** (primate area 14) to the medial and ventral striatum, coupled with afferent projections from midbrain dopaminergic nuclei, mediodorsal thalamus and piriform cortex and a paucity of connections with sensorimotor areas (Barreiros et al. 2021; Le Merre et al. 2021) might suggest an important role in calculating *stimulus value correlates*. By contrast, the ACC, part of the dmPFC, is integrated in a network of sensorimotor and limbic areas but also the mOFC (Le Merre et al. 2021), potentially imparting it with the ability to weight stimulus values by their associated *costs,* especially physical ones. The medial striatum, ventral pallidum and subthalamic nucleus are all recipients of topographically organised convergent medial frontal cortical input (Parent and Hazrati 1995b; Mailly et al. 2013; Averbeck et al. 2014). This might, in theory, enable them to represent reward, cost and integrated value, depending

on the subsampled region. However, owing to the recurrent nature of the cortico-basal ganglia circuit, regional parcellation of function might not be very straightforward and estimates of reward, cost and integrated value might be found throughout the network.

### 2.4.2 Indexing into nervous tissue activity

There is a dazzling variety of methods available for gauging nervous system activity. However, none of them are "silver bullets". For example, there is a trade-off between the methods' spatial resolution and their invasiveness. Further, there is variation of orders of magnitude in the spatial and temporal resolution (fig. 16). The neuroscientific toolbox also has two fundamentally different ways of indexing into brain activity: observational (or correlative) and causational.

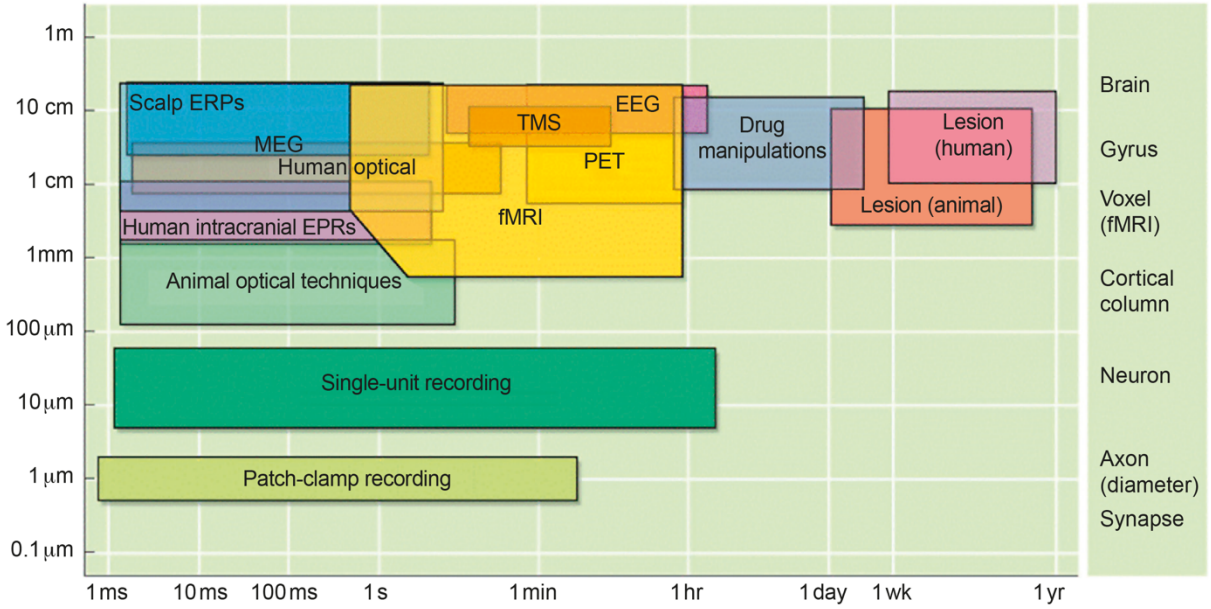



**Figure 16 | Neuroscientific methods employ different scales of spatial and temporal resolution.** The spatial resolution (y-axis) of some of the most commonly used techniques are plotted against their temporal resolution (x-axis). ERPs: event-related potentials; MEG - magnetencephalography; TMS: transcranial magnetic stimulation; fMRI - functional magnetic resonance imaging; PET - positron emission tomography. Adapted from Glimcher and Fehr 2014.

For the purposes of this thesis, I will only focus on the method of invasive intracerebral electrophysiological recording, which goes at the heart of chapters 2-3 of this thesis. This technique, referred to as single-unit recording on fig. 16, traditionally employs around ~50 microns diameter tungsten wire microelectrodes bundled into groups of 4 or 6 which are lowered into target areas with the help of a stereotaxic frame (Hubel 1957; Buzsáki 2004). While in the brain, electrodes are commonly either driven in and out of their recording areas (McNamara et al. 2014), typically in the hippocampal recording field, or lowered progressively deeper on a daily basis to sample new cells. There is a debate in the field about whether new cells are sampled even if electrodes are not lowered to unique positions and while some studies diligently sample unique positions to build up a representative sample size (Hillman and Bilkey 2010; Porter et al. 2019), others stay put and assume that new cells are observed on each recording day (Li et al. 2022). Though silicone probes have been introduced with the potential of simultaneous sampling of a much higher numbers of cells, they are limited by their much higher price, linear dimensions, and short (1cm) shank (Jun et al. 2017).

Single-unit recording data are then typically passed through a spike sorting algorithm to extract action time stamps for each identified unit (characterised in depth in chapter 2). Further analyses typically centre around quantifications of *firing rate* - the number of action potentials observed per unit of time. Firing rate changes are then usually analysed relative to time points of interest, e.g. the onset of a reward-predictive stimulus or action. This often points to evidence of the existence of a *rate code,* whereby levels of an experimental variable map onto variation in firing rate. This type of analysis often assumes that each neuron is statistically independent. However, groups of neurons have been shown to co-ordinate their firing within time windows approximately equivalent to the membrane time constant (10-25 ms), as these events are readily stabilised by spike-timing-dependent plasticity. Coincidentally, this dynamic is also the basis of a *temporal code* which can hold information about environmental contingencies that is often not evident in the rate responses of single units (El-Gaby et al. 2021).

### 2.4.3 Identifying value signals in the brain

Before delving deep into the neurobiological correlates of cost-benefit decision-making, it is important to point out a few caveats that can influence observed experimental results. First, though the exact numerical quantities of the experimental variables (e.g. reward magnitude) may constitute an objective reward function, organisms likely do not interpret it the same way. Instead, rewards might be weighed up in terms of their *utility* - a composite measure of their usefulness for biological needs such as homeostasis, reproduction, and pleasure (Schultz 2015) and the amount of that good the decision-maker already owns (marginal value theorem; Glimcher and Fehr 2014). To address this, the neuroeconomics field often estimates a *subjective value* either (i) from participants' ratings of the utility of the good offered during testing or (ii) directly from choice data (Glimcher and Fehr 2014). In paradigms that also feature effort costs for obtaining goods, subjective value is commonly formulated as (effort-)discounted reward value (Hosokawa et al. 2013; Klein-Flügge et al. 2016; Chong et al. 2017). As subjective value is closely correlated to each subject's behaviour, it likely has implications for our ability to infer value-related brain signals (Glimcher and Fehr 2014).

Second, covariates of reward can often disguise as value. One such important covariate is *expectation error*. As organisms navigate through the environment, statistical regularities are learned, however violations of expectation prompt the need for increased vigilance and re-evaluation. Indeed, midbrain dopaminergic and noradrenergic neurons readily respond to changing contingencies (Schultz et al. 1997; Bouret and Sara 2004) and the putamen has been shown to respond to both positive and negative prediction errors (McClure et al. 2003). Another important covariate is salience - a composite of motor preparatory, arousal and attentional responses to the revealed reward identity. For example, reward value can be decoded from electromyographical activity from the neck musculature (Roesch and Olson 2003). This confound can be controlled for in tasks that involve appetitive as well as aversive outcomes. The validity of this approach has been demonstrated empirically (Kahnt et al. 2014), but accounting for this remains an outstanding challenge in most tasks with appetitive outcomes. Other suggested covariates include the informativeness or identity of the reward-predictive stimuli or the particulars of the required reward-guided action (reviewed in Hayden and Niv 2021b).

Third, value-like signals might arise as an epiphenomenon of behavioural or brain processes that actually do not use it. For example, the "chosen value" signal identified in the vmPFC of several studies may be driven by time-varying dynamics of the underlying decision process (reviewed in Hunt and Hayden 2017). Second, value correlates are inextricably linked with behavioural preference which makes the two difficult (but not impossible) to disentangle (Hayden and Niv 2021b). Third, linear models are often used for the detection of value signals. Several studies that have identified linear relationships between stimulus value and firing rate have used just two magnitudes of reward (Hillman and Bilkey 2010; Porter et al. 2019). However, the usage of 3 or more levels of a variable would constitute a more rigorous test of this relationship, leading to fewer type I errors.

### 2.4.4 Cost-benefit decision-making in the frontal cortical basal ganglia circuit: divide and conquer?

One influential account of the dynamics of cost-benefit decision making in the frontal cortical-basal ganglia circuit has been that of serial processing. This framework ascertains that cost-benefit decision-variables are processed in distinct areas and passed between specialised nodes of the frontal cortical-basal ganglia network. The following is one potential distillation of the serialist argument, driven by converging evidence that implicates the mOFC in the processing and comparison of reward value, of ACC in the computation of integrated value based on locally stored action values/costs and of the basal ganglia subsequently shaping action vigour and inhibiting competing motor programmes (fig. 17). In the following section, I will evaluate the evidence in favour of this account and with a specific focus on paradigms that included *effort* cost.

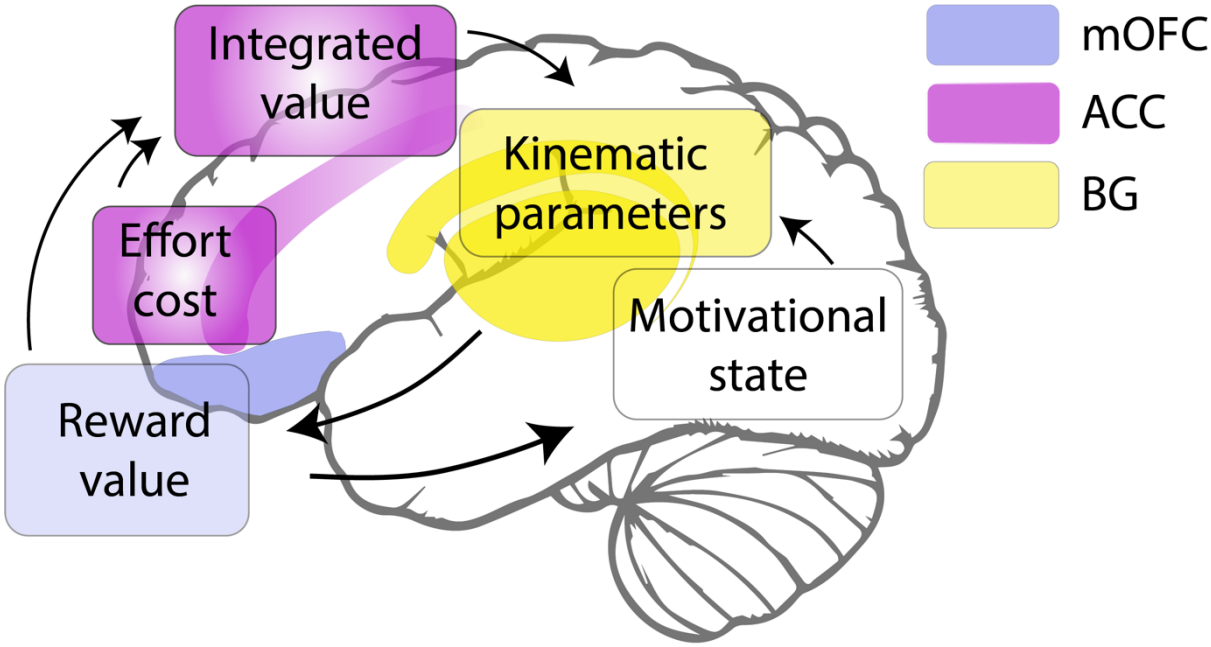



**Figure 17 | A model for serial and localised implementation of cost-benefit decision-making.** Reward (or stimulus) value is learned in the mOFC through reinforcement and are subject to changes in internal state (e.g. satiety). This is passed on to the ACC, which encodes the associated effort costs and arrives at an integrated value as a result of a comparison between the two, informing the organism's subsequent decision. Those parameters are forwarded to the BG which direct and invigorate the appropriate reward-guided action.

#### 2.4.4.1 Medial orbitofrontal cortex and reward value

A seemingly selective role of the mOFC in common-currency stimulus value coding has been presented by a series of studies, even after controlling for expectation error and salience (Kable and Glimcher 2007; Plassmann et al. 2007; Hare et al. 2008, 2009; Boorman et al. 2009, 2013; Chib et al. 2009; Litt et al. 2011). These results are mirrored by an alteration in perceived reward value in humans following lesions to this area (Manohar and Husain 2016). However, relatively little is known about the cellular level mechanisms of this common-currency coding in the mOFC of rats and non-human primates, as this area is difficult to access and most recording studies have instead recorded from vOFC or lOFC. The small amount of available evidence in non-human primates and rats suggests that though stimulus value coding does take place in the mOFC, this signalling is particularly dependent on the organisms' internal state

(Monosov and Hikosaka 2012; Abitbol et al. 2015; Lopatina et al. 2016; San-Galli et al. 2018).

#### 2.4.4.2 The role of ACC in signalling effort, integrated value and decision

The ACC, by contrast, has received special attention for its role in *integrating* reward with effort and shaping decisions. Aside from observations from clinical case reports outlined above, a mechanistic link between ACC and effort was highlighted by a series of studies where rats were asked to make decisions between a high reward-high effort and low reward-low effort option, associated with 4 food pellets and a climb over a barrier vs 2 food pellets and no barrier, respectively. Whereas normal rats nearly always chose the high effort-high reward option, ACC lesions reversed this choice pattern. This was not attributable to an inability to exert the requested amount of effort, demonstrated by lesioned rats' reinstated preference for the high reward when both options were paired with the same barrier climb requirement. Nor was this attributable to a failure to value large rewards over small ones (Walton et al. 2002, 2003; Schweimer and Hauber 2005; Rudebeck et al. 2006). However, the ACC has a specific role in motivating organisms to overcome physical effort, not delay costs (Rudebeck et al. 2006).

Reflections of effort processing have also been documented in studies examining the activation of ACC neural tissue. A 2-deoxy-2-$^{18}$F-fluoro-D-glucose positron emission tomography study showed metabolic activity in the ACC to be significantly higher when rats needed to integrate reward information with effort costs vs when effort was equated (Endepols et al. 2010). Single unit recordings have corroborated this with evidence of local coding of reward and effort in a task which independently manipulated each (with a small fraction of cells coding for both; Porter et al. 2019). ACC unit firing rates further indicate cost/benefit integration by signalling rats' upcoming/ongoing engagement with (Hillman and Bilkey 2010; Cowen et al. 2012; Porter et al. 2019) or switching to (Elston et al. 2019) the optimal option in tasks incorporating reward and effort trade-offs.

Recent research has also shed light on the neural mechanisms by which choice is shaped in the ACC. However, it has also raised a question of exactly what choice is and how it

might be represented on the neural level. In 2-alternative choice scenarios, the answer might be that the representation consists of a mapping between firing rate and the tendency of an animal to choose one of the two options (e.g. left or right). One notable report in which non-human primates needed to serially sample visual cues pertaining to left/right options showed the evolution of ramp-to-bound type signal predicting subsequent choice in the ACC (Hunt et al. 2018). By contrast, another approach has been to frame choice in a context resembling naturalistic foraging scenarios. In such scenarios, subjects typically encounter one option at a time and need to either accept or reject it. To solve such problems, subjects may need to track the value of the encountered option (*encounter value*), environmental richness (*search value*) and the cost of sampling the environment vs the current option (*search cost*). One such report showed the ACC to be particularly sensitive to *search value*, in essence signalling when it might be beneficial to switch from exploiting a current option to exploring the environment (Kolling et al. 2012).

#### 2.4.4.3 The role of the striatum in cost-benefit decision-making

Accounts of the striatum in cost-benefit decision-making have largely focused on the NAc, specifically in the context of dopaminergic neuromodulation. Levels of DA in the core sub-region of the NAc (NAcC) correlate with the amount of effort animals put in and depletion thereof reduce rats' ability to sustain a fixed ratio (FR) 5 schedule of responding for reward (Sokolowski and Salamone 1998). When given a choice between pressing a lever 5 times to obtain a preferred reward or obtaining less preferred food without a cost, rats with reduced intra-NAcC DA tone prefer the latter despite a preserved appetitive preference for the former outcome (Salamone et al. 1991). Further, bilateral excitotoxic (Hauber and Sommer 2009), and DA depletions (Salamone et al. 1994), shifted preferences toward the low-effort/low-reward option on an effort T-maze paradigm. Subsequent research has demonstrated integrated value to be encoded in the human ventral striatum during a cost-benefit decision-making task (Müller et al. 2021). However, a serialist proposition where relevant cost-benefit information gets passed from the ACC to the NAcC cannot be ruled out, as disconnecting the two via contralateral lesioning of each was sufficient to produce an impairment in an effort T-maze paradigm (Hauber and Sommer 2009). Unfortunately, little is known about the

local computations in NAcC cells and their temporal evolution compared to other frontal cortical basal ganglia regions in circumstances where costs and benefits need to be weighed up against each other to make decisions.

The rodent DMS receives frontal cortical and dopaminergic input that partially overlaps with the NAc, thus setting it up as a potential key player in the representation of cost-benefit decision variables. Nevertheless, the local electrophysiological work to date has largely focused on the neural substrate of execution of reward-guided action, rather than cost-benefit decision-making. For example, DMS cells shape and encode the vigour of reward-guided nose-poking while taking into account the environmental richness (Wang et al. 2013). Further, DMS (but also dorsolateral striatal) cells control and represent the vigour of movements in a task where a joystick needed to be displaced a certain amount to earn reward (Panigrahi et al. 2015).

#### 2.4.4.4 A diverse role of the subthalamic nucleus in cost-benefit decision-making

The evidence to date, though scant, prove the STN to be a reliable reporter - if not mediator - of cost-benefit decision-making. For instance, bilateral fibre-sparing lesions of the rat STN disinhibit reward pursuit, leading to the elevated locomotor reactions to food presentation and a higher PR break-point (Baunez et al. 2002). However, in the absence of disconnection lesion studies with other afferent regions, it is hard to adjudicate the *independent* role of the STN in such behaviours. The primate STN is also contains cells that reliably report on the level of prospective reward or required effort (with very small number showing an interaction) in a task where non-human primate were asked to press on a lever with varying amount of force to obtain variable rewards (Nougaret et al. 2022). These results were also reflected in human STN local field potentials (LFPs) as a modulation of power in the 1-10Hz range in response to reward and effort cues (Zénon et al. 2016b).

#### 2.4.4.5 The role of the ventral pallidum in cost-benefit decision-making

To my knowledge, VP local circuitry has never been studied in the context of cost-benefit decision-making. However, important evidence in its role in the activational

component of motivation can be gleaned from reports showing a dramatic reduction in reward-guided behaviour following lesions to this region (reviewed in Smith et al. 2009). Further, VP neurons indicate the identity (Ottenheimer et al. 2018) and magnitude (Tachibana and Hikosaka 2012) of reward, track changes in reward value based on internal state and guide reward-oriented action selection (Ottenheimer et al. 2020) and vigour (Tachibana and Hikosaka 2012).

### 2.4.5 Cost-benefit decision-making in the frontal cortical basal ganglia circuit: everything, everywhere, all at once?

A significant challenge to the serialist perspective has been posed by literature showing the evolution of cost-benefit decision variables simultaneously in multiple frontal cortical nodes. For example, population activity representing an animal's spatial focus of attention and, cue value difference and subsequent decision emerged in parallel across a non-human primate OFC, dorsolateral PFC and ACC network in an attention-guided decision-making task (Hunt et al. 2018). Similarly, examples of neurons coding the levels of probability, payoff and cost in a choice task that presented pairs of options varying in one dimension at a time were found the ACC, OFC and lateral PFC neurons (though ACC neurons exhibited the highest degree of multiplexing Kennerley and Wallis 2009).

These findings can be reconciled within the framework of a distributed system for cost-benefit decision-making. In such a system, functions such as reward or decision representation arise from the collective action of populations of neurons, potentially scattered across multiple brain areas (Hunt and Hayden 2017). This raises of the question of how the brain arrives at a consensus if each neuron is simply "casting a vote". One proposed mechanism for a consensus is competition via inhibition, where distinct populations of neurons are selectively driven by their preferred options but also inhibit each other. In the process, activity in the "winning" population persists, giving a read-out of the decision variable. This model has proven to be biophysically plausible human PFC and parietal cortex where neural activity transformed from a sum of the chosen/unchosen values to their value d¿ifference (Hunt et al. 2012). An

electrophysiological proof of the implementation of such a principle in the brain came from a study where macaques had to choose between non-simultaneously presented offers in a gambling task. While their vmPFC neurons initially represented the values of each offer through linear firing rate changes, the tunings for each option became anticorrelated during the subsequent comparison epoch (Strait et al. 2014), consistent with competition via inhibition.

Distributed coding of decision variables is also consistent with the loop-like frontal cortical basal ganglia network architecture, which could in theory spread decision variable representations through the whole network. However, some regional specialisation might nevertheless be retained through specific anatomical connections, perhaps explaining why the vmPFC vs dmPFC have dissociable roles in representing stimulus- vs action-reward associations (discussed above). Likewise, even when two regions might be representing something like stimulus value using a similar proportion of neurons, the time scales of such representations can vary area-by-area by virtue of their feedback connectivity. The rodent PFC has been demonstrated to have a highly recurrent pattern of connectivity compared to e.g. sensory areas which predominantly give rise to feedforward connections (Le Merre et al. 2021). This correlates with a slower time scale of variation of intrinsic activity patterns, perhaps setting the PFC up specifically for cognitive functions like working memory and decision-making (Murray et al. 2014).

## 2.5 Summary

In the second half of this chapter, I have given a cursory overview of the anatomy, physiology and functions of frontal cortical-basal ganglia regions implicated in the dysregulation of motivation. Though these regions may lie some way away from each other to the eye, they are in fact heavily interconnected through cortico-basal ganglia thalamic loops. Further, though the loops are organised by topographic patterns according on the affective and effective ends, they are heavily overlapping. Each region within, though, is characterised by different connectivity to other brain regions. This has important implications for their function in behaviour. First, there is a degree of

regional specialisation in cost-benefit decision making that favours a serial function perspective, as evidenced in lesion, imaging and electrophysiological recording studies. However, cost-benefit decision variables are also highly pervasive, giving the impression of a distributed system that nevertheless has some regional variation depending on anatomical connection patterns. The read-outs of these variables in the brain must be interpreted with some caution, as they reflect parallel phenomena such as salience, expectation error, and cue identity.

## 3 Towards a brain network-level understanding of cost-benefit decision-making

Though considerable progress has been made in understanding the mechanisms underlying cost-benefit decision-making in individual frontal cortical-basal ganglia regions, relatively little is still known about their collective function in enabling this behaviour within the confines of a single behavioural paradigm. Are cost and benefit processed with separate time courses in individual areas or instead concurrently represented in each? How do these signals shape the eventual decision to pursue the reward of choice? Do these signals arise on the level of individual units or populations within each region, or are they an emergent property of the entire frontal cortical basal ganglia network? What are the fractions of individual units that each region employs to represent decision variables? Finally, is there evidence of functional specialisation in the frontal cortical basal ganglia network or are decision variable computations distributed?

Piecing together a unifying theory of the mechanisms of cost-benefit decision-making in the frontal cortical-basal ganglia network has complicated by the relatively small selection of studies that independently manipulated both reward and effort. Moreover, few studies have concurrently sampled areas representative of the whole network with high time resolution. In this thesis, a novel behavioural paradigm along with a multisite recording setup was developed to address these shortcomings and answer the outlined questions in rats. The paradigm features an orthogonal manipulation of food reward and physical effort to disentangle the neural signatures of each. Crucially, rats need to

either accept or reject individually presented offers. This has the effect of likening the decision problem to a naturally occurring foraging scenario. By recording individual neurons across the frontal cortical-basal ganglia circuit while rats are making cost-benefit decisions in this carefully tailored setup, this thesis aims to offer a unified description of the underlying neural mechanisms at service of this intricate behaviour.

# 2. Setting up the stage for the examination of cost-benefit decision-making in rats


## ABSTRACT

Adaptive value-guided decision-making often requires weighing-up of the costs and benefits of pursuing an available opportunity. Little is currently known about how this process unfolds in the frontal cortical basal ganglia network. This is because few studies to date have (i) independently manipulated costs and benefits within the confines of a single behavioural paradigm and (ii) recorded from a network of interconnected frontal cortical basal ganglia areas while subjects are engaged in cost-benefit decision-making. The body of this thesis aims to address both of these issues. In the first part of the chapter, the general methods that the following experimental chapters employ will be outlined, including an overview and development of the behavioural paradigm, statistical analyses and electrophysiological as well as pharmacological techniques. The second part will be a summary of the main behavioural findings which will lay the foundation for interpreting neurophysiological findings in chapters 3 and 4.

## 1. Introduction

Over the past decades, much progress has been made in the way of developing behavioural paradigms for studying the neural correlates of value-guided decision-making. This has resulted in ever more control over decision variables, allowing us to develop theories about value representation - i.e. good- vs action-based (Cisek 2006; Padoa-Schioppa 2011) - and the mechanisms by which these shape decisions (Hare et al. 2011; Cisek 2012).

Though the field has come far in shedding light on the mechanisms of representation of benefits, the flipside of the equation - costs - has received much less attention (Walton and Bouret 2019). Insofar as nearly every naturally occurring value-guided decision is paired with them, this gap in our understanding significantly undermines our ability to encapsulate the mechanisms of value-guided decision-making. This gap has arisen because many paradigms to date have not orthogonalized reward and effort. Further, effort in many paradigms is a mixture of different types of cost (e.g. physical and time).

These endeavours are also complicated by the use of distinct types of paradigm for studying value-guided decision-making. While the simultaneous presentation of offers of different values has been a convenient go-to method that allows the examination of individual as well as comparative value representations (e.g. Kennerley et al. 2009; Philiastides et al. 2010; Hare et al. 2011; Knudsen and Wallis 2020), cost-benefit decisions in the wild seldom take this form. Instead, animals often make sequential foraging-like decisions based on comparisons of the foreground and background utility, corresponding to the reward rates offered by the currently exploited option vs returns in elsewhere in the environment if the decision-maker were to explore (Charnov 1976). Moreover, there is converging evidence that decision-makers default to sequential decision-making even when faced with multiple-option choices (Kacelnik et al. 2011) and that there is a brain circuit dedicated to the implementation of such an algorithm (Kolling et al. 2012; Rushworth et al. 2012).

Neuroscience has also made considerable progress in the ability to monitor single cells and neural populations in awake and behaving animals, allowing detailed documentation of the neurophysiological multilogue that underpins value-guided behaviour. However, the progress has mainly boosted our ability to record a bigger yield of neurons from one area, e.g. with 2-photon calcium imaging (Helmchen et al. 2001; Dombeck et al. 2007; Flusberg et al. 2008), or from multiple areas along one anatomical dimension, e.g. with silicon probes in head-fixed animals (Jun et al. 2017) or microwire electrodes in freely moving animals (Buzsáki 2004). Crucially, cost-benefit decision-making is known to engage an interconnected system of frontal cortical-basal ganglia areas which have hitherto been difficult to probe simultaneously in freely moving animals. Simultaneous, rather than serial, recordings across the network offer the advantage of being able to probe inter-area interactions and functional connectivity (Semedo et al. 2019; Gocken et al. 2022).

This chapter is a description of the development of a behavioural task and a recording approach that has aimed to address some of these problems by:
(i) designing a behavioural paradigm design where cost and benefit are manipulated independently, and where sequentially presented offers can be either accepted or rejected;
(ii) setting up a recording preparation that can acquire clean extracellular biopotentials in 5 frontal cortical-basal ganglia areas - mOFC, ACC, DMS, VP and STN - in freely moving rats performing this task. These regions were chosen due to their widely demonstrated role in regulating motivation (chapter 1).

I will show, in a final version of the task, that rats can flexibly make decisions about whether to pursue reward based on the amount on offer versus the required effort, and that this behaviour is sensitive to changes in the value of the outcome.

# 2. Methods

## 2.1. Animals

All procedures were carried out in accordance with the UK Animals Scientific Procedures Act (1986) under the project licence PB9E6B4B1. For the reported set of experiments, we used 13 male Lister-Hooded rats (Charles River) that weighed 300-390g on ad libitum food at the beginning of procedures. Rats were maintained on a 12-hour light/dark cycle (lights on at 07:00 AM) and group housed except after having received a microwire electrode implant. Rats were allowed to habituate to the animal unit for at least 1 week after their arrival, before the start of experimental procedures. During training and testing periods, rats were maintained at 85-90% of their free-feeding weight by means of restriction to access of solid foods, whereas water was available ad libitum. Only naive animals were used.

## 2.2 Apparatus

The behavioural task was developed in two stages - v1 and v2 - with some differences between the apparatus design, outlined in table 1. The operant box was designed in Fusion360 (Autodesk) and consisted of two 17.6cm diameter custom-designed 3D-printable plastic semi-circular arenas, connected by a 130 cm-long Perspex corridor with walls on either side (fig. 1). Each arena was able to accommodate one bespoke 3D-printed reward dispenser and one nose-poke port. However, v1 of the apparatus was an asymmetric design with a nose-poke port only one side (table 1). V2, by contrast, was symmetrical and had a nose-poke and reward port on either side. Nose-poke ports and reward ports were fitted with an infrared (IR) light-gate and a white LED house light. Reward ports were connected with food pellet dispensers (Med-associates) fitted outside the walls of the arenas. In contrast to V1, V2 also featured white light LEDs above both arenas, 6 light-gates, 3 barrier slots distributed evenly along the corridor, and moveable floors that were able to block the reward ports.

**Table 1 | Overview of the core apparatus feature differences between the two versions of the cost-benefit decision-making task.**

| Feature | Task v1 | Task v2 |
|---|---|---|
| **Nose-poke ports** | Only on side A | Side A and B |
| **Corridor light-gates** | At ends of corridor (2x) | Throughout corridor (6x) |
| **Lighting** | 60W red light bulbs on either end | Far-red LEDs throughout the corridor |
| **Video recording** | Spontaneous The Imaging Source camera controlled by Bonsai at 30fps. | Triggered Ximea camera controlled by StreamPix 8 (Norpix) at 100fps |
| **Sucrose rewards** | 45mg pellets | 20mg pellets |
| **Arena lights** | None | One above each arena |
| **Microcontroller** | Arduino Nano | Teensy 3.6 |
| **Barriers** | None | 3 barrier slots in corridor |
| **Moveable reward ports floors** | None | Yes |
| **Faraday cage** | None | Yes |

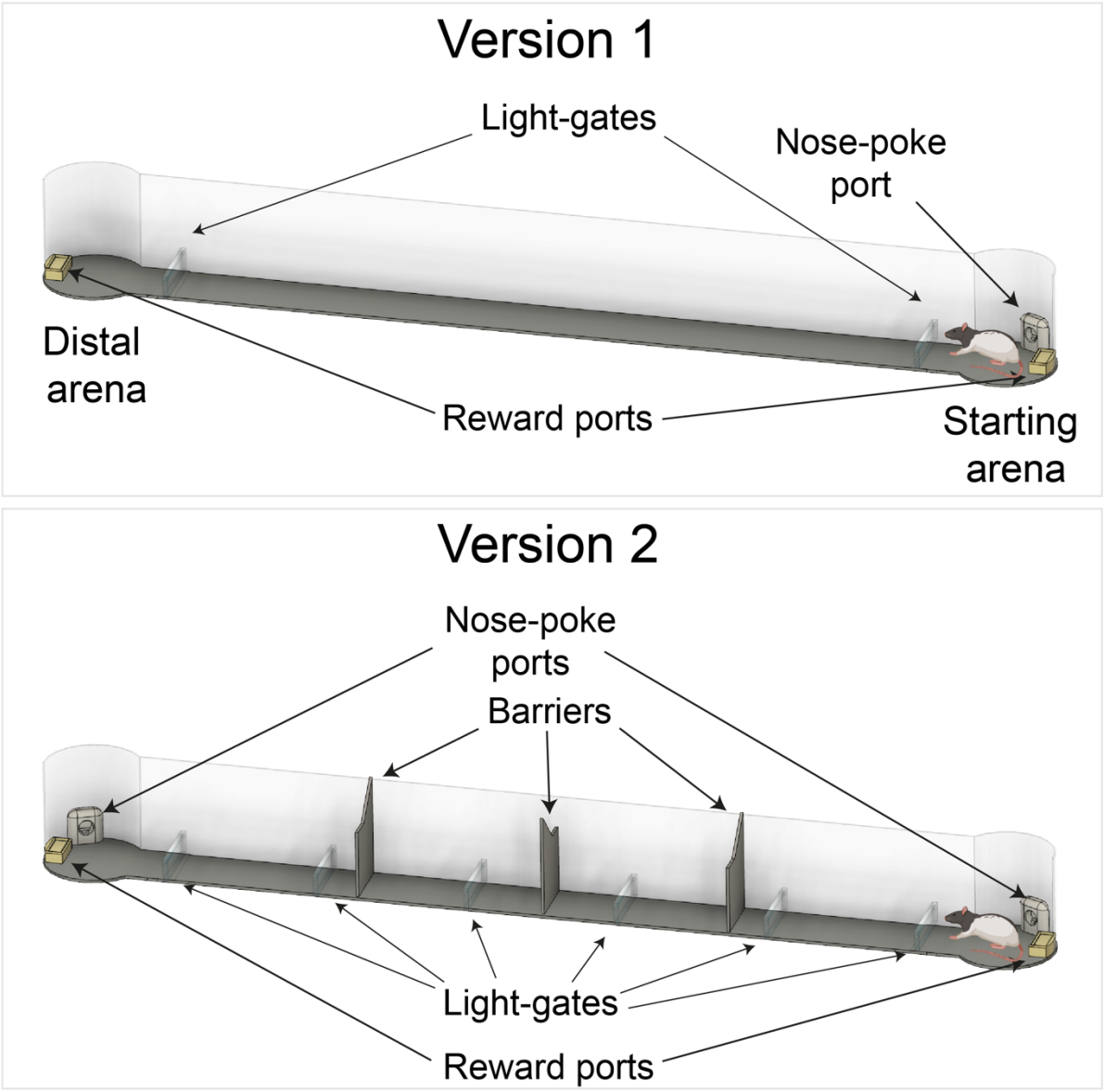



**Figure 1 | Comparison of the two versions of the behavioural apparatus**

## 2.3 Behavioural task

The aim was to set up a paradigm where, upon initiating and sustaining a nose-poke, rats would hear an auditory stimulus predictive of the reward on offer if they chose the effortful option on running along the corridor on that trial. However, there were 2 major differences between the two versions of the paradigm. First, in version 1 but not in version 2, a nearby small default reward was available. Second, in version 2 but not in

version 1, effort was varied orthogonally to reward. Other differences between the two versions are summarised in table 2. Third, in version 1 the starting arena was fixed on one side of the apparatus, whereas in version 2, an animal could start at either end of the apparatus.

**Table 2 | Overview of the main differences between task versions**

| Feature | V1 | V2 |
|---|---|---|
| Reward cues | pulsed tone, siren, white noise | pulsed tone, pure tone, siren, sweep |
| Reward sizes | 0, 1, 2 or 4 sucrose pellets | 0, 1, 3 or 6 sucrose pellets |
| Rewards used | 45 mg flavourless sucrose pellets | 20 mg flavourless sucrose pellets |
| Reward available in the… | starting or distal arena | distal arena only |
| Reward offer variation | random | pseudo-random |
| Effort levels | none | 0-1, 1-2 or 2-3 barriers in corridor, varied in blocks of ~20 trials |
| Forced trials present | yes | no |
| Return to starting arena required | no | yes |
| ITI signalled | no | yes |
| Time allowance for retrieving distal reward | 4-5s | 3s (non-implanted) or 15s (implanted) choice allowance + 5s (non-implanted) or 10s (implanted) running allowance |
| Reward rate adjustment across the sides | Pre-reward delay on starting side equalised to the cohort median distal reward port entry latency | ITI on accept trials reduced by the cohort median distal reward port entry latency |

### 2.3.1 Testing

*V1 (two-alternative choice) testing.* A schematic of this version of the task can be seen in fig.2. Rats initiated trials by poking and holding in the nose-poke port of the starting arena ("pre-cue", fig. 1, upper panel). After a jittered interval of 1.0-1.2s, an auditory

reward cue ("offer") was played from the distal side of the apparatus if the nose-poke was successfully sustained. After the onset of the reward cue, rats were immediately free to leave the nose-poke to either pursue the cued reward in the distal arena or, on choice trials, opt instead to receive a small reward in the proximal arena (fig. 2, upper panel). Rats executed their choices by poking into the proximal reward port or running along the corridor and poking in the distal reward port. Unsuccessful nose-pokes - defined as a nose-poke interruption for more than 100ms (the grace period) - elicited a 2.5s-long error signal (flashing LED in the nose-poke port), followed by a 2.5s ITI. Distal reward prospects were pseudo-randomly varied trial-by-trial (3 sizes: nil, medium, high corresponding to zero, two or four 45-mg sucrose pellets, associated with either a pulsed 3 Khz Hz tone, a siren alternating between 5 and 10 KHz or a 15 KHz pure tone, delivered at ~80 dB with associations counterbalanced between animals). To balance delays to reward on both sides of the corridor, reward delivery in the proximal reward port was delayed by the previous day's cohort median distal reward port entry latency.

10% of medium and high reward trials were “forced” trials, where the proximal reward would not be made available (indicated by the activation of a LED light cue in the distal reward port only). If rats had chosen the distal reward, they needed to run back to the starting arena to initiate the next trial. The data reported in this thesis was collected over 3 consecutive days of behavioural recordings.

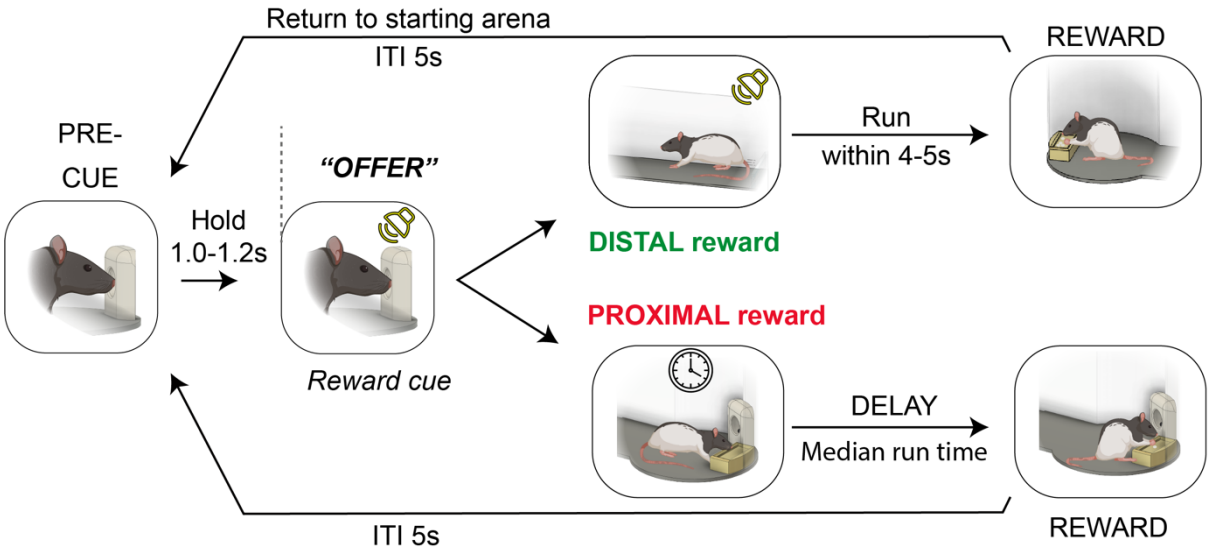



**Figure 2 | Overview of version 1 of the behavioural paradigm.**

*V2 (accept-reject choice) testing.* A schematic of this version of the task can be seen in fig. 3. Rats initiated trials by poking and holding in the nose-poke port of the starting arena ("pre-cue", fig. 1, lower panel). The starting arena was randomly allocated on the 1st trial. After a pre-cue period (0.3-0.4s), an auditory cue ("offer") was again presented from the distal end of the apparatus to signal the available reward at end of the track from where they initiated (fig. 2, lower panel). The available reward was pseudo-randomly varied trial-by-trial (4 sizes: nil, low, mid, high, corresponding to zero, one, three or six 20mg sucrose pellets), signalled by a distinct auditory cue (Fig. 2, lower panel). Rats were required to maintain nose-poking for another 0.6-0.8s until the onset of a visual Go cue, signalled by a white light LED. Premature exiting of the nose-poke either during the pre-cue or offer period would elicit a 5-s error time-out, signalled by a flashing house-light in the nose-poke, followed by a 5s ITI. A grace period was also used in this version of the task to accommodate for minor adjustments in body position during nose-poking. To accept the offer, rats needed to enter the corridor (cross a light-gate 24cm away from the arena) within a time window; otherwise, that offer was rejected. In this version of the task, the cost of pursuing this reward was also pseudo-randomly varied across blocks of 17-23 trials (3 levels: low, moderate, high), implemented using different numbers of barriers in the corridor. The barriers were 0.5 cm thick and were 12 cm and 18 cm tall at their lowest and highest points respectively,

with a 45-degree slant mid-way between to force the rats to run in a zig-zag pattern and increase travel cost (fig. 3).

To ensure that average reward rate was minimally affected by rats' accept/reject choices, ITIs were adjusted on 'accept' trials by subtracting the previous day's cohort median far reward port entry latency. If the offer was rejected, the rat could start the next trial from the starting arena. Alternatively, if distal reward was retrieved, rats could start the next trial from the arena they ran to. This was indicated to the rats by the reward arena floor being elevated 2.5s after reward delivery, resulting in that port being blocked until the next accepted trial.

Non-implanted recording sessions lasted 60 minutes or 120 trials, whichever came sooner. Recordings with implanted animals lasted for 180 trials or 270 minutes, whichever came sooner. The baseline behaviour data reported in this thesis was collected over 10 consecutive days of recordings in each of the two cohorts (n=6 per cohort). Tethered behaviour data was collected over 14-21 days in each animal depending on how quickly electrodes were advanced through the structures of interest.

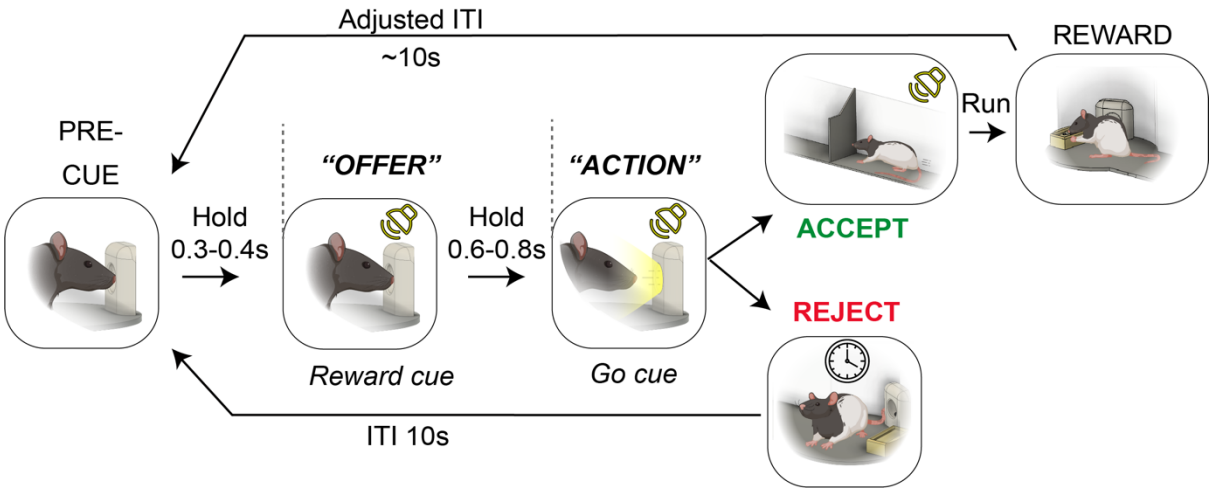



**Figure 3 | Overview of version 2 of the behavioural paradigm.**

### 2.3.2 Habituation and training

Rats were handled daily for a week prior to training. Training consisted of 5 steps in both versions of the task, summarized in table 3.

Step 1 ("magazine training") was similar in both cases and consisted of rats being introduced into the apparatus with bedding from their home cages and receiving 20 rewards in each food dispenser, which they were free to consume in 20 minutes.

In Step 2 ("cue-reward association training") - entry of the starting arena as detected by a light-gate between the arena and the corridor (v1) or entry/presence in the arena of latest reward delivery (v2) would trigger one of 3 (v1) or 4 (v2) auditory cues, predicting the magnitude of the promised reward in the distal reward port. In v1, the offer cue was played and distal reward port LED activated for 5s, regardless of the animal's actions. After this, the offer cue was stopped and reward port LED inactivated, followed by delivery for the corresponding reward. In v2, by contrast, the offer cue was played until the animal made an entry into the distal reward port. Next, the reward port LED was activated for a 1s pre-reward delay, followed by delivery of the reward corresponding to offer cue. 2.5s after reward delivery, the moveable reward port floor was elevated in that same arena, resulting in the blocking of that port to indicate that reward would not be available there on the next trial. After the rats had reached a criterion of initiating 100 (v1) or 120 (v2) trials in ≤ 60 minutes, rats were taken to step 3.

In step 3 ("nose-poke training"), in v1, rats had to learn to sustain a nose-poke with a jittered 1-1.2s length in order to elicit the offer cue. The jittering was implemented to dissociate motor planning from the subsequent processing of the offer cue in electrophysiological recordings. Rats were free to leave the nose-poke port immediately after the onset of the cue and would receive a reward if they ran over to the distal reward port within a jittered 4-5s time window. Jittering here was implemented to create uncertainty about the time allowed for running and speed up distal reward pursuit. Premature exiting of the nose-poke port elicited a 2.5s error time-out signalled by a flashing white light LED in the nose-poke port, after which a new trial could be initiated straight away. However, in order to allow for small changes in body posture that may have caused the nose-poke sensor to stop registering an entry, a 0.1s grace period was introduced which did not result in a nose-poke error. The nose-poke hold requirement was increased by 0.2s each day until the target (1s) was reached.

By contrast, in v2, the reward cue was played 0.3-0.4s after the start of the nose-poke ("pre-cue"), and rats were not free to leave until another 0.6-0.8s later ("offer") once a visual Go cue, signalled by a white light LED in the nose-poke port, had turned on. This was implemented to ensure a movement-free period following offer presentation for electrophysiological analyses. The pre-cue nose-poke requirement was incremented from 0.1s to 0.4s in increments of 25 ms after every successful trial. Once animals had learned to sustain successful pre-cue nose-pokes, the cue nose-poke requirement was incremented from 0.0s to 0.8s in steps of 20 ms after every successful trial. Once animals had learned to complete 120 trials with the full 1.2s nose-poke requirement, jitters of 0.1s ("pre-cue") and 0.2s ("offer") were introduced. Importantly, if a nose-poke was interrupted for more than 0.1s (the grace period) during the "offer" window, the same offer cue would be repeated during the next trial. This was not an intentional feature, but rather the side-effect of using a pre-determined sequence of offers. Once rats had reached a criterion of receiving 100 rewards (v1) or initiating 120 trials (v2) within 60 minutes, they were taken on to step 4.

Step 4 consisted of either "retrieve proximal reward" in v1 or "barrier training" in v2. In "retrieve proximal reward" (v1), rats learned that the house-light in the reward port of the starting arena, which illuminated after a variable interval, predicted the delivery of 1 pellet upon entry within 10s. At this stage the distal reward port as well as the nose-poke port were blocked. The promise of a reward to come, indicated by a white light LED in the reward port, was signalled throughout a pre-reward delay, defined as the interval between reward port head entry and reward delivery. This delay was incremented from 0 to 2.3s within one training session.

In "barrier training" in v2, rats needed to nose-poke to elicit offer cues as outlined previously, but now needed to additionally learn to scale barriers when traversing the corridor between the arenas. Barriers were added in one by one in blocks of 40 trials, starting from 1 barrier and going up to a maximum of 3 barriers. Once rats had successfully managed to retrieve 50 rewards from the reward port of the starting arena (v1) or complete 120 trials within 60 mins (v2), they were taken on to step 5.

In step 5 ("final training"), the rats were required to hold a 1-1.2s long nose-poke in order to trigger the offer cue. In v1, this signalled the promise of either a nil, medium or large reward, with the order of presentation varied randomly in proportions of 0.33/0.33/0.33. Note that the small reward offer (corresponding to one 45 mg sucrose pellet) was no longer used in this version, as this was now used as proximal reference option instead. However, all offer cues were used in v2 and were presented pseudo-randomly in proportions of 0.1/0.3/0.3/0.3. In the following, I outline the details of the rest of the trial separately for each version.

*V1 final training* Rats were trained to choose between running to the distal arena to retrieve the offered reward or obtaining a fixed small reward from the reward port of the starting arena (fig. 2). Entry of the distal arena during the cue interval on a medium or large reward trial triggered immediate reward delivery, followed by a 5s ITI. On choice trials, entry into the reward port of the starting arena within 4-5s triggered delivery of 1 pellet after a pre-reward delay (signalled by a house light in that food dispenser) that, for each training day, was equalised to the previous day's cohort median food dispenser B entry latency. Crucially, 25% of medium and large reward trials in v1 were "forced", meaning that reward could only be retrieved from the distal arena and the proximal reward port was unresponsive. Forced trials were signalled by the activation of the white light LED in the distal reward port. If neither entry to arena B nor food dispenser A was detected during the cue period, the cue would turn off and a 5s ITI would begin. Rats were considered ready for testing if they had managed to collect 100 rewards within 60 min and were at least 60% successful in eliciting reward delivery on every trial type.

*V2 final training - non-implanted animals* Rats were required to either accept or reject a distal reward offer available at the other side of the corridor (fig. 2). After the onset of the Go cue, to accept the offer, rats needed to enter the corridor within a 3s time window as determined by a light-gate placed near the border of the arena; otherwise, that offer is rejected. The effort cost which was fixed across blocks of ~20 trials (3 levels: low – 1 barrier; moderate – 2 barriers; high – 3 barriers; table 3). Upon corridor entry, they were allowed a further 5s to traverse the corridor and reach the distal reward port. To ensure

average reward rate was minimally affected by rats' accept/reject choices, ITIs were adjusted on 'accept' trials by subtracting the median run time from 10s (the ITI length on reject trials).

*V2 post-surgery re-training* To account for the additional constraints on implanted rats with a recording tether, the corridor entry and traverse times were increased to 15s and 10s respectively post-surgery. Additionally, to avoid collisions between the implant and the barriers, the number of barriers per each level of effort was reduced by 1. Rats were considered ready for the recording experiment once they managed to initiate 120 trials within 60 minutes.

**Table 3 | Overview of training steps and criteria for moving onto the next step**

| **Version 1** | | **Version 2** | |
|---|---|---|---|
| Training step | Criterion for passing | Training step | Criterion for passing |
| **1: Magazine training** | Each reward dispenser visited ≥ 1 time(s). | **1: Magazine training** | None |
| **2: Cue-reward association training** | 100 pellets retrieved in ≤ 60 minutes. | **2:** Cue-reward association training | 120 trials completed in ≤ 60 minutes. |
| **3: Nose-poke training** | 100 pellets rewards in ≤ 60 minutes. Ability to reliably sustain a 1-1.2s nose-poke. | **3. Nose-poke training** | 120 trials completed in ≤ 60 minutes. Ability to reliably sustain a 1-1.2s nose-poke. |
| **4: Retrieve proximal reward** | 50 pellets retrieved in arena A in ≤ 60 minutes. | **4. Barrier training** | 120 trials completed in ≤ 60 minutes. |
| **5: Final training** | 100 pellets retrieved in ≤ 60 minutes and ≥ 60% forced trial success rate. | **5. Final training** | 120 trials completed in ≤ 60 minutes. |

## 2.4 Outcome devaluation

Following standard testing, outcome devaluation was performed on a subset of rats (n=6, first cohort) in a within-subjects regular Latin square design. In devaluation

sessions, subjects were placed in a solitary "feeding" cage with *ad libitum* lab chow for 1h prior to the start of the testing session. Rats were placed in a solitary "control" cage for 1h on control sessions. The outcome devaluation lasted 3 days altogether, out of which days 1 and 3 were testing days and day 2 was a baseline training day.

## 2.5 Microwire electrode implant preparation

The electrode drive used in this thesis was a re-design of an implant developed by Colin McNamara and Max Rothwell. It is a bespoke ~20g device that allows the implantation of moveable microelectrodes to enable recordings across any combination of brain regions. The implant was comprised of a lid which housed a 128-pin connector designed for Intan 64-channel head-stages (Intan Technology), a plastic corpus that housed the electrode movement mechanisms, glass tubing and a base (fig. 4). All components were assembled in house. Each implant took ~2-3 weeks to build. For assembly, the elements of the corpus were first 3D printed and post-processed with UV radiation and propanol. Second, holes were drilled in the corpus to allow threading of electrode wires and the driving mechanisms. Third, driving support mechanisms were created using metal rods, plastic shuttles and screws. Fourth, glass tubing was introduced into the driving mechanisms as electrode supports. Fifth, strands of 50-micron diameter tungsten wire (California Fine Wire) were twisted and heated to create 4- and 6-strand electrodes (tetrodes and hextrodes, respectively). Sixth, these were threaded through the glass tubing and the rest of the implant corpus. Seventh, ground wire stubs (connected to ground wires on the animal during surgery, see below) were introduced into the implant using a dedicated set of stationary holes. Finally, the implant was sealed and the electrodes cut to length by hand.

Three of the subjects (gh001, gh009, gh010) reported in this work received an implant that targeted all 5 regions, whereas one subject (oh011) received a modified drive without the VP electrodes (table 4). Electrodes in the mOFC, ACC, and VP were arranged as a row along one sagittal plane. DMS electrodes were as rows along two sagittal planes and STN electrodes were arranged in triangular fashion around a centre point (fig. 5). The ML and AP positions of each electrode in the drive were planned

around a centre-point co-ordinate obtained as an average between readings in two different versions of a commonly used rat brain atlas (Paxinos and Watson 1998, 2006) that visually offered the most exposure in each structure, while also avoiding ventricles very superficial cortical positions (fig. 4). Not all 128 channels were used in the implants for this thesis due to spatial geometry constraints within the device. Relative electrode positions in the AP and ML dimensions were planned around centroid co-ordinates

**Table 4 | Overview of the number of electrodes and channels in each subject and targeted recording area.**

| | **Number of electrodes (channels) in targeted area** | | | | |
|---|---|---|---|---|---|
| | MOFC | ACC | DMS | VP | STN |
| **oho11** | 4 (32) | 4 (16) | 5 (30) | 0 (0) | 3 (18) |
| **gho01** | 4 (32) | 4 (16) | 4 (24) | 2 (12) | 3 (18) |
| **gho09** | 4 (32) | 4 (16) | 5 (30) | 2 (12) | 3 (18) |
| **gho10** | 4 (32) | 4 (16) | 5 (30) | 2 (12) | 3 (18) |

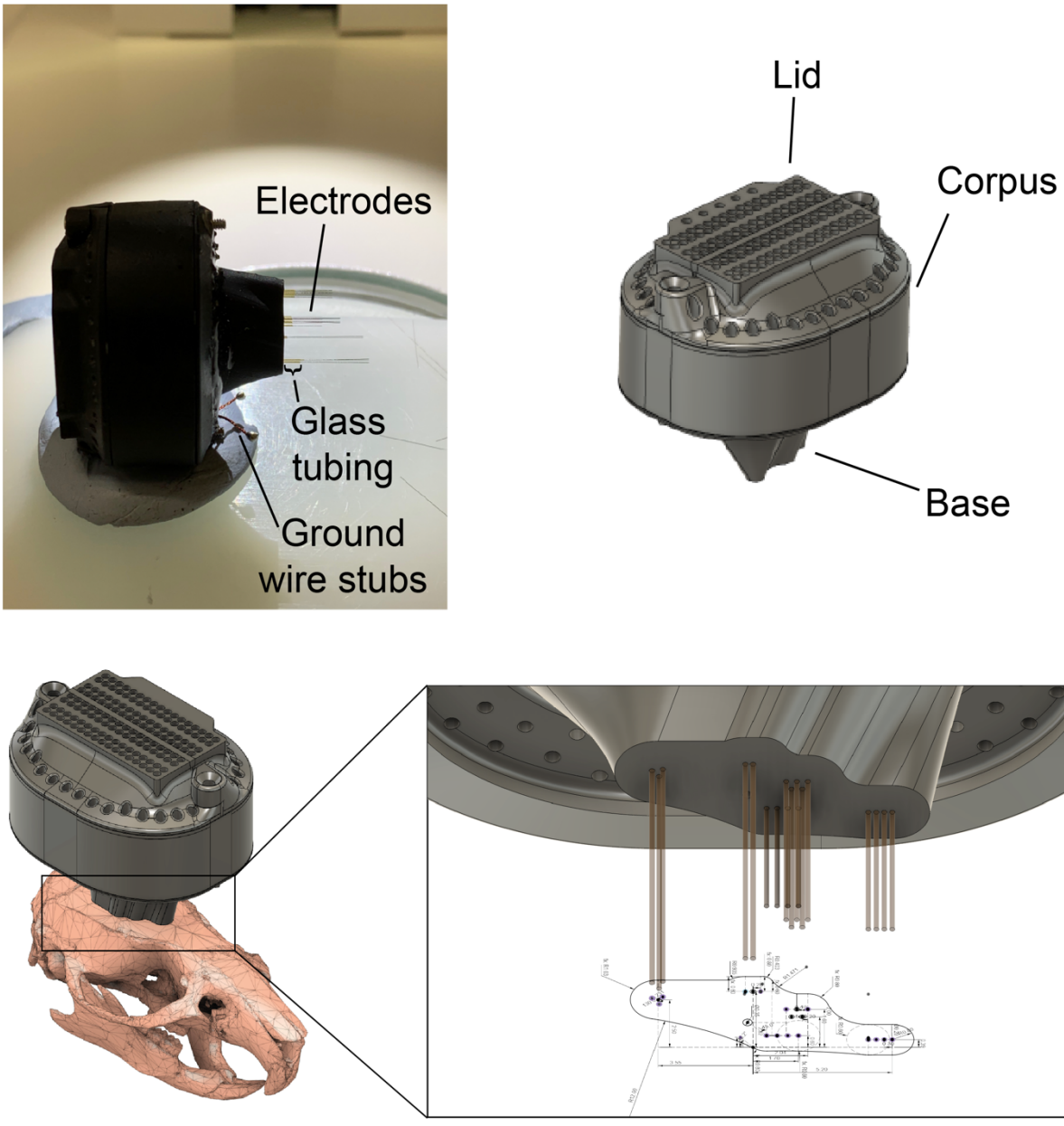



**Figure 4 | Overview of the drive design**. **Top left,** photograph of the implant with electrodes, glass tubing and ground wire stubs marked. **Top right,** image of the 3D design of the implant with the main constituents of the plastic body highlighted. **Bottom left,** schematic of the implant in comparison to an average rat skull model. **Bottom right,** schematic of the drive bottom along with the sketch used to design electrode placements.

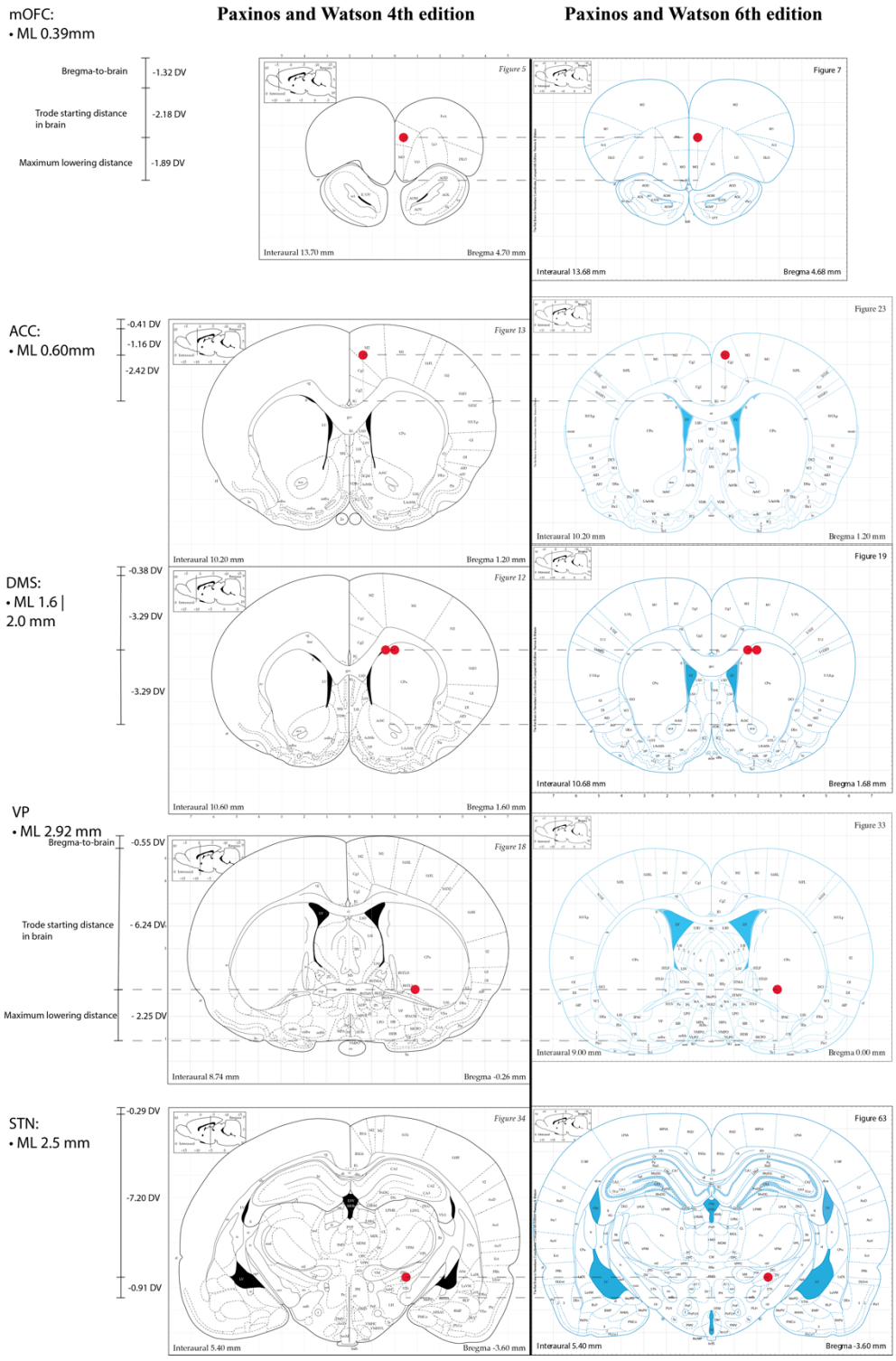



**Figure 5 | Centre-point co-ordinates used for the configuration of implant electrode positions**. Centroids (red dots) and the planned electrode tracts (dotted vertical lines beneath) are marked on the schematics of each structure from two rat brain atlas versions (Paxinos and Watson 1998, 2006).

## 2.6 Microwire electrode implant surgery

Electrode implantations were performed in a subset of animals (n=5). Surgeries were performed under isoflurane anaesthesia (5% for induction, 1.5-2.5% maintenance). Additionally, 2% lidocaine and 1 mg/kg meloxicam were administered subcutaneously prior to scalp incision. Craniotomies were created above each of the targeted areas and *dura mater* thoroughly removed from the visible brain surface. The drive was aligned on the animals' heads by using the centre-point of the STN electrode bundle as a reference. Next, the drive was lowered to target depth on a stereotaxic frame (David Kopf Instruments) based on the DV reading of the STN electrode bundle from brain surface. The implant was then secured to skull-attached screws using adhesive dental cement and dental acrylic. Implant ground wire stumps were soldered to supra-cerebellar skull-attached ground screws. Neuronal recordings began at least 2 weeks after surgery. Electrode placement was confirmed using histology (below). 0.03 mg/kg burprenorphine was administered on the day of the surgery as well as 2 days post-operatively for pain relief.

## 2.7 Histology

To verify electrode placements, after the experiments were completed, animals were deeply anaesthetised with intraperitoneal 200 mg/kg sodium pentobarbitone and perfused trans-cardially with phosphate buffered saline (PBS) followed by 10% formaldehyde solution (PFA, vol/vol). Heads were submerged in 10% PFA for 24h, following which brains were extracted and washed 3 times in PBS and stored in a PBS solution containing 0.02% sodium azide until sectioning. Brains were sectioned at 0.1mm thickness using a vibratome (Leica), immunohistologically stained and mounted in Vectashield. However, only unfiltered "bright-field" epifluorescence microscope imaging were used for this corpus of work, obtained using 5x magnification and Zen Blue Pro (Zeiss).

## 2.8 Data acquisition and storage

*Behaviour V1* For each day and animal, a novel sequence of trial types was created using the random number generator on Matlab (Mathworks). This was saved into a text file

and uploaded to the Arduino Nano to administer the behavioural session. Behavioural tracking data was recorded on the Arduino IDE communication port dialogue window in the form of a text log, later saved as text files and processed in Matlab.

*Behaviour V2* A bespoke graphical user interface (GUI) was written to operate custom-built operant box circuitry. The operant box circuitry was designed to allow flexible programming of novel behavioural tasks and control of modules such as speakers, reward dispensers, light-gates, LED lights and reward dispenser floors. The GUI, in turn, offered a user-friendly interface for the (i) creation and storage of session meta-data, including jittered pre-cue and cue nose-poke requirement lengths, sequences of effort and reward offers (fig. 4a, b) and (ii) logging and storage of raw behavioural data in binary format (fig. 4c), ensuring reproducible session data acquisition across multiple experimenters.

*Video* For v1, video was recorded at 30 fps on Bonsai. By contrast, video on v2 was recorded using triggerable Ximea cameras at 100fps on Streampix 8 (Norpix). Trigger signals originated from the Teensy 3.6 microcontrollers. Files were acquired in monochrome format and saved as mp4.

*Electrophysiology* Wideband and accelerometer data were recorded, amplified and digitised on 2 tethered Intan 64-channel headstages at 20,000 Hz using an Intan acquisition board (Intan Technologies). No hardware filtering was applied during recording. Files were saved in binary format. Each electrode was advanced to a new DV position at least 2h ahead of each recording session to ensure identification of unique units every day.

## 2.9 Data pre-processing and spike sorting

Clustering and curation of units were carried out in KiloSort2 (Pachitariu et al. 2016) and phy (available at https://phy.readthedocs.io/en/latest/), respectively. Both are open-source libraries optimised for high-density multielectrode recordings. Individual units were isolated from the broadband signal based on waveform morphologies and

principal component analysis. Any units that were not stable over the recording session (as indexed by the principal component weightings over time) were discarded. No multi-units were used in the reported set of analyses. Neurons were not grouped based on background firing rates of waveform morphologies.

## 2.10 Data analyses

*Behavioural measures: v1* All data were analysed using Matlab (Mathworks). The primary measure of interest was proportion of **distal reward choice**, defined as the percentage of trials where rats went to the distal reward port within the available response window. To determine the predictors of distal reward choice, we carried out a binomial mixed effects' regression analysis (using the function "fitglme" in Matlab) where task variables were regressed onto distal reward choice data on a subject-by-subject basis. The regression models for each animal took the following format:

$$\text{distal reward choice} = \text{distal reward prospect} + \text{choice/forced trial} + \text{reward rate} + \text{reward count} + (1 \mid \text{date})$$

Here, distal reward prospect was coded as 0, 2 or 4. The choice/forced nature of the trial was coded as 0 for choice trials and 1 for forced trials, To determine the length of reward history rats might be influenced by, rewards accrued on trial t-1, the preceding 5, 10, 15, …, 30 trials and 1 minute (coded as the number of pellets received over those intervals of interest) were fit onto distal reward choice. The predictor which was significant on the group level and had the biggest absolute cohort average regression coefficient $\bar{\beta}$ was entered into the final regression. Finally, reward count was coded as the number of pellets received over the session. All predictors were standardised to facilitate visualisation of their regression weights side-by-side. To account for potential day-by-day differences in behaviour, date was used as a grouping variable (Austin et al. 2017). Fitting was allowed to undergo 10,000 rounds Laplace estimation. The group level significance of the obtained regression coefficients was analysed using a two-sided t-test against zero. Averaged coefficients are represented using the symbol $\bar{\beta}$.

*Behavioural measures: v2.* All data were analysed in Python. The primary measure of interest was the **proportion of accepted trials**. Secondary outcomes of interest were the **action initiation latency** - defined as the interval between the onset of the Go cue and subsequent nose-poke head-exit - and **corridor entry latency** - defined as the interval between action initiation and entry into the corridor and **premature head-exits,** defined as the number of unsuccessful nose-poke attempts during the offer cue before the trial was successfully initiated. A binomial mixed effects model using the pymer4 library (Jolly 2018) was used to regress the following fixed effects onto trial-by-trial choice:

- *reward*, coded as 0, 1, 3 or 6
- *effort*, coded as 1, 2, or 3 (0, 1, or 2 in implanted animals)
- *reward x effort interaction*, coded as the multiplication of those vectors
- *accept t-1* (acceptance of an offer on the previous trial), coded as -0.5 or 0.5 for rejected vs accepted trials respectively
- *reward count*, coded as the number of pellets earned since the start of the session.

To safeguard against type I errors, we aimed to fit a maximal random effects structure whenever possible. This entailed modelling each predictor as a random, in addition to a fixed, slope (Barr et al. 2013). To account for covariance resulting from animals' potential side biases, a random slope was included for the side of the apparatus when starting a trial (coded as 0.5 or -0.5 for the right and left side, respectively). To control for any biases resulting from rats' prediction of the offer (following unsuccessful nose-poke attempts during the offer cue before the trial was successfully initiated), a random slope was also introduced for premature head-exits (coded as 0 or 1 where none or any instances had been registered before the onset of the trial, respectively). All predictors were standardised.

To account for day-by-day and subject-by-subject covariation in the data, each of these random slopes was paired with a random intercept for either the testing day or the subject ID. Together, the initial baseline behaviour model for both non-implanted and implanted cohorts had the following formula:

$$\text{accept} = \text{reward} \bullet \text{effort} + \text{accept}_{t-1} + \text{reward count} + (\text{random slopes} \mid \text{date}) + (\text{random slopes} \mid \text{subject})$$

The same set of fixed and random effects were also regressed onto premature head-exits, coded as 0 or 1 as mentioned above.

For the outcome devaluation experiment, we added pre-feeding and its interaction with reward and effort as an additional predictor, resulting in the following initial formula:

$$\text{accept} = \text{reward} \bullet \text{effort} \bullet \text{pre-feeding} + \text{accept}_{t-1} + \text{reward count} + (\text{random slopes} \mid \text{date}) + (\text{random slopes} \mid \text{subject})$$

To analyse the factors influencing action initiation and corridor entry latencies, the same fixed and random effects' structures were fitted onto the latency outcomes using a linear mixed effects' model from the pymer4 library (Jolly 2018).

Whenever model fitting failed to converge, random effects that fit close to zero variance were removed to simplify the random effects' structure. Model results are plotted as coefficient weightings ± standard errors. P-values were obtained using the likelihood ratio test.

## 3. Results

### 3.1 Paradigm v1

We first set up a behavioural paradigm where rats needed to decide between a variable distal large reward payoff or a proximal small reward (see methods, fig. 6A-B). Rats nose-poked to receive an auditory cue associated with the promise of either a nil, medium or large reward (zeros, two or four sucrose pellets, respectively) if they ran a 1.3m-long corridor to the distal reward port within 4-5s. On choice trials, rats could also opt for a small reward (1 sucrose pellet) in the proximal reward port. Distal reward offers were had a substantial numerical influence on rats' (n=6) choice behaviour (fig. 6C-D).

However, they did not always choose to go to the distal food dispenser on forced trials despite that being the only rewarded option on those trials, resulting in an average success rate of just 60%. This was driven by a significantly higher number of proximal food dispenser head entries on forced medium vs large reward trials (main effect of forced choice trial type: $F_{1,5}$=7.362, p=0.042), potentially indicating that on medium reward trials rats were not paying as much attention to the distal reward port light cue that signalled a forced trial. There was no effect of testing day during the 3 days of data acquisition, ($F_{2,8}$=0.24, p=.789, data not shown), indicating stable performance over time.

To quantify the effect that the distal reward promise had on rats' decisions to exert effort, we regressed two kinds of predictors against the probability of running. These variables included ones that were explicitly manipulated (distal reward, forced/choice trial), as well as ones that resulted from the rats' interaction with the task, such as the history of rewards and reward count. To determine how long a reward history to consider in my main regression, we first iteratively regressed the sum of rewards earned over the preceding 1, 5, 10, 15…30 trials and 1 minute against the probability of running. This revealed that the preceding 10 trials weighed in most heavily on rats' choice to run (Fig. 6E, $\bar{\bar{\beta}}$= 0.074 ± 0.011, $t_5$=6.993, p < 0.001). The main regression revealed that running probability was significantly modulated by distal reward ($\bar{\bar{\beta}}$ = 1.785 ± 0.600, $t_5$=2.974 p=0.031) and whether the trial was a forced or a choice trial ($\bar{\bar{\beta}}$ = 0.589 ± 0.128, $t_5$=4.600, p=0.006) (fig. 6F).

Together, these results show that rats' motivation to pursue reward can be examined in a paradigm that involves a decision between a proximal small vs a potentially larger distant reward. However, rats were heavily biased to seek reward in the proximal reward port under all circumstances, resulting in distal reward choice rates of only 25% and 50% of medium and large reward trials and a reward retrieval rate of just 60% on forced choice trials.

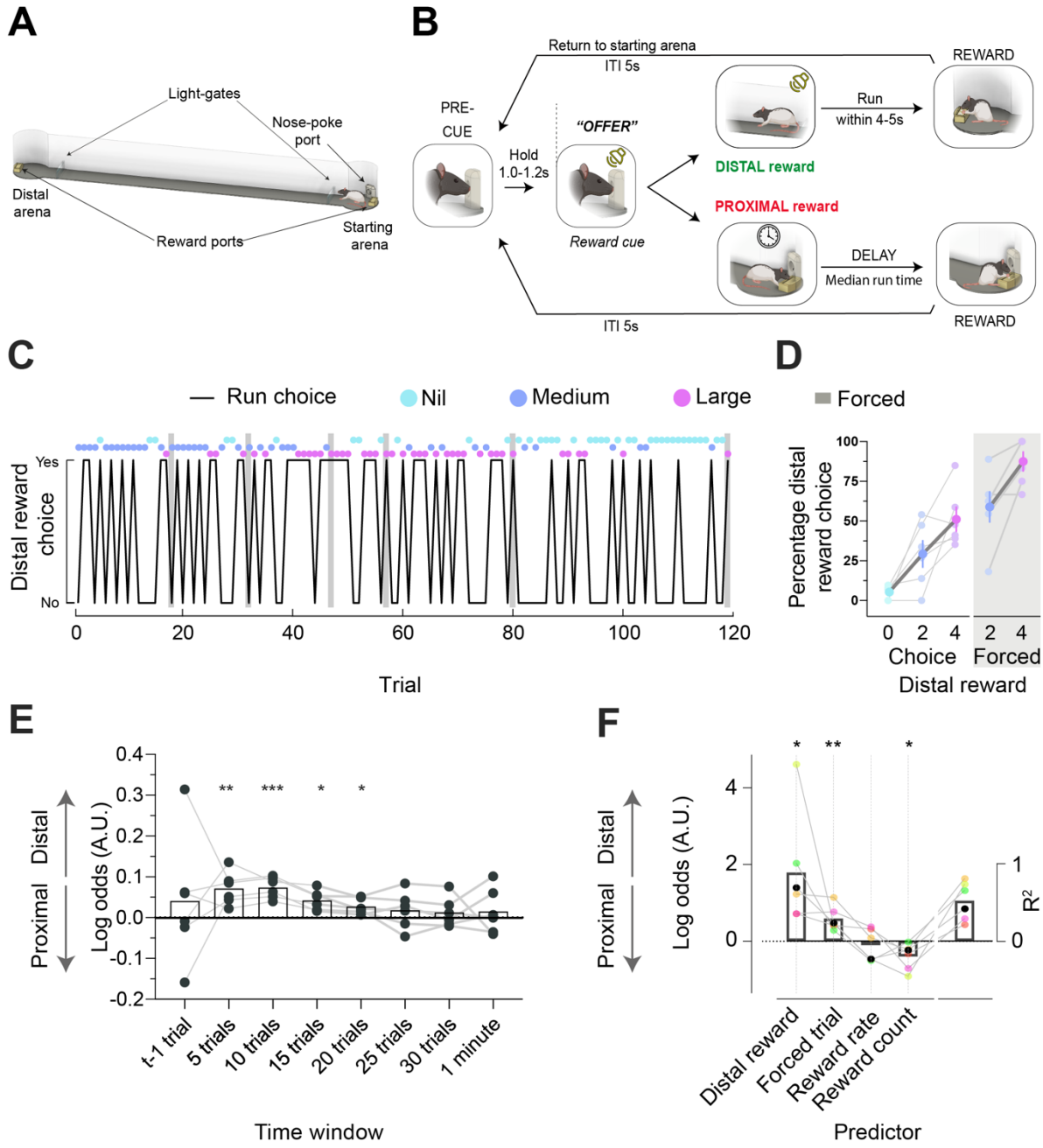



**Figure 6 | Analysis of the determinants of effortful choice on a 2 alternative choice paradigm (n=6). A-B,** rats choose between reward available at the proximal vs distal reward port. **C,** trial-by-trial distal reward choice (solid line) and reward prospect variation in an example recording session. Distal reward prospects are depicted as coloured dots and forced trials are highlighted with dark shading. **D,** percentage of trials where the distal reward was chosen, depicted as a function of reward prospect. Individual datapoints (thin lines and dots) are superimposed by means and SEMs (thicker lines and larger dots). **E,** regression weights of a set of reward rates predictors in iterative binomial regressions predicting the probability of choosing distal reward. Thin lines and dots depict individual rat data, superimposed bars depict the group means. **F,** Left: regression weights of predictors entered into the main regression,

depicted for each rat (thin lines and dots) and superimposed by the group means (bars). Right: $R^2$ values obtained from the regression model. *p < 0.05, **p < 0.01, ***p < 0.001, t-test against null-distribution.

### 3.2 Paradigm v2

The low rate of distal reward preference on choice trials, potential ambiguity about the choice/forced nature of the trial and the inability to modulate effort cost prompted us to re-evaluate the task design. In v2, we thus sought to complement the trial-by-trial variation of reward prospects with an orthogonal manipulation of effort and abolish the prospect of a small reward in the proximal food port (fig. 7A-B, methods). Rats initiated trials by poking and holding in a nose-poke port (located at both ends of the track in this version) while an auditory cue was presented, which signalled the reward available at the far end of the track. After a Go cue, to accept the offer, rats needed to enter the corridor within a time window; otherwise, that offer was rejected. effort cost was fixed across blocks of ~20 trials (3 levels: low, moderate, high), implemented using different numbers of barriers in the corridor.

Rats' (n=12) choices were sensitive to the benefits and costs of acting: they were more likely to pursue reward when it is larger or when the effort cost is lower (Fig. 7D, left panel). To quantify the sizes of these effects, gauge the influence of decision history and factor in a proxy of motivational state, a binomial mixed effects' regression was carried out. This revealed a strong positive influence of reward ($\beta = 1.31$, $p = 6.20 \times 10^{-23}$) and a negative influence of effort ($\beta = -0.38$, $p = 2.86 \times 10^{-9}$) over choice (fig. 7D, right panel). There was no significant effect of reward ∗ effort interaction, reward count or previous decision (all $|\beta| < 0.08$ , $p > .107$). Though the history of decisions was not a robust determinant of choice, a separate binomial mixed effects' lag regression revealed that the size of the previous two reward offers had a *negative* influence over current choice (fig. 7E).

The lack of effect of reward count might suggest that rats were either (i) not titrating their choices to changing outcome values or (ii) not accruing enough calories to reach satiety. To rule out the former, we manipulated outcome value in a counterbalanced

Latin square design by giving a subset of rats (n=6) unlimited access to food chow for 1h prior to testing. This robustly reduced reward pursuit (Fig.7F, effect of pre-feeding, $\beta = -0.74$, $p = 7.16 \times 10^{-13}$).

Latencies provide a further proxy of offer evaluation on this task. Indeed, action initiation and corridor entry latency both numerically scaled with the level of effort, however while action initiation was slowed by increasing reward, the opposite was true for corridor entry latency (fig. 7G-H). This was reflected in a positive influence of effort on both latencies ($\beta = 21.72$, $p = 2.64 \times 10^{-10}$ and $\beta = 107.72$, $p = 8.60 \times 10^{-117}$ respectively for action initiation and corridor entry latencies), but opposing effects of reward ($\beta = 66.98$, $p = 1.22 \times 10^{-84}$ and $\beta = -21.63$, $p = 2.68 \times 10^{-6}$) in a linear mixed effects model. To investigate this pattern of effects on the action initiation latency may have resulted from increased inhibitory drive as a function of reward, a regression to analyse the determinants of premature nose-poke errors was carried out. This revealed reward to be positively ($\beta = 0.11$, $p = 4.21 \times 10^{-7}$) and effort negatively ($\beta = -0.09$, $p = 1.57 \times 10^{-4}$) associated with premature nose-poke errors during the offer window.

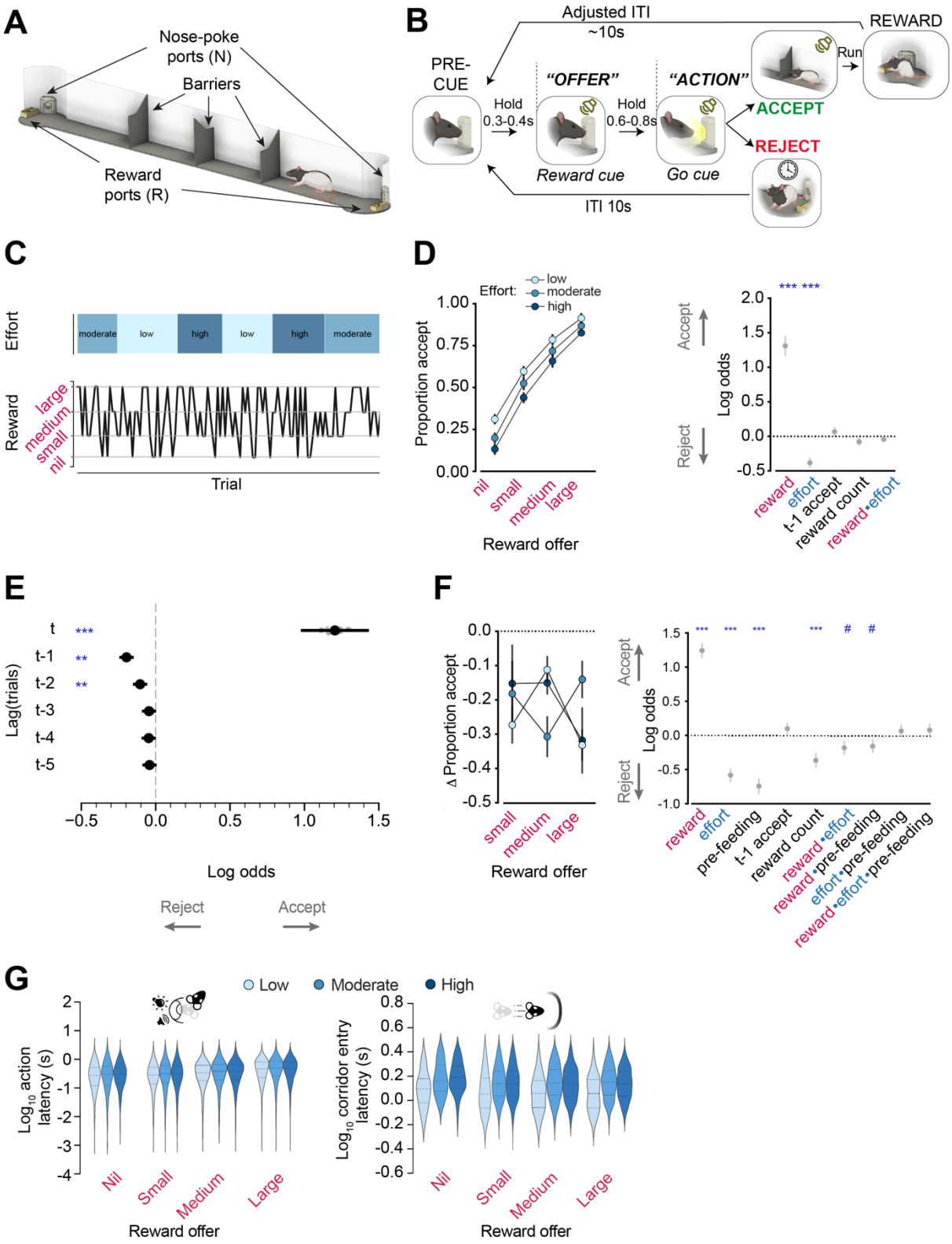



**Figure 7 | Analysis of the determinants of effortful choice on an accept/reject choice paradigm. A-C,** Rats choose between reward available at the far end of a linear track ('accept') vs opting out ('reject'). Reward magnitude (4 levels, lower panel C) is signalled trial-by-trial by an auditory cue, whereas effort (3 levels, upper panel C), is expressed in blocks as the number of barriers in the corridor. **D,** Accept rates (left panel) and choice predictor weightings from a

binomial mixed effects model (right panel) in food-restricted rats (n=12). **E,** Predictor weightings from a lagged binomial mixed effects model. **F,** Accept rate differences when giving rats (n=6) *ad libitum* access to their regular diet for 1 hour before testing vs. not pre-feeding (left panel) and predictor weightings from a binomial mixed effects' model (right panel). Coloured dots in D, F indicate mean ± SEM, adjusted for the within-subject variance across all levels of effort. Grey dots indicate mean ± SEM. **G,** Log-transformed action initiation (left panel) and corridor entry (right panel) latencies, grouped by reward and effort. Data are depicted as kernel density estimates (violin bodies) of all latency instances, superimposed by the quartiles of the data (dotted horizontal lines). # .09 < p < .05; ** p < .01; *** p < .001.

### 3.3 Paradigm v2 behaviour is resilient to electrophysiology-related modifications

To ensure that the pattern of cost-benefit choices in implanted rats (n=4) remained similar despite (i) tethering (ii) potential tissue damage in areas important for value-guided decision-making (iii) changes in the decision latency (3s in non-implanted animals vs 15s in implanted animals) and run allowance (5s in non-implanted animals vs 10s in implanted animals), and (iv) unit decrease in all levels of effort, we re-analysed choices in all electrophysiological recording sessions. This revealed a preserved pattern of accept choices (fig. 8A-B), though implanted animals' decisions were more sensitive to effort ($t_2$ = 4.42, p = .048, paired t-test between coefficients in non-implanted vs implanted animals) and were further influenced by a reward x effort interaction (β = -0.089, p = .001).

Action latencies on the task provide a valuable resource to analyse time-evolving neurophysiology signals in relation to ongoing behaviour. The median action initiation latency was 0.76 ± 0.01s, which together with the Go cue onset latency (0.6-0.8s) provided ample time to analyse cost-benefit processing without the presence of gross movement (fig. 8C, left panel). The median corridor entry latency on accepted trials was 5.94 ± 0.03s (mean ± SEM, fig 7C, right panel) - a sizeable window of opportunity to study the neural antecedents of preparation for reward pursuit. Similar to non-implanted rats, prospective effort slowed both latencies (β = 61.46, p = 2.60 x $10^{-6}$ and β = 530.17, p = 1.63 x $10^{-65}$ respectively for action initiation and corridor entry), whereas

reward slowed action initiation ($\beta$ = 150.94, p = 5.63 x $10^{-32}$) but speeded corridor entry ($\beta$ = -258.70, p = 9.71 x $10^{-18}$).

Together, these data demonstrate the feasibility of an accept-reject type cost-benefit decision-making paradigm to study motivation in rats.

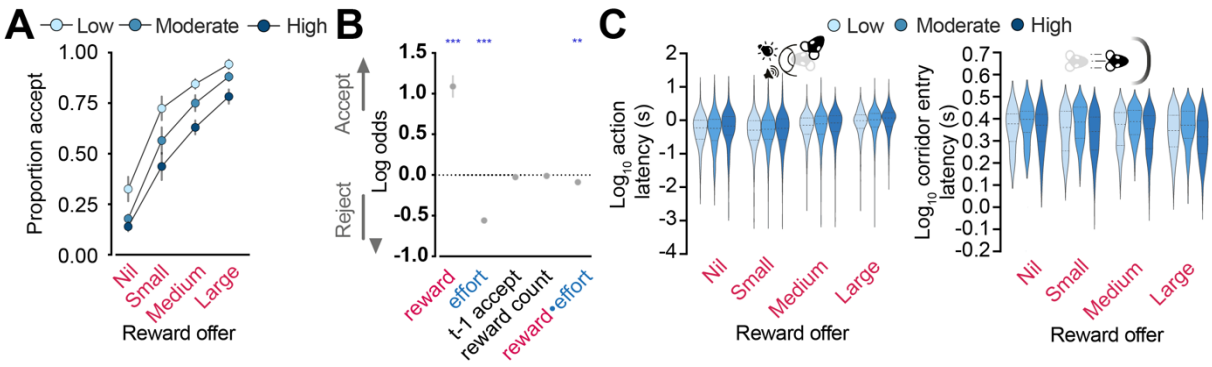



**Figure 8 | Core characteristics of cost-benefit decision-making persist after surgical implantation of a microelectrode device A,** Accept rates as a function of reward and effort. Coloured dots indicate mean ± SEM, adjusted for the within-subject variance across all levels of effort. **B,** choice predictor weightings from a binomial mixed effects' model. Data are depicted as mean ± SEM. ** p < .01; *** p < .001. **C,** Log-transformed action initiation (left panel) and corridor entry (right panel) latencies, grouped by reward and effort. Data are depicted as kernel density estimates (violin bodies) of all latency instances, superimposed by the quartiles of the data (dotted horizontal lines).

## 3.4 Electrophysiological signal quality

To index into the quality of the recorded electrophysiology data, we carried out spike sorting to extract individual neurons (fig. 9A, C) and analysed local field potentials fig. (8B). Analysis of power spectral densities obtained using the Welch method in one randomly picked animal-session revealed smooth distributions compatible with a power law-like shape in each recording structure (Podvalny et al. 2015) and no elevations indicative of alternating current contamination. These signals were conducive to the extraction of large numbers of individual units in each area (fig. 9C), including 603.75 ±

215.76 in the DMS (mean ± SEM over subjects),246 ± 90.42 in the VP, 244 ± 74.17 in the ACC , 157.50 ± 70.80 in the MOFC and, and 122.25 ± 44.22 in the STA.

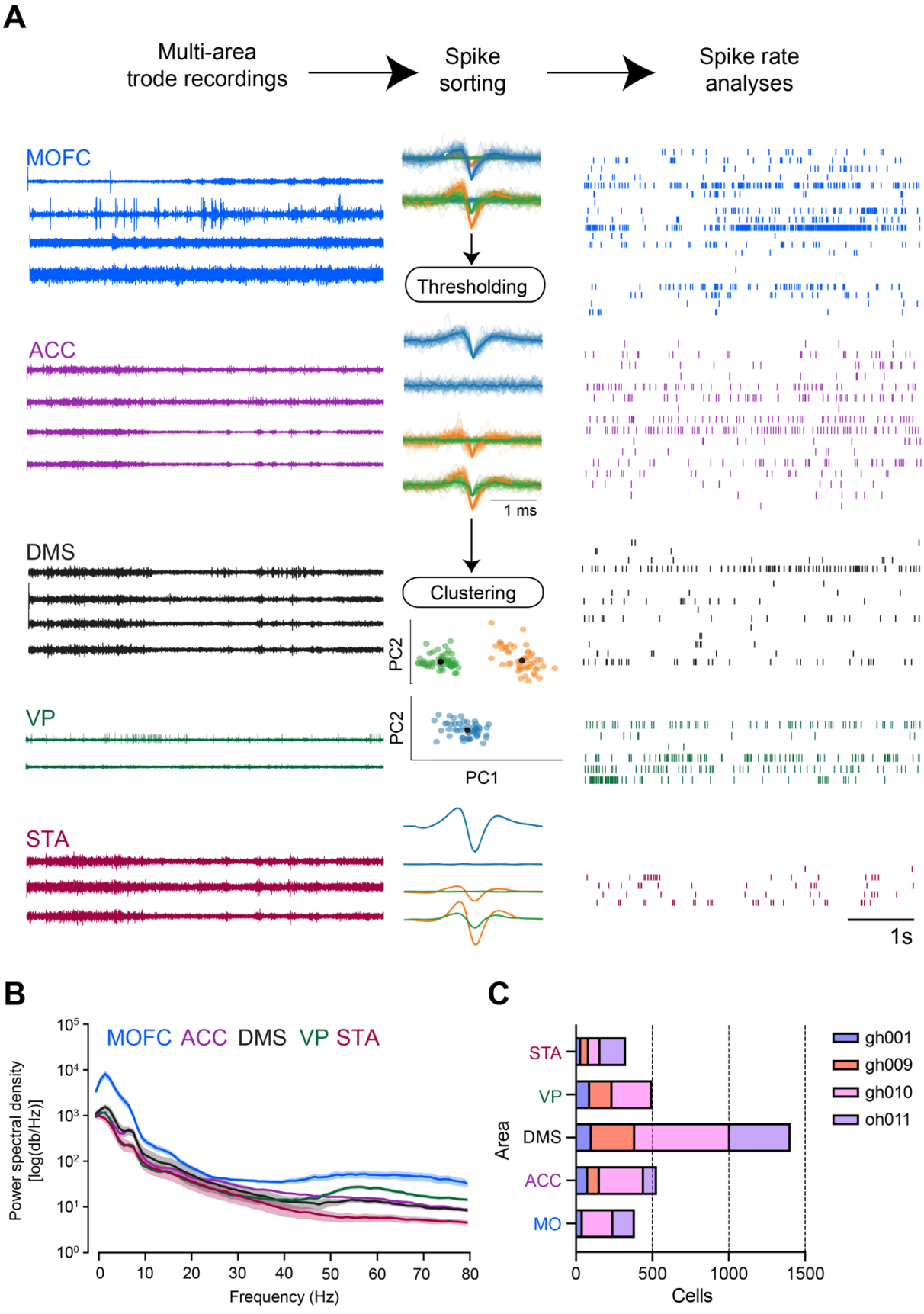


**Figure 9 | Overview of quality checks on recorded electrophysiological signals. A,** Wide-band signals (left) were recorded and monitored daily for noise contamination. These were later entered into a spike sorting pipeline (see methods). **B,** Power Spectral Densities calculated using the Welch method. **C,** yields of cells in each recording area (horizontal bars) broken down by subject (coloured subsections).

## 4. Discussion

We have hereby shown results from two different versions of a paradigm to measure cost-benefit decisions and motivation to pursue reward in rats and demonstrated electrophysiological signal quality in recordings from rats that are performing the final version of the task. In the first version of the paradigm, while trial-by-trial variation of distal reward prospects strongly influenced rats' decisions to overcome a fixed amount of effort for a potentially higher payoff, the rats also exhibited a strong bias to take the proximal reward. Therefore, in the second version of the paradigm, we removed the proximal option and also included an explicit manipulation of the cost of pursuing the distal reward. Now, rats' willingness to accept offers was significantly increased both by prospects of larger distal reward *and* by smaller effort cost. Patterns of choice on this latter version of the paradigm were resilient to surgical implantation and tethering. Importantly, extracellular field potentials were largely noise-free and yielded hundreds of cells per animal per structure. Therefore, we concluded that version 2 constitutes a feasible platform for the investigation of cost-benefit decision variables in the rodent frontal cortical-basal ganglia network.

The development of the final version of the paradigm was closely informed by results from the first version. Here, rats' probability of distal reward pursuit was not just significantly increased by higher amounts of reward on offer across the corridor, but was also decreased by increasing reward count and additionally modulated by recent reward history in a separate set of regressions. This effect is in accordance with foraging theory which predicts that background reward rate guides effortful behaviour (Stephens et al. 1986) and findings from newer reports (Hamid et al. 2016; Mohebi et al. 2019).

Interestingly the rats' probability of effortful action diverged from a policy of reward maximisation, reflected by "medium" and "large" reward choice on average only 25% and 50% of trials respectively. It is plausible that modelling running probability as a function of marginal energy intake rate, as has been done previously in foraging literature (Krebs et al. 1977), could offer a more parsimonious explanation for effortful choice in this task.

However, an alternative explanation is that choice behaviour was influenced by the Pavlovian draw towards the proximal reward site, which animals would need to inhibit in order to reach the distal reward port. In support of this, the rats were heavily biased to retrieve reward from the proximal reward port, even on forced trials. Coupled with the persistent tendency to enter the proximal reward port on medium reward forced trials, this could mean that rats had a policy of mostly trying to retrieve the proximal small reward and, failing that, would sometimes choose the distal reward. Normatively, a strategy of staying in the starting arena might have delivered a high enough reward turnover to render distal reward pursuit undesirable. Additionally, although rats' probability of running was significantly higher on forced trials, this was accompanied by more entries into the proximal reward port. This could have been a reflection of false reward expectancy, Pavlovian drive, a failure to attend to the visual cue that signalled a forced trial or the result of a policy such as the one mentioned above.

Version 2 was therefore instituted to remedy the drawbacks of version 1 and introduce an orthogonal manipulation of physical effort. Acceptance of offers was positively linked to the size of the reward offers and negatively linked to prospective effort cost. However, in addition, it was negatively impacted by reward offers on the previous 2 trials (regardless of whether this subsequently led to reward pursuit), similar to cross-trial correlative effects seen in previous reports (Jiang et al. 2019; Macke and Nienborg 2019; Akam et al. 2021). It is possible that temporally sustained representations such as these set up a frame of reference against which the goodness of future offers are evaluated (Martino et al. 2006; De Martino et al. 2009). Though it is likely that the history of effort prospects also has some influence on the rats' upcoming choices, this was not analysed this due extended time scale of variation of this parameter. Future

studies where effort is manipulated on a trial-by-trial basis will be needed to elucidate this effect.

We further found that acceptance of offers was sensitive to the incentive value of the outcome, evidenced when rats fed to satiety prior to testing. This shows that reward pursuit decisions on this task are not habitual, but rather depends on the rats' internal state of satiety. Two key latencies - action initiation and corridor latency - were also sensitive to prospects of reward and effort, and could thus be considered proxies of rats' evaluation of the offers. However, while changes in corridor entry latency may have reflected underlying estimation of net value (note the positive influence of effort and negative impact of reward), action initiation latencies were actually *increased* by reward prospects. One possibility is that this may reflect the increasing opportunity cost associated with exiting the nose-poke prematurely on high reward trials. Indeed, the probabilities of premature head-exits were mirrored by similar positive and negative influences of reward and effort, suggesting that rats employed increased restraint when exposed to more valuable offers. Further, this pattern of results also demonstrated the speed of corridor entry and rate of premature head-exits to be additional proxies of offer valuation.

The overall patterns of choice and action latencies on this task were similar in implanted and tethered rats, proving the pattern of decision-making to be resilient in despite major surgery, the time off task post-surgery and the change in bodyweight associated with the implant. The requirement to keep nose-poking while hearing the offer cue, the long latencies to initiate action and reward pursuit and the high cellular yields obtained from implanted animals provide a valuable platform for studying the neural evolution of neural correlates of cost-benefit decisions. However, repeated presentations of the same cue following premature head-exits need to be accounted for in neural analyses to factor out variation resulting differences in expectation error in such circumstances.

Though the regression approaches presented here have provided an invaluable means of quantifying the weightings of the main task variables on choice, they were limited in their ability to capture time-varying interactions between task variables or other

algorithmic operations that could explain the pattern of decisions. For example, it is plausible that decisions on this task are instead formulated as a comparison between the utility of the upcoming vs offer vs the utility rate over some number of trials in the recent history, mirroring a foraging-type computation (Stephens and Krebs 1986; Hayden and Walton 2014). Alternatively, rats could be keeping track of recoverable and/or unrecoverable components of fatigue over the course of a testing session and considering them alongside the value of the current offer (Müller et al. 2021). Additionally, subject-specific effort discounting parameters were not fitted here, although they could help reduce the error of our regression models. A re-evaluation of the approach to modelling these behavioural data in line with these considerations could provide a closer fit to the data and shed light on what constitutes normative decision-making on this task.

# 3. Single-unit correlates of cost-benefit decision-making in the frontal cortical-basal ganglia network


## ABSTRACT

The single unit firing dynamics underlying cost-benefit evaluation and subsequent value-guided choice in the frontal cortical-basal ganglia network is elusive. This gap in our understanding is driven by the lack of a clear overview of *network-wide* decision variable-related firing dynamics in relation to *ongoing behaviour*. We targeted the mOFC, ACC, DMS, VP and STA for single-unit recordings while testing animals on a cost-benefit decision-making where action initiation was temporally separated from offer presentation. We uncover time-and sign-varying contributions of individual neurons in all areas to signalling reward, effort and decision. Together, these results point to temporally varied but distributed processing of decision variables in the frontal cortical-basal ganglia network.

## 1. Introduction

Pathological motivation is a highly prevalent transdiagnostic symptom with poor functional outcomes and limited therapeutic options (Husain and Roiser 2018). How costs and benefits are weighed up by neurons and utilised to inform decisions about reward pursuit is thus a pressing research question. Our full understanding of this process has been hampered by, *first*, an incomplete view of the temporal dynamics of decision variable signal processing in relation to offer evaluation, action initiation and reward pursuit and, *second*, piecemeal understanding of simultaneous activity in the frontal cortical basal ganglia network within the confines of one single cost-benefit decision-making task.

Converging evidence to date supports the view that variables needed for value-guided decisions emerge in parallel across a range of frontal cortical-basal ganglia regions, but with some regional differences. In human fMRI studies, simultaneous engagement of the prefrontal-striatal circuit has been found to correspond to processing of expected reward (dACC/putamen), effort (putamen/vmPFC) and cost-benefit weighing (dACC/vmPFC) during offer evaluation (Bonnelle et al. 2016; Klein-Flügge et al. 2016; Chong et al. 2017). Similarly, firing rate activity scaling with the payoff and physical effort cost of offers was found to simultaneously evolve across the macaque OFC and ACC in the offer evaluation stage, with the latter structure exhibiting the highest degree of within-neuron multiplexing (Kennerley et al. 2009; Hosokawa et al. 2013).

Despite these advances, it remains unclear how these decision variable correlates emerge in single units within the interconnected basal ganglia network and what transformations the signals undergo in relation to ongoing behaviour. To address this, we designed a bespoke behavioural paradigm where rats need to evaluate whether to accept or reject singly presented effort-reward offers (reviewed in chapter 1) while targeting the mOFC, ACC, DMS, VP and STN for single-unit recordings. While the level of effort was varied across blocks and were not externally cued, reward offers were signalled explicitly using auditory cues and varied trial-by-trial. Importantly, as rats needed to hold still while hearing the reward offer cues and could only initiate reward

pursuit after a distinct Go cue, this allowed us to disentangle *offer evaluation*, *action initiation* and *initiation of reward pursuit*. Based on previous work, we hypothesised firing rate changes corresponding to reward magnitude to develop in single units concurrently in all regions during offer evaluation and to be sustained throughout trials (Lau and Glimcher 2008; Kennerley et al. 2009; Hillman and Bilkey 2010; Hosokawa et al. 2013; Wang et al. 2013; Klein-Flügge et al. 2016; Zénon et al. 2016b; Hunt et al. 2018; Ottenheimer et al. 2018). We expected individual unit activity to be tuned to effort starting from action initiation in the ACC (Kennerley et al. 2009; Hillman and Bilkey 2010; Cowen et al. 2012; Hosokawa et al. 2013; Porter et al. 2019) and STN (Zénon et al. 2016b; Nougaret et al. 2022). We further anticipated firing rate differences in all regions at action initiation to map onto whether rats would prospectively pursue reward or not (Lau and Glimcher 2008; Kennerley et al. 2009; Hillman and Bilkey 2010; Hosokawa et al. 2013; Wang et al. 2013; Klein-Flügge et al. 2016; Zénon et al. 2016b; Hunt et al. 2018; Ottenheimer et al. 2020).

## 2. Methods

### 2.1. Anatomical data analysis

To elucidate the anatomical origins of the cellular yield, we used SHARP-Track - a tool that was originally developed to map silicone probe trajectories onto the Allen Mouse Brain Atlas (Shamash et al. 2018). First, bright-field images from each animal were down-sampled and their contrast optimised. Second, electrode tracts were marked and matched to corresponding reference coronal sections. Third, the end-depths of each identified electrode tract as well as the DV heights of any penetrated regions were obtained on a reference mouse brain atlas. Fourth, mouse brain DV readings were converted to rat brain equivalents by scaling them by the rat/mouse brain depth ratios from template coronal sections used by SHARP-Track and averaging the results across the 4 recorded animals (table 1). However, the dataset analysed in this chapter was not restricted based on the anatomical findings.

**Table 1 | Overview of the averaged scalars used to convert mouse brain DV readings into rat equivalents in each targeted area.**

| Area | MOFC | ACC | DMS | VP | STA |
|---|---|---|---|---|---|
| Scalar | 1.59 | 1.57 | 1.50 | 1.46 | 1.47 |

## 2.2 Neurophysiological data analysis

All data were analysed using custom-written python 3.7.7 code.

**Data manipulations (fig. 3-6)** All units with an averaged firing rate of < 0.1Hz throughout the recording data were discarded, as per previous reports (Porter et al. 2019; Steinmetz et al. 2019). I began the neuronal firing rate analyses by first constructing peri-event firing rate histograms and raster plots. For this I quantified each neuron's firing rate in 100 ms bins -0.4s to +0.4s around the initiation of the nose-poke (termed "Pre-Cue"), -0.4 to +0.4s around the onset of the offer cue (termed "Offer"), -0.5 to +0.5s around the onset of the Go cue (termed "Go Cue") and -.0.5s pre- to +1.0s post-action initiation (termed "Action"). Data was clipped at 1s after Action to minimise the influence of motor covariates on the firing rate analyses. These bins were sampled every 10 ms and hence overlapped. For visualisation, I divided the data by either the level of offered reward, the magnitude of required effort or by the animal's eventual decision, each time averaging across the other variables. To facilitate the alignment of data on the task, I sought to transform the time on trial into a uniform bin space. For this, I divided the Pre-Cue epoch (-0.4s before Pre-Cue up to Offer) into 16 bins, the Offer epoch (from Offer up to Go Cue) into 16 bins, the Go Cue epoch (from Go Cue up to Action) into 20 bins and the Action epoch (from Action to 1s post-action initiation) into 20 bins. To facilitate analyses on the level of neuronal populations, I next z-scored the firing rate data by centring and standardising by the mean and standard deviation of values obtained from a concatenated series of ITI data from within the ongoing effort block.

**General linear model (GLM) construction (fig. 6)** To measure the influence of task variables on firing rates, I constructed GLMs by regressing reward, effort and decision, as well as the number of recent cue nose-poke errors, previous trial reward offer, previous trial decision and cumulative reward on z-scored firing rates within each neuron and time bin. Each time bin used for regression approaches integrated over 500

ms of firing rate data. Though I was primarily interested in quantifying the relationship between reward, effort and decision on firing rates, I entered the rest of the regressors to separate out any unique variance that might have been attributed to them. I obtained the loading of each variable on the firing rate response using the following closed form least squares solution (McCullagh and Nelder 1983):

$$\beta = (X^T X)^{-1} X^T Y$$

where X is a design matrix of shape [trials x variables] and Y is a vector of z-scored firing rates of shape [trials x 1]. The 7 regressors used in the design matrix X were constructed as follows:

- Reward: a vector containing the values 1, 3 or 6 corresponding to small, medium or large reward trials.
- Effort: a vector containing the values 1, 2 or 3 corresponding to low, moderate or high effort trials.
- Decision: a vector comprised of the values -0.5 and +0.5 corresponding to rejected or accepted trials.
- Recent nose-poke errors: a vector containing the values -0.5 or +0.5 relating to whether (+0.5) or not (-0.5) any premature nose-poke exits were detected immediately prior to the initiation of the trial.
- Previous trial reward offer: a vector containing the values 1, 3 or 6 corresponding to a small, medium or large reward offer on the previous trial.
- Previous trial decision: a vector containing the values -0.5 or +0.5 corresponding to either a reject or accept decision respectively on the previous trial.
- Cumulative reward: a vector containing the sum of accrued rewards in the ongoing testing session.

All regressors were then z-scored to themselves before obtaining the coefficient loadings.

**Quantification of task-selective neurons (fig. 6a)** To identify task-modulated neurons, I searched each fitted neuron-time bin GLM for a significant ($p < .0.01$) regression slope for the variable of intertest. To mitigate for the false discovery rate

(FDR) resulting from large number of units being tested in each area, p-values were adjusted using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). To compare the maximum percentages of units tuned to either of the 3 primary task variables, I employed a 2-way ANOVA in JASP (University of Amsterdam) with main factors of area (MO, ACC, DMS, VP, STA) and variable (reward, effort, decision). The Tukey method was applied to correct for FDR in *post-hoc* tests.

**Quantification of coefficients of partial determination (fig. 6b-d)** To examine the extent of the unique variance explained by each task variable, coefficients of partial determination (CPD) were calculated. The CPD reflects the percentage of variance in neuronal activity that is explained by a full regression model vs a model with the variable of interest excluded. This is calculated as follows:

$$\mathrm{CPD_i} = 100 \times \left(\frac{\mathrm{SSE_{\sim i}} - \mathrm{SSE_i}}{\mathrm{SSE_{\sim i}}}\right)$$

For each neuron, I generated 500 shuffled datasets by circularly shifting the entries of the design matrix in the range [0, N], where N was the total number of trials in the testing session. Circular shifting, rather than random shuffling, preserves the autocorrelative structure of the dataset and is thus a more realistic index into the chance-level outcome. Individual neuron CPD p-values for each time point (reported in fig. 6b-c) were calculated as the fraction of (500) shuffled models which yielded a higher CPD than the corresponding non-permuted model. Population CPD p-values (reported in fig. 6d) were expressed as the fraction of permutations for which the (500) shuffled population average CPDs were larger than non-permuted population mean CPD.

# 3. Results

## 3.1. Anatomy

24 out of a total of 51 individual implanted electrodes in the 4 recorded subjects were successfully reconstructed as tissue tracts *post-mortem,* with at least one tract per animal per structure captured successfully (fig. 1a). The low reconstruction yield was caused by electrodes' tendency to clump together and appear as merged tracts. Aside from the 5 targeted brain regions, individual units were spuriously recorded from an additional 7 adjacent brain regions (fig. 1b). Whereas just 4% ACC cell yield originated from secondary motor cortex (M2), 46% of targeted MOFC cells originated from either pre-limbic (PrL, 33%) or ventrolateral orbitofrontal cortex (vlOFC, 13%), 14% of targeted DMS cells originated from NAc, 21% of targeted VP cells originated from caudate-putamen (CPu), and 51% of targeted STN cells originated from either midbrain (MB, 25%) or Zona Incerta (ZI, 26%). Based on these findings, the targeted STN will henceforth be named as the subthalamic area (STA). Targeted mOFC will keep on being referred to as mOFC, despite the inclusion of cells from PrL and vlOFC.

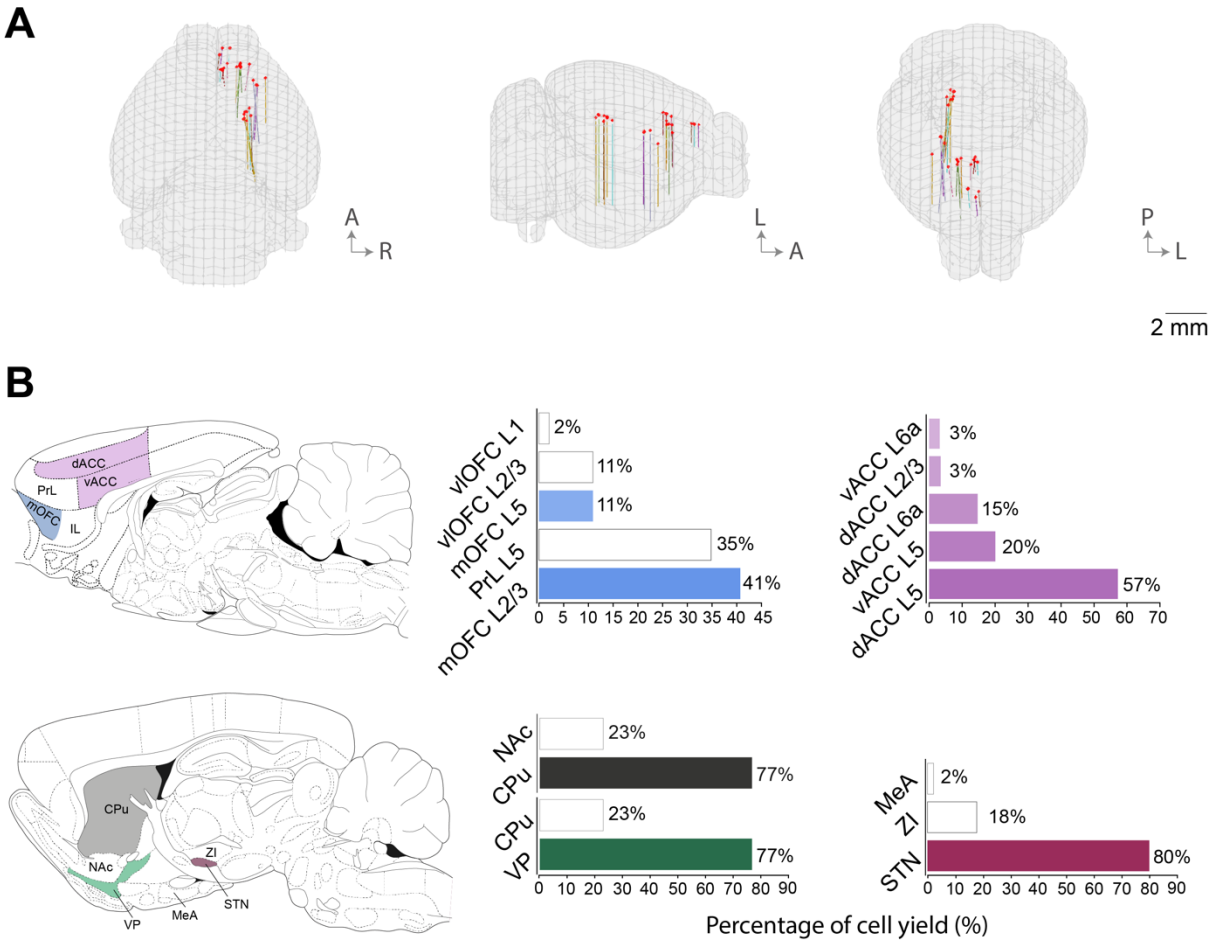



**Figure 1 | Quantification of origins of the cellular yield. A,** summary of trode locations on 3D meshes as registered to the Allen Common Coordinate Framework 3D space as viewed from the back (left), right side (middle) and front (right). A, anterior; R, right; L, left; P, posterior. **B,** breakdown of the cellular yield by targeted area. Left, brain diagrams depicting the targeted areas as well as immediately surrounding areas that were recorded from, adapted from Paxinos and Watson 1998. Right, graphs depicting the percentages of cells from accidentally recorded (empty bars grey) as well as targeted (filled bars). vl - ventrolateral; m - medial; d - dorsal, v - ventral; L - layer; OFC, medial orbito-frontal cortex; ACC - anterior cingulate cortex; PrL - prelimbic cortex; M2 - secondary motor cortex; NAc - nucleus accumbens; CPu - caudate-putamen; VP - ventral pallidum; ZI - zona incerta; STN - subthalamic nucleus; MB - midbrain.

## 3.2. Overview of neuronal yield

To index into the functional makeup of the cell populations recorded in each region, their firing rates, inter-spike intervals and coefficients of variation computed. The regions contained a highly variable population of neurons as judged by averaged firing

rates, inter-spike interval (ISI) distributions and coefficients of ISI variation (fig. 2b-d, table 2).

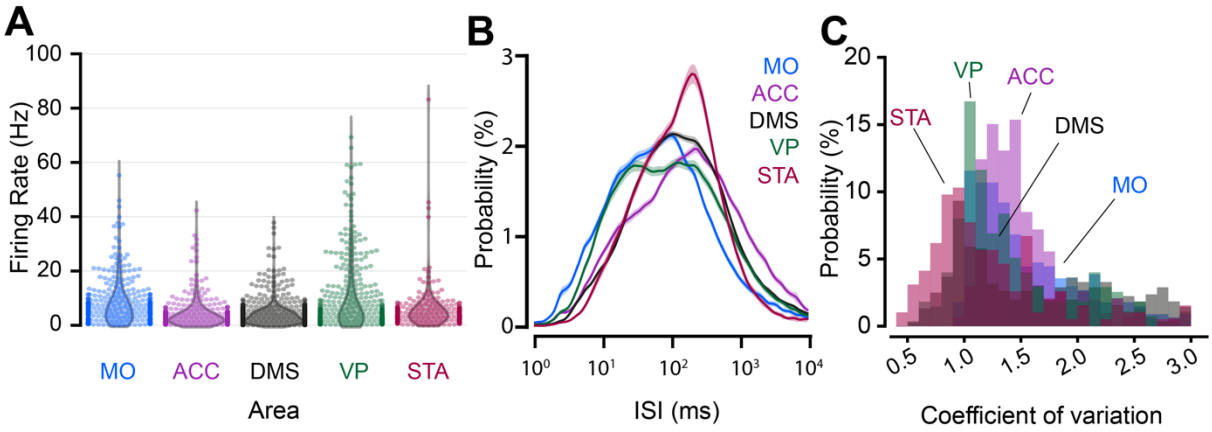



**Figure 2 | Heterogenous functional characteristics in identified neurons in the prefrontal cortical basal ganglia network. A,** average firing rates of each identified neuron (dots) and the neuronal population (violin bodies) coloured by the targeted area. **B,** histogram of average ISI values of neurons in each region. Values are depicted as the population mean probabilities (dark lines) ± SEM (shaded areas). Histograms are smoothed with a gaussian kernel with a standard deviation of 1.0. **C,** probability distribution of individual neuron coefficients of ISI variation (bars), coloured by region.

**Table 2 | Summary of functional characteristics of identified neurons.**

| **Area** | Median Firing Rate (Hz) | Median Inter-Spike Interval (ms) | Mean Coefficient of Variation |
|---|---|---|---|
| **MO** | 5.79 ± 0.45 | 172.59 ± 216.70 | 2.16 ± 0.11 |
| **ACC** | 2.53 ± 0.27 | 395.77 ± 269.34 | 1.58 ± 0.03 |
| **DMS** | 2.71 ± 0.13 | 369.56 ± 136.05 | 2.39 ± 0.06 |
| **VP** | 3.99 ± 0.62 | 250.68 ± 120.42 | 1.55 ± 0.04 |
| **STA** | 3.91 ± 0.49 | 255.81 ± 277.31 | 1.83 ± 0.09 |

## 3.3. Firing rate relationships to task variables in identified individual units

We next sought to characterise individual neuron firing rate dynamics in each recording area in relation to the cost-benefit decision-making task. To do this, we examined firing

rate activity time-locked to 4 trial events: (i) entry into the nose-poke port (termed "Pre-Cue"), (ii) offer cue onset (termed "Offer"), (iii) Go cue onset (termed “Go Cue”), and (iv) nose-poke exit (termed “action”). To further analyse the relationship between firing rate changes and task factors, the data was split by levels of prospective reward, prospective effort and the animal's decision, henceforth collectively termed the primary task variables. Assessment of the peri-event PSTHs and raster plots of hand-picked units revealed example cells in each area whose firing rate linearly scaled with these 3 task variables, albeit with varying time dynamics (fig. 3). Additionally, variable patterns of task variable tuning were found. First, levels of the variables of interest were sometimes found to be signalled through a tonic firing rate baseline throughout the trial (e.g. MOFC gh010_#0020, top middle panel), whereas in other cases there were phasic firing rate variations time-locked to particular trial events (e.g. DMS gh001_#0070). Second, while in some cases neurons had a positive tuning to increases in a task variable (e.g. STA gh001_#0026), the relationship was negative in others (e.g. VP gh010_#0031). Positive and negative tuning was sometimes even found alternative within single cells during the time course of the trial (e.g. DMS gh001_#0070).


Reward
small
medium
large
Effort
low
med
high
Decision
reject
accept
MO
oh011_#0123
gh010_#0020
gh001_#0036
ACC
gh010_#0157
gh010_#0003
gh001_#0011
DMS
gh001_#0070
gh009_#0023
gh010_#0496
VP
gh010_#0031
gh010_#0143
gh009_#0117
STA
gh001_#0026
gh009_#0045
oh011_#0095
Firing rate (Hz)
Pre-Cue
Offer
Go Cue
Action
Time from event (s)

**Figure 3 | Examples of task variables encoding could be found in individual cells in each area.** Data on each level of the variables (left - reward; middle - effort; right - decision) are depicted as firing rates and spike raster plots, divided into peri-event (pre-cue, offer, Go cue, action) intervals of variable duration from left to right, Firing rates are depicted as line plots, colour-coded by the level of the variable, averaged across other conditions (lines) ± SEM (shaded areas). Raster plots denote firing events on each trial and are colour-coded and sorted along the y-axis by the level of the task variable. All peri-stimulus time histogram curves are smoothed with gaussian kernel with a standard deviation of 1.3.

To ensure these effects were not driven by sampling distortions caused by examining activity in overlapping time windows, I next warped the time axis by dividing each inter-event epoch into a set number of bins and extracted an aligned bin space starting 0.4s before nose-poke initiation up to 1s after action initiation. As demonstrated on figure 4, this resulted in very little appreciable distortion of the original data (fig. 3), while allowing me to capture the full extent of each trial epoch.

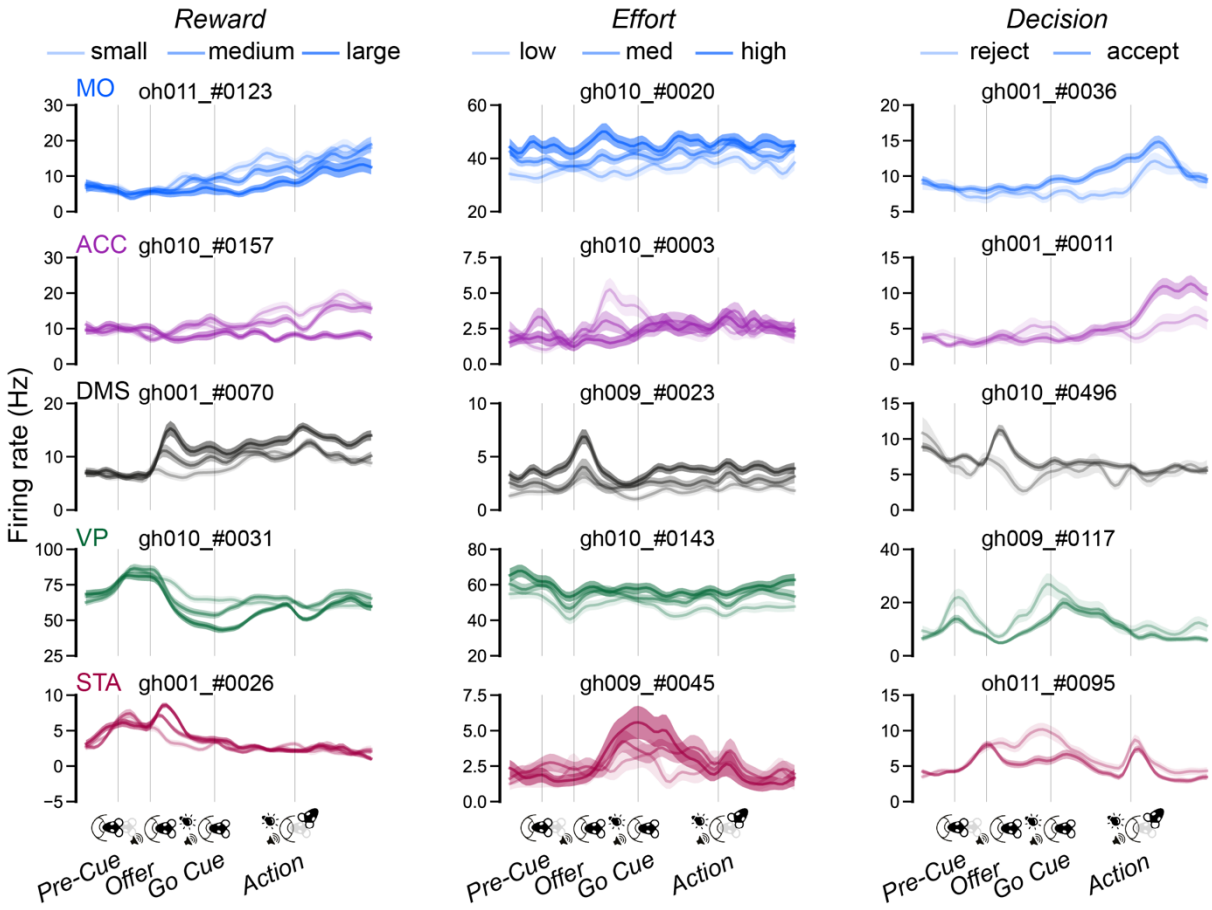


**Figure 4 | Coding of levels of reward, effort and decision in selected exemplary units in warped time.** Firing rates for each chosen neuron are depicted as means (lines) ± SEM (shaded areas) over levels of these task variables within one session averaged across other conditions (dark lines). The x-axis denotes warped time from -0.4s before the initiation of a nose-poke up to 1s after action initiation. Vertical grey lines denote the time-points of initiation of the nose-poke, onset of offer cue, onset of Go cue and action initiation. All curves are smoothed with gaussian kernels with a standard deviation of 1.3.

To compare effects across neurons, firing rates of each neuron within a structure were z-scored within each block of ~20 trials (each constituting its own level of effort) relative to the total concatenated per-block ITI data. The population average z-scored firing rate profiles that were borne out of this (fig. 5) revealed mainly phasic reward- and decision-related activity in the VP and STA, but prolonged signals in the MOFC and ACC, while the DMS showed evidence of both types of time-changes. Effort, by contrast, explained only subtle amounts of variation in the neurons' firing rates. While population-averaged reward tuning had a positive relationship to reward in VP and STA, it was negative in the mOFC and ACC. DMS, on the other hand, switched the polarity of its coding from positive in the Offer window to negative in the Go cue and Action windows. There properties were largely consistent across animals, with some variation of the strength and timing of task tuning between animals (supplementary figure 1).

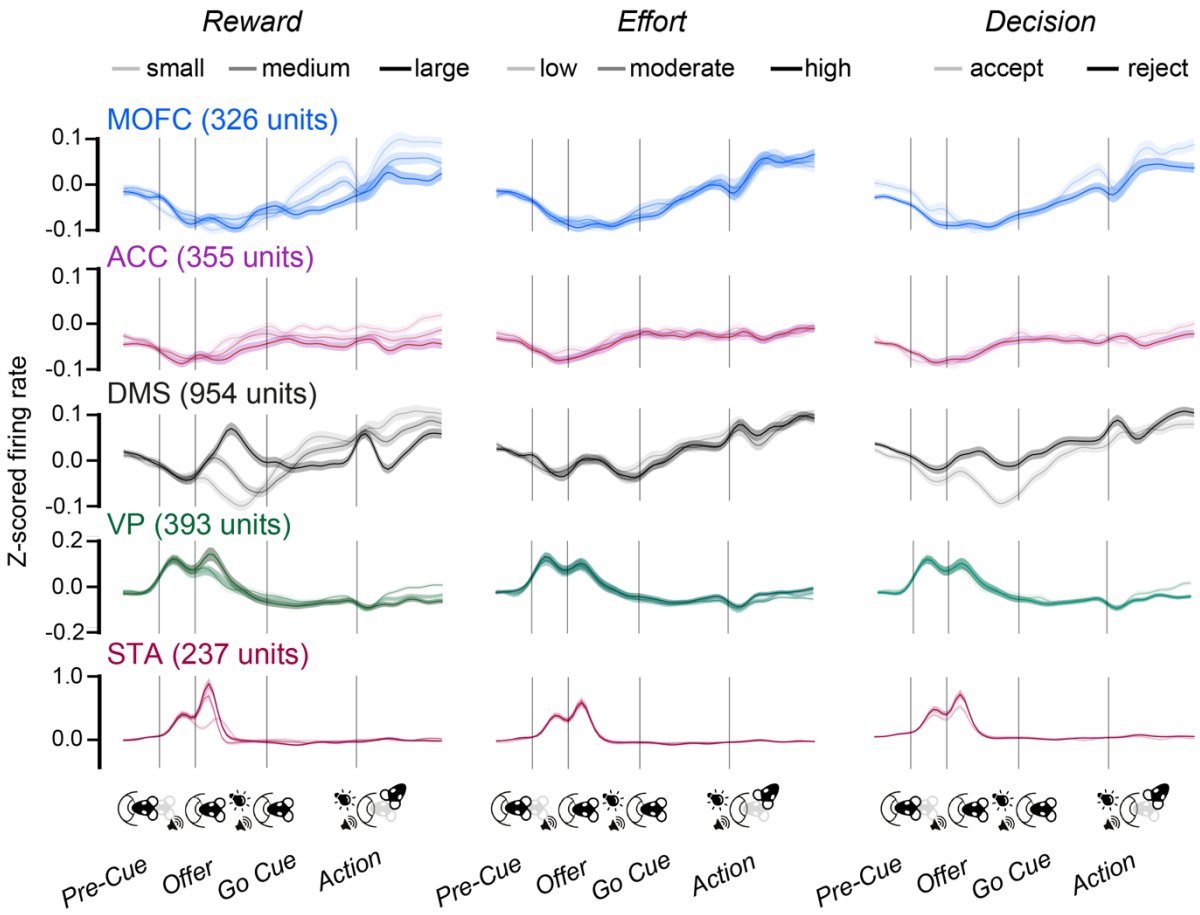



**Figure 5 | Diverse and distributed averaged representation of reward, effort and decision in the prefrontal cortical basal ganglia network**. Values are depicted as average z-scored population firing rates, averaged across other conditions (dark lines) ± SEM (shaded areas). All curves are smoothed with gaussian kernel with a standard deviation of 1.3.

To gauge the percentage of neurons in each area whose firing rates were contributing to signalling these variables, I employed general linear models (GLMs) that considered the influence of reward, effort and decision while controlling for recent cue nose-poke errors, previous trial reward offer, previous trial decision and cumulative reward (fig. 6a, left). I then identified the percentage of neurons in each area across all recorded animals that displayed a significant ($p < .0.01$) regression slope to the primary task variables *at any time point* within the -0.4s before Pre-cue up to 1s post-Action. To mitigate for the false discovery rate (FDR) resulting from large number of units being tested in each area, p-values were adjusted using the Benjamini-Hochberg procedure. This revealed that considerable proportions of neurons in basal ganglia areas (58.6%, 52.0% and 15.0% in the STA, DMS and VP respectively compared to 4.9% in the MOFC

and 1.7% in the ACC) reliably signalled the magnitude of reward, with substantially lesser representation of effort and decision in any area (fig. 6a, right). Reward signalling engaged units most frequently in the STA > DMS > VP > MOFC > ACC. Effort coding, by contrast, was most representative of unit signalling in DMS > MOFC > VP > ACC > STA. Decision was coded by neurons most frequently in MOFC > DMS > VP > STA > ACC. Very few cells (< 1.7% of cells in total) conjointly encoded multiple variables and interestingly such examples could only be found in the DMS and STA, not in the cortical areas.

The order in which cost, benefit and decision-related activity emerge in the frontal cortical-basal ganglia network is still contentious. One popular framework advocates that these quantities are processed in a serial and localised fashion (Gold and Shadlen 2007; Kable and Glimcher 2009; Rangel and Hare 2010; Padoa-Schioppa 2011; Dolan and Dayan 2013). Alternatively, decision variables may be processed in parallel in multiple brain regions with the final behavioural output determined by a mutually competitive process (Cisek 2012; Rushworth et al. 2012; Strait et al. 2016; Hunt and Hayden 2017). To adjudicate between these possibilities, I used GLMs as above and examined the time courses of (i) the shares of neurons in each region that instantaneously signaled these variables and (ii) the unique variance explained by each task variable. Reward signalling instantaneously recruited up to 21.31% (STA) and 14.39% (DMS) of units during the offer window, but 8.13% (VP), 6.39% (MOFC) and 2.85% of ACC neurons after action initiation (fig. 6b, left). This was reflected in a task variable x area interaction in an ANOVA on the maximal proportions of cells recruited ($F_{2,4} = 2.74$, $p = .017$), driven by a significant difference in reward signalling neuron percentages between the STA vs ACC ($p = .003$), with no other significant differences between the regions ($p > .066$, fig. 6c, left). These dynamics of neuronal recruitment were mirrored by the emergence of a robust and sustained population-level reward read-out in all areas during the offer window (fig. 6d, left). Effort coding recruited a maximum of 1.57% of STA cells during the pre-cue period, 1.55% (DMS), 1.51% (ACC) and 1.44% of MOFC neurons during offer cue period, and 2.70% of VP neurons during the Go Cue period with no significant differences between the maximal percentages of neurons recruited between regions (all $p = 1.000$). This was reflected by scattered periods of significant population-level

signalling activity in the DMS throughout most of the trial (fig. 6d, middle). The pools of decision-specific units expanded by 1.70% in the ACC during the pre-cue period, 3.15% in the STA during the offer cue period, and 3.16% (VP) , 2.89% (DMS), and 2.88% on the FDR in the MOFC after action initiation, with no significant differences in these percentages between the areas (all p = 1.000). Intervals of robust decision-related read-out were found most frequently in the DMS, where they tended to group in the Offer, Go cue and Action windows, though significant read-outs were also seen in all the other structures at differing time points in the trial. Together, this indicates the absence of clear regional specialisation, but considerable temporal differences in processing reward, effort and decision.

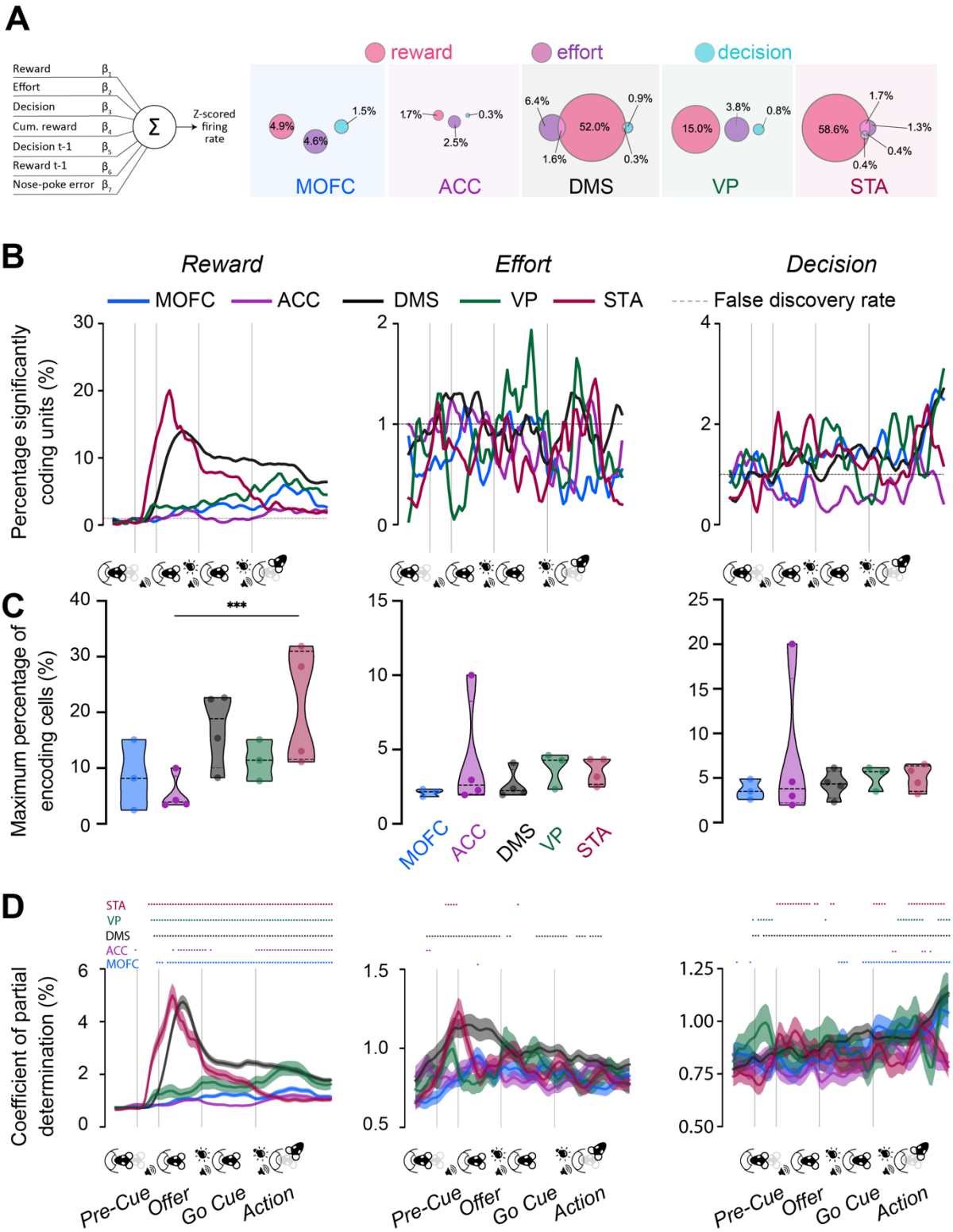



**Figure 6 | Single unit mechanisms of reward, effort and decision signalling. A,** a schematic of the construction of GLMs (left) and Venn diagrams illustrating the percentages of task modulated units as identified using regression coefficient p-values, broken down by each primary variable (coloured dots) in each area (shaded rectangles). **B,** percentages of units significantly coding for reward (left), effort (middle) and decision (right) over time in each

region (coloured lines), identified on using CPD p-values. **C,** maximum percentages of neurons tuned to the primary task variables (organised into panels as above) in each animal (dots) and across the cohort (violin bodies), colour-coded by region; *** $p < .01$. **D,** population average CPD traces, depicted as means (solid lines) and SEM (shaded areas), for each task variable (individual panels) and area (colour-coded lines and shading). Points above indicate time points where the population average CPD is significantly higher than expected by chance in each region using a significance threshold of $p < .001$. Curves in **B** and **D** are smoothed with a gaussian kernel with standard deviations of 1.0 for illustration purposes.

# 4. Discussion

We have hereby demonstrated that it is feasible to simultaneously target areas across the frontal cortical-basal ganglia network for single-unit recordings and have uncovered heterogenous firing rate dynamics corresponding to the magnitudes of reward/effort and to accept/reject decisions in a diverse set of cell populations from each region. Strikingly, examples of neurons coding for each of the three decision variables with varying polarities and timings could be found in each region.

This was corroborated by population averages across all recorded animals, revealing mostly offer cue-driven phasic and prolonged reward offer-oriented responses respectively in the basal ganglia and cortex, coupled with time-distributed decision and subtle effort-related responses across the network. Further, significant proportions of units coding for each decision variable were subsequently found in each recorded region. However, while they were phasically recruited to signal reward in the STA and DMS, the recruitment occurred more gradually after Go cue onset in the VP, mOFC and ACC. Effort was coded by a significant proportion of units in all regions during the Offer/Go cue windows, while robust correlates of prospective decision were found in all areas and were mainly restricted to the action window. Together, these data suggest that there is no clear evidence for regional specialisation or exclusive time windows of coding one decision variable over others.

The literature to date is largely in agreement with this ubiquity of reward-predictive stimulus value signals across the network (Kable and Glimcher 2007; Kennerley et al. 2009; Hillman and Bilkey 2010; Vickery et al. 2011; Cowen et al. 2012; Hosokawa et al. 2013; Wang et al. 2013; Ottenheimer et al. 2018; Nougaret et al. 2022). However, the current data refine our current understanding of the underlying processing dynamics in one important way. While the reward signal, as revealed by peri-stimulus time histograms and regressions of the firing rates, had a strong phasic component in the basal ganglia, it was time-distributed in the PFC. This characteristic of the PFC signal, also commonly seen in other cognitive tasks (Ozdemir et al. 2020) may be a reflection of the patterns of intrinsic connectivity

which set the PFC up for slower-varying activity patterns (Murray et al. 2014). This extended time course of signalling also explains why the proportions of units participating in the signalling of reward in the cortical areas *at any point* tended to be quite low (less than 5% in ACC and mOFC, compared ≥ 15% in the basal ganglia).

However, we note that the proportions of cortical neurons tuned to this variable appear much lower than in other reports with a reward size manipulation (typically in the range of 10-15%, Kennerley et al. 2009; Bouret and Richmond 2010; Lopatina et al. 2016; Hunt et al. 2018). This discrepancy might be driven by differences in the compared species, approaches to quantifying reward tuning (e.g. linear regression vs. ANOVA vs. area under the curve-type analyses, the method used for correcting for family-wise error rate in multiple comparisons or anatomical targeting (note that up to 35% of putative mOFC cells here might have originated from PrL).

Strikingly, with the exception of the ACC in one animal, very few units in any structure engaged with effort encoding, resulting in weak and patchy time periods of significant population level signalling of this variable. Notably, though the proportion of effort-specific units did reach a threshold during the offer window in the ACC, there was no robust evidence of population-level effort coding in this region as judged by CPD permutation testing. The proportion of effort-selective neurons in the ACC at any time point was found to be just 2.5% of the entire recorded neuronal pool. Though this is similar to that found in one report (Kennerley et al. 2009) , it is at odds with findings from some other reports (40-65%, Cowen et al. 2012; Porter et al. 2019).

This discrepancy might be driven by the considerations mentioned above for reward coding. However, these findings might also be a consequence of the block-wise manipulation of effort in our task. Compared to the reward magnitude variation, effort was not also explicitly cued at the beginning of trials and thus had to be memorised throughout blocks. This raises four possibilities for the processing of effort in this task. First, effort signalling might have more predominantly involved other limbic structures with dense connectivity to the hippocampus, such as the

basolateral amgydala (Floresco and Ghods-Sharifi 2007). Second, it may been coded more strongly during the ITI or during reward pursuit (Quilodran et al. 2008; Hayden et al. 2009). Third, confidence about knowing the level of effort might have been highest toward the end of blocks, diluting the effects seen on the neural level in earlier trials. Fourth, effort might have predominantly been coded over longer-term firing rate changes apparent on a block-by-block rather than trial-by-trial firing rate changes.

Prospective accept/reject decision was, on average, coded by a similarly small proportion of units in all areas, despite effect sizes comparable to reward in the MOFC and DMS. This could be consistent with a small minority of neurons driving the firing rate changes on accepted vs rejected trials. However, our ability to detect decision-dedicated neurons could have been curtailed by the limitation that reward/effort and decision were behaviourally heavily correlated (chapter 2). Alternatively, a decision read-out might have instead been contained within interactions with reward and/or effort - a possibility which was not examined here. Notably, we report the ramp-up of recruitment of decision-tuned neurons in the DMS, VP and STA starting in the Action window. This is consistent with a race-to-bound type mechanism (Churchland and Ditterich 2012; Tajima et al. 2016, 2019) which might drive decision-making in these regions.

Although our anatomical targeting approach was largely successful in capturing neurons from our targeted areas. the targeted mOFC, VP and STN populations also included a considerable number of neurons from neighbouring areas. Targeting success in future experiments may be improved by optimising electrode co-ordinates along with the surgery technique, by DV electrode position adjustment peri- or post-surgery based on firing properties of recorded neurons or by computer tomographic/MRI imaging post-surgery. Additionally, Neuropixel probes, which have recently been optimised for prolonged recording, may be of some help in recording superficial regions with precise DV positioning requirements (Jun et al. 2017; Luo et al. 2020). Here, we used SHARP-Track to register our histological slices on reference brain data and obtain electrode tract end-depths (Shamash et al. 2018).

However, as this tool was developed for use on mouse brains, we acknowledge that our results may have been biased by differences between the mouse and rat reference atlases. Usage of an anatomical quantification pipeline dedicated to rat brain data, e.g. BrainReg (available at https://github.com/brainglobe/brainreg) or Atlas Fitter (Kopec et al. 2011) may yield more accurate targeting results.

Finally, in this work, only simple linear mappings between task/behavioural variables and z-scored firing rates were examined. However, other types of codes could be playing an important part in shaping the patterns of cost-benefit decisions seen on our task. One alternative code which has not been considered here is spatial. Accordingly, neurons with spatial selectivity for positions on the apparatus could be indicating variables like effort (which has a strong mapping onto the physical configuration of the testing apparatus) and decisions, which can involve an excursion to the other side of the corridor. Indeed, activation of spatially tuned neurons to signal upcoming goal pursuit has been demonstrated in the OFC (Basu et al. 2021) and ACC (Mashhoori et al. 2018). Alternatively, task variables could be signalled through co-activity of groups of neurons, possibly across the entire frontal cortical-basal ganglia network (El-Gaby et al. 2021) or changes in the variation of firing rates (Kriegeskorte and Diedrichsen 2019; Ebrahimi et al. 2022). Lastly, the network might be making use of population coding, where each neuron's orthogonal signalling dimensions are used to make up a representation on the level of groups of neurons. This will be the topic of the next chapter.

# *4. Ensemble-level correlates of cost-benefit decision-making in the frontal cortical-basal ganglia network*


## Abstract

Population-level signalling can expand the representation capacity of individual units. However, it is unclear to what degree areas in the frontal cortical-basal ganglia network

harness population coding for cost-benefit decision-making. Here, using linear classifiers trained on two levels of each decision variable in sliding windows, we show that while there is little representation of decision in any area, reward can robustly be read out from population activity after offer cue onset in the basal ganglia but not the mPFC. Interestingly, effort was signalled DMS ensembles across all the analysed trial epochs and was dependent on the number of cells in those ensembles. This reveals a dichotomy between the population-level signalling of reward and effort, the former represented uniformly in the basal ganglia, but the latter coded in diverse activity profiles in the DMS.

## 1. Introduction

The ability to simultaneously record from ensembles of neurons within subjects, as we have done, has opened up the possibility of indexing into decision variable representations on the level of groups of neurons. Determining whether decision variables are linearly readable (i.e. decodable) from groups of neurons in the frontal cortical-basal ganglia network would help us get closer to a mechanistic understanding of cost-benefit decision-making. Biophysically, such a processing scheme might require a population of *upstream neurons* which, in concert, signal a quantity such as offer value and *downstream neurons* which linearly read this information off the population code. Indeed, evidence for the plausibility of such a mechanism comes from the finding that neocortical pyramidal cells' dendrites can weight inputs based on distance from the soma (Andersen et al. 1980) and implement linear summation (Cash and Yuste 1999).

Population- rather than individual unit-level responses have successfully been used to infer reward-predictive stimulus value and choice from the OFC (Rich and Wallis 2016; Balewski et al. 2022), value-guided orienting responses in the striatum (Balewski et al. 2022), upcoming reward size and social competition success in the ACC (Li et al. 2022), sensory discrimination-based choice in the striatum and frontal cortex (Steinmetz et al. 2019), costs and benefits in the STN (Nougaret et al. 2022), outcome valence in the striatum (Bloem et al. 2022), time and sensorimotor state in the striatum (Mello et al. 2015), reward preference in the VP (Ottenheimer et al. 2020), environmental contingencies (El-Gaby et al. 2021) or spatial position in the hippocampal CA1 layer (Wilson and McNaughton) and motor intentions in motor cortical areas for neural prosthetics (Willett et al. 2021). Importantly, such population responses often rely on the concert dynamics of each neuron and the idiosyncratic cross-correlations between them, rather than one single member (Rich and Wallis 2016; Ottenheimer et al. 2018; Panzeri et al. 2022).

The extent to which the frontal cortical-basal ganglia network makes use of such dynamics to support cost-benefit evaluation and decision-making is unclear. Here, we examined whether prospective reward, effort and decision were decodable from

populations of neurons in the mOFC, ACC, DMS, VP and STA using the same input dataset as in chapter 3. However, compared to the findings in that chapter, here we expected that effort and decision, which had a weak effect on single unit firing rates, instead would be readable from population dynamics across the network.

## 2. Methods

### 2.1 Data pre-processing and selection

All source data analysed in this chapter are the same as that analysed in chapter 3.

### 2.2 Decoder models

Linear regression was used to quantify the degree to which levels of task variables could be predicted from the per-trial ensemble responses. For each warped time bin, we calculated optimal sets of weights (one per ensemble member) on the full set of trial data from each structure *on each recording day* using the following least-squares solution (McCullagh and Nelder 1983):

$$\beta = (D^T D)^{-1} D^T C$$

where D is a n(trials) x n(neurons) design matrix of z-scored firing rates and C is a n(trials) x 2 category matrix. Category matrices were contrast coded with extremes of each task variable (reward: small vs large, effort: low vs high, decision: accept vs reject). To avoid decoding biases resulting from unequal numbers of accept vs reject trials, input data were randomly subsampled to include the equal numbers of trials in each category. Decoder predictions were estimated via multiplication of the vector of weighs by the design matrix and accuracy calculated by quantifying the proportion of trials that were correctly categorised around a threshold of zero. For plotting, decoder accuracies were averaged across each modelled ensemble, *population* traces.

To analyse the relationship between ensemble size and decoder prediction accuracy, subsets of cells were randomly sampled from the original ensemble in percentiles ranging from 20% to 100% in increments of 20%. Each subsample was made up of 20 randomly sampled subsets containing the same number of cells. Subsampled decoder accuracies were averaged across the 20 subsets.

### 2.3 Statistical tests

Shuffled control data were obtained by circularly shifting the reward, effort or decision levels before model fitting. 500 shuffled models were trained per each ensemble. Decoder p-values were estimated as the fraction of shuffles (averaged over ensembles) that yielded a higher decoding accuracy than non-shuffled population data. These p-values were corrected for multiple time-wise comparisons using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). For plotting, chance-level decoding accuracy was depicted as the mean ± SEM of ensemble shuffle-averaged decoding accuracies. Shuffled control data were acquired in a similar fashion for subsampled datasets. A statistical threshold of $\alpha = 0.01$ was used in all permutation tests.

To analyse the effect of subsample size on decoding accuracy and the relationship between recording area and peak decoder accuracy, one-way ANOVAs were run per each region and task variable. Where-ever there was a significant main effect, post-hoc tests with Tukey's HSD adjustment of the critical value were used.

### 2.4 Data depiction

Data was smoothed with a gaussian filter of a width of 1 standard deviation (SD) for illustration purposes. For figure 3, a 5 SD gaussian smoothing filter was used.

# 3. Results

## 3.1 Reward and effort are robustly coded on the population level

To determine the degree to which decision variables were readable from the population-level response in each region, extreme levels of each task variable (reward: small vs large, effort: low vs high, decision: accept vs reject) were regressed onto trial-by-trial z-scored firing rates for each day's yield of neurons within each animal in sliding windows (fig. 1A, see methods). We henceforth refer to each such group of neurons as an *ensemble*. Ensembles used for this analysis consisted of 5 (STA), 7, (MOFC), 8 (ACC), 9 (VP) and 21 (DMS) units on average (fig. 1B).

Consistent with reward representation in single units (chapter 2), significant reward-related activity emerged sequentially in STA, DMS and VP ensembles, was time-locked to offer cue onset and sustained until the end of the observation period (1s after nose-poke exit fig. 1C-D). Sparse time points of significant reward-predictive activity in the MOFC and ACC, by contrast, evolved only after Go cue presentation. Maximum baseline-adjusted reward decoding accuracy was dependent on the recording region ($F_{4, 187} = 9.19$, $p = 8.47 \times 10^{-7}$), driven by higher accuracy in the DMS vs MOFC ($p = .001$), ACC ($p = .001$) and VP ($p = .005$). Higher maximum baseline-adjusted decoding accuracy was also seen in the STA compared to ACC ($p = .003$, fig. 1D). There were no other significant differences between the areas (all $p > .122$).

By contrast to the rapid stimulus-driven change in reward decoding accuracy, robust effort-related activity emerged gradually in the DMS following nose-poke initiation and was sustained through to the end of the observation period. Accept/reject decision was readable at discrete time points in the DMS and STA, but this signal was not consistent over time (fig. 1 C-D). There were no significant main effects of area on either maximum baseline-adjusted effort or decision decoding accuracy (all $F < 1.43$, $p > .226$).

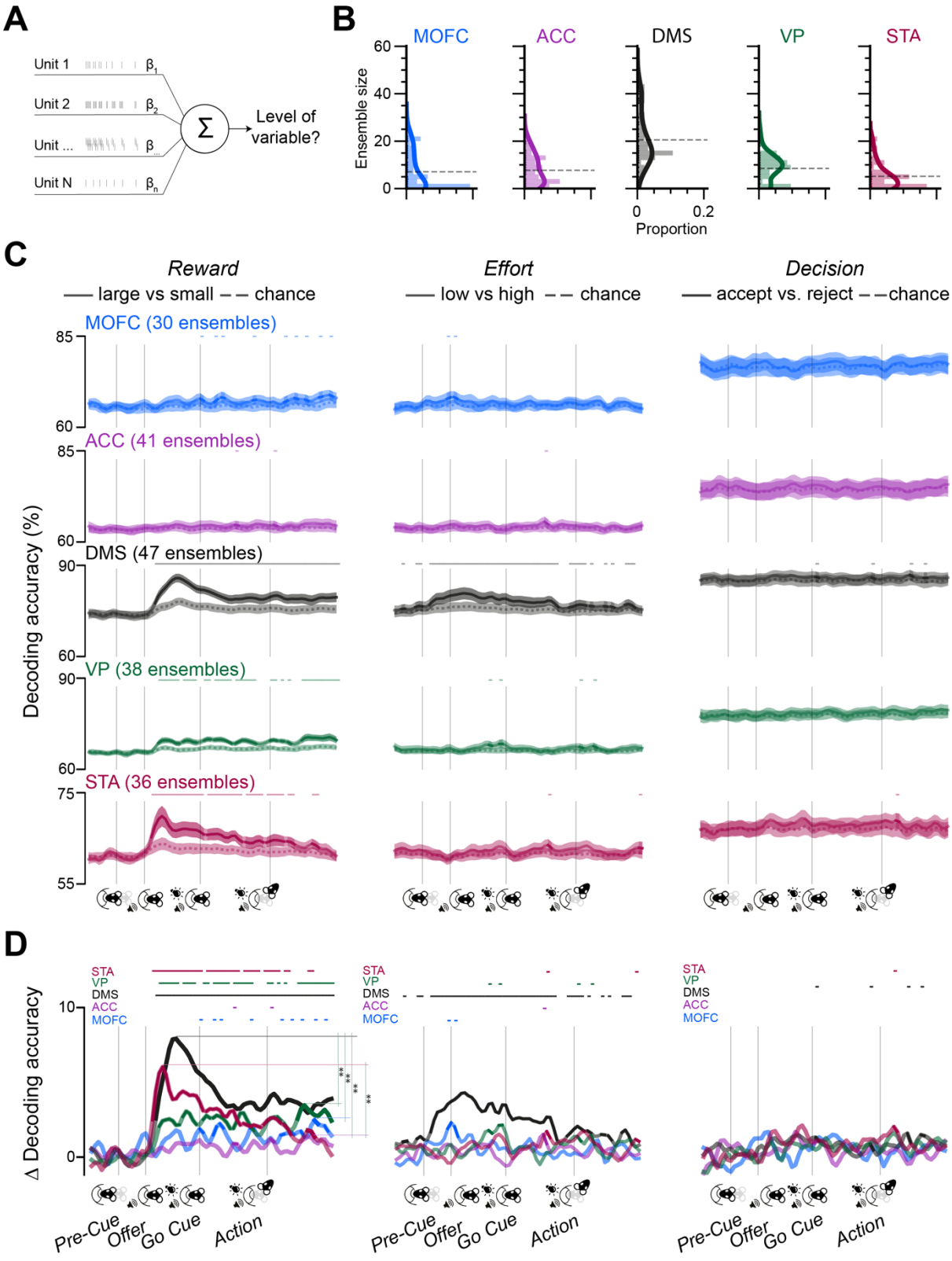



**Figure 1 | Robust coding of reward and effort in population code. A,** daily ensemble firing rate data were regressed onto category labels to yield prediction accuracy estimates. **B,** Distributions of ensemble sizes, depicted separately for each area (columns). Coloured solid lines depict kernel density estimates of the probability density functions; dotted horizontal lines

mark the mean of the distribution. **C,** Ensemble decoder prediction accuracy traces for non-shuffled and shuffled data (solid and dotted shaded lines, respectively). Data are depicted as means (solid lines) and SEM (shaded areas) over individual ensemble data. **D,** Difference traces between average shuffled and non-shuffled decoder accuracies overlaid for each area (coloured solid lines). Coloured straight lines and asterisks indicate significant pairwise differences between the decoder accuracy maxima (**p < .01). **C-D,** Opaque line plot segments and horizontal lines on top of each figure panel depict periods of significant population prediction accuracy over shuffled data (p < .01). Solid vertical lines denote the boundaries of task epochs.

### 3.2 The relationship between ensemble size and decoding accuracy

Given the remarkable similarity of reward and effort representations on the single unit (chapter 2) and ensemble levels, an open question is to what degree their representation depends on concert population activity. If the unit responses to task variables are low-dimensional, then large ensembles would not be expected to outperform small ones in our linear regression approach after adjusting for the shuffled baseline. To address this, we analysed how decoding accuracy in each region covaried with subsampled ensemble sizes. To this end, each original ensemble was subsampled over 20 repetitions and the resulting decoding accuracy averaged. This process was run iteratively with sample sizes varying between 20% up to 100% of the original set in 20% increments.

While baseline-subtracted reward decoder accuracy was only mildly dependent on subsample size, effort representation in the DMS was numerically improved by the proportion of included cells (fig. 2A). To quantify this, we compared accuracies of each subset size within the central time bin of the pre-cue and offer windows. This revealed a significant main effect of subset size on effort decoding accuracy in the DMS during the *offer window* ($F_{4, 230} = 3.44$, p = .009, driven by differences between the 20% vs 60% (p = .038), 20 vs 80% (p = .023) and 20 vs 100% (p = .024) subsets (fig. 2b). A trend main effect of subset size on DMS effort decoding performance was also seen during the pre-cue window ($F_{4, 230} = 2.12$, p = .08) but *post-hoc* comparisons were not carried out here. No robust effect of subset size was detected for any variable in any other region (all F < 1.95, p > .105). Together this suggests that awareness of the level of effort is an emergent property of population-level activity in the DMS.

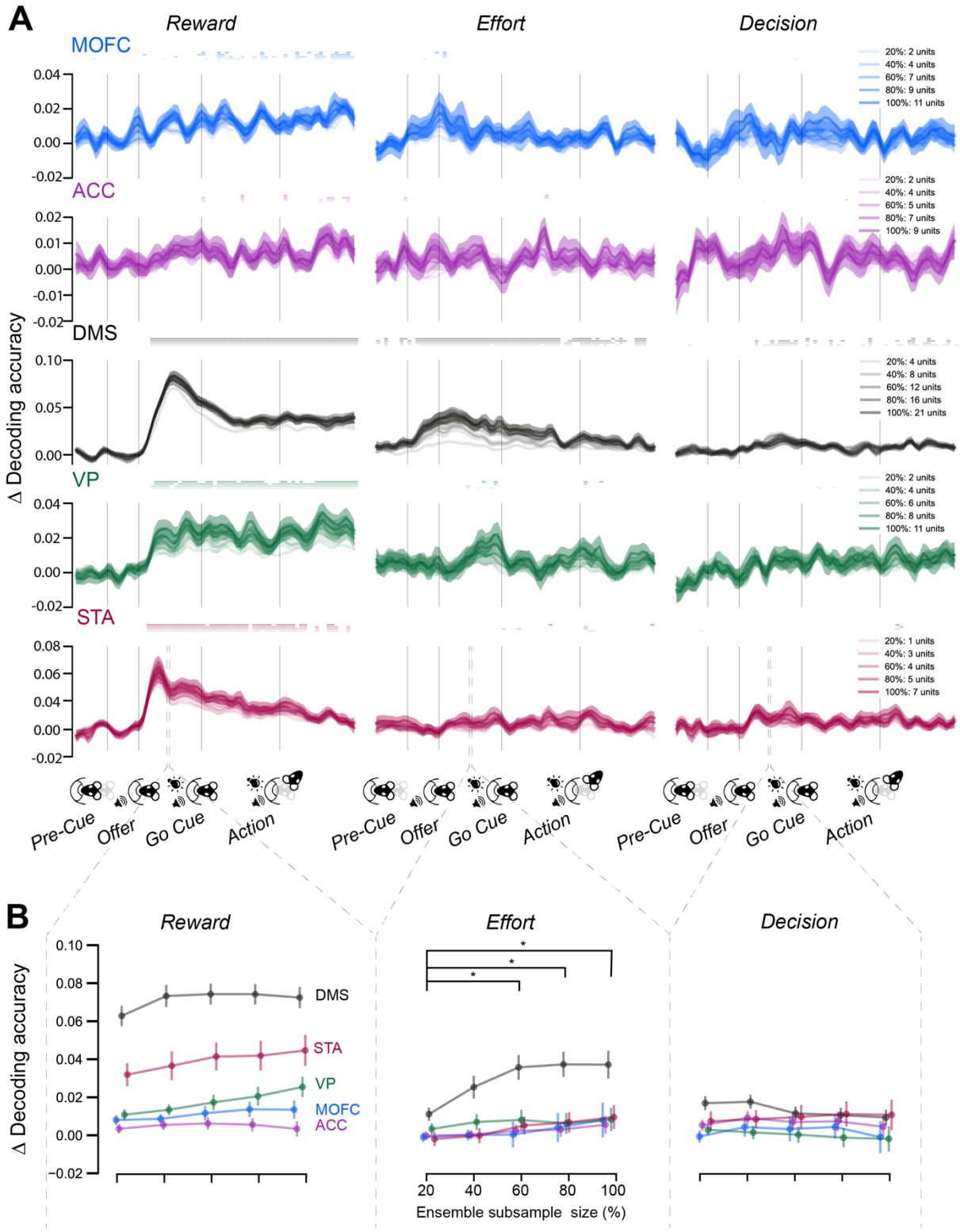



**Figure 2 | The relationship between ensemble size and decoding accuracy. A,** Baseline-adjusted ensemble decoder prediction accuracy traces for subsamples of different sizes (lines of varying opacity). Data are depicted as means (solid lines) and SEM (shaded areas) over individual ensemble data. Horizontal lines on top of each figure panel depict periods of significant population prediction accuracy over shuffled data ($p < .01$). **B,** The relationship

between subset size and decoding accuracy for reward, effort and decision (left, middle, right panels, respectively) quantified in the central bin of the offer window. Coloured lines depict the individual areas; dots and error-bars depict mean and SEM. * $p \leq 0.05$

### 3.3 Time-distributed multiplexing of reward and effort in DMS ensembles

This dichotomy between the mechanisms of reward and effort signalling in the DMS further calls to question whether groups of neurons can simultaneously exhibit *both* the low- and high-dimensional code needed to represent reward and effort. Further, it is unclear what time course such a process might exhibit. An alternative possibility is that some groups might be more devoted to reward, while others effort, signalling. We addressed this by plotting reward vs effort decoding precision in DMS ensembles as trajectories over time during the ITI, pre-cue and offer windows. To visualise the dependence of these trajectories on ensemble subset size, we further partitioned the data along this dimension. This revealed pronounced excursions along the effort decoding axis during the pre-cue window, but marked dynamics along the reward decoding axis during the offer interval (fig. 3). Trajectories along the effort decoding axis were also more pronounced in bigger subsets. Together, this suggests that reward and effort representations are multiplexed over time, rather than simultaneously, in DMS ensembles.

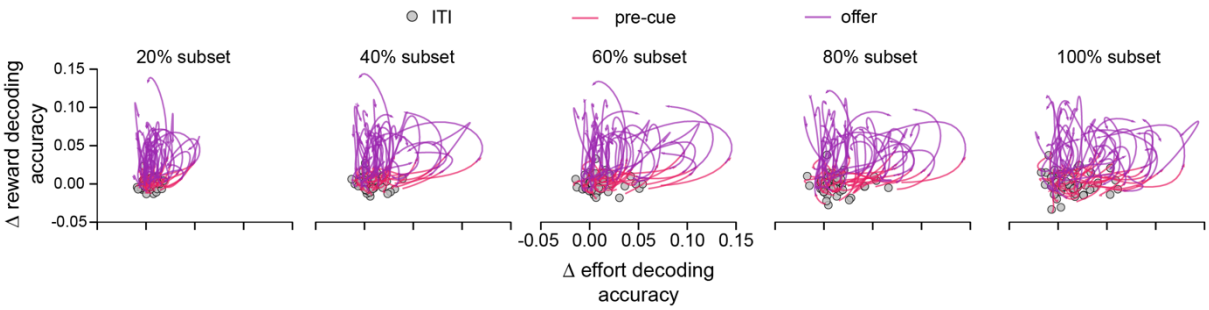



**Figure 3 | Time-distributed multiplexing of reward and effort in DMS ensembles.** Baseline-adjusted reward (y-axis) and effort (x-axis) decoding accuracy in each subset percentile (panels from left to right) are plotted against each other over the ITI (grey dots), pre-cue (pink lines) and offer (purple lines) windows. Each dot or line represents one ensemble. Time direction is depicted with arrows at the tips of the lines.

## 4. Discussion

Population codes have previously been successfully used to infer cognitive and motor variables from primate and rodent brains and can significantly expand the representational capacity of individual units (Rich and Wallis 2016; Ottenheimer et al. 2018; Panzeri et al. 2022). However, a simultaneous portrayal of the population level representation of cost-benefit decision-variables in the frontal cortical-basal ganglia network has not yet been offered. Here, using a linear classifier trained on two levels of each decision variable, we demonstrate that while prospective reward magnitude is robustly decodable from population-level signals in the basal ganglia but not the prefrontal cortex, accept/reject decisions are only very weakly reflected in any region. By contrast, effort was represented in DMS ensembles in a temporally distributed and ensemble-size dependent manner. Further, larger DMS ensembles allowed the multiplexing of reward and effort representations across the pre-cue and offer windows, respectively.

This finding taps into a key mechanistic difference in the representation of reward and effort. While reward in our task is signalled by an explicit and well learned auditory cue, it could reliably be represented by a mapping between the sensory properties of the cue (e.g. tonal frequencies) and an episodic memory representation of the associated outcome value. Further, over the course of training, animals may have learned a cached value pertaining to each cue, rendering it quickly retrievable and correctible by the internal state (Jones et al. 2012; Lee et al. 2012; Hollon et al. 2014). Effort representations in our task, by contrast, rely on divergent streams of information pertaining to episodic and procedural memory, as well as feedback about fatigue from medial and lateral frontal executive and motor systems (Müller et al. 2021). Such integration from different modalities might render effort representations especially dependent on population code.

Interestingly, we report the multiplexing of reward and effort within ensembles over time. However, it is important to stress that what we refer to as ensembles in this work constitute daily samples of cells as the electrodes were advanced through brain tissue.

Whether reward and effort are actually multiplexed by groups of neurons that form *functional* ensembles (e.g. as measured by co-firing patterns in El-Gaby et al. 2021) or share a common anatomical or physiological identity (e.g. MSNs vs TANs) remains to be elucidated.

One caveat that might have undermined our ability to demonstrate similar effort effects in other regions (e.g. in the ACC where strong signalling of this variable was expected (Hillman and Bilkey 2010; Cowen et al. 2012; Porter et al. 2019) was the ensemble size, which tended to be highest in the DMS and numerically lower in the other regions. Consequently, our ability to detect a robust effort read-out and dependencies on subset size in other regions might have been limited purely by the low ensemble sizes. Notably, other reports that have used similar decoding approaches have often achieved large ensemble sizes by constructing *pseudo-ensembles* out of neuronal yields collected over consequent days (Ottenheimer et al. 2018; Li et al. 2022). We specifically avoided this approach, as it carries the risk of biasing the cross-correlational structure between simultaneously recorded neurons (Panzeri et al. 2022; S. Manohar, personal communication). Future work aimed at comparing decoding output characteristics in simultaneously vs sequentially recorded neurons or higher-yield population recordings will help better elucidate the representations of cost-benefit decision variables in the frontal cortical-basal ganglia network.

A further caveat which may have restricted our ability to detect decision variable signals especially in the PFC was the choice of bin size. As was covered in some detail in the discussion of chapter 3, the selection of ~50 ms time-bins may have inappropriately restricted the dynamics that was used as input to the decoders. Consequently, this could have led to disproportionate impairments in our ability to decode variables in mOFC/ACC which are known to display slower-varying activity patterns (Murray et al. 2014) and that can tile entire task epochs (Ozdemir et al. 2020). A comparative examination of these representations across multiple bin sizes or population decoding based on co-activity patterns (El-Gaby et al. 2021) might shed further light on the true underlying patterns of decision variable representation in this network.

# 5. General discussion


## ABSTRACT

Understanding the neurobiological and psychological mechanisms underlying cost-benefit decision-making is important for two reasons. First, variations in cost and/or benefits are ubiquitous in experimental and naturalistic settings, making them indispensable factors to consider when interpreting neuroscience findings. Second, weighing up benefits versus costs is impaired in several neurological and psychiatric diseases, posing a heavy caregiver burden and a hitherto unresolved therapeutic challenge (Pagonabarraga et al. 2015; Husain and Roiser 2018). In this chapter, we will summarise the development a novel rat cost-benefit decision-making task together with an electrophysiological recording setup (chapter 2), used for the subsequent examination of single-unit (chapter 3) and population-level (chapter 4) correlates of variables potentially used by decision-makers to adjudicate between passing up an opportunity versus pursuing a distant reward. We will draw on this experimental work to highlight how this has advanced our understanding of the processing of cost-benefit decision variables across relevant brain networks, as well as underline the most important caveats and directions for future research.


## 1. A novel cost-benefit decision-making task

The neuroscientific literature is replete with different paradigms for the examination of value-guided decision-making. However, there is a striking asymmetry between the number of tasks involving a manipulation of reward as compared to those that vary levels of effort, with the former overshadowing the latter (Walton and Bouret 2019). Further, subjects are often presented with multiple simultaneous offers, which potentially conflicts with the sequential way in which animals often experience offers in more naturalistic settings (Charnov 1976; Kacelnik et al. 2011; Hayden and Walton 2014).

One of the cornerstones of this thesis has thus been the eventual development of a novel cost-benefit *accept-reject choice* paradigm in rats, inspired by equivalent tasks in humans (Chong et al. 2015; Le Heron et al. 2018) and non-human primates (Varazzani et al. 2015; San-Galli et al. 2018). This paradigm features an orthogonal manipulation of payoffs (4 sizes: nil, low, mid, high; varied trial-by-trial; signalled by a distinct auditory cue) and costs (3 levels: low, moderate, high; fixed across blocks of ~20 trials; implemented using different numbers of barriers *en route* to reward). Further, it requires the rats to either move to collect physically distant reward or reject the offer and thus not expend as much energy until the next trial can be initiated.

Crucially, this task was carefully designed to address a few prevalent confounds. First, rats were asked to maintain a nose-poke to initiate trials, ensuring that rats were fully engaged with the task before the presentation of subsequent offer information. Second, rats needed to continue to maintain nose-poking for another 0.6-0.8s after the presentation of the offer cue to disentangle value judgements from reward pursuit. Third, the pre-cue and offer nose-poke duration requirements were jittered to avoid the contamination of decision variable representations by motor preparatory signals. Fourth, to ensure that average reward rate is minimally influenced by accept/reject decisions, ITIs following accepted trials were adjusted by subtracting the cohort median run time.

Here, acceptance of offers was positively impacted by offered reward and negatively influenced by prospective effort cost as well as the reward offered on the previous two trials. Crucially, reductions in the incentive value of the outcome (induced by pre-feeding) robustly caused rats to accept less offers within just one session of testing, proving decisions on this task to be sensitive to rats' internal state. Additionally, patterns of nose-poke interruptions during offer cue presentations and latencies of entering the corridor on accepted trials mirrored the pattern of influences that reward and effort had on decisions to accept, offering further proxies of decision-making on this task. These behavioural patterns were reproduced in implanted animals. It is possible that other proxies of accept/reject decisions can also be identified on this task.

For example, one hitherto unexplored possibility is that the length of time spent off task (in the ITI) was influenced by the level of effort (and possibly cross-trial rates of reward).

The predecessor to this final version of the paradigm was a *two-alternative choice* paradigm. Here, rats instead needed to make a choice between a small "reference" reward that was available nearby the starting nose-poke and a variably large reward available across the corridor. This version lacked an orthogonal effort manipulation and did not require the rats to keep on nose-poking during the presentation of the offer cue. The lack of an effort manipulation, coupled with the rats' persistent bias to attempt retrieval of the small reference reward (even on forced trials where this was not made available) led to the subsequent discontinuation of this paradigm.

Binomial mixed effects' regression models were used throughout to quantify task variable influences on accept/reject choices. Several refinements could be utilised in future research to improve model fits and to glean more mechanistic insights into the determinants of cost-benefit choices. First, care should be taken to model the effort discounting properties that influence offer evaluation within each animal. Aside from reducing residual error, such an approach might offer better a more accurate and individually appropriate index of effort and net value representations in the brain (Hosokawa et al. 2013; Klein-Flügge et al. 2015, 2016; Chong et al. 2017). Second, it is unclear what the most appropriate fixed effects' formula is to capture patterns of cost-benefit choices on our task. Formal model comparisons using an information criterion (as has been done in Akam et al. 2021) to find out the best balance between fixed effect terms and overfitting could help address this shortcoming in the future. Third, cross-trial variations in levels of fatigue or reward rate could have had an additional effect on rats' willingness to accept which was not taken into account in this work. Notably, one recent human study found that fatigue can be partitioned into recoverable (relieving after a trial of rest) and unrecoverable (accumulating across a session) components that successfully capture patterns of engagement with effort for reward in a cost-benefit decision-making task (Müller et al. 2021). Moreover, reward rate has consistently been shown to influence behavioural vigour in rodent experiments (Hamid et al. 2016; Mohebi et al. 2019). Attempts at factoring these latent variables in on our task might

lead to considerable improvements in model performance and may additionally shed light on neuronal activation patterns in the frontal cortical-basal ganglia network on this task.

## 2. Neural signatures of cost-benefit decision-making in a frontal cortical-basal ganglia network

Having set up a suitable platform for the electrophysiological investigation of the neural signatures of cost-benefit decision-making, we next implanted and analysed single-unit recordings from a subset of rats (n=4) in chapter 3. We were able to identify hundreds of neurons per animal in each targeted structure over the course of multiple weeks of daily recording within each animal, with the highest per-animal and overall yield originating from the DMS and the lowest yield from the STA. This was closely mirrored by the amount of neurons identified per structure per day, which was within the range of 5-21. However, neurons in some targeted regions (most notably mOFC and STN) were also spuriously captured from neighbouring regions. As we did not discriminate the yield based on baseline firing rates, waveform characteristics or depth in target area, we obtained a heterogenous population of neurons in each region.

In line with previous findings of ubiquitous and parallel coding of decision variables across the frontal cortical-basal ganglia circuit (Kennerley et al. 2009; Sul et al. 2010; Hosokawa et al. 2013; Stott and David Redish 2014; Hunt et al. 2018; Elston et al. 2019; Hashemnia et al. 2020), we detected examples of neurons signalling reward, effort and decision with heterogenous timing and polarity in each targeted area. This was echoed by population- averaged firing rate responses and regression analyses which showed mostly rapid cue-driven coding of offered reward in the basal ganglia, while cortical reward offer coding was more gradual and less strong. This coincided with slightly less marked but nevertheless statistically robust neuronal recruitment to signal effort in all regions during the offer window, followed by decision coding mainly in the action window in all areas. However, aside from these trends, there was a notable degree of spatial and temporal parallelism in the processing of these variables.

Three dichotomies are apparent when evaluating this pattern of results. First, there is a regional difference in the dynamics with which units are recruited to signal reward between the mPFC and basal ganglia in our task. While the onset of the offer cue was marked by a marked increase in the number of units coding for promised reward in the basal ganglia, the recruitment was slower in the mPFC. This could be reflective of intrinsic connectivity differences between the regions, as frontal cortical areas are known to feature a high degree of interconnectivity (Murray et al. 2014), suitable for the prolonged representation of decision variables such as reward. Second, while reward promise coding was delayed until the onset of the offer cue, effort was signalled by a significant number of units in the ACC, DMS, STA and VP already in the ITI/pre-cue windows, suggesting that awareness of the level of prospective effort sets the stage for the subsequent interpretation of the value of the offer cue in our task. Third, while reward and effort evaluation tend to be concentrated near the presentation of the offer cue, the recruitment of decision-coding neurons in DMS, VP and MOFC is delayed until action initiation. This is suggestive of a process where offer value evaluation subsequently shapes the evolution of a ramp-to-bound type mechanism of decision.

We note, however, that the proportions of units dedicated to the signalling of any of these decision variables tended to be low when compared to most other reports in non-human primates (Kennerley et al. 2009; Bouret and Richmond 2010; Hunt et al. 2018) and rodents (Lopatina et al. 2016). Several differences in methodology, e.g. the choice of the bin size used here for sliding window regressions, the method used for the adjustment of family-wise error rate in multiple comparisons, the anatomical positioning and technique used for the verification of histological placement, the choice of the time window for analyses, and the context of the behavioural task, where reward is physically distant at cue presentation, may underlie this. Notably, in our task effort was varied in blocks of ~ 20 trials and was not signalled by an explicit cue but had to be internally recalled, rendering this manipulation quite different to most other cost-benefit decision-making paradigms. Additionally, here we conservatively clipped the time window of analysis at 1s post-action initiation to limit contamination of decision variable representation by the execution of motor plans. Future analyses that employ a

more extended time window, use a wider bin size, examine changes in the variation of firing rates and verify our anatomical results against other histological quantification methods might cast a different light on the single unit contributions of cost-benefit decision-variable coding on this task. Additionally, regression models that factor in the nesting of neurons within subjects might provide a more appropriate population-level overview of the activity profiles corresponding to decision variables in each area.

We next asked how these decision variables were represented on the level of groups of neurons in the frontal cortical-basal ganglia network. Using linear regressions that predicted the level of each variable from *ensemble* firing rate data in chapter 4, we again found robust read-outs of promised reward information in the basal ganglia, but little evidence for population-level coding of the subsequent accept/reject decision in any area. Interestingly, effort was readable from DMS ensembles across most of the analysed trial time window, and this representation degraded proportionally to the number of units excluded in the training phase. This indicated that while reward signals can be accurately represented on the level of individual units, effort representations might emerge from the divergent activity profiles of groups of neurons. Further, while DMS ensembles were able to multiplex reward and effort information, this happened gradually over time with effort representations preceding reward processing. This feature was again dependent on the size of the ensemble, with bigger ensembles displaying more pronounced multiplexing.

However, a few important limitations may have curbed our ability to describe the true pattern of ensemble-level representations of decision variables in this task. Notably, our ensembles were not functionally coupled groups of neurons, but rather daily samples of cells as the electrodes were advanced through the structures. Moreover, ensemble sizes were relatively small and varied considerably between regions. This, coupled with the drawbacks discussed above in relation to the single-unit analyses, might have significantly limited our ability to draw conclusions about, for example, effort and decision representation in ACC or MOFC ensembles.

## 3. Conclusions and future directions

Together, we demonstrate diverse single-unit and ensemble-level firing rate dynamics underlying cost-benefit decision-making in the rat frontal cortical-basal ganglia network. There is a striking degree of parallelism and distribution in the processing of decision variables across this network, consistent with recently proposed frameworks for value-guided decision-making (Walton et al. 2007; Rushworth et al. 2012; Hunt and Hayden 2017).

Future work will be aimed at addressing the drawbacks discussed above as well as developing a better understanding of the types of decision variables represented in the frontal cortical-basal ganglia network, the network interactions that these processes engage and the differences in the operation of this network between different behavioural contexts. Importantly, we aim to break away from the first-pass grouping of decision variables outlined in this work and move toward a more holistic description of the short- and long-time scale determinants of cost-benefit decisions, with possible implications for the elucidation of the supportive neural mechanisms. To provide a more sophisticated view of the dynamism of this network during motivated behaviour, we also aim to record single-unit data from other important nodes of this network, including the basolateral amygdala and nucleus accumbens (Salamone et al. 1994; Floresco and Ghods-Sharifi 2007; Floresco et al. 2008a; Gan et al. 2010).

Capitalising on this rich set of concurrently recorded single-unit data, we next aim to characterise the patterns of network-level interactions and functional connectivity underlying cost-benefit decision-making, including any potential serial mechanisms of information flow within the network (Elston et al. 2019; Knudsen and Wallis 2020; Gocken et al. 2022). We aim to contrast frontal cortical-basal ganglia network dynamics underlying the behavioural context reported here to one where animals are behaving spontaneously without any concurrent reward and effort associations. To this end, we have recorded a parallel set of open-field single-unit and video tracking data in a subset

of the implanted rats (n=3). Finally, we will also aim to elucidate the causal importance of this network in cost-benefit decision-making, for example by inactivating the targeted brain regions or stimulating the anterior limb of the internal capsule in the rat (Coizet et al. 2017), homologous to the cingulum bundle in humans (Bonnelle et al. 2016; Bubb et al. 2018). Such work can hopefully act as a prelude to enable future studies to examine the effects of pharmacological or electrical neuromodulation to promote willingness to overcome effort constraints, a necessary step towards developing treatments for apathy.

# *Bibliography*


Abitbol R, Lebreton M, Hollard G, et al (2015) Neural mechanisms underlying contextual dependency of subjective values: Converging evidence from monkeys and humans. J Neurosci 35:2308–2320. https://doi.org/10.1523/JNEUROSCI.1878-14.2015

Adam R, Leff A, Sinha N, et al (2013) Dopamine reverses reward insensitivity in apathy following globus pallidus lesions. Cortex 49:1292–1303. https://doi.org/10.1016/j.cortex.2012.04.013

Akam T, Rodrigues-Vaz I, Marcelo I, et al (2021) The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. Neuron 109:149-163.e7. https://doi.org/10.1016/j.neuron.2020.10.013

Albin RL, Young A., Penney JB (1989) The functional anatomy of basal ganglia disorders. Trends Neurosci 12:366–375

Alexander G (1986) Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. Annu Rev Neurosci 9:357–381. https://doi.org/10.1146/annurev.neuro.9.1.357

Andersen P, Silfvenius H, Sundberg SH, Sveen O (1980) A comparison of distal and proximal dendritic synapses on CA1 pyramids in guinea-pig hippocampal slices in vitro. J Physiol 307:273–299

Ang YS, Manohar S, Plant O, et al (2018) Dopamine Modulates Option Generation for Behavior. Curr Biol 28:1561-1569.e3. https://doi.org/10.1016/j.cub.2018.03.069

Association AP (2013) Diagnostic and statistical manual of mental disorders : DSM-5, Fifth edit. Washington, DC

Averbeck BB, Lehman J, Jacobson M, Haber SN (2014) Estimates of projection overlap and zones of convergence within frontal-striatal circuits. J Neurosci 34:9497–9505. https://doi.org/10.1523/JNEUROSCI.5806-12.2014

Bailey MR, Simpson EH, Balsam PD (2016a) Neural substrates underlying effort, time, and risk-based decision making in motivated behavior. Neurobiol Learn Mem 133:233–256. https://doi.org/10.1016/j.nlm.2016.07.015

Bailey MR, Williamson C, Mezias C, et al (2016b) The effects of pharmacological modulation of the serotonin 2C receptor on goal-directed behavior in mice. Psychopharmacology (Berl) 233:615–624. https://doi.org/10.1007/s00213-015-4135-3

Balewski ZZ, Knudsen EB, Wallis JD (2022) Fast and slow contributions to decision-making in corticostriatal circuits. Neuron 110:2170-2182.e4. https://doi.org/10.1016/j.neuron.2022.04.005

Balleine BW, Dickinson A (1998a) Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419. https://doi.org/10.1016/S0028-3908(98)00033-1

Balleine BW, Dickinson A (1998b) The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. Anim Learn Behav 26:46–59. https://doi.org/10.3758/BF03199161

Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang 68:255–278.

https://doi.org/10.1016/j.jml.2012.11.001
Barreiros I., Hironori I, Walton ME, Panayi MC (2021) Defining an Orbitofrontal Compass: Functional and Anatomical Heterogeneity Across Anterior-Posterior and Medial-Lateral Axes. Behav Neurosci 135:166–173
Basu R, Gebauer R, Herfurth T, et al (2021) The orbitofrontal cortex maps future navigational goals. Nature 599:449–452. https://doi.org/10.1038/s41586-021-04042-9
Baunez C, Amalric M, Robbins TW (2002) Enhanced food-related motivation after bilateral lesions of the subthalamic nucleus. J Neurosci 22:562–568. https://doi.org/10.1523/jneurosci.22-02-00562.2002
Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. J R Stat Soc 57:289–300
Benke T, Delazer M, Bartha L, Auer A (2003) Basal ganglia lesions and the theory of fronto-subcortical loops: Neuropsychological findings in two patients with left caudate lesions. Neurocase 9:70–85. https://doi.org/10.1076/neur.9.1.70.14374
Berger M, Gray JA, Roth BL (2009) The expanded biology of serotonin. Annu Rev Med 60:355–366. https://doi.org/10.1146/annurev.med.60.042307.110802
Berns GS, Laibson D, Loewenstein G (2007) Intertemporal choice - toward an integrative framework. Trends Cogn Sci 11:482–488. https://doi.org/10.1016/j.tics.2007.08.011
Berridge KC (2004) Motivation concepts in behavioral neuroscience. Physiol Behav 81:179–209. https://doi.org/10.1016/j.physbeh.2004.02.004
Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res Rev 28:309–369. https://doi.org/10.1016/S0165-0173(98)00019-8
Bhatia KP, Marsden CD (1994) The behavioural and motor consequences of focal lesions of the basal ganglia in man. Brain 117:859–876. https://doi.org/10.1093/brain/117.4.859
Białaszek W, Marcowski P, Ostaszewski P (2017) Physical and cognitive effort discounting across different reward magnitudes: Tests of discounting models. PLoS One 12:1–25. https://doi.org/10.1371/journal.pone.0182353
Björklund A, Dunnett SB (2007) Dopamine neuron systems in the brain: an update. Trends Neurosci 30:194–202. https://doi.org/10.1016/j.tins.2007.03.006
Bloem B, Huda R, Amemori K ichi, et al (2022) Multiplexed action-outcome representation by striatal striosome-matrix compartments detected with a mouse cost-benefit foraging task. Nat Commun 13:. https://doi.org/10.1038/s41467-022-28983-5
Bolles RC (1972) Reinforcement, expectancy, and learning. Psychol Rev 79:394–409. https://doi.org/10.1037/h0033120
Bonnelle V, Manohar S, Behrens T, Husain M (2016) Individual Differences in Premotor Brain Systems Underlie Behavioral Apathy. Cereb Cortex 26:807–819. https://doi.org/10.1093/cercor/bhv247
Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS (2009) How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. Neuron 62:733–743. https://doi.org/10.1016/j.neuron.2009.05.014
Boorman ED, Rushworth MF, Behrens TE (2013) Ventromedial prefrontal and anterior

cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. J Neurosci 33:2242–2253. https://doi.org/10.1523/JNEUROSCI.3022-12.2013
Botvinick MM, Huffstetler S, McGuire JT (2009) Effort discounting in human nucleus accumbens. Cogn Affect Behav Neurosci 9:16–27. https://doi.org/10.3758/CABN.9.1.16
Bouret S, Ravel S, Richmond BJ (2012) Complementary neural correlates of motivation in dopaminergic and noradrenergic neurons of monkeys. Front Behav Neurosci 6:1–12. https://doi.org/10.3389/fnbeh.2012.00040
Bouret S, Richmond BJ (2010) Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. J Neurosci 30:8591–8601. https://doi.org/10.1523/JNEUROSCI.0049-10.2010
Bouret S, Sara SJ (2005) Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. Trends Neurosci 28:574–582. https://doi.org/10.1016/j.tins.2005.09.002
Bouret S, Sara SJ (2004) Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur J Neurosci 20:791–802. https://doi.org/10.1111/j.1460-9568.2004.03526.x
Brandstätter E, Gigerenzer G, Hertwig R (2006) The priority heuristic: Making choices without trade-offs. Psychol Rev 113:409–432. https://doi.org/10.1037/0033-295X.113.2.409
Brehm JW, Self EA (1989) The intensity of motivation. Annu Rev Psychol 40:109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545
Bubb EJ, Metzler-Baddeley C, Aggleton JP (2018) The cingulum bundle: Anatomy, function, and dysfunction. Neurosci Biobehav Rev 92:104–127. https://doi.org/10.1016/j.neubiorev.2018.05.008
Burke CJ, Tobler PN (2011a) Reward skewness coding in the insula independent of probability and loss. J Neurophysiol 106:2415–2422. https://doi.org/10.1152/jn.00471.2011
Burke CJ, Tobler PN (2011b) Coding of reward probability and risk by single neurons in animals. Front Neurosci 5:1–9. https://doi.org/10.3389/fnins.2011.00121
Buzsáki G (2004) Large-scale recording of neuronal ensembles. Nat Neurosci 7:446–451. https://doi.org/10.1038/nn1233
Cabanac M (1971) Physiological Role of Pleasure A stimulus can feel pleasant or unpleasant depending upon its usefulness as determined by internal signals . Science (80- ) 173:1103–1107. https://doi.org/https://doi.org/10.1126/science.173.4002.1103
Cairns H, Oldfield RC, Pennybacker JB, Whiteridge D (1941) Akinetic mutism with an epidermoid cyst of the 3rd ventricle. Brain 64:273–290. https://doi.org/https://doi.org/10.1093/brain/64.4.273
Cannon WB (1980) The wisdom of the body. New York
Cash S, Yuste R (1999) Linear summation of excitatory inputs by CA1 pyramidal neurons. Neuron 22:383–394. https://doi.org/10.1016/S0896-6273(00)81098-3
Charnov EL (1976) Optimal foraging, the marginal value theorem. Theor Popul Biol 129–136
Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common

representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. J Neurosci 29:12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009
Chong TTJ, Apps M, Giehl K, et al (2017) Neurocomputational mechanisms underlying subjective valuation of effort costs. PLoS Biol 15:1–28. https://doi.org/10.1371/journal.pbio.1002598
Chong TTJ, Bonnelle V, Manohar S, et al (2015) Dopamine enhances willingness to exert effort for reward in Parkinson's disease. Cortex 69:40–46. https://doi.org/10.1016/j.cortex.2015.04.003
Churchland AK, Ditterich J (2012) New advances in understanding decisions among multiple alternatives. Curr Opin Neurobiol 22:920–926. https://doi.org/10.1016/j.conb.2012.04.009
Cisek P (2006) Integrated neural processes for defining potential actions and deciding between them: A computational model. J Neurosci 26:9761–9770. https://doi.org/10.1523/JNEUROSCI.5605-05.2006
Cisek P (2012) Making decisions through a distributed consensus. Curr Opin Neurobiol 22:927–936. https://doi.org/10.1016/j.conb.2012.05.007
Cocker PJ, Hosking JG, Benoit J, Winstanley CA (2012) Sensitivity to cognitive effort mediates psychostimulant effects on a novel rodent cost/benefit decision-making task. Neuropsychopharmacology 37:1825–1837. https://doi.org/10.1038/npp.2012.30
Coizet V, Heilbronner SR, Carcenac C, et al (2017) Organization of the Anterior Limb of the Internal Capsule in the Rat. J Neurosci 37:2539–2554. https://doi.org/10.1523/JNEUROSCI.3304-16.2017
Cools R, Nakamura K, Daw ND (2011) Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacology 36:98–113. https://doi.org/10.1038/npp.2010.121
Cowan RL, Wilson CJ, Emson PC, Heizmann C (1990) Parvalbumjn-Containing GmAergic hterneuromin theRat Neostriatmn. 205:
Cowen SL, Davis GA, Nitz DA (2012) Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. J Neurophysiol 107:2393–2407. https://doi.org/10.1152/jn.01012.2011
Crespi LP (1942) Quantitative Variation of Incentive and Performance in the White Rat. Am J Psychol 55:467. https://doi.org/10.2307/1417120
De Martino B, Kumaran D, Holt B, Dolan RJ (2009) The neurobiology of reference-dependent value computation. J Neurosci 29:3833–3842. https://doi.org/10.1523/JNEUROSCI.4832-08.2009
DeLong MR (1990) Primate models of movement disorders of basal ganglia origin. Trends Neurosci 13:281–285. https://doi.org/10.1016/0166-2236(90)90110-V
Der-Avakian A, Barnes SA, Markou A, Pizzagalli D. (2016) Translational Assessment of Reward and Motivational Deficits in Psychiatric Disorders. Curr Top Behav Neurosci 28:231–262. https://doi.org/10.1007/7854_2015_5004
Dolan RJ, Dayan P (2013) Goals and habits in the brain. Neuron 80:312–325. https://doi.org/10.1016/j.neuron.2013.09.007
Dombeck DA, Khabbaz AN, Collman F, et al (2007) Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron 56:43–57. https://doi.org/10.1016/j.neuron.2007.08.003
Dreher JC, Kohn P, Berman KF (2006) Neural coding of distinct statistical properties of

reward information in humans. Cereb Cortex 16:561–573. https://doi.org/10.1093/cercor/bhj004
Ebrahimi S, Lecoq J, Rumyantsev O, et al (2022) Emergent reliability in sensory cortical coding and inter-area communication. Nature 605:713–721. https://doi.org/https://doi.org/10.1038/s41586-022-04724-y https://doi.org/10.1038/s41586-022-04724-y
El-Gaby M, Reeve HM, Lopes-dos-Santos V, et al (2021) An emergent neural coactivity code for dynamic memory. Nat Neurosci 24:694–704. https://doi.org/10.1038/s41593-021-00820-w
Elston TW, Croy E, Bilkey DK (2019) Communication between the Anterior Cingulate Cortex and Ventral Tegmental Area during a Cost-Benefit Reversal Task. Cell Rep 26:2353-2361.e3. https://doi.org/10.1016/j.celrep.2019.01.113
Endepols H, Sommer S, Backes H, et al (2010) Effort-Based Decision Making in the rat: An [18F]Fluorodeoxyglucose Micro Positron Emission Tomography Study. J Neurosci 30:9708–9714. https://doi.org/10.1523/jneurosci.1202-10.2010
Erö C, Gewaltig MO, Keller D, Markram H (2018) A cell atlas for the mouse brain. Front Neuroinform 12:1–16. https://doi.org/10.3389/fninf.2018.00084
Fallon JH, Moore RY (1978) Catecholamine innervation of the basal forebrain. I. The septal area. J Comp Neurol 177:665–683. https://doi.org/10.1002/cne.901770408
Fawcett J, Clark DC, Scheftner WA, Gibbons RD (1983) Assessing Anhedonia in Psychiatric Patients: The Pleasure Scale. Arch Gen Psychiatry 40:79–84. https://doi.org/10.1001/archpsyc.1983.01790010081010
Fletcher PJ, Tampakeras M, Sinyard J, Higgins GA (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 (Berl) 195:223–234. https://doi.org/10.1007/s00213-007-0891-z
Floresco SB, Ghods-Sharifi S (2007) Amygdala-prefrontal cortical circuitry regulates effort-based decision making. Cereb Cortex 17:251–260. https://doi.org/10.1093/cercor/bhj143
Floresco SB, St. Onge JR, Ghods-Sharifi S, Winstanley CA (2008a) Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. Cogn Affect Behav Neurosci 8:375–389. https://doi.org/10.3758/CABN.8.4.375
Floresco SB, Tse MTL, Ghods-Sharifi S (2008b) Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. Neuropsychopharmacology 33:1966–1979. https://doi.org/10.1038/sj.npp.1301565
Flusberg BA, Nimmerjahn A, Cocker ED, et al (2008) High-speed, miniaturized fluorescence microscopy in freely moving mice. Nat Methods 5:935–938. https://doi.org/10.1038/nmeth.1256
Gan JO, Walton ME, Phillips PEM (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat Neurosci 13:25–7. https://doi.org/10.1038/nn.2460
Gardner MPH, Conroy JS, Shaham MH, et al (2017) Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. Neuron 96:1192-1203.e4. https://doi.org/10.1016/j.neuron.2017.10.026
Glimcher PW, Fehr E (2014) Neuroeconomics : decision making and the brain, Second edi. Amsterdam ; Boston
Gocken E, Jasper AI, Semedo JD, et al (2022) Disentangling the flow of signals between

populations of neurons. Nat Comput Neurosci 2:512–525. https://doi.org/10.1038/s43588-022-00282-5
Goethe JW von (1862) Faust. G.H. Wigand, Cassel
Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038
Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. Psychol Bull 130:769–792. https://doi.org/10.1037/0033-2909.130.5.769
Guitart-Masip M, Beierholm UR, Dolan R, et al (2011) Vigor in the face of fluctuating rates of reward: An experimental examination. J Cogn Neurosci 23:3933–3938. https://doi.org/10.1162/jocn_a_00090
Haber SN (2016) Corticostriatal circuitry. Neurosci 21st Century From Basic to Clin Second Ed 1721–1741. https://doi.org/10.1007/978-1-4939-3474-4_135
Haber SN, Tang W, Choi EY, et al (2020) Circuits, Networks, and Neuropsychiatric Disease: Transitioning From Anatomy to Imaging. Biol Psychiatry 87:318–327. https://doi.org/10.1016/j.biopsych.2019.10.024
Habib M, Poncet M (1988) Perte de l'élan vital, de l'intérêt et de l'affectivité (syndrome athymhormique) au cours de lésions lacunaires des corps striés [Loss of vitality, of interest and of the affect (athymhormia syndrome) in lacunar lesions of the corpus striatum]. Rev Neurol 144:751–577
Hamani C, Saint-Cyr JA, Fraser J, et al (2004) The subthalamic nucleus in the context of movement disorders. Brain 127:4–20. https://doi.org/10.1093/brain/awh029
Hamid AA, Pettibone JR, Mabrouk OS, et al (2016) Mesolimbic dopamine signals the value of work. Nat Neurosci 19:117–126. https://doi.org/10.1038/nn.4173
Hare TA, Camerer CF, Rangel A (2009) Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. Science (80- ) 324:646–648
Hare TA, O'Doherty J, Camerer CF, et al (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci 28:5623–5630. https://doi.org/10.1523/JNEUROSCI.1309-08.2008
Hare TA, Schultz W, Camerer CF, et al (2011) Transformation of stimulus value signals into motor commands during simple choice. Proc Natl Acad Sci U S A 108:18120–18125. https://doi.org/10.1073/pnas.1109322108
Härmson O, Grima LL, Panayi MC, et al (2022) 5-HT2C receptor perturbation has bidirectional influence over instrumental vigour and restraint. Psychopharmacology (Berl) 239:123–140. https://doi.org/10.1007/s00213-021-05992-8
Harris JA, Mihalas S, Hirokawa KE, et al (2019) Hierarchical organization of cortical and thalamic connectivity. Nature 575:195–202. https://doi.org/10.1038/s41586-019-1716-z
Harris KD, Shepherd GMG (2015) The neocortical circuit: Themes and variations. Nat Neurosci 18:170–181. https://doi.org/10.1038/nn.3917
Hashemnia S, Euston DR, Gruber AJ (2020) Amphetamine reduces reward encoding and stabilizes neural dynamics in rat anterior cingulate cortex. Elife 1–51
Hauber W, Sommer S (2009) Prefrontostriatal circuitry regulates effort-related decision making. Cereb Cortex 19:2240–2247. https://doi.org/10.1093/cercor/bhn241

Hayden BY (2016) Time discounting and time preference in animals: A critical review. Psychon Bull Rev 23:39–53. https://doi.org/10.3758/s13423-015-0879-3
Hayden BY, Niv Y (2021a) The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). Behav Neurosci 135:192–201. https://doi.org/10.1037/bne0000448
Hayden BY, Niv Y (2021b) The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). Behav Neurosci 135:192–201. https://doi.org/10.1037/bne0000448
Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the anterior cingulate cortex. Science (80- ) 324:948–950. https://doi.org/10.1126/science.1168488
Hayden BY, Walton ME (2014) Neuroscience of foraging. Front Neurosci 8:1–3. https://doi.org/10.3389/fnins.2014.00081
Heath RG (1972) PLEASURE AND BRAIN ACTIVITY IN MAN: Deep and Surface Electroencephalograms During Orgasm. J Nerv Ment Dis 154:3–18. https://doi.org/10.1097/00005053-197201000-00002
Helmchen F, Fee MS, Tank DW, Denk W (2001) A miniature head-mounted two-photon microscope: High-resolution brain imaging in freely moving animals. Neuron 31:903–912. https://doi.org/10.1016/S0896-6273(01)00421-4
Hernádez-López S, Tkatch T, Perez-Garci E, et al (2000) D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLCβ1-IP3-Calcineurin-signaling cascade. J Neurosci 20:8987–8995. https://doi.org/10.1523/jneurosci.20-24-08987.2000
Hernández-López S, Bargas J, Surmeier DJ, et al (1997) Receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J Neurosci 17:3334–3342. https://doi.org/10.1523/jneurosci.17-09-03334.1997
Herzog J, Reiff J, Krack P, et al (2003) Manic episode with psychotic symptoms induced by subthalamic nucleus stimulation in a patient with Parkinson's disease. Mov Disord 18:1382–1384. https://doi.org/10.1002/mds.10530
Higgins GA, Fletcher PJ (2003) Serotonin and drug reward: Focus on 5-HT2C receptors. Eur J Pharmacol 480:151–162. https://doi.org/10.1016/j.ejphar.2003.08.102
Hillman KL, Bilkey DK (2010) Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. J Neurosci 30:7705–7713. https://doi.org/10.1523/JNEUROSCI.1273-10.2010
Hodos W (1961) Progressive Ratio as a Measure of Reward Strength. Sci (American Assoc Adv Sci 134:943–944. https://doi.org/10.1126/science.134.3483.943
Holec V, Pirot HL, Euston DR (2014) Not all effort is equal: The role of the anterior cingulate cortex in different forms of effort-reward decisions. Front Behav Neurosci 8:1–17. https://doi.org/10.3389/fnbeh.2014.00012
Hollon NG, Arnold MM, Gan JO, et al (2014) Dopamine-associated cached values are not sufficient as the basis for action selection. Proc Natl Acad Sci U S A 111:18357–18362. https://doi.org/10.1073/pnas.1419770111
Hosking JG, Cocker PJ, Winstanley CA (2014) Dissociable contributions of anterior cingulate cortex and basolateral amygdala on a rodent cost/benefit decision-making task of cognitive effort. Neuropsychopharmacology 39:1558–1567. https://doi.org/10.1038/npp.2014.27

Hosking JG, Floresco SB, Winstanley CA (2015) Dopamine antagonism decreases willingness to expend physical, but not cognitive, effort: a comparison of two rodent cost/benefit decision-making tasks. Neuropsychopharmacology 40:1005–15. https://doi.org/10.1038/npp.2014.285

Hosokawa T, Kennerley SW, Sloan J, Wallis JD (2013) Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J Neurosci 33:17385–17397. https://doi.org/https://doi.org/10.1523/JNEUROSCI.2221-13.2013

Hubel DH (1957) Tungsten Microelectrode for Recording from Single Units. Science (80- ) 125:549–550

Hunt LT, Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. Nat Publ Gr 18:172–182. https://doi.org/10.1038/nrn.2017.7

Hunt LT, Kolling N, Soltani A, et al (2012) Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci 15:470–476. https://doi.org/10.1038/nn.3017

Hunt LT, Malalasekera WMN, de Berker AO, et al (2018) Triple dissociation of attention and decision computations across prefrontal cortex. Nat Neurosci 21:1471–1481. https://doi.org/10.1038/s41593-018-0239-5

Husain M, Roiser JP (2018) Neuroscience of apathy and anhedonia: A transdiagnostic approach. Nat Rev Neurosci 19:470–484. https://doi.org/10.1038/s41583-018-0029-9

Ishiwari K, Weber SM, Mingote S, et al (2004) Accumbens dopamine and the regulation of effort in food-seeking behavior: Modulation of work output by different ratio or force requirements. Behav Brain Res 151:83–91. https://doi.org/10.1016/j.bbr.2003.08.007

Jean-Richard-Dit-Bressel P, Killcross S, McNally GP (2018) Behavioral and neurobiological mechanisms of punishment: Implications for psychiatric disorders. Neuropsychopharmacology 43:1639–1650. https://doi.org/10.1038/s41386-018-0047-3

Jiang W, Liu J, Zhang D, et al (2019) Short-Term Influence of Recent Trial History on Perceptual Choice Changes with Stimulus Strength. Neuroscience 409:1–15. https://doi.org/10.1016/j.neuroscience.2019.04.010

Jolly E (2018) Pymer4: Connecting R and Python for Linear Mixed Modeling. J Open Source Softw 3:862. https://doi.org/10.21105/joss.00862

Jones JL, Esber GR, McDannald MA, et al Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. Sci (American Assoc Adv Sci 338:953–956. https://doi.org/10.1126/science.1227489

Judge SJ, Gartside SE (2006) Firing of 5-HT neurones in the dorsal and median raphe nucleus in vitro shows differential α1-adrenoceptor and 5-HT1A receptor modulation. Neurochem Int 48:100–107. https://doi.org/10.1016/j.neuint.2005.09.003

Jun JJ, Steinmetz NA, Siegle JH, et al (2017) Fully integrated silicon probes for high-density recording of neural activity. Nature 551:232–236. https://doi.org/10.1038/nature24636

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633. https://doi.org/10.1038/nn2007

Kable JW, Glimcher PW (2009) The Neurobiology of Decision: Consensus and Controversy. Neuron 63:733–745. https://doi.org/10.1016/j.neuron.2009.09.003

Kacelnik A, Vasconcelos M, Monteiro T, Aw J (2011) Darwin's "tug-of-war" vs. starlings' "horse-racing": How adaptations for sequential encounters drive simultaneous choice. Behav Ecol Sociobiol 65:547–558. https://doi.org/10.1007/s00265-010-1101-2
Kahnt T, Park SQ, Haynes JD, Tobler PN (2014) Disentangling neural representations of value and salience in the human brain. Proc Natl Acad Sci U S A 111:5000–5005. https://doi.org/10.1073/pnas.1320189111
Kalis A, Mojzisch A, Schweizer TS, Kaiser S (2008) Weakness of will, akrasia, and the neuropsychiatry of decision making: An interdisciplinary perspective. Cogn Affect Behav Neurosci 8:402–417. https://doi.org/10.3758/CABN.8.4.402
Kemp JM, Powell TPS (1971) The structure of the caudate nucleus of the cat: light and electron microscopy. Phil Trans R Soc Lond B 262:383–401. https://doi.org/https://doi.org/10.1098/rstb.1971.0102
Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision variables. J Cogn Neurosci 21:1162–1178. https://doi.org/10.1162/jocn.2009.21100
Kennerley SW, Wallis JD (2009) Evaluating choices by single neurons in the frontal lobe: Outcome value encoded across multiple decision variables. Eur J Neurosci 29:2061–2073. https://doi.org/10.1111/j.1460-9568.2009.06743.x
Kita T, Kita H, Kitai ST (1984) Passive electrical membrane properties of rat neostriatal neurons in an in vitro slice preparation. Brain Res 300:129–139. https://doi.org/10.1016/0006-8993(84)91347-7
Klein-Flügge MC, Kennerley SW, Friston K, Bestmann S (2016) Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. J Neurosci 36:10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016
Klein-Flügge MC, Kennerley SW, Saraiva AC, et al (2015) Behavioral Modeling of Human Choices Reveals Dissociable Effects of Physical Effort and Temporal Delay on Reward Devaluation. PLoS Comput Biol 11:1–31. https://doi.org/10.1371/journal.pcbi.1004116
Knight FH (2002) Risk, uncertainty and profit. BeardBooks, Washington, D.C.
Knudsen EB, Wallis JD (2020) Closed-Loop Theta Stimulation in the Orbitofrontal Cortex Prevents Reward-Based Learning. Neuron 106:537-547.e4. https://doi.org/10.1016/j.neuron.2020.02.003
Kolling N, Behrens TEJ, Mars RB, Rushworth MFS (2012) Neural mechanisms of foraging. Science (80- ) 335:95–98. https://doi.org/10.1126/science.1216930
Kopec CD, Bowers AC, Pai S, Brody CD (2011) Semi-automated atlas-based analysis of brain histological sections. J Neurosci Methods 196:12–19. https://doi.org/10.1016/j.jneumeth.2010.12.007
Kreitzer AC (2009) Physiology and pharmacology of striatal neurons. Annu Rev Neurosci 32:127–147. https://doi.org/10.1146/annurev.neuro.051508.135422
Kriegeskorte N, Diedrichsen J (2019) Peeling the Onion of Brain Representations. Annu Rev Neurosci 42:407–432. https://doi.org/10.1146/annurev-neuro-080317-061906
Lau B, Glimcher PW (2008) Value Representations in the Primate Striatum during Matching Behavior. Neuron 58:451–463. https://doi.org/10.1016/j.neuron.2008.02.021
Laubach M, Amarante LM, Swanson K, White SR (2018) What, if anything, is rodent

prefrontal cortex? eNeuro 5:. https://doi.org/10.1523/ENEURO.0315-18.2018
Le Heron C, Plant O, Manohar S, et al (2018) Distinct effects of apathy and dopamine on effort-based decision-making in Parkinson's disease. Brain 141:1455–1469. https://doi.org/10.1093/brain/awy110
Le Merre P, Ährlund-Richter S, Carlén M (2021) The mouse prefrontal cortex: Unity in diversity. Neuron 109:1925–1944. https://doi.org/10.1016/j.neuron.2021.03.035
Lee D, Seo H, Jung MW (2012) Neural basis of reinforcement learning and decision making. Annu Rev Neurosci 35:287–308. https://doi.org/10.1146/annurev-neuro-062111-150512
Lein ES, Hawrylycz MJ, Ao N, et al (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445:168–176. https://doi.org/10.1038/nature05453
Li SW, Zeliger O, Strahs L, et al (2022) Frontal neurons driving competitive behaviour and ecology of social groups. Nature 603:661–666. https://doi.org/10.1038/s41586-021-04000-5
Litt A, Plassmann H, Shiv B, Rangel A (2011) Dissociating valuation and saliency signals during decision-making. Cereb Cortex 21:95–102. https://doi.org/10.1093/cercor/bhq065
Lopatina N, McDannald MA, Styer C V., et al (2016) Medial orbitofrontal neurons preferentially signal cues predicting changes in reward during unblocking. J Neurosci 36:8416–8424. https://doi.org/10.1523/JNEUROSCI.1101-16.2016
Lorenz K, Leyhausen P (1973) Motivation of human and animal behavior : an ethological view. Van Nostrand Reinhold Co., New York
Loughlin SE, Foote SL, Grzanna R (1986) Efferent projections of nucleus locus coeruleus: Morphologic subpopulations have different efferent targets. Neuroscience 18:307–319. https://doi.org/10.1016/0306-4522(86)90156-9
Louie K, Khaw MW, Glimcher PW (2013) Normalization is a general neural mechanism for context-dependent decision making. Proc Natl Acad Sci U S A 110:6139–6144. https://doi.org/10.1073/pnas.1217854110
Luo TZ, Bondy AG, Gupta D, et al (2020) An approach for long-term, multi-probe neuropixels recordings in unrestrained rats. Elife 9:1–54. https://doi.org/10.7554/eLife.59716
Macke JH, Nienborg H (2019) Choice (-history) correlations in sensory cortex: cause or consequence? Curr Opin Neurobiol 58:148–154. https://doi.org/10.1016/j.conb.2019.09.005
Mailly P, Aliane V, Groenewegen HJ, et al (2013) The Rat Prefrontostriatal System Analyzed in 3D: Evidence for Multiple Interacting Functional Units. J Neurosci 33:5718–5727. https://doi.org/10.1523/JNEUROSCI.5248-12.2013
Manohar SG, Husain M (2016) Human ventromedial prefrontal lesions alter incentivisation by reward. Cortex 76:104–120. https://doi.org/10.1016/j.cortex.2016.01.005
Marani E, Heida T, Lakke EAJ., Usunoff K. (2008) The subthalamic nucleus Part I: Development, cytology, topography and connections - Abstracts. Springer Nature, BERLIN
Martino B De, Kumaran D, Seymour B, Dolan RJ (2006) Decision-Making in the Human Brain. Science (80- ) 313:684–688
Mashhoori A, Hashemnia S, McNaughton BL, et al (2018) Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. Elife

7:1–18. https://doi.org/10.7554/eLife.29793
McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. Neuron 38:339–346. https://doi.org/10.1016/S0896-6273(03)00154-5
McCullagh P, Nelder JA (1983) Generalized Linear Models, 2nd Edition. Chapman and hall, London New York
McGuire JT, Botvinick MM (2010) Prefrontal cortex, cognitive control, and the registration of decision costs. Proc Natl Acad Sci U S A 107:7922–7926. https://doi.org/10.1073/pnas.0910662107
McNamara CG, Tejero-Cantero Á, Trouche S, et al (2014) Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. Nat Neurosci 17:1658–1660. https://doi.org/10.1038/nn.3843
Mello GBM, Soares S, Paton JJ (2015) A scalable population code for time in the striatum. Curr Biol. https://doi.org/10.1016/j.cub.2015.02.036
Miller NE, Kessen ML (1952) Reward effects of food via stomach fistula compared with those of food via mouth. J Comp Physiol Psychol 45:555–564. https://doi.org/10.1037/h0060113
Mingote S, Weber SM, Ishiwari K, et al (2005) Ratio and time requirements on operant schedules: Effort-related effects of nucleus accumbens dopamine depletions. Eur J Neurosci 21:1749–1757. https://doi.org/10.1111/j.1460-9568.2005.03972.x
Mohebi A, Pettibone JR, Hamid AA, et al (2019) Dissociable dopamine dynamics for learning and motivation. Nature 570:65–70. https://doi.org/10.1038/s41586-019-1235-y
Molnár Z, Cheung AFP (2006) Towards the classification of subpopulations of layer V pyramidal projection neurons. Neurosci Res 55:105–115. https://doi.org/10.1016/j.neures.2006.02.008
Monosov IE, Hikosaka O (2012) Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. J Neurosci 32:10318–10330. https://doi.org/10.1523/JNEUROSCI.1801-12.2012
Muhammed K, Manohar S, Ben Yehuda M, et al (2016) Reward sensitivity deficits modulated by dopamine are associated with apathy in Parkinson's disease. Brain 139:2706–2721. https://doi.org/10.1093/brain/aww188
Müller T, Klein-Flügge MC, Manohar SG, et al (2021) Neural and computational mechanisms of momentary fatigue and persistence in effort-based choice. Nat Commun 12:. https://doi.org/10.1038/s41467-021-24927-7
Münster A, Hauber W (2018) Medial orbitofrontal cortex mediates effort-related responding in rats. Cereb Cortex 28:4379–4389. https://doi.org/10.1093/cercor/bhx293
Murray JD, Bernacchia A, Freedman DJ, et al (2014) A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17:1661–1663. https://doi.org/10.1038/nn.3862
Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal "hyperdirect" pathway. Neurosci Res 43:111–117. https://doi.org/10.1016/S0168-0102(02)00027-5
Németh G, Hegedüs K, Molnâr L (1988) Akinetic mutism associated with bicingular lesions: Clinicopathological and functional anatomical correlates. Eur Arch Psychiatry Neurol Sci 237:218–222. https://doi.org/10.1007/BF00449910
Niv Y, Joel D, Dayan P (2006) A normative perspective on motivation. Trends Cogn Sci

10:375–381. https://doi.org/10.1016/j.tics.2006.06.010
Nougaret S, Baunez C, Ravel S (2022) Neurons in the Monkey's Subthalamic Nucleus Differentially Encode Motivation and Effort. J Neurosci 42:2539–2551. https://doi.org/10.1523/JNEUROSCI.0281-21.2021
Ortiz C, Navarro JF, Jurek A, et al (2020) Molecular atlas of the adult mouse brain. Sci Adv 6:1–14. https://doi.org/10.1126/sciadv.abb3446
Ortoleva P (2013) The price of flexibility: Towards a theory of Thinking Aversion. J Econ Theory 148:903–934. https://doi.org/10.1016/j.jet.2012.10.009
Ostlund SB, Balleine BW (2007) Orbitofrontal cortex mediates outcome encoding in pavlovian but not instrumental conditioning. J Neurosci 27:4819–4825. https://doi.org/10.1523/JNEUROSCI.5443-06.2007
Ottenheimer DJ, Richard JM, Janak PH (2018) Ventral pallidum encodes relative reward value earlier and more robustly than nucleus accumbens. Nat Commun 9:. https://doi.org/10.1038/s41467-018-06849-z
Ottenheimer DJ, Wang K, Tong X, et al (2020) Reward activity in ventral pallidum tracks satiety-sensitive preference and drives choice behavior. Sci Adv 6:27–30. https://doi.org/e
Ozdemir AT, Lagler M, Lagoun S, et al (2020) Unexpected Rule-Changes in a Working Memory Task Shape the Firing of Histologically Identified Delay-Tuned Neurons in the Prefrontal Cortex. Cell Rep 30:1613-1626.e4. https://doi.org/10.1016/j.celrep.2019.12.102
Pachitariu M, Steinmetz N, Kadir S, et al (2016) Fast and accurate spike sorting of high-channel count probes with KiloSort. Adv Neural Inf Process Syst 4455–4463
Padoa-Schioppa C (2011) Neurobiology of economic choice: A good-based model. Annu Rev Neurosci 34:333–359. https://doi.org/10.1146/annurev-neuro-061010-113648
Pagonabarraga J, Kulisevsky J, Strafella AP, Krack P (2015) Apathy in Parkinson's disease: Clinical features, neural substrates, diagnosis, and treatment. Lancet Neurol 14:518–531. https://doi.org/10.1016/S1474-4422(15)00019-8
Panayi MC, Killcross S (2018) Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. Elife 7:1–27. https://doi.org/10.7554/elife.37357
Panigrahi B, Martin KA, Li Y, et al (2015) Dopamine Is Required for the Neural Representation and Control of Movement Vigor. Cell. https://doi.org/10.1016/j.cell.2015.08.014
Panzeri S, Moroni M, Safaai H, Harvey CD (2022) The structures and functions of correlations in neural population codes. Nat Rev Neurosci 23:. https://doi.org/10.1038/s41583-022-00606-4
Parent A, Hazrati L-N (1995a) Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res Rev 20:91–127
Parent A, Hazrati LN (1995b) Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidium in basal ganglia circuitry. Brain Res Rev 20:128–154. https://doi.org/10.1016/0165-0173(94)00008-D
Parvizi J, Rangarajan V, Shirer WR, et al (2013) The will to persevere induced by electrical stimulation of the human cingulate gyrus. Neuron 80:1359–1367. https://doi.org/10.1016/j.neuron.2013.10.057
Paxinos G, Watson C (2006) The rat brain in stereotaxic coordinates, 6th ed.

Academic, Amsterdam ; Boston ; Oxford
Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates, 4th ed. Academic Press, San Diego
Philiastides MG, Biele G, Heekeren HR (2010) A mechanistic account of value computation in the human brain. Proc Natl Acad Sci U S A 107:9430–9435. https://doi.org/10.1073/pnas.1001732107
Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. J Neurosci 27:9984–9988. https://doi.org/10.1523/JNEUROSCI.2131-07.2007
Podvalny E, Noy N, Harel M, et al (2015) A unifying principle underlying the extracellular field potential spectral responses in the human cortex. J Neurophysiol 114:505–519. https://doi.org/10.1152/jn.00943.2014
Porter B, Hillman KL (2019) A Novel Weight Lifting Task for Investigating Effort and Persistence in Rats. Front Behav Neurosci 13:1–13. https://doi.org/10.3389/fnbeh.2019.00275
Porter BS, Hillman KL, Bilkey DK (2019) Anterior cingulate cortex encoding of effortful behavior. J Neurophysiol 121:701–714. https://doi.org/10.1152/jn.00654.2018
Postuma RB, Dagher A (2006) Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. Cereb Cortex 16:1508–1521. https://doi.org/10.1093/cercor/bhj088
Prange S, Pagonabarraga J, Krack P, et al (2018) Historical crossroads in the conceptual delineation of apathy in Parkinson's disease. Brain 141:613–619. https://doi.org/10.1093/brain/awx362
Quilodran R, Rothé M, Procyk E (2008) Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. Neuron 57:314–325. https://doi.org/10.1016/j.neuron.2007.11.031
Ramanathan S, Hanley JJ, Deniau JM, Bolam JP (2002) Synaptic convergence of motor and somatosensory cortical afferents onto GABAergic interneurons in the rat striatum. J Neurosci 22:8158–8169. https://doi.org/10.1523/jneurosci.22-18-08158.2002
Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. Curr Opin Neurobiol 20:262–270. https://doi.org/10.1016/j.conb.2010.03.001
Ratcliff R, McKoon G (2008) The diffusion decision model: Theory and data for two-choice decision tasks. Neural Comput 20:873–922. https://doi.org/10.1162/neco.2008.12-06-420
Ratcliff R, Tuerlinckx F (2002) Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. Psychon Bull Rev 9:438–481. https://doi.org/10.3758/BF03196302
Ressler N (2004) Rewards and punishments, goal-directed behavior and consciousness. Neurosci Biobehav Rev 28:27–39. https://doi.org/10.1016/j.neubiorev.2003.10.003
Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19:973–980. https://doi.org/10.1038/nn.4320
Robert P, Onyike CU, Leentjens AFG, et al (2009) Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. Eur Psychiatry 24:98–104. https://doi.org/10.1016/j.eurpsy.2008.09.001

Robinson TE, Berridge KC (1993) The neural basis of drug craving: An incentive-sensitization theory of addiction. Brain Res Rev 18:247–291. https://doi.org/10.1016/0165-0173(93)90013-P
Roesch MR, Olson CR (2003) Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. J Neurophysiol 90:1766–1789. https://doi.org/10.1152/jn.00019.2003
Root DH, Melendez RI, Zaborszky L, Napier TC (2015) The ventral pallidum: Subregion-specific functional anatomy and roles in motivated behaviors. Prog Neurobiol 130:29–70. https://doi.org/10.1016/j.pneurobio.2015.03.005
Rudebeck PH, Walton ME, Smyth AN, et al (2006) Separate neural pathways process different decision costs. Nat Neurosci 9:1161–1168. https://doi.org/10.1038/nn1756
Rushworth MFS, Kolling N, Sallet J, Mars RB (2012) Valuation and decision-making in frontal cortex: One or many serial or parallel systems? Curr Opin Neurobiol 22:946–955. https://doi.org/10.1016/j.conb.2012.04.011
Rushworth MFS, Noonan MAP, Boorman ED, et al (2011) Frontal Cortex and Reward-Guided Learning and Decision-Making. Neuron 70:1054–1069. https://doi.org/10.1016/j.neuron.2011.05.014
Salamone JD (1996) The behavioral neurochemistry of motivation: Methodological and conceptual issues in studies of the dynamic activity of nucleus accumbens dopamine. J Neurosci Methods 64:137–149. https://doi.org/10.1016/0165-0270(95)00125-5
Salamone JD, Cousins MS, Bucher S (1994) Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. Behav Brain Res 65:221–229. https://doi.org/10.1016/0166-4328(94)90108-2
Salamone JD, Kurth PA, McCullough LD, et al (1993) The role of brain dopamine in response initiation: effects of haloperidol and regionally specific dopamine depletions on the local rate of instrumental responding. Brain Res 628:218–226. https://doi.org/10.1016/0006-8993(93)90958-P
Salamone JD, Steinpreis RE, Mccullough LD, et al (1991) Haloperidol and nucleus accumbens dopamine depletion suppress lever pressing for food but increase free food consumption in a novel food choice procedure. Psychopharmacology (Berl) 104:515–521
Salgado M (2011) Choosing to have Less Choice. SSRN Electron J. https://doi.org/10.2139/ssrn.889311
Samejima K, Ueda Y, Doya K (2005) Representation of Action-Specific Reward Values in the Striatum. 310:1337–1341
San-Galli A, Varazzani C, Abitbol R, et al (2018) Primate ventromedial prefrontal cortex neurons continuously encode the willingness to engage in reward-directed behavior. Cereb Cortex 28:73–89. https://doi.org/10.1093/cercor/bhw351
Sara SJ, Bouret S (2012) Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. Neuron 76:130–141. https://doi.org/10.1016/j.neuron.2012.09.011
Sato F, Lavallée P, Lévesque M, Parent A (2000a) Single-axon tracing study of neurons of the external segment of the globus pallidus in primate. J Comp Neurol 417:17–31. https://doi.org/10.1002/(SICI)1096-9861(20000131)417:1<17::AID-CNE2>3.0.CO;2-I

Sato F, Parent M, Levesque M, Parent A (2000b) Axonal branching pattern of neurons of the subthalamic nucleus in primates. J Comp Neurol 424:142–152. https://doi.org/10.1002/1096-9861(20000814)424:1<142::AID-CNE10>3.0.CO;2-8
Schlaepfer TE, Cohen MX, Frick C, et al (2008) Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. Neuropsychopharmacology 33:368–377. https://doi.org/10.1038/sj.npp.1301408
Schultz W (2006) Behavioral theories and the neurophysiology of reward. Annu Rev Psychol 57:87–115. https://doi.org/10.1146/annurev.psych.56.091103.070229
Schultz W (2016) Dopamine reward prediction-error signalling: A two-component response. Nat. Rev. Neurosci. 17:183–195
Schultz W (2007) Multiple dopamine functions at different time courses. Annu Rev Neurosci 30:259–288. https://doi.org/10.1146/annurev.neuro.28.061604.135722
Schultz W (2015) Neuronal reward and decision signals: From theories to data. Physiol Rev 95:853–951. https://doi.org/10.1152/physrev.00023.2014
Schultz W, Dayan P, Montague PR (1997) Neural Substrate of Prediction and Reward. Science (80- ) 275:1593–1599. https://doi.org/10.1126/science.275.5306.1593
Schweimer J, Hauber W (2006) Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. Learn Mem 13:777–782. https://doi.org/10.1101/lm.409306
Schweimer J, Hauber W (2005) Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. Learn Mem 12:334–342. https://doi.org/10.1101/lm.90605
Sebanz N, Prinz W, Ainslie G, et al (2006) Disorders of Volition. MIT Press, Cambridge
Semedo JD, Zandvakili A, Machens CK, et al (2019) Cortical Areas Interact through a Communication Subspace. Neuron 102:249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026
Seymour B, Singer T, Dolan R (2007) The neurobiology of punishment. Nat Rev Neurosci 8:300–311. https://doi.org/10.1038/nrn2119
Shadmehr R, Reppert TR, Summerside EM, et al (2019) Movement vigor as a reflection of subjective economic utility. Trends Neurosci 42:1–24. https://doi.org/10.1016/j.tins.2019.02.003
Shamash P, Carandini M, Harris K, Steinmetz N (2018) A tool for analyzing electrode tracks from slice histology. bioRxiv 447995. https://doi.org/10.1101/447995
Shipp S (2017) The functional logic of corticostriatal connections. Brain Struct Funct 222:669–706. https://doi.org/10.1007/s00429-016-1250-9
Silveira MM, Wittekindt SN, Mortazavi L, et al (2020) Investigating serotonergic contributions to cognitive effort allocation, attention, and impulsive action in female rats. J Psychopharmacol 34:452–466. https://doi.org/10.1177/0269881119896043
Simpson EH, Kellendonk C, Ward RD, et al (2011) Pharmacologic rescue of motivational deficit in an animal model of the negative symptoms of schizophrenia. Biol Psychiatry 69:928–935. https://doi.org/10.1016/j.biopsych.2011.01.012
Sinha N, Manohar S, Husain M (2013) Impulsivity and apathy in Parkinson's disease. J Neuropsychol 7:255–283. https://doi.org/10.1111/jnp.12013
Smith KS, Tindell AJ, Aldridge JW, Berridge KC (2009) Ventral pallidum roles in reward and motivation. Behav Brain Res 196:155–167.

https://doi.org/10.1016/j.bbr.2008.09.038
Smith Y, Bevan MD, Shink E, Bolam JP (1998) Microcircuitry of the direct and indirect pathways of the basal ganglia. Neuroscience 86:353–387. https://doi.org/10.1016/S0306-4522(98)00004-9
Snaith RP, Hamilton M, Morley S, et al (1995) A scale for the assessment of hedonic tone. The Snaith-Hamilton Pleasure Scale. Br J Psychiatry 167:99–103. https://doi.org/10.1192/bjp.167.1.99
Sokolowski JD, Salamone JD (1998) The role of accumbens dopamine in lever pressing and response allocation: Effects of 6-OHDA injected into core and dorsomedial shell. Pharmacol Biochem Behav 59:557–566. https://doi.org/10.1016/S0091-3057(97)00544-3
Soubrié P (1986) Reconciling the role of central serotonin neurons in human and animal behavior. Behav Brain Sci 9:319–335. https://doi.org/10.1017/s0140525x00022871
St. Onge JR, Floresco SB (2009) Dopaminergic modulation of risk-based decision making. Neuropsychopharmacology 34:681–697. https://doi.org/10.1038/npp.2008.121
Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD (2019) Distributed coding of choice, action and engagement across the mouse brain. Nature 576:1–8. https://doi.org/10.1038/s41586-019-1787-x
Stephens DW (2008) Decision ecology: Foraging and the ecology of animal decision making. Cogn Affect Behav Neurosci 8:475–484. https://doi.org/10.3758/CABN.8.4.475
Stephens DW, Krebs JR (1986) Foraging theory. Princeton
Stott JJ, David Redish A (2014) A functional difference in information processing between orbitofrontal cortex and ventral striatum during decision-making behaviour. Philos Trans R Soc B Biol Sci 369:. https://doi.org/10.1098/rstb.2013.0472
Strait CE, Blanchard TC, Hayden BY (2014) Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron 82:1357–1366. https://doi.org/10.1016/j.neuron.2014.04.032
Strait CE, Sleezer BJ, Blanchard TC, et al (2016) Neuronal selectivity for spatial positions of offers and choices in five reward regions. J Neurophysiol 115:1098–1111. https://doi.org/10.1152/jn.00325.2015
Sul JH, Kim H, Huh N, et al (2010) Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron 66:449–460. https://doi.org/10.1016/j.neuron.2010.03.033
Sutherland WJ, Stephens DW, Krebs JR (1988) Foraging Theory.
Tachibana Y, Hikosaka O (2012) The Primate Ventral Pallidum Encodes Expected Reward Value and Regulates Motor Action. Neuron 76:826–837. https://doi.org/10.1016/j.neuron.2012.09.030
Tajima S, Drugowitsch J, Patel N, Pouget A (2019) Optimal policy for multi-alternative decisions. Nat Neurosci 22:1503–1511. https://doi.org/10.1038/s41593-019-0453-9
Tajima S, Drugowitsch J, Pouget A (2016) Optimal policy for value-based decision-making. Nat Commun 7:1–12. https://doi.org/10.1038/ncomms12400
Tanno T, Silberberg A (2014) Inter-response-time reinforcement and relative reinforcer frequency control choice. Learn Behav 43:54–71.

https://doi.org/10.3758/s13420-014-0161-y
Tasic B, Yao Z, Graybuck LT, et al (2018) Shared and distinct transcriptomic cell types across neocortical areas. Nature 563:72–78. https://doi.org/10.1038/s41586-018-0654-5
Taylor CR, Schmidt-Nielsen K, Raab JL (1970) Scaling of energetic cost of running to body size in mammals. Am J Physiol 219:1104–1107. https://doi.org/10.1152/ajplegacy.1970.219.4.1104
Tepper JM, Bolam JP (2004) Functional diversity and specificity of neostriatal interneurons. Curr Opin Neurobiol 14:685–692. https://doi.org/10.1016/j.conb.2004.10.003
Thomsen KR (2015) Measuring anhedonia: impaired ability to pursue, experience, and learn about reward. Front Psychol 6:1–11. https://doi.org/10.3389/fpsyg.2015.01409
Thomsen KR, Whybrow PC, Kringelbach ML (2015) Reconceptualizing anhedonia: Novel perspectives on balancing the pleasure networks in the human brain. Front Behav Neurosci 9:1–23. https://doi.org/10.3389/fnbeh.2015.00049
Toates FM (1986) Motivational systems. Cambridge University Press, Cambridge
Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. Science (80- ) 307:1642–1645. https://doi.org/10.1126/science.1105370
Treadway MT, Zald DH, Thomsen KR, et al (2015) Pleasure Systems in the Brain. Neuron 35:646–664. https://doi.org/10.3389/fpsyg.2015.01409
Trillet M, Croisile B, Tourniaire D, Schott B (1990) Perturbations de l'activité motrice volontaire et lésions des noyaux caudés [Disorders of voluntary motor activity and lesions of caudate nuclei]. Rev Neurol 146:338–344
Turner LH, Solomon RL, Stellar E, Wampler SN (1975) Humoral factors controlling food intake in dogs. Acta Neurobiol Exp (Wars) 35:491–498
Tversky A, Simonson I (1993) Context-Dependent Preferences. Manage Sci 39:1179–1189. https://doi.org/10.1287/mnsc.39.10.1179
Valenstein ES, Cox VC, Kakolewski JW (1970) Reexamination of the role of the hypothalamus in motivation. Psychol Rev 77:16–31. https://doi.org/10.1037/h0028581
Van De Werd HJJM, Rajkowska G, Evers P, Uylings HBM (2010) Cytoarchitectonic and chemoarchitectonic characterization of the prefrontal cortical areas in the mouse. Brain Struct Funct 214:339–353. https://doi.org/10.1007/s00429-010-0247-z
Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015) Noradrenaline and Dopamine Neurons in the Reward/Effort Trade-Off: A Direct Electrophysiological Comparison in Behaving Monkeys. J Neurosci 35:7866–7877. https://doi.org/10.1523/JNEUROSCI.0454-15.2015
Vickery TJ, Chun MM, Lee D (2011) Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. Neuron 72:166–177. https://doi.org/10.1016/j.neuron.2011.08.011
Von Neumann J, Morgenstern O (2007) Theory of games and economic behavior, 60th anniversary ... Princeton University Press, Princeton, N.J. ; Woodstock
Wakabayashi KT, Fields HL, Nicola SM (2004) Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. Behav Brain Res 154:19–30. https://doi.org/10.1016/j.bbr.2004.01.013
Walton M, Bannerman D, Alterescu K, Rushworth M (2003) Functional Specialization

within Medial Frontal Cortex ofthe Anterior Cingulate for Evaluating Effort-Related Decisions. J Neurosci 23:6475–6479. https://doi.org/10.1207/s15328015tlm1802_3
Walton ME, Bannerman DM, Rushworth MFS (2002) The role of rat medial frontal cortex in effort-based decision making. J Neurosci 22:10996–1003
Walton ME, Bouret S (2019) What Is the Relationship between Dopamine and Effort? Trends Neurosci 42:79–91. https://doi.org/10.1016/j.tins.2018.10.001
Walton ME, Kennerley SW, Bannerman DM, et al (2006) Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. Neural Networks 19:1302–1314. https://doi.org/10.1016/j.neunet.2006.03.005
Walton ME, Rudebeck PH, Bannerman DM, Rushworth MFS (2007) Calculating the cost of acting in frontal cortex. Ann N Y Acad Sci 1104:340–356. https://doi.org/10.1196/annals.1390.009
Wang AY, Miura K, Uchida N (2013) The dorsomedial striatum encodes net expected return , critical for energizing performance vigor. Nat Neurosci 16:639–647. https://doi.org/10.1038/nn.3377
Willett FR, Avansino DT, Hochberg LR, et al (2021) High-performance brain-to-text communication via handwriting. Nature 593:249–254. https://doi.org/10.1038/s41586-021-03506-2
Wilson CJ, Chang HT, Kitai ST (1990) Firing patterns and synaptic potentials of identified giant aspiny interneurons in the rat neostriatum. J Neurosci 10:508–519. https://doi.org/10.1523/jneurosci.10-02-00508.1990
Wilson CRE, Gaffan D, Browning PGF, Baxter MG (2010) Functional localization within the prefrontal cortex: Missing the forest for the trees? Trends Neurosci 33:533–540. https://doi.org/10.1016/j.tins.2010.08.001
Wilson MA, McNaughton BL Dynamics of the Hippocampal Ensemble Code for Space. Sci (American Assoc Adv Sci 261:1055–1058. https://doi.org/10.1126/science.8351520
Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron 81:267–279. https://doi.org/10.1016/j.neuron.2013.11.005
Wirtshafter D, Davis JD (1977) Set Points, Settling Points, and the Control of Body Weight. Physiol Behav 19:75–78
Wittmann MK, Kolling N, Akaishi R, et al (2016) Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. Nat Commun 7:12327. https://doi.org/10.1038/ncomms12327
Yazbek H, Norton J, Capdevielle D, et al (2014) The Lille Apathy Rating Scale (LARS): Exploring its psychometric properties in schizophrenia. Schizophr Res 157:278–284. https://doi.org/10.1016/j.schres.2014.04.034
Yoon T, Geary RB, Ahmed AA, Shadmehr R (2018) Control of movement vigor and decision making during foraging. Proc Natl Acad Sci U S A 115:E10476–E10485. https://doi.org/10.1073/pnas.1812979115
Zénon A, Devesse S, Olivier E (2016a) Dopamine Manipulation Affects Response Vigor. J Neurosci 36:9516–9525. https://doi.org/10.1523/JNEUROSCI.4467-15.2016
Zénon A, Duclos Y, Carron R, et al (2016b) The human subthalamic nucleus encodes the subjective value of reward and the cost of effort during decision-making. Brain 139:1830–1843. https://doi.org/10.1093/brain/aww075
Zheng T, Wilson CJ (2002) Corticostriatal combinatorics: The implications of

corticostriatal axonal arborizations. J Neurophysiol 87:1007–1017. https://doi.org/10.1152/jn.00519.2001