Neuronal Circuits Underlying Flexible Behaviour



Inês Vitória Barreiros

Christ Church College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term, 2021

Inês Vitória Barreiros

Neuronal Circuits Underlying Flexible Behaviour

Supervisors: Dr Mark Walton and Prof Tim Behrens

University of Oxford

Christ Church College

Oxford, UK

Acknowledgements

Perhaps a PhD should start with writing down our vows: in sickness and in health, in first attempt successes or experiments complete failure, in Science or in bioRxiv... until *viva* do us part. It was a long and sometimes rocky journey, but it was not a lonely one. I would not have come this far without the support of many people who directly or indirectly contributed to the work in this thesis. I would like to take this opportunity to express my deepest gratitude to them.

First and foremost, I would like to thank my supervisors Mark Walton and Tim Behrens. Thank you for accepting me as part of your research groups and for being great mentors and true role models. With you, I have regained my passion for research and learned a great deal about experimental design and how to do thorough and rigorous analyses, from thinking about all the right controls to spotting when something is off because of some annoying bug in the code. But, above all, thank you for always being so kind and encouraging me to do what makes me happy.

Part of my PhD work was also supervised by Marios Panayi, who was an invaluable source of wisdom, helping me begin to understand the magical orbitofrontal cortex and guiding me through the world of statistical testing. Thank you for the fantastic collaboration on the neuroanatomy study and for the helpful discussions and data analysis advice on all other projects. On top of all that, thank you for being a good friend!

Many other members of the Walton lab have contributed to the work in this thesis. First, I must thank Thomas Akam for his help designing and building the operant boxes and for valuable insights into the design of behavioural experiments. I am also grateful to Marta Blanco Pozo and Yves Wiesenberger for help with the modelling of the behavioural data. But not everything in the Walton lab is about work. I take with me many good memories from house parties, “holibobs”, and after-work drinks at the Uni Club. Special thanks to Laura Grima, Lauren Burgeno, and Tom JP for driving the Walton lab social life.

I am also grateful to our collaborators at Eli Lilly (UK), especially Gary Gilmour and John Huxter, for providing data from the really interesting experiments included in Chapters 4 and 5.

I must also thank the Biotechnology and Biological Sciences Research Council and the Interdisciplinary Bioscience Training Programme for the scholarships that funded this work and enabled my studies at Oxford. I am also grateful to Christ Church College for offering me a home along with many good friends and nourishing meals for this past half a decade.

After all is over, we tend to look back with rose-tinted glasses. But I think it is important to not downplay the impact that a PhD, and academic research in general, can have on

one's mental health. For me, things were not always easy. I am very grateful for the help I received from the Oxford University Counselling Services, especially from counsellor Laura Evers. I am very lucky because I had access to these services, had truly kind supervisors, good friends, and a supporting boyfriend. But not everyone is as fortunate; despite all mental health awareness campaigns, there is still much work to be done to make academia a healthier environment.

A good friend is like a four-leaf clover: hard to find and lucky to have. I truly am fortunate because I have found a bunch of them. Some of my dearest memories of my time in Oxford are from evenings spent with the "Cool People" –Daniel DeBrincat, Güneş Ünal, Neven Blaskovic, and Karrar Ali–, often in the Linacre College common room playing pool, poker, Uno, table tennis, or the legendary Mafia. The Linacre ball and the trip to Turkey with them were some of the highlights of my time in Oxford. When some of us moved to other corners of the world and the pandemic turned our lives upside down, we moved the common room to weekend Zoom calls or board games in Gathertown. I am grateful for all those good memories, and I am sure we will keep on creating many more in the years to come. Over the last couple of summers, I had a great time exploring the English countryside on cycle rides with Angad Singh, Deepanshu Singh, and the "Botley Squad" (Andreas Schmid and Shu Ishida) –thank you all for your friendship and the beautiful trips! I am also grateful to Beatriz Godinho, Augustinas Malinauskas, David Brito, Fran van Heusden, Tiago Inácio and Leonor Madeira for their friendship and many good moments over the last few years.

I am also grateful to Nuno Markl, Ricardo Araújo Pereira, and the rest of the team of Manhãs da Comercial for providing me with entertainment and company during yet another milestone in my scientific career. O Homem Que Mordeu o Cão and Mixórdia de Temáticas cheered me up while slicing brains and mounting endless histology slices, as well as in many sleepless nights.

I would also like to thank my family, especially my mother, Leonilde, and my brothers, Marco and André, for their unconditional love and support throughout the PhD despite not always understanding what it is that I do.

Special thanks also to Santiago Vasquez, Antigona Vasquez and Freddy Hernao (and their four-legged family Merlin, Alba, Bruno, Lucio, Frida, and Matilda) for all the love and support and for making me feel so at home over the holidays in Germany.

Last but certainly not least, I would like to thank Sebastian. Throughout this PhD, you filled the shoes of partner, colleague, friend, and mentor. Words cannot express how grateful I am for your incredible support in each of those roles. Thank you for taking a second look at my histology results or helping with the modelling, for all the discussions about data analysis or unexpected results. Thank you for reading and giving me feedback on the thesis; it made writing it so much more bearable. Above all, thank you for all your love and support, and for encouraging me to always follow my dreams.

Declaration

I hereby declare that the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this or any other university. A fraction of the general introduction chapter (chapter 1) of the present thesis, particularly the sections on OFC functional heterogeneity, have recently been published as a review manuscript in a scientific journal, for which I am the main contributing author (Barreiros, Ishii, et al., 2021). Similarly, a considerable fraction of the results chapter on the neuroanatomical tracing study (chapter 3) has previously been published as a research manuscript (Barreiros, Panayi, et al., 2021). I confirm that I wrote and produced the material here reproduced and have obtained permission from my co-authors to include it in my thesis.

All the results presented are the product of my own work with the following exception: collaborators at Eli Lilly (UK) acquired the raw data, which I then analysed and present as part of chapters 4 and 5.

Oxford, UK, October 2021

Inês Barreiros

Inês Vitória Barreiros

Table of contents

Nomenclature ..... *xiii*

Abstract ..... *xvii*

1. General Introduction *1*

1.1 Reward-guided flexible behaviour **1**
1.1.1 Associative learning 1
1.1.2 Model-free and model-based reinforcement learning 3
1.1.3 Experimental frameworks for investigating behavioural flexibility 5
Reinforcer devaluation 5
Reversal learning 6

1.2 Role of dopamine in reward-guided flexible behaviour **7**
1.2.1 Anatomy of dopaminergic pathways 8
1.2.2 Dopamine and reward processing 9
1.2.3 Dopamine and reward prediction errors 11
1.2.4 Dopamine functional heterogeneity 14
1.2.5 Dopaminergic signalling dysfunction & disease 17
1.2.6 Haemodynamic correlates of reward prediction errors 18

1.3 NAc contribution to reward-guided behaviour **20**

1.4 Role of OFC in reward-guided flexible behaviour **22**
1.4.1 Anatomical inputs and outputs of the rat OFC 23
1.4.2 OFC representations supporting reward-guided behaviour 26
Organising knowledge for flexible behaviour 29
1.4.3 OFC functional heterogeneity 31
Classical parcellation of the rat OFC 33
Medial vs lateral distinctions 35
Beyond a dichotomous mediolateral distinction? 39
Anterior vs posterior distinctions 41
Summary 43

1.5 Thesis aims **45**

2. Materials and methods *47*

2.1 Anatomical characterisation of OFC inputs along its axes of organisation **47**
2.1.1 Subjects 47
2.1.2 Surgery 48
2.1.3 Histology 49
2.1.4 Image acquisition and analysis 50

2.2 Investigating haemodynamic correlates of prediction errors in the NAc and OFC **53**
Oxygen amperometry 53
2.2.1 Subjects 54
2.2.2 Behavioural experiments 55
Apparatus 55
Probabilistic instrumental reversal-learning task 55
Behavioural training 56
2.2.3 Electrode construction 58
2.2.4 Electrode calibration 59
2.2.5 Surgery 59
2.2.6 O_2 amperometry data recordings 62
2.2.7 *In vivo* lesion validation (Dopamine Lesion Experiment) 62
2.2.8 Pharmacological manipulations (Hyperdopamine Experiment) 63
2.2.9 Histology 63
Electrode placements 64

6-OHDA lesion (Dopamine Lesion Experiment) ... 66
2.2.10 Data analysis ... 67
Dopamine Lesion Experiment (data analysis) ... 67
Hyperdopamine Experiment (data analysis) ... 84

2.3 Development of a novel behavioural paradigm for studying non-spatial cue-based serial learning. ... **89**
2.3.1 Subjects ... 90
2.3.2 Behavioural experiments ... 90
Apparatus ... 90
Generalisation trials ... 92
Control trials ... 92
Auditory cues and reward contingencies ... 93
Behavioural training ... 94
2.3.3 Data analysis ... 99
Performance on generalisation trials ... 99
Performance on control trials ... 101

3. Anatomical characterisation of OFC inputs along its axes of organisation ... ***103***

3.1 Introduction ... **103**

3.2 Methods summary ... **105**
3.2.1 Surgery ... 105
3.2.2 Histology ... 105
3.2.3 Image acquisition and analysis ... 105

3.3 Results ... **107**
3.3.1 Prefrontal cortex afferents ... 122
Intra-OFC projections ... 122
Medial prefrontal projections ... 124
Motor cortex ... 124
Sensory cortex ... 124
3.3.2 Thalamic afferents ... 125
Mediodorsal nucleus ... 125
Submedius nucleus ... 128
Paratenial nucleus ... 131
Nucleus reuniens ... 131
3.3.3 Temporal lobe afferents ... 132
Amygdala ... 133

3.4 Discussion ... **136**
3.4.1 Distinct and topographically organised thalamic inputs to OFC subdivisions ... 137
3.4.2 Multi-sensory nature of OFC inputs ... 139
3.4.3 Distinct amygdala input patterns into OFC subdivisions ... 139
3.4.4 Anterior and posterior LO are part of distinct anatomical and functional circuits ... 141

4. Causal contribution of dopamine to haemodynamic signatures of reward prediction ... ***145***

4.1 Introduction ... **145**

4.2 Methods ... **148**
4.2.1 Probabilistic reversal-learning task ... 149
4.2.2 Surgery ... 149
4.2.3 O_2 amperometry data recording ... 150
4.2.4 *In vivo* lesion validation ... 150
4.2.5 *Histology* ... 150
4.2.6 *Data Analysis* ... 151
Behaviour ... 151
Behaviour modelling ... 151
Amperometry ... 152

4.3 Results **152**
4.3.1 6-OHDA injection in NAc reduces TH labelling in NAc and OFC 152
4.3.2 Unilateral NAc lesion does not affect behavioural performance in probabilistic reversal-learning task 155
4.3.3 *In vivo* validation of NAc 6-OHDA lesion 157
4.3.4 NAc and OFC track outcome but not outcome estimates 159
4.3.5 Lesioning of dopaminergic inputs to NAc attenuates responses in the NAc and potentiates those in the OFC 162
4.3.6 Contrasting effects of expectation in the NAc before and after learning new reward contingencies 165
4.3.7 Lesioning of dopaminergic inputs to NAc has no effect on the encoding of specific trial features after reversal of reward contingencies 167
4.3.8 Lesioning of dopaminergic inputs to NAc potentiates OFC haemodynamic responses 169
4.3.9 Distinct T_{O2} response dynamics in anterior and posterior portions of the OFC 170

4.4 Discussion **173**
4.4.1 Lesioning effect of 6-OHDA injection into NAc extends to OFC dopaminergic inputs 173
4.4.2 Haemodynamic signatures of RPEs in NAc and OFC 175
4.4.3 Effect of lesioning of dopaminergic inputs on NAc and OFC haemodynamic signals 177
4.4.4 Haemodynamic responses across anterior–posterior OFC 179

5. Effect of hyperdopaminergic states in haemodynamic signatures of reward prediction ***183***

5.1 Introduction **183**

5.2 Methods Summary **187**
5.2.1 Probabilistic reversal-learning task 187
5.2.2 Surgery 188
5.2.3 O_2 amperometry data recording 188
5.2.4 Pharmacological manipulations 188
5.2.5 Histology 188
5.2.6 Data analysis 189
Behaviour 189
Behavioural modelling 189
Amperometry 190

5.3 Results **190**
5.3.1 Effect of single and repeated systemic amphetamine injections on behaviour 190
5.3.2 OFC haemodynamic signals are mainly driven by outcome, whereas NAc tracks both outcome and reward estimate 195
5.3.3 Distinct TO2 response dynamics in medial and lateral OFC 199
5.3.4 Amphetamine disrupts NAc haemodynamic signals 200

5.4 Discussion **202**
5.4.1 Effect of single and repeated doses of amphetamine on reward-guided behaviour 202
5.4.2 Haemodynamic signatures of RPEs in NAc and OFC 206
4.4.3 Effect of single and repeated doses of amphetamine on NAc haemodynamic signals 208

6. Development of a novel behavioural paradigm for non-spatial cue-based serial learning ***211***

6.1 Introduction **211**

6.2 Behavioural paradigm design **213**

6.3 Methods summary **215**
6.3.1 Behavioural task 215
Task structure 217
Auditory cues 219
6.3.2 Data analysis 219

6.4 Behavioural paradigm development and learnings **220**

6.4.1 Training rats to hold at initiation port during long periods 220
6.4.2 Performance on the 3-cue serial learning task 221
6.4.3 Performance on the 2-cue serial learning task 225
6.4.4 Performance on the 2-cue serial learning task after normalisation of auditory cues ... 229
6.4.5 Performance on control trials 232
6.5 Discussion **234**
7. General Discussion and Conclusions ***239***
7.1 Summary of findings **239**
7.2 OFC functional and anatomical heterogeneity **241**
7.3 Link between dopaminergic signalling and haemodynamic signatures of RPEs **245**
7.4 Dopaminergic signalling dysregulation and maladaptive learning **250**
7.5 Neuroscience needs histology **252**
7.6 Bridging rodent and human research with haemodynamic signals recordings **253**
7.7 Developing behavioural paradigms for investigating OFC function **255**
7.8 In conclusion **258**
Bibliography ***261***

Nomenclature

6-OHDA	6-hydroxidopamine
a24b	anterior cingulate cortex, area 24b
AF	alexa fluor
Ag	silver
AgCl	silver chloride
AI, d, v	agranular insular cortex, dorsal, ventral parts
ALO	lateral orbitofrontal cortex, anterior part
Amy	amygdala
ANOVA	analysis of variance
AP	anteroposterior
AUC	area under the curve
AVO	ventral orbitofrontal cortex, anterior part
BIC	Bayesian inference criteria
BLA	basolateral amygdala, anterior part
BLP	basolateral amygdala, posterior part
cm	centimetre
CPE	carbon paste electrode
CS	conditioned stimulus
CTB	cholera toxin subunit B
d	day
dB	decibels
DL	dorsolateral
DLO	dorsolateral orbitofrontal cortex
DLS	dorsolateral striatum
DV	dorsoventral
ML	mediolateral
EMA	exponential moving average
FR	fixed ratio
g	gram
GI	gustatory insular cortex

GLM	general linear regression
h	hour
Hz	hertz
IL	infralimbic cortex
IR	infrared
ITI	inter-time interval
kg	kilogram
L	litre
LaDL	lateral amygdala, dorsolateral part
LaVL	lateral amygdala, ventrolateral part
LaVM	lateral amygdala, ventromedial part
LDTg	lateral dorsal tegmentum
LED	light-emitting diode
LH	lateral habenula
LO	lateral orbitofrontal cortex
M1	primary motor cortex
M2	secondary motor cortex
mA	milliamperes
MD, m, c, l	mediodorsal nucleus of the thalamus, medial, central, lateral parts
mg	milligram
min	minute
ml	millilitre
mm	millimetre
MO	medial orbitofrontal cortex
mV	millivolts
n	number
N_2	nitrogen
NA	numerical aperture
nA	nanoampere
NAc, C, S	nucleus accumbens, core, shell
NaOH	sodium hydroxide
NDS	normal donkey serum
nl	nanolitre

O_2	oxygen
OFC	orbitofrontal cortex
PBS	phosphate buffered saline
PBS-T	phosphate buffer saline with triton
PD	dorsal peduncular cortex
PFA	paraformaldehyde
PFC	prefrontal cortex
Pir	piriform cortex
PL	prelimbic cortex
PLO	lateral orbitofrontal cortex, posterior part
PT	paratenial nucleus of the thalamus
PVO	ventral orbitofrontal cortex, posterior part
R	reward
Re	nucleus reuniens of the thalamus
RL	reinforcement learning
RMS	root mean square
RMTg	rostromedial tegmentum
ROI	region of interest
RT	room temperature
s	second
S1, BF, FL, J, Tr	primary somatosensory cortex, barrel field, forelimb region, jaw region, trunk region
S2	secondary somatosensory cortex
sc	subcutaneously
sd	standard deviation
SEM	standard error of the mean
Sub, D, V	submedius nucleus of the thalamus, dorsal, ventral parts
TH	tyrosine hydroxylase
TH^+	TH-expressing
T_{O2}	amperometric tissue oxygen
UK	United Kingdom
v	volume
V2	secondary visual cortex

VI	variable interval
VLO	ventrolateral orbitofrontal cortex
VO	ventral orbitofrontal cortex
VP	ventral pallidum
w	weight
µl	microlitre
µm	micrometre
µM	micromolar

Neuronal circuits underlying flexible behaviour

Ines V. Barreiros
Christ Church College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2021

Abstract

To thrive in an ever-changing world, animals must rapidly and flexibly adapt their behaviour according to their needs. Some problems require an initial period of learning through trial and error; others can be quickly solved by generalising and integrating information from previous experiences. In this thesis, I explore anatomical and functional features of the brain system that support these learning and decision-making processes. In particular, I focus on the role of the orbitofrontal cortex (OFC) and midbrain dopaminergic projections which play a critical role in sustaining flexible behaviour. Behavioural research suggests the OFC is functionally heterogeneous along its medial–lateral and anterior–posterior axes. In Chapter 3, I investigated whether this functional heterogeneity is underpinned by differences in anatomical connectivity by simultaneously characterising the input patterns into the anterior lateral, posterior lateral, and posterior ventral portions of the rat OFC. The results show that these OFC subdivisions receive distinct and topographically organised inputs from various thalamic nuclei, basolateral amygdala, and various cortical areas. Learning and behavioural adaptation are particularly important when expectations are not met, eliciting reward prediction errors (RPEs) –the difference between expected and actual outcomes. In the brain, the encoding of these teaching signals has long been linked with the activity of dopamine neurons. However, there has been no direct demonstration between the behaviourally derived haemodynamic signatures of RPEs measured with functional magnetic resonance imaging (fMRI) and dopamine release. Thus, to establish the contribution of dopamine to haemodynamic signatures of RPEs encoding, Chapter 4 combined recordings of haemodynamic signals in the OFC and nucleus accumbens (NAc) with lesioning of NAc dopaminergic inputs. OFC responses were surprisingly larger after lesioning of dopaminergic inputs to NAc, suggesting a potential balance between the activity of the two regions. Dopaminergic signalling dysregulation is thought to drive the neuronal mechanisms underlying psychotic symptoms such as delusions through disruption of RPEs encoding in frontostriatal circuits. Chapter 5 shows that maladaptive behaviour during amphetamine-induced hyperdopaminergic states is linked to disruption of NAc haemodynamic responses. Then, in Chapter 6, I describe the development of a novel rodent behavioural paradigm in which rats must develop a model of the world to maximise reward, for investigating how the brain encodes information during generalisation. Chapter 7 summarises the findings from the previous chapters and discusses their implications for our understanding of how the brain encodes information for supporting flexible behaviour.

This work was supported by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship.

1. General Introduction

Suppose the coffee vending machine in your building has broken down and a new one has temporarily been installed. Even if this machine is different from any you have used in the past, you would probably quickly figure out how to get your coffee. From your previous experience with coffee vending machines, you expect there would be buttons for choosing the type of drink, somewhere where you could place a cup to collect it, and perhaps a card reader or coin slot for payment. And, even if the buttons were unlabelled, you could rapidly learn through trial and error which one was associated with the drink you wanted. This thesis concerns the brain mechanisms that support this ability to flexibly modify our behaviour in response to changes in the environment, both through generalisation and trial-and-error learning. In this general introduction, I put into context the work presented in the remaining chapters by giving an overview of what is known about brain circuits that play a critical role in sustaining flexible behaviour, focusing on the orbitofrontal cortex (OFC) and the mesocorticolimbic dopaminergic system.

1.1 Reward-guided flexible behaviour

1.1.1 Associative learning

Reward-guided behaviour is a process that involves action selection or response to stimuli with the purpose of maximising reward. To this end, an individual must be able to perceive contingent relationships and predict what might happen after a specific stimulus occurs or an action is taken. In a first instance, this ability requires learning associations between events or actions and their outcomes. One of these forms of associative learning was first described by Ivan Pavlov, who observed that by sounding a

bell before each time a dog was fed, the dog eventually started salivating when hearing it (Pavlov, 1927). In this learning process, known as *Pavlovian* or *classical conditioning*, a biologically significant stimulus (e.g., food), the unconditioned stimulus, is paired with a previously neutral one (e.g., bell), the conditioned stimulus, which occurs independently of the individual's behaviour. After learning, the conditioned stimulus anticipates the unconditioned stimulus and elicits a response similar to that produced by the latter. Pavlov termed this strengthening of the association between conditioned and unconditioned stimuli with repeated paired presentation as 'reinforcement' (Pavlov, 1927). Thus, unconditioned and conditioned stimuli are sometimes also referred to as primary and secondary reinforcers.

Not all rewards are obtained independently of one's behaviour; some require the individual to perform specific voluntary actions. This type of reward-guided behaviour is known as *operant* or *instrumental conditioning* and was first studied by Edward Thorndike. He placed a hungry cat in a puzzle box containing mechanisms that could be used for exiting the box and accessing a bowl of food placed outside. Thorndike observed that initially the cat was able to exit the box through trial and error but, after each successful escape, the cat took less time to activate the correct mechanism by becoming more selective in its responses (Thorndike, 1898). These observations form the basis for response reinforcement: actions followed by a favourable outcome become more likely to be repeated, whereas those punished will occur less frequently. A favourable outcome might be a reward such as food (positive reinforcement) or the termination of an unpleasant stimulus such as a loud noise (negative reinforcement). These theories were expanded by B. F. Skinner, who developed the operant conditioning box –also known as Skinner box– based on Thorndike's cat box, and is now regarded as the father of instrumental conditioning (Skinner, 1953).

1.1.2 Model-free and model-based reinforcement learning

In order to react appropriately to an event or follow the most advantageous course of action, it is necessary to maintain value estimates for each stimulus or action. These estimates might be obtained by exploring the environment through trial and error and computing the average reward amount each stimulus or action has produced in the past. In an ever-changing world, value estimates must be continuously monitored and updated by comparing expectations with what actually happens. Hence, one way to determine when to update these expectations is to detect when these are not met, which elicit *reward prediction errors* (RPEs) – the difference between expected and actual outcomes (Bush & Mosteller, 1951; Rescorla & Wagner, 1972; Sutton & Barto, 1998). Positive and negative RPEs signal stimuli or actions that are more or less rewarded than expected, respectively. RPEs promote updating of *cached values* –the weighted average value estimates of rewards which are transferred from the reward to the stimulus or actions predicting it–, thus making it more or less likely that specific actions are repeated. Since this learning method does not rely on a knowledge of the structure of the environment but instead just updates immediately past events based on RPEs, it is known as *model-free*. It requires extensive exploration of the environment to assess the most advantageous course of action and cached values are acquired based on direct experience and change slowly over time (Daw et al., 2005; Huys et al., 2014; Sutton & Barto, 1998).

An alternative class of learning, known as *model-based*, is one in which the individual makes predictions about future rewards based on what it knows about the rules of the environment and its own motivational status (Daw et al., 2005; Sutton & Barto, 1998). In a series of influential experiments, Edward Tolman and colleagues observed that rats trained to find rewards on a maze were able to take shortcuts or find

new routes when familiar ones were blocked (Tolman et al., 1946; Tolman & Honzik, 1930). Tolman concluded that, through exploration of the mazes, rats had developed a *cognitive map* of their environment (Tolman, 1948). These map-like mental representations form an internal model of the world that accounts for the relationships between events and outcomes and might be used for orientation within an environment (Behrens et al., 2018; Kitchin, 1994). Such world models might be developed by extracting relevant features from sensory experiences and deducing causal relations between them. Then, by applying this abstract structure to new situations, it is possible to predict the outcome of hypothetical actions (Gershman & Niv, 2010; Kemp & Tenenbaum, 2009). This was first demonstrated experimentally by Harry Harlow, who trained non-human primates to choose the rewarded stimulus from a series of stimuli pairs (Figure 1.1-A). Harlow noticed that, initially, the monkeys learned the value of each object slowly and through trial and error. However, as the monkeys experienced more sets of objects, their learning curves became steeper (Figure 1.1-B); they were able to extract the abstract structure of the task (i.e., one of the two objects was always rewarded) and generalise it to new sets of objects. Harlow termed the abstract knowledge underlying this ability to quickly solve similar tasks a *learning set* (Harlow, 1948).

Model-free and model-based behaviour are often seen as the two ends of a spectrum of learning methods. In practice, however, the two are thought to coexist in the brain, interact with each other, and compete for promoting behavioural flexibility (Akam & Walton, 2021; Daw et al., 2005; Miranda et al., 2020; Pezzulo et al., 2013).

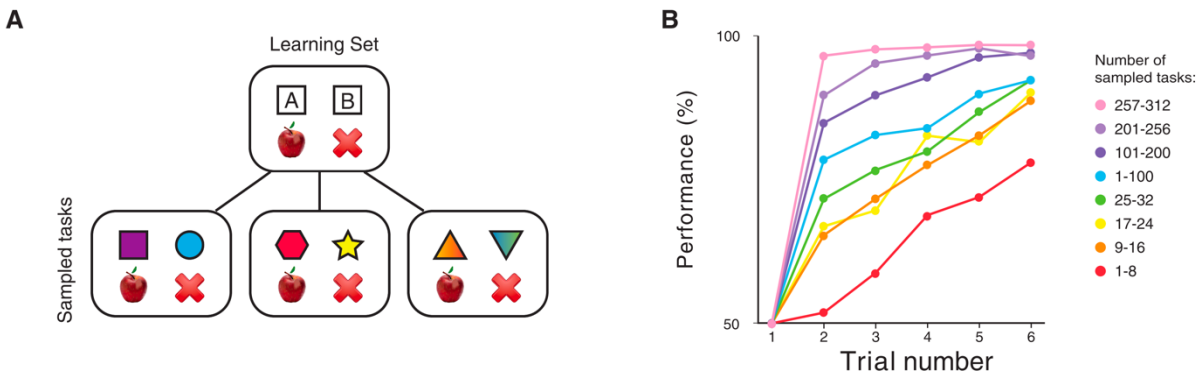


Figure 1.1 Formation of learning sets. (**A**) Schematic of Harlow's task. Monkeys were trained on a series of problems with a common underlying structure: only one of the pair of objects presented was associated with a reward. (**B**) Discrimination learning curves on successive blocks of problems (Harlow, 1949). Reprinted from Behrens et al. (2018), Copyright (2018) with permission from Elsevier.

1.1.3 Experimental frameworks for investigating behavioural flexibility

Our understanding of the brain circuits underlying behavioural flexibility has been greatly aided by combining certain behavioural paradigms with neuronal activity manipulations or recordings (Hamilton & Brigman, 2015; Krakauer et al., 2017). Below, I describe two tasks commonly used for investigating reward-guided behavioural flexibility in neuroscience research.

Reinforcer devaluation

In reinforcer devaluation paradigms, following an initial period of discriminative learning, certain stimuli or actions are devaluated by having their outcomes (e.g., food) consumed to satiety or paired with illness (e.g., via lithium chloride injection). This results in decreased responding to the devalued stimulus or action while leaving responses for non-devalued outcomes unchanged (Colwill & Rescorla, 1985, 1990). In a first instance, this behavioural adaptation requires individuals to register the change in the valence of the primary reinforcer during the devaluation procedure. Then, they must

recall which stimulus or action was paired with the devalued reinforcer and to update its value estimate. Later, when presented again with that stimulus or action, individuals must be able to retrieve its updated value estimate. When pairing an outcome with lithium chloride injection, testing is performed after the animal experiences illness which causes long-term updating of the outcome associated value in future tests. Thus, behavioural adaptation to reinforcer devaluation involves: (1) forming associations between specific stimuli or actions and respective primary reinforcers; (2) registering the change in the appetitive value of the primary reinforcer during devaluation; (3) integrating the updated value of the primary reinforcer with the stimulus or action predicting it; (4) and retrieving the updated value of the conditioned reinforcer during subsequent decisions. In sensory-specific satiety procedures, testing typically occurs shortly after consumption, though often in a different context, and reveals a reduction in the actions that gain access to the devalued outcome. Interestingly, if testing is carried out the next day, this effect is not observed. This means that the animals' decisions are guided by their *current* motivational state and not by that at the time they last encountered the outcome. Since animals do not experience the changed value of outcome in the context of making a choice in either of the devaluation procedures, this behavioural adaption fits better with model-based than model-free mechanisms.

Reversal learning

Reversal learning is another paradigm commonly used in neuroscience research for assessing behavioural flexibility (see Izquierdo et al., 2017 and Izquierdo & Jentsch, 2012 for reviews). In standard versions, subjects learn that one stimulus or action among at least two is reinforced and, after a performance criterion is reached, the reward contingencies are reversed. The contingencies can be deterministic or probabilistic. For

example, in an initial phase, approaching stimuli A or B might yield a reward 75 or 25% of the time, respectively; after reversal, the probabilities of obtaining a reward after approaching stimuli A or B would be 25 and 75%. Reversal learning problems might be solved using either model-free or model-based strategies (Izquierdo et al., 2017; Masumi & Sato, 2021). After a reversal, an individual might gradually and separately learn the new value of each stimulus by integrating experiences across trials. Alternatively, and after experiencing several reversals, the individual might understand the complementary relationship between the reward contingencies of the options and know that these periodically switch. Thus, when there is enough evidence that the contingencies have changed, an individual will be able to infer the value of a stimulus after experiencing the outcome associated with the other. In sum, successful behavioural adaptation to a reversal of reward contingencies requires: (1) forming initial associations between relevant stimuli or actions and respective reward contingencies; (2) updating the value estimates of the stimuli or actions after reversal of the contingencies, either gradually as new trials are experienced (model-free) or as soon as there is enough evidence the contingencies have been reversed (model-based); (3) retrieving the updated reward contingencies during subsequent decisions and, based on these, withhold responding to previously preferred stimulus or action and increase responding to the alternative one.

1.2 Role of dopamine in reward-guided flexible behaviour

Dopamine is one of the most widely studied neurotransmitters. It has been shown to play an essential role in several brain functions including motor control, motivation, reward signalling, and reinforcement (Wise, 2004). In this thesis, I focus on its contribution to reward-guided flexible behaviour.

1.2.1 Anatomy of dopaminergic pathways

The cell bodies of dopamine-producing neurons in the brain are mainly concentrated in the ventral tegmental area (VTA), the retrorubral field, and the substantia nigra pars compacta (Dahlstrom & Fuxe, 1965; Moore & Bloom, 1978). Those located in the VTA project mainly to the nucleus accumbens (NAc) and to the prefrontal cortex (including medial frontal and, to a lesser extent, orbitofrontal cortex), among other regions (see Figure 1.2) (Oades & Halliday, 1987; Russo & Nestler, 2013). These two dopaminergic pathways are often referred to as the *mesolimbic* and the *mesocortical dopaminergic pathways*; together, they form the *mesocorticolimbic dopaminergic system* (Wise, 2004).

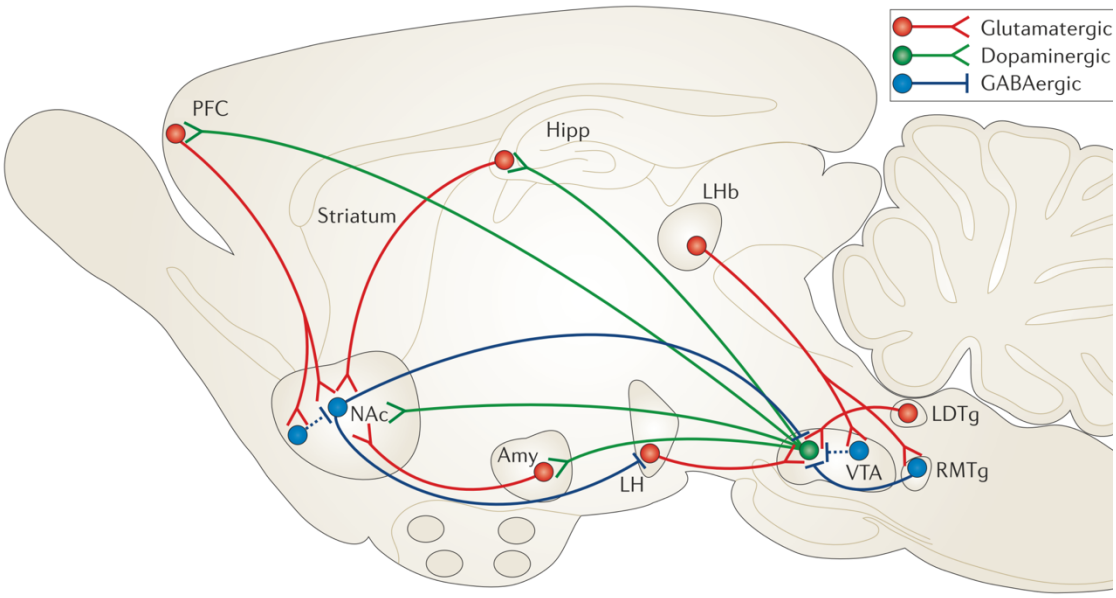


Figure 1.2 Anatomical circuit of the mesocorticolimbic dopaminergic system. Simplified schematic of the major dopaminergic, glutamatergic and GABAergic projections of the mesocorticolimbic dopaminergic system in the rodent brain. Adapted from Russo & Nestler (2013), Copyright (2013) with permission from Springer Nature. Amy: amygdala; Hipp: hippocampus; LH: lateral habenula; LDTg: lateral dorsal tegmentum; NAc: nucleus accumbens; PFC: prefrontal cortex; RMTg: rostromedial tegmentum; VTA: ventral tegmental area.

VTA neurons projecting to the NAc and prefrontal cortex pass through the medial forebrain fibre bundle en route to the striatum and, from here, mesocortical projections

extend to the prefrontal cortex (Kalsbeek et al., 1988; Oades & Halliday, 1987). Despite being anatomically close, dopaminergic projections from the mesolimbic and mesocortical pathways arise from largely separate populations of VTA neurons (Deniau et al., 1980; Sobel & Corbett, 1984; Swanson, 1982) with distinct electrophysiological and molecular properties (Lammel et al., 2008, 2011; see Lammel et al., 2014 for a review). VTA dopaminergic projections synapse onto GABAergic medium spiny neurons in the NAc (Surmeier et al., 2007), and pyramidal neurons and GABAergic interneurons in the prefrontal cortex (Tritsch & Sabatini, 2012). In target areas, dopamine release modulates information processing by modifying membrane excitability and synaptic transmission (Nicola et al., 2000; Tritsch & Sabatini, 2012). These regions project back to VTA via glutamatergic (prefrontal cortex) or GABAergic projections (NAc) (Nauta et al., 1978; Swanson, 1982; Swanson & Cowan, 1975), thus influencing the dopaminergic input they receive (Einhorn et al., 1988; Grace, 1991; Lodge, 2011; Rahman & McBride, 2000; Takahashi et al., 2011).

1.2.2 Dopamine and reward processing

In the 1950s, James Olds and Peter Milner implanted electrodes in several brain areas of rats and used them to deliver small electric currents when the animals pressed a lever. They noticed that when electrodes were implanted in some areas, but not others, rats would keep pressing the lever for electrical stimulation (Olds, 1958; Olds & Milner, 1954). Later studies showed that areas whose electric stimulation appeared to reinforce self-stimulation contained dopamine neurons (Corbett & Wise, 1980). Moreover, studies also showed that intracranial self-stimulation was attenuated following lesions of the dopaminergic system or systemic administration of dopamine antagonists (Fibiger et al., 1987; Fouriezos & Wise, 1976; Wise et al., 1978). Together, these formed the foundation

to suggest a link between dopamine transmission and reward processing. In parallel, this link was further supported by studies showing that natural rewards or addictive psychostimulants such as cocaine or amphetamine increase extracellular dopamine levels (Carlson, 1970; van Rossum et al., 1962; Wise & Rompre, 1989).

As the link between dopamine and reward was dissected further, it became apparent that while this neurotransmitter is essential for aspects of reward-guided behaviour, it does not itself mediate the hedonic experience of reward (Berridge & Robinson, 1998; Spanagel & Weiss, 1999). For example, food consumption elicited dopamine release consistently in the NAc of food-restricted but not of *ad libitum* fed rats (Bassareo & Di Chiara, 1997; Wilson et al., 1995), suggesting dopamine might play a role in encoding motivation for pursuing rewards. Additional evidence linking dopamine to incentive motivation or cost/benefit decisions came mainly from pharmacological and lesion studies showing increased dopamine transmission made animals more likely to choose the option associated with the more valued reward even if it required more effort whereas its disruption had the opposite effect (see Salamone et al., 2018 for a review). For example, dopaminergic antagonists made animals more likely to choose smaller or less palatable rewards over large or more palatable ones when they required waiting or additional effort to obtain rewards, such as pressing a lever several times or overcome a physical barrier (Cousins et al., 1996; Denk et al., 2005; Salamone et al., 1994, 2007).

Arguably the most influential theory on the role of dopamine in modulating reward-guided behaviour focuses on its contribution to learning about reward-predicting stimuli and in signalling reward prediction errors (as reviewed in Berke, 2018; Schultz, 2019). In this thesis, I explore the neuronal mechanisms of this theory of dopamine function, starting with a summary of the main ideas supporting it below.

1.2.3 Dopamine and reward prediction errors

The encoding of RPEs in the brain has long been linked with the activity of neurons in the VTA and substantia nigra pars compacta. In seminal work, Read Montague and Peter Dayan, working with data from Wolfram Schultz's group, showed that the activity of midbrain dopamine neurons correlates with an RPE term in reinforcement learning models (Montague et al., 1996; Schultz et al., 1997). They observed that before monkeys learned that a conditioned stimulus was followed by a reward, dopamine neurons would increase their firing upon reward delivery (positive RPE). However, after several days of training, the increased firing rate was temporally shifted to the presentation of the reward-predicting stimulus and, if a predicted reward was omitted, the firing rate would decrease at the time the reward would have been delivered (negative RPE; Schultz et al., 1997; see Figure 1.3). Therefore, the firing rate of midbrain dopamine neurons may be used as a teaching signal modulating plasticity mechanisms in regions receiving dopaminergic inputs (Schultz, 1998; Schultz et al., 1997).

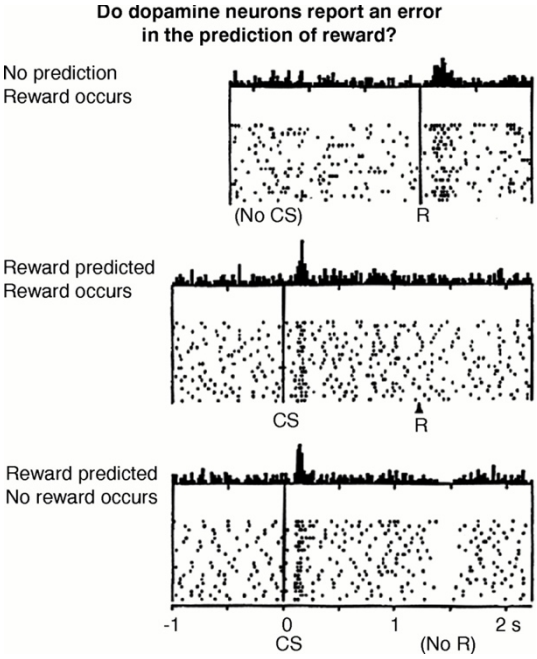


Figure 1.3 Reward prediction error signalling in midbrain dopamine neurons. Electrophysiology recordings from midbrain dopamine neurons in response to an unpredicted reward (top), a predicted reward that is anticipated by a conditioned stimulus (middle), and omission of a predicted reward (bottom). Reprinted from Schultz et al. (1997), Copyright (1997) with permission from AAAS. CS: conditioned stimulus; R: reward.

As discussed earlier, RPEs are a key component of reinforcement learning theories (Rescorla & Wagner, 1972; Sutton & Barto, 1998). The phasic activity of midbrain dopamine neurons encodes RPEs at the time of reward and also at the time of events predicting it (e.g., conditioning stimulus). Thus, these RPEs fit well with model-free reinforcement learning models, particularly those from the class of temporal difference (TD) learning models. These models take into account the timing of the events, allowing for value updating not only at time of reward delivery or omission but also to previous moments in time when reward-predicting events occur (Daw & Tobler, 2013; Sutton, 1988; Sutton & Barto, 1998).

To be considered correlates of RPEs, neuronal signatures must satisfy three key criteria: (1) a positive influence of reward estimate on reward-predicting cue-elicited signals, with cues that predict larger rewards (or with higher probability) eliciting greater responses; (2) a negative influence of reward estimate at the time of reward delivery, with unexpected rewards producing greater responses; and (3) a positive influence of outcome at the time of delivery, with the delivery of a valued reward producing a greater response than its omission (Caplin & Dean, 2008; Rutledge et al., 2010). Since the initial observations by Schultz and colleagues, a link between dopamine activity and the encoding RPEs has been consistently observed using a range of different approaches, including measurement of dopamine release in NAc terminals or monitoring changes in blood oxygenation dependent signals (BOLD) in the ventral striatum, and in different species (e.g., humans: Behrens et al., 2008; D'Ardenne et al., 2008; Rutledge et

al., 2010; monkeys: Bayer & Glimcher, 2005; Enomoto et al., 2011; rodents: Blanco-pozo et al., 2021; Cohen et al., 2012; Day et al., 2007; Eshel et al., 2015; Flagel et al., 2011; Hart et al., 2014; Roesch et al., 2007). It is important to note, however, that while numerous functional resonance magnetic imaging (fMRI) studies using reward-guided behavioural paradigms have reported encoding of correlates of RPEs estimates (e.g. Berns et al., 2001; Knutson et al., 2001; Lohrenz et al., 2007; O'Doherty et al., 2004; O'Doherty et al., 2003), only a few have used the stringent criteria described above (e.g., Behrens et al., 2008; Niv et al., 2012; Rutledge et al., 2010). This means that, as pointed out by Behrens et al. (2008), some of the neuronal signatures reported as RPEs in some fMRI studies could be representations of reward signals as the two are often correlated. The link between the encoding of RPEs in these haemodynamic responses and the activity of dopamine neurons will be discussed in more detail later in the section on *Haemodynamic correlates of reward prediction errors* (section 1.2.6).

In addition to encoding the difference between expected and received outcomes, the activity of midbrain dopamine neurons and striatal dopamine release appears to also be influenced by the reward magnitude, probability, and timing (Bayer et al., 2007; Bayer & Glimcher, 2005; Fiorillo et al., 2003, 2008; Hollerman et al., 1998; Takahashi et al., 2017; Tobler et al., 2005). Dopamine neurons are also sensitive to motivational state, reflecting subjective reward values (Lak et al., 2014; Papageorgiou et al., 2016). However, the activity of dopamine neurons also encodes value-neutral information, signalling mismatches between expected and observed sensory properties of specific events (Stalnaker et al., 2019).

Moreover, there is increasing evidence that the involvement of dopamine in reward-guided learning is not restricted to model-free learning. For example, Sadacca and colleagues (2016) trained rats on a sensory preconditioning task in which one

neutral cue was paired with a second one that predicted whether a reward would be delivered. They observed that the activity of VTA dopamine neurons encoded RPEs when either the paired or the unpaired cue was presented. This means that the activity of these neurons reflected an ability to infer the relationship between the unpaired cue and reward even though the two had not been previously experienced, thus, showing that VTA neurons encode predictions of both cached (previously experienced) and inferred value predictions (not experienced directly) (Sadacca et al., 2016). Additionally, a series of recent studies have shown that midbrain dopamine activity accounts for RPEs encoding of belief states –the probability of being in a specific state of the environment in which certain rules or conditions are in place (Babayan et al., 2018; Blanco-Pozo et al., 2021; Sharpe et al., 2017; Starkweather et al., 2017); see Akam & Walton, 2021 and Lerner et al., 2021 for recent reviews). Together, these studies suggest dopamine contributes to reward-guided behaviour through both model-free and model-based adaptation mechanisms.

1.2.4 Dopamine functional heterogeneity

It is important to recognise that not all dopamine neurons signal reward or reward-predicting cues. Instead, dopamine neurons are rather functionally heterogeneous, with some subpopulations being responsive to aspects of movement (Da Silva et al., 2018; Dodson et al., 2016; Howe & Dombeck, 2016; Hughes et al., 2020; Jin & Costa, 2010) or aversion (Brischoux et al., 2009; Lammel et al., 2012, 2014; Matsumoto & Hikosaka, 2009; Zweifel et al., 2011).

Moreover, the theories on dopamine function I have described thus far relate mainly to the activity of dopamine neurons in the mesolimbic pathway. As mentioned in the section on *Anatomy of dopaminergic pathways* (section 1.2.1), mesolimbic and

mesocortical projections arise from largely separate populations of VTA neurons with distinct properties (see Lammel et al., 2014 for a review). Whereas stimulation of mesolimbic dopaminergic projections has a reinforcement effect (e.g., Kim et al., 2012; Phillips et al., 2003; see Wise, 2004 for a review), stimulation of the mesocortical pathway does not (Ellwood et al., 2017). Instead, cortical dopamine has been shown to influence a range of cognitive functions, including working memory and attentional control (see Cools & D'Esposito, 2010; Robbins & Roberts, 2007 for reviews). For example, manipulation of dopamine levels in the prefrontal cortex either pharmacologically or through selective lesioning (e.g., 6-OHDA neurotoxin) has repeatedly been shown to impair performance in working memory tasks such as in the spatial delayed alternation task in which an animal has to choose the alternative option to that selected on the previous trial to receive a reward (Brozoski et al., 1979; Sawaguchi et al., 1990; Sawaguchi & Goldman-Rakic, 1991; Zahrt et al., 1997). Either increased or reduced levels of cortical dopamine also affect performance on attentional set-shifting tasks in which it is necessary to selectively shift attention between cues on different dimensions (e.g., odours or visual stimuli) (Ellwood et al., 2017; Floresco et al., 2006; Tunbridge et al., 2004).

The relationship between cortical dopamine levels and performance on working memory and attentional tasks is nonlinear, and it somewhat resembles an inverted-U-shaped relationship (Figure 1.4). The effects of manipulating the levels of dopamine in the prefrontal cortex will depend on the individual and task-dependent baseline levels, with both too high or too low levels impairing cognitive function (see Cools & D'Esposito, 2011 for a review). For example, high or low levels of cortical dopamine were observed in rats performing well on a working memory task with a 30 min delay or in rats performing poorly on a version of the same task with longer delays, respectively (Phillips

et al., 2004). Moreover, pharmacologically increasing cortical dopamine activity impaired performance on the task version with short delays but improved it on the version with long delays (Floresco & Phillips, 2001).

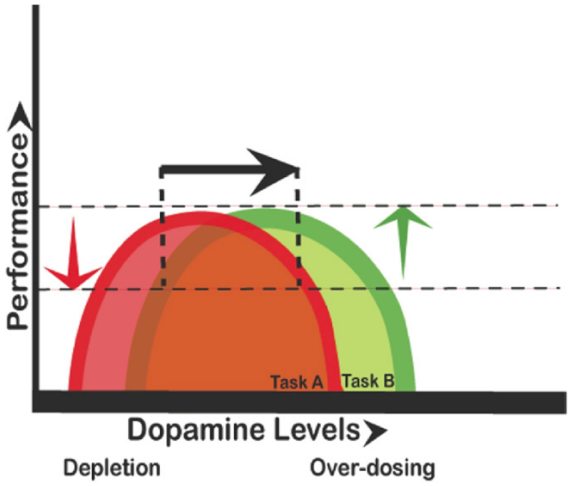


Figure 1.4 Relationship between behavioural performance and cortical dopamine levels follows an inverted-U-shaped relationship. Both insufficient and excessive levels of cortical dopamine have detrimental effects on cognitive performance. Moreover, the effects of dopamine level manipulations will depend on individual and task-dependent baseline levels. For example, an increase in cortical dopamine levels (black arrow) might impair performance in some tasks (red arrow) but improve it in others (green arrow). Reprinted from Cools & D'Esposito (2011), Copyright (2011) with permission from Elsevier.

The somewhat paradoxical relationship between cortical dopamine levels and cognitive performance might be linked to the dopamine levels in the cortex versus those in the striatum. Optimal cognitive flexibility requires a balance between stability or perseverance and flexibility or switching behaviour, and this balance might depend on the levels of dopamine in both the cortex and the striatum (as reviewed in Cools & D'Esposito, 2011). Moreover, a growing body of evidence suggests that dopamine levels in the two regions might be functionally coupled, albeit in an antagonistic manner. Decreasing cortical dopamine levels, either pharmacologically or through lesioning of cortical dopaminergic terminals, has been shown to result in elevated striatal dopamine

(Clarke et al., 2014; Pycock, Carter, et al., 1980; Pycock, Kerwin, et al., 1980; Roberts et al., 1994; Scornaiencki et al., 2009; Thompson & Moss, 1995). Conversely, an increase in prefrontal cortex dopamine levels has been linked to decreased striatal dopamine (Doherty & Gratton, 1996; Kayser et al., 2012; Kolachana et al., 1995). In neuropsychiatric conditions, such as in schizophrenia, increased mesolimbic dopaminergic signalling coexists with reduced cortical dopamine (Davis et al., 1991; Weinberger, 1987). Thus, an adequate balance between the two appears to be important for normal brain function.

1.2.5 Dopaminergic signalling dysfunction & disease

Dopaminergic signalling dysfunction has been implicated in several neurological and neuropsychiatric conditions, including Parkinson's disease, schizophrenia, attention deficit hyperactivity disorder (ADHD), and anhedonia (Deserno et al., 2016; Gilmour et al., 2015; Heinz & Schlagenhauf, 2010; Marsden, 2006; Zald & Treadway, 2017). The contribution of dopamine to psychosis is supported by the clinical efficacy of dopamine antagonists in ameliorating schizophrenia symptoms (Creese et al., 1976; Seeman et al., 1976). Considering the links between the activity of dopamine neurons and encoding of RPEs in reinforcement learning models and between dopamine and psychosis, it has been suggested that disrupted reinforcement learning may contribute to the formation of psychotic delusions (Deserno et al., 2016; Kapur, 2003; Smith et al., 2007). Moreover, amphetamine-like psychostimulants increase the extracellular concentration of dopamine (Amara & Sonders, 1998; Sulzer et al., 2005), and methamphetamine administration has been shown to induce psychotic symptoms and to disrupt the encoding of RPEs in the NAc of healthy volunteers (Bernacer et al., 2013). While there is evidence linking dopamine dysregulation and consequent disruption of RPEs encoding to psychotic symptoms, the neuronal mechanisms underlying these processes in

neuropsychiatric disorders remain unclear. In Chapter 5, I investigate the neuronal mechanisms underlying maladaptive learning during amphetamine-induced hyperdopaminergic states similar to those thought to elicit psychotic symptoms in neuropsychiatric conditions.

1.2.6 Haemodynamic correlates of reward prediction errors

While most of the data on RPEs encoding in the brain and their link to dopamine activity comes from studies in animal models, it is important to look at the mechanisms underlying these representations in humans, especially considering their potential link to neuropsychiatric conditions as discussed above. Studies of RPEs encoding in the human brain have often used fMRI to measure BOLD signals during reward-guided behavioural tasks having reported correlates of RPEs estimates (e.g. Berns et al., 2001; Knutson et al., 2001; Lohrenz et al., 2007; O'Doherty et al., 2004; O'Doherty et al., 2003). While BOLD signals represent a net increase in neuronal activity (Hall et al., 2016), several studies support a link between the correlates of RPE estimates observed in these signals and dopamine. For example, Schott and colleagues observed a positive correlation between fMRI signals and measures of dopamine release obtained separately with positron emission tomography (PET) during reward anticipation (Schott et al., 2008). Further evidence linking dopamine release to BOLD changes in the NAc comes from pharmacological studies (Knutson & Gibbs, 2007). For instance, BOLD signals in the NAc increase following systemic administration of dopamine reuptake inhibitors or drugs that potentiate dopamine transmission (e.g., cocaine or amphetamine) (Chen et al., 1997; Kufahl et al., 2005; Marota et al., 2000; Völlm et al., 2004). Moreover, this effect is abolished if NAc dopamine neurons are selectively lesioned (Chen et al., 1997) or

following administration of antagonists of dopamine receptors which stimulation increases cellular excitability (dopamine receptors type 1, D1) (Marota et al., 2000).

Despite growing evidence pointing in this direction, there has been no direct demonstration of the link between behaviourally derived haemodynamic signatures of RPEs and dopamine. Lohrenz and colleagues (2016) compared datasets from two previous human studies in which participants were trained on a sequential gambling task, deciding how much to invest in a fictional market based on its value history and then seeing how much they had lost or earned. One of the studies observed correlates of RPEs in the ventral striatum using fMRI (Lohrenz et al., 2007). The other used fast-scan cyclic voltammetry to measure the levels of dopamine release in the same region of Parkinson's disease patients trained on the same task (Kishida et al., 2016). Comparison between the data from the two studies points towards a role of dopamine release in modulating BOLD signals encoding RPEs; however, the authors did not observe a 1-to-1 relationship between the two (Lohrenz et al., 2016).

Unlike in animal models, it is often challenging to directly measure dopamine release or specifically disrupt its transmission in human subjects. Moreover, while fMRI studies have the advantage of being non-invasive, they require subjects to remain motionless, largely limiting fMRI imaging in rodents to non-behavioural studies. For example, recent rodent fMRI studies have investigated how the activity of VTA dopaminergic neurons affects haemodynamic signals in the brain by combining this technique with optogenetic stimulation of tyrosine hydroxylase-expressing VTA neurons, observing signal changes across several regions, not all of them associated with dopamine neurotransmission (Decot et al., 2017; Lohani et al., 2017). *In vivo* oxygen amperometry, a technique that allows for measuring haemodynamic signals comparable to the BOLD signals of fMRI in behaving animals (Francois et al., 2012), might help bridge

the gap between fine-scale invasive studies of neural activity in rodents and human fMRI research (Gilmour et al., 2015). Using this technique, a recent study observed haemodynamic signatures of signed RPEs in the NAc and a surprise or salience signal in the OFC of rats trained on a probabilistic Pavlovian task (Werlen et al., 2019). In Chapter 4, I combine *in vivo* oxygen amperometry recordings with lesioning of NAc dopaminergic inputs to investigate the link between behaviourally derived haemodynamic signatures of RPEs and dopamine.

1.3 NAc contribution to reward-guided behaviour

The NAc is a crucial component of the brain circuits underpinning reward-guided motivation and behavioural flexibility. It is localised in the ventral striatum and has been anatomically divided into two main portions: the core and the shell, with the shell bordering the core medially and ventrally (Zahm & Brog, 1992). The NAc has classically been seen as an interface between limbic and motor systems, receiving and integrating motivational information from multiple brain regions, including the amygdala, hippocampus, VTA, and prefrontal cortex areas, and influencing the direction and intensity of behaviour via output projections to subcortical motor structures (Figure 1.5; Alexander, 1986; Mogenson et al., 1980). Its output projections originate from GABAergic medium spiny neurons, which make up the majority (~90%) of the neurons in the NAc (Kemp & Powell, 1971; Meredith, 1999); the remainder are aspiny cholinergic or GABAergic local interneurons (Kawaguchi et al., 1995; Marín et al., 2000).

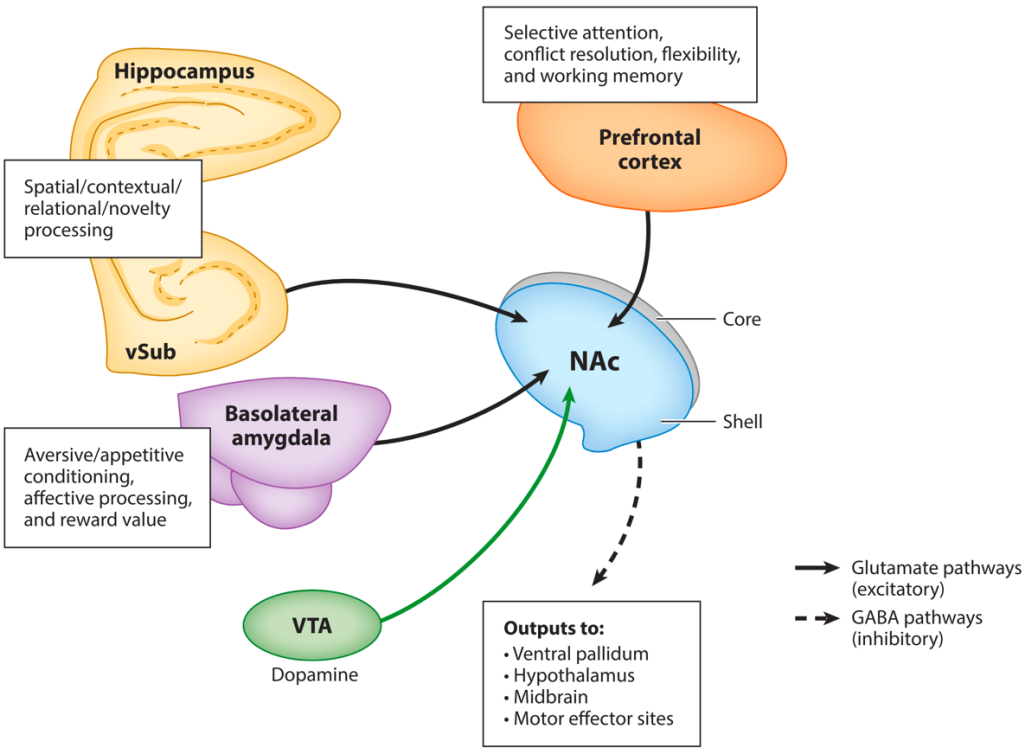


Figure 1.5 Main anatomical inputs and outputs to NAc. The NAc receives glutamatergic inputs from the basolateral amygdala, hippocampus (predominantly ventral subiculum), and prefrontal cortex, and dopaminergic inputs from VTA. It then influences behaviour through GABAergic outputs to motor structures. Reprinted from Floresco (2015), Copyright (2015) with permission from Springer Nature. NAc: nucleus accumbens; vSub: ventral subiculum; VTA: ventral tegmental area.

A significant proportion of our understanding of NAc function derives from studies that targeted local dopaminergic transmission (as discussed in previous sections). Consequently, there is often substantial overlap between theories of the function of NAc with that of meso-accumbens dopamine, particularly in relation to the encoding of rewards and reward-guided learning (Berridge & Robinson, 1998). This is particularly prominent in neuroimaging studies, where activation of NAc is often used as a proxy for a role of dopamine and not of NAc (see Knutson & Gibbs (2007) for a review on the link between dopamine and BOLD signals in the NAc). Despite this tendency, it is important to consider that dopamine is only one of the many inputs being integrated in the NAc

(Figure 1.5). Accordingly, manipulation of dopaminergic inputs to this region has been shown to produce different behavioural effects than direct manipulation of NAc neurons (e.g., Yun et al., 2004). It has been argued that a more accurate description of the NAc function is that it mediates the influence that rewarding behaviour has on action selection, particularly when there is high ambiguity about the most appropriate course of action to achieve a goal (Floresco, 2015). This approach reconciles the two views of this region being involved in reward processing and locomotor behaviour, and accounts for the diversity of anatomical inputs that converge onto NAc neurons.

Moreover, the NAc might not only be involved in biasing behaviour towards rewarded stimuli, but also have an equally important role in inhibiting behaviours that might interfere with goals currently being pursued. Lesioning or inactivating the NAc core reduces approach responses to reward-predictive cues, whereas similar manipulations to the NAc shell increase responding to non-rewarding ones (Blaiss & Janak, 2009; Di Ciano et al., 2008; Parkinson et al., 2000; Saunders & Robinson, 2012). Thus, these two contrasting NAc functions of approach or inhibition might be anatomically separated, with the core of the NAc mediating approach behaviours, while the NAc shell is more involved in behavioural inhibition (see Floresco, 2015 for a review). In sum, this region plays a critical role in reward-guided behavioural flexibility by promoting approach to stimuli or actions with a desirable outcome and suppressing irrelevant behaviour, particularly in changing or uncertain conditions.

1.4 Role of OFC in reward-guided flexible behaviour

The beginning of research on the function of the OFC traces back to 1848 and the case of Phineas Gage, a railroad worker who had his skull pierced by an iron rod during an

accidental explosion. The accident resulted in extensive damage of his orbital and midline prefrontal regions (Damasio et al., 1994). Even though Gage's cognitive capacities were mainly unaffected, he was reported to lack self-control and social inhibition after the lesion. Moreover, according to the medic who treated him, he became unable to "[…] estimate size or money accurately, though he [had] memory as perfect as ever. He would not take $1000 for a few pebbles […]" (Harlow, 1848). These were the first observations hinting at a key role of the OFC in decision-making. Since then, the OFC has been the focus of intense research looking at its contribution to the regulation of flexible behaviour (Murray et al., 2007; Stalnaker et al., 2015; Wikenheiser & Schoenbaum, 2016). In this section, I focus on the role of the OFC in supporting decision-making processes required for reward-guided behavioural flexibility.

1.4.1 Anatomical inputs and outputs of the rat OFC

While the structural organisation of the primate OFC has been extensively investigated (Carmichael & Price, 1994; Ongur et al., 2003; Price, 2007), the precise boundaries of rodent OFC are still under debate. The rodent OFC is situated in the dorsal bank of the rhinal sulcus and, in contrast to both the human and non-human primate OFC, which can be cytoarchitectonically defined using granular, dysgranular and agranular areas (Price, 2006), all its regions are agranular (Van De Werd & Uylings, 2008; Wise, 2008). This led Broadman to conclude that the prefrontal cortex is unique to primates and that other species were devoid of homologous regions (Brodmann, 1909). However, Rose & Woolsey (1948) later proposed that homologous regions of the primate prefrontal cortex can be identified in other mammals by the reciprocity of projections from the mediodorsal (MD) thalamus. Even though this definition has survived (Carlén, 2017; Groenewegen, 1988; Uylings et al., 2003), additional criteria for designating prefrontal

areas and identifying homologous regions in other species have since been proposed. These include ontogenetic and phylogenetic relationships, cytoarchitecture, the pattern and relative density of projections, and electrophysiological, neurochemical and behavioural properties (Uylings et al., 2003). Based on these criteria, the rodent OFC has been subdivided into different portions which will be discussed in detail under the section on *OFC functional heterogeneity* (section 1.4.3).

The OFC establishes robust anatomical connections with the amygdala, striatum, thalamus, and sensory cortices (Figure 1.6), amongst other regions. Sensory inputs reach the OFC directly from primary and secondary sensory cortices, including piriform, gustatory insula, somatosensory, visual, and perirhinal cortex (Agster & Burwell, 2009; Datiche & Cattarelli, 1996; Kondo et al., 2005; Price, 2006; Reep et al., 1996). The OFC receives sensory and other subcortical information also indirectly via the thalamus, including from its mediodorsal, reuniens, submedius, and paratenial nuclei (Alcaraz et al., 2015; Coffield et al., 1992; Corwin et al., 1994; Krettek & Price, 1977a; Kuramoto et al., 2017; Reep et al., 1996; Vertes & Hoover, 2008; Yoshida et al., 1992). Thus, the OFC receives strong multi-modal sensory information.

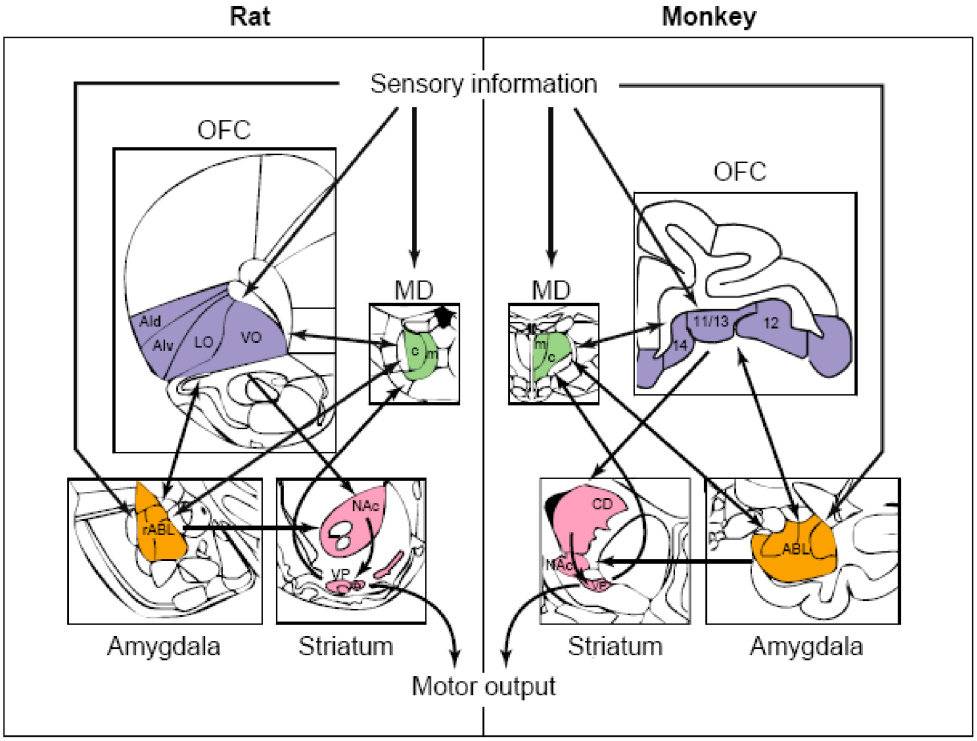


Figure 1.6 Anatomical connectivity of the OFC in the rat and monkey. Each panel represents a snippet of a coronal section, and arrows represent the direction of connectivity between the orbitofrontal cortex and the amygdala (orange), the striatum (pink) and the mediodorsal thalamus (MD, green). In both species, the OFC receives sensory input directly from sensory cortices and indirectly via the thalamus. The OFC also receives associative information from the amygdala and sends outputs to the motor system through the striatum. AId: agranular insular cortex, dorsal part; AIv: agranular insular cortex, ventral part; c: central; CD: caudate; MD: mediodorsal thalamus; m: medial; LO: lateral OFC; NAc: nucleus accumbens; rABL, rostral basolateral amygdala; VO: ventral OFC; VP: ventral pallidum. Reprinted from Schoenbaum et al. (2006). Copyright (2006) with permission from Elsevier.

With the amygdala, the OFC establishes reciprocal and topographically organised connections (Gabbott et al., 2005; Kita & Kitai, 1990; Mcdonald, 1991, 1998; Reep et al., 1996). The amygdala plays an important role in associative learning through the integration of sensory and limbic information (Everitt et al., 1999; Johnson et al., 2009; Ono et al., 1995) and its lesioning results in behavioural impairments in outcome devaluation and reversal learning tasks (see Balleine & Killcross, 2006; Sharpe & Schoenbaum, 2016 for reviews).

In addition to the amygdala, the OFC sends outputs to the striatum, including the NAc and caudate-putamen (Berendse et al., 1992; Heilbronner et al., 2016; Schilman et al., 2008), and thalamus (Groenewegen, 1988; Krettek & Price, 1977b). The striatum itself sends GABAergic projections to the thalamus, which in turn projects back to OFC and adjacent cortical regions. Together, these pathways form a cortical-striatal-thalamic loop central for the functional coupling of cortical and basal ganglia circuits in behavioural control (Alexander, 1986; Bonelli & Cummings, 2007).

The OFC receives dopaminergic inputs from the VTA (Descarries et al., 1987; Murphy & Deutch, 2018) and noradrenergic inputs from the locus coeruleus (Cerpa et al., 2019; Gerfen & Clavier, 1979). These inputs provide the OFC with information necessary for updating stimuli- and action-outcome associations (see Cerpa et al., 2021 for a review).

In sum, OFC connectivity provides a means to integrate sensory, motivational, and associative information to predict the value of stimuli or actions, impact behavioural output, and update expectations based on observed outcomes.

1.4.2 OFC representations supporting reward-guided behaviour

Inspired by the behavioural changes observed in Phineas Gage after his brain lesion (Harlow, 1848), the first and most influential theory on OFC function focused on emotional regulation and inhibitory control, particularly during changing circumstances. This was supported by subsequent observations from patients and animal models with OFC lesions using reversal learning paradigms (e.g., Chudasama & Robbins, 2003; Jones & Mishkin, 1972; McAlonan & Brown, 2003; McEnaney & Butter, 1969; Mishkin, 1964; Roberts & Wallis, 2000; Rolls et al., 1994; Schoenbaum et al., 2003). However, while in many of these studies lesioning of the OFC slowed response switching after a change in

reward contingencies, it did not impair response inhibition or discrimination during initial learning. Thus, it became apparent that the OFC does not play a general role in response inhibition.

The development of research techniques allowing for neuronal activity measurements furthered the refinement of theories of OFC function. Using single-cell neuronal recordings, a number of studies have shown that OFC neurons fire in anticipation to reward during the presentation of reward-predictive cues, similarly to what was discussed earlier for midbrain dopaminergic neurons (see section 1.2.3 on *Dopamine and reward prediction errors*), in addition to firing during the delivery of the reward itself (e.g., Gottfried et al., 2003; O'Doherty et al., 2002; Schoenbaum et al., 1998; Schoenbaum, Setlow, Saddoris, et al., 2003; Tremblay & Schultz, 1999). The idea that the OFC encodes the expected outcome value is further supported by the impairments observed in reinforcer devaluation paradigms after OFC lesion (Critchley & Rolls, 1996; Gallagher et al., 1999; Izquierdo & Murray, 2004; Machado & Bachevalier, 2007). Interestingly, an impairment in this task is observed even if lesioning of the OFC occurs after outcome devaluation (Murray et al., 2015; Pickens et al., 2003, 2005). Single-unit recordings have also revealed that OFC neurons adapt their responses to stimuli after reversal of contingencies or change in motivational status, thus, reflecting the changes in their incentive value (Critchley & Rolls, 1996; Gottfried et al., 2003; Scott et al., 2005; Thorpe et al., 1983). Thus, the OFC encodes expected outcome values during anticipation and seems necessary for accessing the updated expected value of cues and actions or outcomes, such as in reversal learning or reinforcer devaluation paradigms.

For selecting an action or deciding how to respond to a stimulus, an organism must represent the subjective value of each action or stimulus and then be able to compare them. Representations supporting these processes have also been observed in the OFC.

For example, Padoa-Schioppa and Assad (2006) recorded the activity of OFC neurons while presenting monkeys with a series of choices of two different juices in varying quantities. The behavioural choice pattern helped assess which juice was the monkey's preferred one and its relative value (e.g., a monkey might value one unit of juice A as much as four times that amount in juice B). Subsequent analysis of the neuronal responses revealed that separate populations of OFC neurons encoded the subjective value of each of the options available, the identity of the chosen juice, and the subjective value of the selected option. Interestingly, the value representations were independent of specific sensory properties or the spatial location of the rewards (Padoa-Schioppa & Assad, 2006). Thus, the OFC might promote optimal decision-making by encoding the subjective value of available options on a common scale (see Padoa-Schioppa & Conen, 2017 for a review).

In addition to the subjective value of expected reward and information on the events predicting it, the OFC has been shown to encode information about its timing, probability, magnitude, and identity (Kennerley & Wallis, 2009; Lipton et al., 1999; Ramus & Eichenbaum, 2000; Roesch & Olson, 2005; Stalnaker et al., 2014; van Duuren et al., 2007, 2009). Moreover, several studies have reported that, at outcome delivery, neuronal signals in the OFC are also scaled by reward expectation (e.g., O'Doherty et al., 2003; Sul et al., 2010; Tobler et al., 2006; Werlen et al., 2019; but see Stalnaker et al., 2018). However, whereas, for example, responses in the NAc are positively influenced by delivery of reward and negatively influenced by expectation, those in the OFC are scaled by how surprising the outcome is irrespective of valence (i.e., increases in activity to both surprising rewards or reward omissions), potentially reflecting a salience signal (e.g., Ogawa et al., 2013; Werlen et al., 2019). While OFC representations do not fulfil the criteria for RPE-encoding, the OFC sends information back to midbrain structures and

might, therefore, contribute to RPEs computations (Lodge, 2011; Swanson, 1982; Takahashi et al., 2011). Thus, OFC representations might be used for learning the associations between specific stimuli or actions and their outcomes. Moreover, the OFC might facilitate *credit assignment*, i.e., attribution of an outcome to a specific stimulus or choice (Noonan et al., 2010; Walton et al., 2010, 2011). For example, it has been shown that OFC lesions result in misattribution of rewards to temporally close choices that are not necessarily responsible for that outcome, such as to an alternative choice made on a previous or subsequent trial instead of to the contingent choice (Walton et al., 2010). This reinforces the key role of the OFC in flexibly adapting reward-guided behaviour, particularly in changing environments with multiple options or with delayed outcomes.

Organising knowledge for flexible behaviour

One prominent recent theory of OFC function proposes that this region mediates behavioural flexibility by contributing to the representation of cognitive maps. How information is represented and organised into cognitive maps is best understood in the spatial domain (Manns & Eichenbaum, 2009; O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1979). However, a growing number of studies provide support for a similar encoding of maps for abstract mental knowledge, such as relationships between stimuli, actions and outcomes, and other non-spatial conceptual information, in the absence of a continuous organising dimension such as space (Constantinescu et al., 2016; Garvert et al., 2017; Tolman & Brunswik, 1935; Wikenheiser & Schoenbaum, 2016).

This view that the OFC contributes to the representations of cognitive maps attempts to unify the multiplicity of cognitive functions in which this region seems to be involved. Specifically, it has been argued that the OFC represents task state (Behrens et al., 2018; Niv, 2019; Schuck et al., 2016; Wikenheiser & Schoenbaum, 2016; Wilson et al.,

2014; Zhou, Gardner, et al., 2019; Zhou, Montesinos-Cartagena, et al., 2019), i.e., abstract "latent" representations that help us know which set of rules (implicit or inferred information) are currently in place that go beyond what is available simply from taking into account only the available, observable information. For example, in reversal learning paradigms, a subject might learn that, at times, cue A is rewarded while B is not, but there are periods in which the opposite is true. By categorising the world into these two states, a subject might be able to infer the reward contigency of each cue by asessing what is the current state, and to infer the value of one cue after experiencing the outcome of the other (Izquierdo et al., 2017). Thus, after a reversal, OFC-lesioned subjects would need to "unlearn" the previous contingencies and update the reward expectations for each stimulus separately and without relying on prior knowledge of the task structure. Wilson and colleagues (2014) modelled the behaviour of control and OFC-lesioned animals on several tasks in which lesions of the OFC impair behavioural performance, such as reversal learning and reinforcer devaluation paradigms. They observed that reinforcement-learning models that included task state representations explained well the behaviour of control animals, whereas those models that did not include those parameters behaved similarly to OFC-lesioned animals (Wilson et al., 2014).

Encoding of task state might involve learning and representing the abstract relationships between stimuli or events. This also fits well with findings from OFC single-unit recordings. As mentioned in the previous section, subpopulations of OFC neurons reverse their responses to stimuli after a reversal of their reward contingencies (e.g., Thorpe et al., 1983). Moreover, encoding might be made more efficient by representing relevant stimuli or events abstracted from its sensory features (as reviewed in Behrens et al., 2018). This would allow for rapid and flexible adaptation to new contexts with similar structures, such as in the task used by Harlow (1948) to investigate the formation

of learning sets (described earlier in the section on *Model-free and model-based reinforcement learning,* section 1.1.2). Interestingly, it has been shown that if monkeys are offered to choose between different amounts of juices A and B (see the study by Padoa-Schioppa & Assad, 2006 described in the previous section), and then these juices are replaced by two different ones, C and D, OFC neurons encoding the value of offered and chosen juices maintain their functional role across the different sets of juices (Xie & Padoa-Schioppa, 2016).

Even though a growing number of studies support the involvement of the OFC in encoding aspects of cognitive maps, it remains to be explored how different subpopulations of OFC neurons contribute to the representation of the various components of task structure. With this in mind, in Chapter 6, I attempt to establish a novel rodent behavioural paradigm for investigating the contribution of the OFC to the encoding of cognitive maps, specifically during generalisation.

1.4.3 OFC functional heterogeneity

As interest and the number of publications on the function of the OFC increase (Izquierdo, 2017; Murray et al., 2007), inconsistencies are often found, especially when comparing data from different species. Several reviews comparing cross-species findings have emphasised that not all areas of the OFC have been equally studied (Izquierdo, 2017; Wallis, 2011) and the most targeted areas in each species are not necessarily homologous (Heilbronner et al., 2016; Rudebeck & Murray, 2011; Wallis, 2011). For example, rodent studies frequently focus on lateral OFC, which often extends to include adjacent anterior insular cortex (AI), both of which are agranular. By contrast, studies in primates typically focus on area 13, an anterior and agranular sector of central OFC. Aligning the findings

from different studies is made more complicated as studies often do not explicitly specify which subregion within the OFC was the focus of the investigation (Izquierdo, 2017; Wallis, 2011).

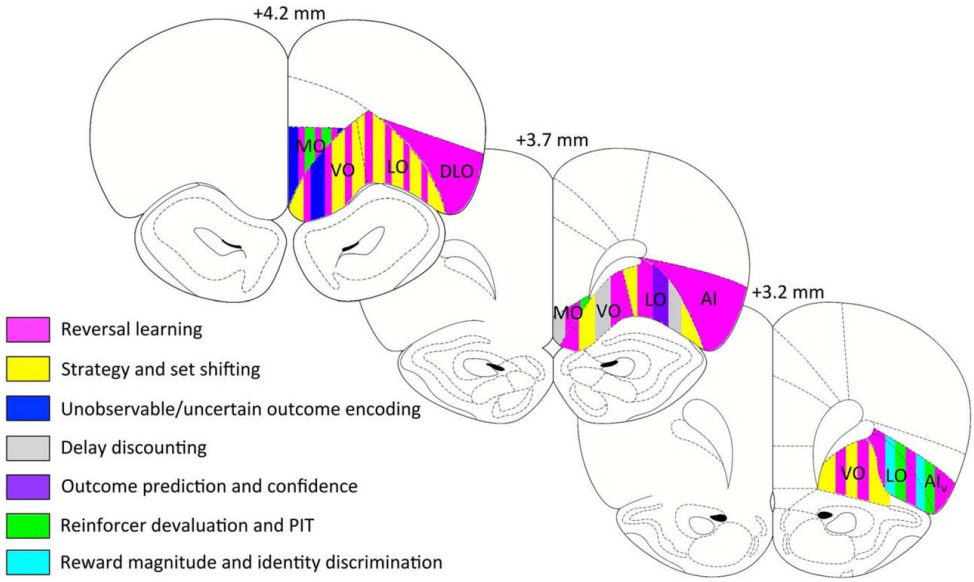


Figure 1.7 Coronal sections depicting functional heterogeneity within rat OFC. Behavioural task effects are mapped onto coronal sections of the rat OFC. Numbers on top of each section indicate the AP distance (in millimetres) from bregma (Paxinos and Watson, 2007). PIT: Pavlovian-instrumental transfer. Reproduced from Izquierdo (2017).

Increasingly researchers are considering how the functional diversity within the OFC aligns with its anatomical subdivisions. This has mostly focused on established anatomical subdivisions along the medial-lateral axis (Figure 1.7; Bradfield & Hart, 2020; Izquierdo, 2017; Kringelbach & Rolls, 2004; Rudebeck & Murray, 2014; Walton et al., 2015). More recent data suggests that there may also be important functional distinctions along the anterior-posterior axis (Bradfield et al., 2018; Panayi & Killcross, 2018). Indeed, sensitivity to these anatomical distinctions will likely be critical to progress in understanding the functional role of the OFC. These anatomical considerations raise the question of where, within the OFC, should functional boundaries be drawn? And are there

distinct functional divisions, or should we consider these differences functional gradients instead?

One lens through which to view such questions is by focusing on anatomical studies (Devlin & Poldrack, 2007). For example, as described by Dick Passingham, "The brain is not porridge"; instead, we ought to think of the brain as 'spaghetti' (Passingham, 2017). In other words, connectivity matters. The function of a brain region is, to a great extent, determined by the pattern of neuroanatomical connections it makes with other areas; the more dissimilar the connectivity pattern of two brain regions, the easier it should be to differentiate their functions (Passingham et al., 2002). Therefore, next, I review the evidence for functional heterogeneity first along the medial-lateral and then the anterior-posterior axes of OFC organisation, describing how functional differences relate to the patterns of anatomical inputs and outputs. I primarily focus on what has been found in the rodent OFC and selectively relate this to evidence of similar functional and organisational principles in the human and non-human primate OFC.

Classical parcellation of the rat OFC

While borders of the various portions of the OFC are indistinct cytoarchitectonically, several subdivisions of the rodent OFC are currently recognised (Figure 1.8). Originally, Krettek & Price (1977b) subdivided the rat orbital region into medial (MO), ventral (VO), ventrolateral (VLO) and lateral (LO) OFC, and the agranular insular cortex, lateral to LO, into dorsal and ventral sectors (AIv, AId). In their original description, the delimitation of the agranular insula extended rostrally, lateral to the orbital areas. However, the AId portion rostral to the claustrum was later reclassified as dorsolateral orbital area (DLO) due to its distinct thalamocortical projections and cytoarchitecture (Price, 2006; Ray & Price, 1992). Based on structural and connectional similarities, it has been proposed that

MO, VLO, LO and DLO may be comparable to areas 14, 13a, 13m/l and 12o respectively in monkeys, while AIv and AId may correspond to monkey insula areas Iam/Iapm and Iai (Price, 2006).

In practice, which regions are included under the umbrella of OFC proper has not been agreed upon in rodents or primates. For example, in rodents, AI is often targeted alongside adjacent LO (e.g. Ogawa et al., 2013) and has been considered an orbital structure, whereas MO is not always considered part of rodent OFC (e.g. Schoenbaum et al., 2006). Similarly, in primates, there is debate over the status of parts of area 12 (Murray & Rudebeck, 2018; Price, 2007; Sallet et al., 2020). Here I will not attempt to reach any conclusions as to which regions should be considered part of the OFC, and instead will inclusively consider the functional and anatomical distinctions of all potential OFC subregions along the orbital surface in rodents, which for simplicity I will refer to as MO, VO, LO, and AI.

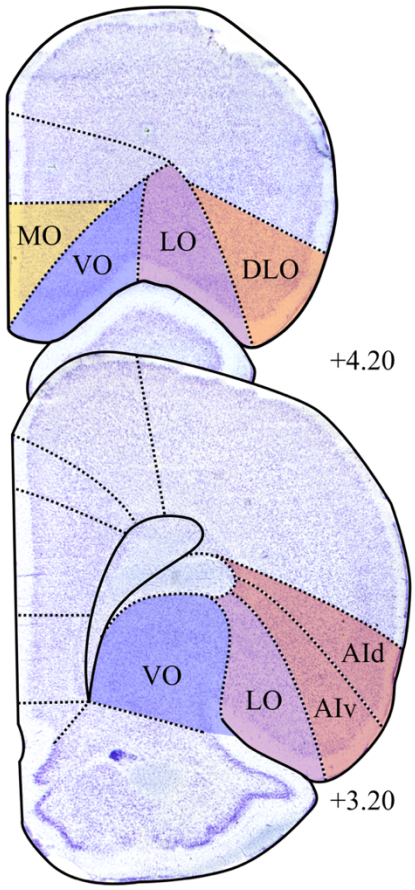


Figure 1.8 Classical parcellation of rat OFC. Currently recognised OFC subdivisions include MO, VO, LO, DLO and AI. Represented is the right hemisphere of Nissl-stained coronal rat brain slices at two different anterior-posterior levels of the OFC. Numbers on the right-hand side of the slices indicate the distance in millimetres from bregma, consistent with Paxinos & Watson (1998). An intermediate region, VLO (not shown), has also been identified, located halfway between VO and LO. In this thesis, VLO will not be considered. Photomicrographs adapted from Panayi & Killcross (2018). AId: agranular insular cortex, dorsal part; AIv: agranular insular cortex, ventral part; DLO: dorsolateral OFC; LO: lateral OFC; MO: medial OFC; VO: ventral OFC; VLO: ventrolateral OFC.

Medial vs lateral distinctions

A series of influential neuroanatomical studies in primates gave rise to the idea that the anatomical organisation of medial and orbital parts of the prefrontal cortex can be aggregated into two functional networks with partially overlapping but distinct connectivity patterns (revised in Ongur & Price, 2000; Price, 2007). Specifically, there is a 'medial visceromotor' network comprising predominantly of parts of the medial OFC (area 14) and the medial frontal cortex (as well as some caudolateral parts of OFC), and

an 'orbital sensory' network largely comprising the central and lateral OFC in primates (centred on areas 11, 13 and 12). While these divisions are primarily based on studies of non-human primates (NHPs), based on putative cytoarchitectonic and connectional similarities between rodent OFC and caudal parts of primate OFC, a similar division has also been proposed to be present in rodents (Ongur & Price, 2000; Price, 2007).

Therefore, it is not surprising that there has been a focus on examining whether there are functional correlates of this medial-lateral distinction in primates and, more recently, in rodents. In rodents, there is now mounting causal evidence that MO and LO play distinct roles in a range of functions, including cost-benefit decision making (St. Onge & Floresco, 2010; Stopper et al., 2014), reversal learning (probabilistic spatial reversal learning, Dalton et al., 2016; Verharen et al., 2020; delay-discounting with spatial reversal, Mar et al., 2011; serial visual reversal learning, Hervig et al., 2019), and drug reinstatement (Arinze & Moorman, 2020; Fuchs et al., 2004). While there are many differences between the studies, it appears that MO disruption particularly affects choices when values have to be integrated and compared, whereas LO disruption has a bigger effect on value updating in similar circumstances. In addition, it is notable that LO disruption only impairs behavioural flexibility in Pavlovian stimulus-outcome learning (Gallagher et al., 1999; Gardner et al., 2017; Panayi & Killcross, 2018; Schoenbaum, Setlow, Nugent, et al., 2003), whereas MO disruption instead impairs behavioural flexibility in instrumental action-outcome learning (Bradfield et al., 2015; Gardner et al., 2017, 2018; Gourley et al., 2016; Ostlund & Balleine, 2007; Panayi & Killcross, 2018). This pattern of dissociable medial-lateral functions is also supported by electrophysiological recordings of neuronal activity in LO and MO showing distinct representations of task parameters (e.g. size of predicted reward in a delay discounting task, Burton et al., 2014; Roesch et al., 2006; trial type, epoch within a trial and value state representation in a

Pavlovian unblocking task, Lopatina et al., 2017), with neurons in MO tracking value representation independent of context and LO tracking contextual features independent of value.

In their recent review, Bradfield & Hart (2020) conciliate these functional differences by suggesting that MO and LO have complementary, interdependent roles in the representation of cognitive maps of task spaces (Wilson et al., 2014). Specifically, they argue that LO represents the animal's position within the task, determining which actions are available for selection at that point. On the other hand, MO encoding of unobservable rewards suggests a role in the representation of the future position within the task space after selection of each of the available actions. Together, these two OFC subdivisions might thus guide the animal in selecting the actions necessary to achieve a particular goal.

Functional dissociations between medial and lateral OFC subdivisions have also been observed in both NHP and human studies (e.g. Kringelbach & Rolls, 2004). Again, the most consistent differences have been observed in tasks where subjects make cost-benefit decisions or need to exhibit behavioural flexibility (Rudebeck & Murray, 2011; Walton et al., 2011). There appear to be encouraging commonalities between these and rodent studies. For instance, medial OFC is more commonly associated with value integration and comparison and with internally-guided choices, whereas lateral OFC appears more closely linked with sensory processing and updating specific stimulus-reward associations (Howard et al., 2015; Noonan et al., 2010, 2017; Rudebeck & Murray, 2011; Walton et al., 2010).

There is also evidence for an anatomical underpinning to the medial-lateral distinction in rodents, which follows the connectivity patterns observed in primates. For example, Heilbronner et al. (2016) point out that rodent MO, like primate medial OFC, is well placed to form part of a medial motivational/visceromotor network based on

projections to medial and ventral striatum while VO/LO, like NHP central/lateral OFC, projects to the central striatum and thus could comprise a sensorimotor network. Similarly, just as in NHPs, there is a gradient of distinct mediodorsal thalamic inputs in rodents that suggest separate visceromotor and sensory networks from medial to lateral OFC, respectively (Krettek & Price, 1977a; Price, 2006; Ray & Price, 1993; Reep et al., 1996; Shi & Cassell, 1998). Moreover, VO and LO, but not MO, are reciprocally connected to primary and secondary somatosensory areas (Reep et al., 1996), and both, along with AI, receive projections from the piriform cortex, suggesting a strong association between these lateral regions and a sensory network (Datiche & Cattarelli, 1996).

However, there are a number of findings that raise questions about how straightforward this homology is. For example, while perirhinal cortex connections are primarily restricted to the more central/lateral orbital network in NHPs, all OFC subregions in the rat share reciprocal projections with the perirhinal cortex (Agster & Burwell, 2009; Kondo, Saleem, & Price, 2005; Price, 2006). Likewise, projections to the periaqueductal grey and lateral hypothalamus in the rat are not only from MO and VO, corresponding to similar projections from a medial visceromotor network in NHPs, but also from AI (Floyd et al., 2000; Hoover & Vertes, 2011; Öngür et al., 1998). Furthermore, the exact boundary of a medial and lateral network in rodents is under debate. For example, while VO is often included in the lateral network (Heilbronner et al., 2016), Hoover & Vertes (2011) instead suggest that rodent VO should be linked more closely with MO as part of a medial network given that they share similar projection patterns and both regions have only weak projections to LO.

Therefore, while rodent MO and LO appear to map neatly onto a medial-lateral connectional distinction, adjacent VO and AI regions do not. Note, however, that even in NHPs, some medial and lateral OFC regions are classified as belonging to both the medial

and orbital networks (Carmichael & Price, 1995; Kondo et al., 2005; Price, 2007). Notably, several of these "intermediate interface" areas in primates are agranular, which have been proposed to be the most likely homologues with rodent agranular OFC regions (Wise, 2008). Therefore, while further systematic tracing studies are needed to map out rodent OFC networks, the extant evidence suggests that broadly similar medial and lateral orbital networks might exist in rodents and NHPs (Rudebeck & Murray, 2011; Wallis, 2011).

Beyond a dichotomous mediolateral distinction?

Despite the compelling overlap between anatomical and functional evidence for separate medial and lateral orbital networks, it remains an open question whether this is the most appropriate level of granularity to consider OFC functions. For example, as described above, while a lateral orbital network might include VO, LO, and AI, there is increasing evidence that these VO and AI regions may have distinct and even dissociable functions compared to LO.

Even though LO and AI are putatively classified as part of the lateral orbital connectivity network, there is evidence that they have different functions. For example, Ishii and colleagues showed that LO and AI have bidirectional effects on risky decision-making: LO inactivation increases risk preference in rats during a gambling task, whereas AI inactivation decreases it (Ishii et al., 2012). In a subsequent study, they showed that these effects are mimicked by 5-HT1a receptor antagonists in LO or AI (Ishii et al., 2015). Given the interconnections between these areas, Ishii and colleagues speculate that the areas may interact to ensure an appropriate balance of behavioural policies after gains or losses for AI and LO, respectively. Moreover, chemogenetic-induced inhibition of posterior AI impaired the expression of goal-directed behaviour, whereas VO/LO

inhibition only disrupted goal-directed behaviour when there was a change in the identity of the outcome (Parkes et al., 2018). This suggests that the recruitment of VO/LO in action-outcome learning depends on the specific task structure, whereas AI plays a more general role in supporting goal-directed behaviour.

Unfortunately, to date, few studies have directly compared the functions of VO and LO. However, lesions of rat VO but not LO were found to impair performance in a spatial reference memory task (Corwin et al., 1994) and cause attentional deficits (King et al., 1989), suggesting that there may also be a functional difference distinction between VO and LO (see Balleine et al., 2011).

Recently in NHPs, in addition to distinguishing between medial and lateral OFC (Rudebeck & Murray, 2011; Walton et al., 2010, 2011), a finer-grained distinction is emerging within lateral OFC networks. Several studies have specifically implicated the more central areas 13/11 as necessary for processing of current reward value and identity but not for behavioural flexibility (e.g. Burke et al., 2014; Rudebeck et al., 2017), whereas parts of more lateral areas, such as orbital area 12 and adjacent anterior insula are involved in credit assignment of reward value to cues and actions and implementing flexible behavioural strategies (e.g. Chau et al., 2015; Rudebeck et al., 2017; Sallet et al., 2020; Wittmann et al., 2020).

Such functional differences appear underpinned by distinct differences in inputs to these orbital regions, as well as distinct cortico-striatal projections organised topographically along the medial-lateral axis. In rodents, for instance, in support of the dissociations observed by Ishii and colleagues (Ishii et al., 2012, 2015), AI and MO send strong projections to the nucleus accumbens, whereas LO and VO project more densely to the central striatum (Berendse et al., 1992; Heilbronner et al., 2016; Schilman et al., 2008). Additionally, AI and MO but not LO or VO receive strong dopaminergic inputs

(Descarries et al., 1987). Moreover, these distinctions extend beyond a separation of AI from other lateral OFC sectors. For instance, both inputs from prelimbic and infralimbic cortices (Vertes, 2004) and outputs to the basolateral amygdala (BLA) (Gabbott et al., 2005) are strong for MO, VO, and AI, but only weak for LO. There is also a medial-lateral distinction in inputs from the thalamus to the posterior portions of VO and LO, with both receiving receive distinct, topographically organised projections from the mediodorsal and submedius thalamic nuclei (Reep et al., 1996).

Similarly, in primates, there are also a number of anatomical distinctions between central (area 11/13) and more lateral (area 12) regions that might support functional differences. For example, in primates, polymodal inputs from the superior temporal gyrus more strongly target lateral than central OFC regions (Carmichael & Price, 1995; Price, 2007), whereas the perirhinal cortex projects predominantly to central OFC (Ongur & Price, 2000). Moreover, there are gradations of strength of inputs from the amygdala across the mediolateral axis, with the strongest projections targeting medial and lateral regions but not central OFC.

Taken together, these studies suggest that rather than a simple lateral versus medial functional distinction, it matters where exactly in the mediolateral axis OFC function is investigated (Izquierdo, 2017; Murphy & Deutch, 2018).

Anterior vs posterior distinctions

Another important functional distinction, although so far less recognised in the field, lies along the anterior-posterior axis (see also Murray & Rudebeck, 2018). For instance, Panayi & Killcross (2018) found that, while both anterior (ALO) and posterior LO (PLO) lesions, defined anterior and posterior to the genu of the corpus callosum, impaired Pavlovian outcome devaluation, only PLO but not ALO lesions disrupted Pavlovian

reversal learning. This suggests that, within the anterior-posterior axis of the LO, there is a dissociation of rapid, flexible behavioural control and long-term updating of outcome value. Bradfield and colleagues (2018) have also revealed a similar anterior-posterior functional dissociation within rodent MO. They found that only lesions of the anterior but not posterior portion of MO disrupt instrumental devaluation, selectively implicating this anterior region in the control of instrumental actions based on inferring information about future unobservable outcomes (Bradfield et al., 2018).

While an anterior-posterior functional organisation has also been reported in primates, this has tended to highlight a gradation in the complexity of reinforcers, moving from more primary representations in evolutionarily older posterior regions to more abstract secondary reinforcers in more anterior regions in both medial and lateral OFC (Klein-Flugge et al., 2013; Kringelbach & Rolls, 2004; McNamee et al., 2013; Sescousse et al., 2010, 2013; Smith et al., 2010). Of more relevance, and paralleling the observations by Panayi & Killcross (2018), Murray et al. (2015) found a double dissociation after selectively inactivating either anterior (area 11) or posterior (area 13) OFC in NHPs at different points of an outcome devaluation task (Murray et al., 2015). They concluded that posterior OFC appears to be required for updating outcome valuations during the selective satiety devaluation procedure, whereas anterior OFC is necessary for translating this knowledge into action for flexible behavioural control.

Given the relative lack of cytoarchitectonic distinctions in the rodent (see Van De Werd & Uylings, 2008), there is increasing interest in whether anatomical differences underpin these functional distinctions. Bradfield and colleagues (2018) showed that their observed pattern of impairments was supported by anatomical differences, with anterior MO being more strongly connected to the nucleus accumbens (NAc) core while posterior MO has a higher density of projections to the BLA (Bradfield et al., 2018). Taken together

with previous studies on the effects of lesions to these target regions on flexible behaviour in instrumental devaluation (Corbit et al., 2001; Shiflett & Balleine, 2010), this also suggests the existence of dissociable functional circuits from anterior and posterior MO. It remains to be investigated whether similar anatomical distinctions between anterior and posterior LO underlie the functional differences observed by Panayi & Killcross (2018). I address this question in Chapter 3 of the present thesis.

Anterior-posterior distinctions such as those I have highlighted here are likely to be present within other subdivisions of the OFC. For instance, the anterior and posterior portions of VO receive different noradrenergic inputs from the locus coeruleus (Cerpa et al., 2019). Taken together, this suggests that separate anterior-posterior functional distinctions might be present throughout rodent OFC. Such functional differences might also exist between DLO and AI, which are anatomically separate along the anterior-posterior axis. Future studies selectively targeting DLO, for which there is currently a paucity of both functional and connectivity data, might help test this hypothesis.

Summary

Recent theoretical reviews have attempted to unify the significant functional differences between medial and lateral OFC regions (e.g. Bradfield & Hart, 2020; Wilson et al., 2014). However, it is increasingly clear that, within this medial-lateral dichotomy, there are complementary but distinct functional roles for all the classically defined orbital sub-regions. Therefore, it is also important to understand how the complementary functions of these subregions contribute to the overall function of the OFC. Indeed, here I also highlight emerging evidence which points to further distinctions along the anterior-posterior axis within these classically defined OFC subregions.

While sensitivity to these anatomical distinctions is increasing in the field, there are currently few studies available that explicitly target multiple subregions to enable direct comparisons of anatomical connectivity, neuronal activity, or dissociable function. This issue is not unique to OFC: for example, there is also evidence for anterior-posterior distinctions within the prelimbic and infralimbic cortex (e.g. Reynolds & Zahm, 2005), which may also underlie functional differences. A recent review by Laubach and colleagues (Laubach et al., 2018) examining similar issues when comparing studies of medial frontal structures called for increased use of precise anatomical terms (once they are agreed upon) along with reporting of estimated distances from agreed landmarks such as bregma; it could be argued that a similar rigour could also benefit OFC research to facilitate clarity and reproducibility.

It is also apparent that these anatomical considerations are relevant across species and may aid in clarifying a number of outstanding questions in the literature regarding homology between rodent and primate OFC (Ongur & Price, 2000; Preuss, 1995; Rudebeck & Murray, 2011; Wise, 2008). A fine-grained anatomical consideration of functional differences between OFC subregions has also recently helped refine questions of whether the OFC is indeed necessary for reversal learning, normally considered a hallmark deficit following OFC lesions, in NHPs (Burke et al., 2014; Rudebeck et al., 2013; Sallet et al., 2020). For example, Rudebeck and colleagues (Rudebeck et al., 2013; Rudebeck & Murray, 2014) have shown that, unlike aspiration lesions, fibre-sparing excitotoxic OFC lesions in macaques do not cause reversal deficits. In contrast, excitotoxic lesions of OFC in rodents and marmosets impair reversal learning. Recently, Wittman et al. (2020) have implicated macaque agranular insular regions in reversal learning. They propose that the discrepancies between reversal deficits in macaque, marmoset and rodent studies may be resolved by considering the macaque agranular insular, posterior

to OFC, as homologous to marmoset and rodent OFC. Indeed in rodents, deficits in adaptive responding, including reversal learning, have often been caused by targeting more posterior regions of LO and adjacent AI regions (Groman et al., 2019; Lichtenberg et al., 2017; Panayi & Killcross, 2018). An approach sensitive to these anatomical considerations could also help resolve questions about whether the rodent OFC encodes neuroeconomic value (Gardner et al., 2018, 2019; Levy & Glimcher, 2012; Padoa-Schioppa & Assad, 2008). Taken together, the evidence reviewed here indicates that as OFC subregions are targeted with increasing specificity both across the medial-lateral and anterior-posterior axis, many surprising functional dissociations are likely to emerge within the OFC. In Chapter 3, I investigate whether OFC functional heterogeneity along its anterior–posterior and medial–lateral axes is underpinned by differences in anatomical connectivity.

1.5 Thesis aims

In my doctoral thesis, I aimed to investigate how the brain represents information to facilitate behavioural flexibility during decision-making. In particular, I focus on the role of the OFC, the NAc and the dopaminergic reward system, which play a critical role in reward-guided learning and flexible behaviour. I start by investigating whether OFC functional heterogeneity is underpinned by differences in anatomical connectivity (Chapter 3). Then, I investigate the contribution of dopamine to haemodynamic signatures of RPEs encoding by combining amperometry recordings of tissue oxygen signals in the OFC and NAc with lesioning of NAc dopaminergic inputs during a probabilistic reversal-learning task (Chapter 4). In another study, I combine the same behavioural paradigm and brain activity monitoring technique with injections of

amphetamine to investigate the neuronal mechanisms underlying maladaptive learning in hyperdopaminergic states similar to those thought to elicit psychotic symptoms in neurological and neuropsychiatric conditions (Chapter 5). Finally, I attempt to develop a novel rodent behavioural paradigm for investigating how the brain encodes information during generalisation (Chapter 6).

2. Materials and methods

All animal research procedures described here were approved by the local ethical review board at the University of Oxford and carried out in accordance with the UK Home Office regulations and under the Animals (Scientific Procedures) Act 1986.

2.1 Anatomical characterisation of OFC inputs along its axes of organisation

Chapter 3 describes the results of a neuroanatomical tracing study in which we characterised the retrograde projections of different subdivisions to the rat orbitofrontal cortex (OFC), using the tracer cholera toxin subunit B (CTB). CTB enters cells through endocytosis and travels to the endoplasmatic reticulum (Wernick et al., 2010). Therefore, injection of CTB into a brain region allows for visualisation of the cell bodies of neurons projecting to it (Stoeckel et al., 1977). When coupled to dyes such as Alexa Fluor, CTB can be visualised with a fluorescence microscope once the brain sections have been mounted onto slides without additional processing steps (e.g., antibody staining). This, combined with rapid transport (~7 days), makes CTB an attractive tracer for characterising retrograde projections. Importantly for our study, there are several commercially available CTB compounds coupled to different Alexa Fluor conjugates, allowing for simultaneously labelling retrograde projections from various brain subdivisions (Conte et al., 2009).

2.1.1 Subjects

A total of 14 male Lister Hooded rats (Envigo, UK), ~275-300 g in weight at the start of the experiment (8-12 weeks old), were used in this study. Animals were housed in groups

of 3 in polycarbonate cages, with *ad libitum* access to water and food. The housing room was maintained at a room temperature of 23 ± 1°C, humidity of 40 ± 10% and on 12-hour light/dark cycles beginning at 7 A.M., with lights during the day.

2.1.2 Surgery

Rats were placed in induction chambers and anaesthetised initially with 4% isoflurane (2 L/min O_2) and maintained on 1–2% isoflurane (2 L/min O_2) for the rest of the procedure. After induction, the head was shaved, and the rat was secured in a stereotaxic frame (Kopf Instruments). Body temperature was maintained at 37 ± 0.5°C with the use of a homoeothermic heating blanket. Corneal dehydration was prevented with the application of ophthalmic ointment (Lacri-Lube, Allergan). The scalp was cleaned with diluted hibiscrub and 70% ethanol, and a local anaesthetic, bupivacaine (2 mg/kg), was applied to the area. The skin on top of the head was retracted, and holes were then drilled in the skull for the injections. 50 nl of retrograde tracer CTB coupled to Alexa Fluor 488 (green), 555 (yellow) or 648 (red) conjugates (C34775, C34776, C34778, Life Technologies) was microinjected through micropipettes with ~10-µm Ø pulled from thin-walled borosilicate glass capillaries (outer Ø: 1.14 mm, inner Ø: 0.53 mm; 3-000-203-G/X, Drummond Scientific) using a vertical micropipette puller (PE-2, Narishige) via an automated injector (Nanoject III, Drummond Scientific). Tracers were injected unilaterally, into the right hemisphere, at a rate of 1 nl/s. Each animal received injections of one tracer into the anterior lateral (ALO), another into posterior lateral (PLO), and the third one into posterior ventral portions (PVO) of the OFC. Tracer-site combinations (i.e., CTB coupled with a particular conjugate) were counterbalanced across animals. The coordinates used were as indicated in Table 2.1. Before injection, the micropipette was lowered 0.2 mm below the intended injection sites and left in place for 1 min before being

raised to the injection site dorsoventral coordinate. This step was performed to avoid backflow of the tracer along the needle track by creating a larger space for the injection bolus with the thicker portion of the micropipette. Micropipettes were left in place for 5 min after each bolus injection to ensure diffusion of tracer. Micropipettes were then slowly retracted, and the incision closed with vicryl sutures.

All animals were administered buprenorphine (0.1 ml/kg, sc), both pre- and post-surgically, and meloxicam (Metacam, 0.2 ml/kg, sc), post-surgically. Animals were then allowed to recover in thermostatically controlled cages and given palatable food for consumption. Meloxicam was also administered for at least 3 d following surgery and rats were monitored closely for at least 7 d post-surgery.

Table 2.1

Coordinates used for tracer injection into OFC subdivisions

OFC subdivision	AP from bregma (mm)	ML from bregma (mm)	DV from dura (mm)
ALO	+3.7 (+4.2)	+2.5 (+2.6)	-3.3 (-3.5)
PLO	+2.7 (+3.2)	+3.3 (+3.4)	-4.4 (-4.7)
PVO	+2.7 (+3.2)	+1.9 (+2.0)	-4.4 (-4.5)

Note. Coordinates were experimentally adjusted based on data from a pilot study to target the brain areas at the coordinates depicted in brackets in a standard rat brain atlas (Paxinos & Watson, 1998). Based on the previously established functional boundary in lateral OFC (LO) (Panayi & Killcross, 2018), we used the presence of the corpus callosum and the claustrum in the coronal section to mark the anterior-posterior division of LO. AP: anteroposterior, ML: mediolateral, DV: dorsoventral.

2.1.3 Histology

Ten days after tracer injection, animals were administered a terminal dose of pentobarbitone (30-60 mg/kg). After loss of pedal reflex, the animals were perfused transcardially with 150 ml of phosphate buffer saline followed by 400 ml of 4% paraformaldehyde. The brains were removed, kept in paraformaldehyde for 24 h and then transferred to phosphate buffer saline. Before slicing, brains were cryoprotected in

30% sucrose solution. 60-µm thick coronal sections were obtained with a freezing microtome (Leitz). Sections were mounted onto slides and coverslipped with mounting medium with DAPI (H-1500 Vectashield, Vector Labs).

2.1.4 Image acquisition and analysis

Images of whole brain sections were captured using a fluorescence slide scanner (Axio Scan.Z1, Zeiss) equipped with an air 20x/NA 0.8 objective. Images of example single-, double, and triple-labelled neurons in the submedius nucleus of the thalamus were acquired using a confocal microscope (Zeiss LSM 880) equipped with a 20x/NA 0.8 Plan-Apochromat objective. Image colour settings, including brightness, contrast and gamma, were adjusted to aid visualisation of labelled cells.

Data from injections that were off-target (n = 11 injection sites), in which tracer did not diffuse (n = 8 injection sites), were both off target and the tracer did not diffuse (n = 6 injection sites), or in which the tracer diffused into the white matter (n = 2 injection sites) were excluded from all analyses (see Figure 3.3). The data presented in Chapter 3 therefore included 7 brains: 4 triple-labelled (n = 12 injection sites), 3 single-labelled (n = 3 injection sites).

Levels of cell labelling density were manually categorised offline into four levels: 0, absence of labelling; 1, weak (i.e., just a few labelled cells); 2, moderate; 3, strong (see Figure 3.1 for examples). The summary table and figure depicting the average labelling for each injection site are based on the density average across brains, subsequently averaged across slices when a single slice is represented and discretised into four density values: –, absence of labelling; +, weak; ++, moderate; +++, strong (Table 3.2 and Figure 3.14). The weak density level (+) also includes any region consistently labelled, i.e., across at least 50% of the brains, even if only weak labelling was observed. Areas with maximal

tracer deposit, only found around injection sites, were not included in the analyses. If a brain region contained maximal tracer deposit in a coronal plane in more than 50% of the brains included in the analyses, we represented that area as maximal deposit in the average labelling figure; otherwise, we calculated the average labelling for that region excluding those brains. The classification of brain areas was based on a standard rat brain atlas (Paxinos & Watson, 1998). The nomenclature used to classify area 1 of the anterior cingulate cortex was based on (Paxinos & Watson, 2013) to reflect current understanding of the heterogeneity within the medial prefrontal cortex (Laubach et al., 2018). In our analysis, we focused on the quantification of frontal lobe structures, including OFC subdivisions, prelimbic, infralimbic, anterior cingulate cortex and motor cortex; temporal lobe structures, including amygdala, lateral entorhinal cortex and perirhinal cortex; retrosplenial cortex; primary sensory cortices; thalamic nuclei, including paratenial, submedius and mediodorsal. Due to difficulties in obtaining brain slices of consistent quality posterior to -5.30 mm from bregma, we did not quantify labelling in midbrain structures.

The proportion of double- and triple-labelled cells in the submedius nucleus of the thalamus was obtained by counting the number of single-, double- and triple-labelled cells in this region coronal slices at -2.12, -2.30 and -2.56 mm from bregma of successfully triple-labelled brains. These values were then summed across slices and averaged across brains.

In the description and interpretation of our results, we consider anterior and posterior divisions of medial (MO), ventral (VO), and lateral OFC (LO) separately, but not of dorsolateral OFC (DLO). We refer to the portions of LO and VO contained in the coronal slices from +4.70 to + 4.20 mm from bregma as anterior and to those contained in the coronal slices from +3.20 to +2.20 mm as posterior, with the corpus callosum and the

claustrum in the coronal section marking the anterior-posterior division as previously established for LO (Panayi & Killcross, 2018). Based on a previously observed functional dissociation in MO (Bradfield et al., 2018), we refer to the areas in coronal slices at +4.70 and +3.70 mm from bregma as anterior and posterior MO, respectively. These coordinates are based on a standard rat brain atlas (Paxinos & Watson, 1998). While a ventrolateral OFC (VLO) area has been proposed to exist between VO and LO, here, we will only differentiate between areas VO and LO as per the boundaries defined in Paxinos & Watson (1998).

We tested whether differences in labelling strength were statistically different in key areas of interest (i.e., thalamus, amygdala) using the average labelling density across brain sections for each area subregion. We used a nonparametric approach to an ANOVA model to capture the factorial structure of the data analysed using R statistical software (Lenth et al., 2020; R Core Team, 2020). A nonparametric approach was important given that the labelling was quantified with ordinal data. We used an aligned rank transform procedure (Wobbrock et al., 2011) that allows for valid main effect and interaction inferences while maintaining a nominal Type I error rate. Multiple-comparison post-hoc analyses were corrected using a Sidak correction. In all analyses, a mixed-effects model with a random effect of subject was defined at the level of each brain, and tests are reported with Kenward-Roger residual degrees of freedom. We used a similar approach to statistically test differences in labelling strength in the subregions of the submedius nucleus of the thalamus across coronal sections.

Depiction of labelling in key areas of interest, including the amygdala, mediodorsal thalamus and submedius nucleus of the thalamus, was obtained by superimposing hand drawings of the labelling in each brain using an opacity value proportional to the density level of labelling observed.

2.2 Investigating haemodynamic correlates of prediction errors in the NAc and OFC

In Chapters 4 and 5 of this thesis, we investigate the effects of manipulating dopamine levels on the haemodynamic correlates of prediction errors in the nucleus accumbens (NAc) and OFC. These haemodynamic signals were obtained from rats performing a probabilistic reversal-learning task after lesioning of dopaminergic inputs into the NAc (Chapter 4) or after systemic injection of amphetamine, an indirect sympathomimetic known to potentiate dopamine release, on key training sessions (Chapter 5). For simplicity, I refer to these two sets of experiments as 'dopamine lesion' and 'hyperdopamine' experiments throughout this section. This work formed part of a collaboration between my lab and groups at Eli Lilly U.K. The experimental procedures and data collection were performed at Lilly, and I performed all the data analyses.

Oxygen amperometry

Haemodynamic recordings were obtained using oxygen amperometry. This technique is based on the measurement of the current generated by oxidation-reduction reactions of electroactive molecules in contact with a probe when current is applied. Applying a constant potential of -650 mV at carbon paste electrodes induces the reduction of dissolved oxygen. This occurs in two steps, the first of which results in the formation of hydrogen peroxide, which is, in turn, reduced to water:

$$O_2 + 2H^+ + 2e^- \longrightarrow H_2O_2$$

$$H_2O_2 + 2H^+ + 2e^- \longrightarrow 2H_2O$$

Thus, the difference between the current applied and the current measured is proportional to the concentration of oxygen on the surface of the electrode (Bazzu et al., 2009; Lowry et al., 1996). These measurements serve as an index for local blood flow and can be considered as comparable to blood-oxygenation level-dependent signals obtained with functional magnetic resonance (fMRI) (Francois et al., 2012; Li et al., 2011; Lowry et al., 1997, 2010). While fMRI studies have the advantage of being non-invasive, they require subjects to remain motionless during scanning, largely limiting fMRI imaging in rodents to non-behavioural studies. Additionally, oxygen amperometry does not require habituating animals to the scanner noise or correction for motion artefacts in the signals. Moreover, it allows for multi-site recordings (e.g., Francois et al., 2014; Werlen et al., 2019) and can be combined with electrophysiology (e.g., McHugh et al., 2011). This makes oxygen amperometry a potential translational tool to help bridge between fine-scale invasive studies of neural activity in rodents and human fMRI research (Gilmour et al., 2015).

2.2.1 Subjects

Adult male Wistar rats (Charles River, UK) were used: n = 36 and n = 37 in the dopamine lesion and hyperdopamine experiments, respectively. Before surgery, animals (400-450 g) were housed in standard housing conditions: four per cage, 12 h light/dark cycles beginning at 7 A.M. with lights during the day, controlled temperature and humidity, *ad libitum* water. Rats were food restricted and maintained at no less than 85% of their free-feeding weight relative to their normal growth curve. For environmental enrichment, home cages were supplied with wooden blocks and nesting materials. After surgery, animals were singly housed in the same environment.

2.2.2 Behavioural experiments

Apparatus

Operant boxes (length: 30, width: 25, height: 19 cm) housed in sound- and light-attenuating chambers were used (standard rat operant chamber, Med Associates), ventilated with a fan that provided constant background noise. Each chamber contained a houselight (100 mA; ENV 215M, Med Associates) and two retractable levers. The levers were located on either side of a recessed magazine where food pellets (45 mg; Formula P, Noyes) were delivered from an automatic pellet dispenser. Magazine entries were detected with an infrared beam-break detector (IR LED and photodiode pair). Experimental sessions were controlled, and data recorded using in-house programs written with MedPC-IV software (Med Associates).

Probabilistic instrumental reversal-learning task

Animals performed an operant 2-option probabilistic reversal learning task. One option was associated with a 75% probability of reward delivery and the other a 25% reward probability, and this association would reverse each week (4-5 sessions block – see below for details).

At the beginning of each trial, a cue light illuminated above either both levers (choice trials) or just one of the levers (forced trials). After an 8 s interval, the corresponding levers extended, and the rats then had 10 s to respond by pressing a lever. A single lever press caused the retraction of the levers and the food magazine light to illuminate. After a 5 s delay, the lights of levers and the food magazine turned off, and, on rewarded trials, four 45 mg sucrose pellets were delivered to the food magazine. There was then an inter-trial interval (ITI) of 30-50s. If the rats did not make a lever response

within 10 s of lever extension, an omission was recorded. Omissions were followed by a 30 s penalty timeout, followed by the standard 30-50 s ITI.

Each training session consisted of 112 forced and choice trials, delivered in mini-blocks of 16 trials, consisting of 8 forced trials (4 to each lever, pseudorandom order) followed by 8 choice trials. See Figure 2.1 for a schematic summary of the times of the events which each trial involves.

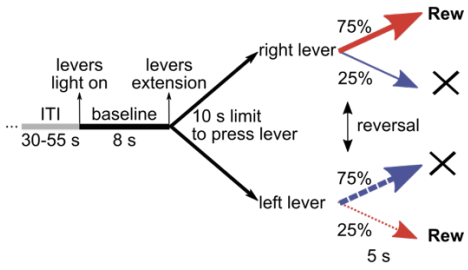


Figure 2.1 Trial structure of probabilistic reversal learning task in choice trials. In forced trials, only one of the levers was made available. Rew: rewarded.

Behavioural training

During all training steps, the session start was signalled by the houselight turning off, and its permanent illumination indicated the end of a session. Animals progressed through several training stages, as schematised in Figures 2.2 and 2.3. First, animals were given a 30 min session to familiarise themselves with the food magazine (magazine training stage). Pellets were delivered on a variable interval (VI) 60 s schedule, and head entries were recorded. During the following two days, animals were trained to lever press for food pellets (45 min) (Stage 1). Only one of the levers was available on the first day and, on the second day, the available lever would be the alternative one. At this stage, animals would receive pellets for lever presses on a fixed ratio (FR) of 1:1 (e.g., 1 lever press would always lead to 1 food pellet). Over the next five training sessions, the delay from lever press to reward delivery gradually increased from 0 to 5 s and the forced trials

number from 0 to 50% of the total trials number (Stages 2 to 5). On the next training stage (Stage 6, 5 sessions), the probability of obtaining a reward after pressing either lever dropped to 50%. During this stage, rats were trained on this schedule one session per week. Then, animals were assigned to either 25/75% or 75/25% reward probability after pressing right/left levers and were trained on this schedule for five sessions (Stage 7). The assignment of high/low reward probability schedule to left/right levers was counterbalanced across animals while counteracting any observed lever preference bias that individual animals developed during the previous sessions. At the end of this training stage, animals underwent surgery. After a recovery period of ~3 weeks, there was one session for O_2 recordings calibration. For the dopamine lesion experiment, there was also one session for testing the goodness of the 6-OHDA lesion with subcutaneous apomorphine injection (0.1 mg/kg).

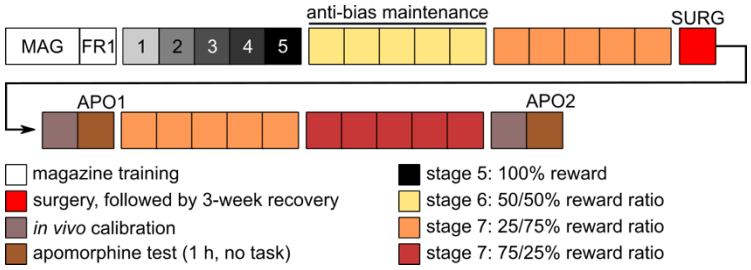


Figure 2.2 Behavioural training stages and key procedures (Dopamine Lesion Experiment). Food restricted rats were randomly assigned to two experimental groups and trained on a probabilistic reversal learning task before and after undergoing surgery for bilateral implantation of electrodes in the OFC and NAc. APO: apomorphine injection; FR1: fixed ratio of 1 lever press:1 reward; Rew: rewarded; SURG: surgery.

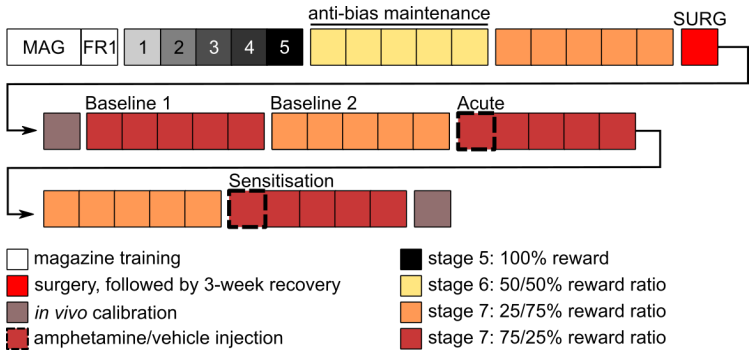


Figure 2.3 Behavioural training stages and key procedures (Hyperdopamine Experiment). Food restricted rats were trained on a probabilistic reversal learning task before and after undergoing surgery for implantation of electrodes in the OFC and NAc. On the first day of the third and fifth post-surgery 5-session blocks, rats received a system injection of amphetamine of vehicle. FR1: fixed ratio of 1 lever press:1 reward; MAG: magazine training; SURG: surgery.

The testing phase consisted of two (dopamine lesion) or five blocks (hyperdopamine) of ~5 days each. In the dopamine lesion experiment, the first 5-day block had the same lever-reward probability contingencies as in pre-surgery Stage 7 and then reversed at the beginning of each subsequent block. However, in the hyperdopamine experiment, the first 5-day block had the *opposite* reward contingencies as in stage 7.

After the last day of training and ~5 weeks after the first post-surgery behavioural session, there was again one session for O_2 recordings calibration and, for the dopamine lesion experiment, another session where apomorphine was administered. For the hyperdopamine experiment, rats received amphetamine (amph) or vehicle (veh; PBS) injections before the first session of the third and fifth 5-day session blocks. Amperometric recordings took place during the training phase after the surgery, and the amperometric tissue oxygen (T_{O2}) signal was recorded for the entire length of each session.

2.2.3 Electrode construction

Carbon paste electrodes (CPEs) were constructed from 8T (200 µm bare diameter, 270 µm coated diameter) Teflon-coated silver wire (Advent Research Materials), as described previously (Francois et al., 2012; McHugh et al., 2011). The Teflon insulation was slid along the wire to create a ~2 mm deep cavity, which was packed with carbon paste. Carbon paste was prepared by thoroughly mixing 7.1 g of carbon graphite powder and 2.5 ml of silicone oil (both from Sigma-Aldrich) (O'Neill et al., 1982). Reference and

auxiliary electrodes were also prepared from 8T Teflon-coated silver wire by removing the Teflon from the tip. All electrodes were soldered to gold connectors, which were cemented into six-pin plastic sockets (both from Plastic One) during surgery.

2.2.4 Electrode calibration

Before implantation, all CPEs were calibrated *in vitro* in a standard three-electrode glass electrochemical cell (C3 cell stand, BASi), with an Ag/AgCl reference electrode and a BASi platinum auxiliary electrode. Calibrations were performed in a 15 ml phosphate buffer saline solution (PBS), pH 7.4, saturated with nitrogen (N_2) gas, atmospheric air (from a RENA air pump), or pure O_2 at room temperature, providing a 3-point calibration of known concentrations of 0 μM (N_2 saturated), 240 μM (air-saturated), and 1260 μM (O_2 saturated) oxygen. CPEs were chosen for implantation if their calibration curves were linear and the measured oxygen values from the saturated solutions were not greatly different from those expected (least-square linear regression, $R^2 \geq 0.98$).

2.2.5 Surgery

Animals were placed in vaporisation chambers and anaesthetised with 5% isoflurane (2 L/min O_2) and maintained on 1–2% isoflurane (2 L/min O_2) for the rest of the procedure. After induction, the head was shaved, and the animal secured in a stereotaxic frame (Kopf Instruments). Body temperature was maintained at 37 ± 0.5°C with the use of a homoeothermic heating blanket. Corneal dehydration was prevented with application of ophthalmic ointment (Lacri-Lube, Allergan). The scalp was cleaned with diluted hibiscrub and 70% alcohol, and a local anaesthetic, bupivacaine (2 mg/kg), was applied to the area. Following midline incision, the skull was cleaned and dried, and sites for

implantation of electrodes were stereotaxically determined using a 33-gauge internal guide cannula (C3131, Plastics One Inc).

Based on Croxson et al. (2014) and pilot experiments, for the dopamine lesion experiment, rats received injections of 6-hydroxidopamine (6-OHDA) or vehicle (phosphate buffer saline, PBS) into multiple injection sites in NAc unilaterally (left hemisphere; see Table 2.2 for coordinates details), 0.5 µl (0.25 µl/min) per site were delivered through a cannula. The cannula was left in place for a further 3 min to ensure diffusion away from the injection site. It was then retracted, moved to the next position, and the above repeated until all six sites were infused.

Table 2.2

Coordinates of 6-OHDA injection sites into the left NAc.

Target brain region	AP from bregma (mm)	ML from bregma (mm)	DV from dura (mm)
Left NAc	+2.2	+1.7	-6.1
		+1.2	-6.9
	+1.6	+2.0	-6.5
		+1.2	-6.9
	+1.1	+2.0	-4.0
		+1.2	-6.4

Note. AP and ML were measured from bregma and DV from dura. 6-OHDA: 6-hydroxidopamine; AP: anteroposterior; DV: dorsoventral; ML: mediolateral.

For the dopamine lesion experiment, CPEs were implanted bilaterally into the NAc and the lateral OFC (see Table 2.3 for coordinates details). For the hyperdopamine experiment, CPEs were implanted bilaterally into the NAc and unilaterally into both lateral and medial OFC (see Table 2.4 for coordinates details).

Table 2.3

Coordinates used for placement of electrodes into lateral and medial OFC and NAc.

Target brain region	AP from bregma (mm)	ML from bregma (mm)	DV from dura (mm)
NAc	+1.4	±1.4	-6.4
LO	+3.8	±2.6	-4.0

Note. AP and ML were measured from bregma and DV from dura. 6-OHDA: 6-hydroxidopamine; AP: anteroposterior; DV: dorsoventral; LO: orbitofrontal cortex, lateral part; ML: mediolateral; NAc: nucleus accumbens; OFC: orbitofrontal cortex.

Table 2.4

Coordinates used for placement of electrodes into lateral and medial OFC and NAc.

Target brain region	AP from bregma (mm)	ML from bregma (mm)	DV from dura (mm)
NAc	+1.4	±1.4	-6.4
LO	+3.8	+2.6	-4.4
MO	+4.4	+0.6	-4.0

Note. AP and ML were measured from bregma and DV from dura. 6-OHDA: 6-hydroxidopamine; AP: anteroposterior; DV: dorsoventral; LO: orbitofrontal cortex, lateral part; ML: mediolateral; MO: orbitofrontal cortex, medial part; NAc: nucleus accumbens; OFC: orbitofrontal cortex.

The skull surface was roughed up with a drill bit ("golf balling") to improve contact between glue/dental cement and the skull. Six screws were inserted into the skull to ensure the longevity of the head cap. The first working electrode (i.e., CPE) to be inserted was placed into an electrode holder attached to the arm of the stereotaxic frame and inserted into the brain under stereotaxic guidance. A hypodermic needle (25-gauge) was used to break the dura before the insertion of the electrode. Once inserted, the electrode was secured with glue (Geristore Syringeable, DenMat). Once all CPEs had been implanted and cemented, the auxiliary electrode was wrapped around one of the posterior skull screws. The reference electrode was inserted into the posterior cortex by 2 mm in depth and secured with glue and cement.

After all the electrodes were cemented into place, the gold sockets of the electrodes were inserted into a six-pin plastic socket ('pedestal'; MS363, Plastics One). Glue was used to secure the gold sockets of the four electrodes within the pedestal, which was then cemented to the skull. Finally, the entire pedestal was encased in dental cement (Meadway Rapid Repair, Mr Dental Suppliers Ltd) to make a robust head cap. A plastic dust cap (363DC, Plastics One) was then screwed onto the pedestal, and the rat removed from the stereotaxic frame. Post-surgically, all animals were administered analgesia (carprofen; 5 mg/kg, sc), a long-lasting antibiotic (Convenia) and 5 ml of saline (sc) and allowed to recover in thermostatically controlled cages. A post-operative recovery period of 3 weeks was allowed before testing commenced. Animals were food-restricted for 7 d before testing began.

2.2.6 O_2 amperometry data recordings

O_2 signals were recorded from the carbon paste electrodes targeted at the OFC and NAc core. Rats were connected to a four-channel potentiostat (Biostat, ACM Instruments) through a six-pin socket (Plastics One) via a flexible screened six-core cable (Plastics One). A PowerLab 8/30 Data Acquisition System was used for analogue/digital conversion, and data were collected on a computer running Chart version 5 software (AD Instruments). O_2 signals were recorded at a sample rate of 200 Hz. For all test sessions where amperometric signals were recorded, animals were tethered, and a constant potential (–650 mV) was applied for the duration of the session.

2.2.7 *In vivo* lesion validation (Dopamine Lesion Experiment)

Lesioning of dopaminergic neurons following 6-OHDA injection leads to sensitisation of any remaining dopamine receptors (Schoenfeld & Uretsky, 1972). Thus, systemic

injection of a dopamine agonist would result in potentiation of responses from the NAc ipsilateral to 6-OHDA injection. Therefore, to measure the effect of the 6-OHDA lesion *in vivo*, all rats received an injection of a non-selective dopamine receptor agonist, apomorphine (0.1 mg/kg, s. c.), on the recording sessions before the first and after the last post-surgery session. After allowing 15 min for O_2 signals to settle, animals received the apomorphine injection and signals were recorded for 1 h.

2.2.8 Pharmacological manipulations (Hyperdopamine Experiment)

Rats received amphetamine (amph) or vehicle (veh; PBS) injections before the first session of the third and fifth 5-day session blocks. D-amphetamine sulphate (Sigma, UK) was dissolved in a 5% (w/v) glucose solution, and pH adjusted towards neutral with the dropwise addition of 1 M NaOH as necessary. Amphetamine was administered at 1 mg/kg (free weight) intraperitoneally, 30 min before the behavioural session. This moderate dose of amphetamine promotes dopamine release in the NAc (e.g. Curran et al., 2004; Sharp et al., 1987) and induces behaviour change without causing exaggerated motor stereotypy that could impair operant behaviours (Grilly & Loveland, 2001). Amphetamine administration around this dose range has been shown to induce positive, psychotic-like symptoms in humans (Wallis et al., 1949). The pharmacological manipulation was blinded during administration.

2.2.9 Histology

At the end of experimentation, animals were deeply anaesthetised with pentobarbital and perfused transcardially with 0.9% saline, followed by 10% buffered paraformaldehyde, and the brains were extracted for sectioning. Before slicing, brains were cryoprotected in 30% sucrose solution. Brains were cut into 50 μm-thick coronal

sections using a freezing microtome (Leitz). One in every two slices was kept for analysing the 6-OHDA lesion and the remainder for assessing electrodes placement. Both electrode placements and the extent of 6-OHDA lesion were assessed with reference to a standard rat brain atlas (Paxinos & Watson, 1998).

Electrode placements

Brain sections were mounted on glass slides and stained with cresyl violet. Sections were then viewed microscopically to assess the placement of electrode tips. Signals obtained from any CPE with improper placement or which track could not be visualised were excluded from all analyses (see Figures 2.4 and 2.5 for placement of all electrodes; see Tables 2.5 and 2.7 for final sample sizes and exclusion details).

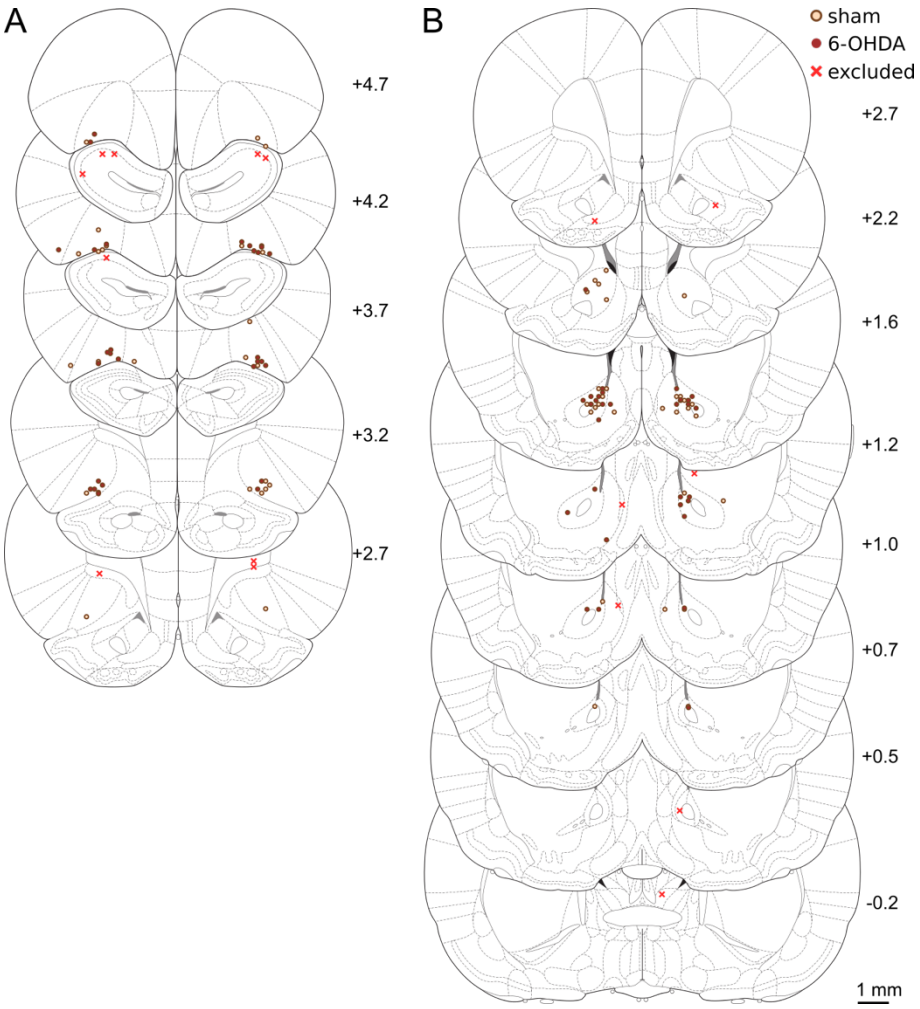


Figure 2.4 Location of all electrodes placed in OFC (A) and NAc (B) of sham and 6-OHDA lesioned subjects (Dopamine Lesion Experiment). OFC electrodes were clustered mainly in the lateral OFC, while NAc electrodes were mainly in the dorsomedial NAc, including dorsomedial core and shell. Excluded electrodes represented in this figure were excluded due to anatomical placement. *n* = 36 subjects (sham = 18, 6-OHDA = 18). 6-OHDA: 6-hydroxydopamine; OFC: orbitofrontal cortex; NAc: nucleus accumbens.

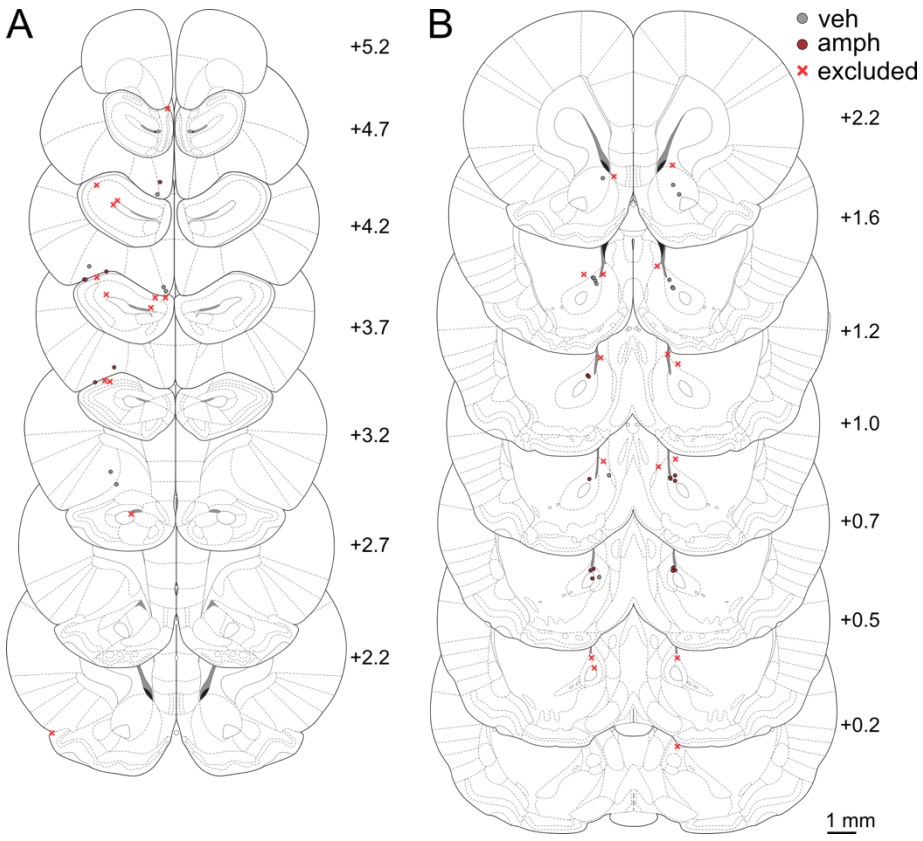


Figure 2.5 Location of all electrodes placed in OFC (A) and NAc (B) of veh and amph group subjects (Hyperdopamine Experiment). OFC electrodes were clustered in medial and lateral OFC, while NAc electrodes were mainly in the dorsomedial NAc core. Excluded electrodes represented in this figure were excluded due to anatomical placement. *n* = 22 subjects (veh = 11, amph = 11). amph: amphetamine; veh: vehicle; OFC: orbitofrontal cortex; NAc: nucleus accumbens.

6-OHDA lesion (Dopamine Lesion Experiment)

For labelling tyrosine hydroxylase (TH) containing neurons, free-floating brain slices were treated with 0.1 M phosphate-buffered saline with 0.3% Triton X-100 (PBS-T) at room temperature (RT), 3 x 10 min. Sections were then blocked with 10% normal donkey serum (NDS), for 2 h, at RT. Brain slices were then incubated overnight at RT with the primary antibody anti-tyrosine hydroxylase (rabbit, 1:1000, polyclonal; AB152, Millipore) in 2% NDS blocking solution. After further washing using PBS (3 x 10 min), sections were incubated for 4 h, at RT, with the secondary antibody anti-rabbit Cy3 (donkey, 1:1000; 711-165-152, Jackson ImmunoResearch) in 1% NDS blocking solution.

After further washing in PBS-T (3 x 10 min), slices were mounted onto slides and coverslipped with mounting medium with DAPI (H-1500 Vectashield, Vector Labs). Images were acquired with an Aperio Scanscope AT slide scanner (Leica Biosystems) at 20x magnification.

2.2.10 Data analysis

All data analysis was performed using custom-written Python scripts (Python Software Foundation, 2013). The experimental units (*n*) reported in the behavioural analyses are individual animals. Those reported in analyses of tissue oxygen amperometry recordings are individual carbon paste electrodes.

Dopamine Lesion Experiment (data analysis)

The table below summarises the sample sizes and exclusions for the data analyses of the dopamine lesion experiment.

Table 2.5

Summary of sample sizes and exclusions for data analyses of the dopamine lesion experiment.

Analysis	*n*	Reasons for exclusions
Electrode placements	36 rats (18 sham, 18 6-OHDA)	**Exclusions for all T_{O2} signals analyses:** -9 LO and 1 NAc electrodes were excluded because their track could not be visualised; -9 LO and 7 NAc electrodes were excluded due to incorrect anatomical placement.
Lesion histological assessment	33 rats (16 sham, 17 6-OHDA)	**Exclusions for this analysis:**

		-3 rats (2 sham, 1 6-OHDA) were excluded due to difficulty obtaining intact brain slices. **Exclusions for all subsequent analyses:** -1 6-OHDA rat for which it was not possible to assess the extent of the lesion histologically (see above). -4 6-OHDA rats for which the lesion was unsuccessful.
Behavioural performance and modelling	31 rats (18 sham, 13 6-OHDA)	
***In vivo* lesion validation**	-63 LO electrodes (20 6-OHDA: 10 ipsi-, 10 contralateral to injection; 25 sham: 13 ipsi-, 12 contralateral to injection) from 27 rats (14 sham, 13 6-OHDA); -56 NAc electrodes (24 6-OHDA: 12 ipsi-, 12 contralateral to injection; 32 sham: 16 ipsi-, 16 contralateral to injection) from 31 rats (18 sham, 13 6-OHDA)	**Exclusions for this analysis:** -3 LO (2 6-OHDA: 1 ipsi-, 1 contralateral to injection; 1 sham, contralateral to injection) and 2 NAc (2 sham: 1 ipsi-, 1 contralateral to injection) electrodes which did not yield meaningful (i.e., flat or noisy) signals in neither of the apomorphine tests.
Trial-by-trial T_{O2} signals: Outlier signals	-39 LO electrodes (19 6-OHDA: 10 ipsi-, 9 contralateral to injection; 20 sham: 10 ipsi-, 10 contralateral to injection) from 26 rats (14 sham, 12 6-OHDA); -38 NAc electrodes (20 6-OHDA: 11 ipsi-, 9 contralateral to injection; 28 sham: 14 ipsi-, 14 contralateral to injection) from 27 rats (15 sham, 12 6-OHDA)	**Exclusions for this and all subsequent analyses:** -7 LO (3 6-OHDA: 1 ipsi-, 2 contralateral to injection; 4 sham: 1 ipsi-, 3 contralateral to injection) and 9 NAc (4 6-OHDA: 1 ipsi-, 3 contralateral to injection; 5 sham: 3 ipsi-, 2 contralateral to injection) electrodes which did not yield meaningful (i.e., flat or noisy) signals. **Exclusions for all subsequent analyses:** -2 NAc electrodes of 6-OHDA rats (1 ipsi-, 1 contralateral to injection) whose mean current was > 2.5 SD from the mean of all channels in that region or where the sign differed from the group.

Trial-by-trial T_{O2} signals: pre-reversal, responses characterisation in *sham* rats	-19 LO electrodes (9 ipsi, 10 contralateral to injection) from 11 rats; -27 NAc electrodes (14 ipsi, 13 contralateral to injection) from 11 rats	**Exclusions for this analysis:** -Only electrodes from *sham* rats that yielded useable signals for any of the relevant training sessions were considered in this analysis.
Trial-by-trial T_{O2} signals: post-reversal, early vs late sessions in *sham* rats	-19 LO electrodes (9 ipsi, 10 contralateral to injection) from 11 rats; -26 NAc electrodes (13 ipsi, 13 contralateral to injection) from 11 rats	**Exclusions for this analysis:** -Only electrodes from *sham* rats that yielded useable signals for any of the early and late post-reversal training sessions were considered in this analysis.
Trial-by-trial T_{O2} signals: pre-reversal, comparison *ipsilateral* signals sham vs 6-OHDA	-18 LO electrodes (9 sham, 9 6-OHDA) from 18 rats (9 sham, 9 6-OHDA); -24 NAc electrodes (14 sham, 10 6-OHDA) from 24 rats (14 sham, 10 6-OHDA)	**Exclusions for this analysis:** -Only electrodes from the *ipsilateral* side to the injection and that yielded useable signals for any of the relevant training sessions were considered in this analysis.
Trial-by-trial T_{O2} signals: pre-reversal, comparison *contralateral* signals sham vs 6-OHDA	-19 LO electrodes (10 sham, 9 6-OHDA) from 19 rats (10 sham, 9 6-OHDA); -21 NAc electrodes (13 sham, 8 6-OHDA) from 21 rats (13 sham, 8 6-OHDA)	**Exclusions for this analysis:** -Only electrodes from the *contralateral* side to the injection and that yielded useable signals for any of the relevant training sessions were considered in this analysis.
Trial-by-trial T_{O2} signals: pre-reversal, comparison *control* (ipsi- and contralateral sham and contralateral 6-OHDA) vs ipsilateral 6-OHDA	-37 LO electrodes (9 ipsi sham, 10 contra sham, 9 ipsi 6-OHDA, 9 contra 6-OHDA) from 22 rats (11 sham, 11 6-OHDA); -45 NAc electrodes (14 ipsi sham, 13 contra sham, 10 ipsi 6-OHDA, 8 contra 6-OHDA) from 25 rats (14 sham, 11 6-OHDA)	
Trial-by-trial T_{O2} signals: post-reversal, early vs late sessions in *6-OHDA* rats	-9 LO electrodes from 9 rats; -10 NAc electrodes from 10 rats	**Exclusions for this analysis:** -Only electrodes from *6-OHDA* rats and implanted *ipsilateral* to the injection site that yielded useable signals for any of the early and late post-reversal training sessions were considered in this analysis.

Trial-by-trial T_{O2} signals: post-reversal, comparison *control* (ipsi- and contralateral sham and contralateral 6-OHDA) vs ipsilateral 6-OHDA	-35 LO electrodes (9 ipsi sham, 10 contra sham, 9 ipsi 6-OHDA, 7 contra 6-OHDA) from 22 rats (10 sham, 12 6-OHDA); -44 NAc electrodes (13 ipsi sham, 13 contra sham, 10 ipsi 6-OHDA, 8 contra 6-OHDA) from 25 rats (14 sham, 11 6-OHDA)	**Exclusions for this analysis:** -Only electrodes that yielded useable signals for any of the early and late post-reversal training sessions were considered in this analysis.
6-OHDA lesion effect on trial-by-trial T_{O2} current: correlation between T_{O2} amplitude in the lesion side OFC and NAc in sham and 6-OHDA lesioned subjects	All electrodes from 13 rats (9 sham, 4 6-OHDA)	**Exclusions for this analysis:** -Only data from subjects whose both NAc and both OFC electrodes yielded useable data were considered in this analysis.
6-OHDA lesion effect on trial-by-trial T_{O2} current: correlation between 6-OHDA lesion and T_{O2} amplitude changes	-20 LO electrodes (10 sham, 10 6-OHDA) from 20 rats (10 sham,10 6-OHDA); -14 NAc (14 sham, 10 6-OHDA) from 14 rats (14 sham, 10 6-OHDA)	**Exclusions for this analysis:** -Only data from electrodes implanted ipsilateral to the lesion (either sham or 6-OHDA) were considered in this analysis.
6-OHDA lesion effect on trial-by-trial T_{O2} current: anterior vs posterior OFC	21 LO electrodes (12 anterior (4 ipsi-, 3 contralateral to sham injection; 5 contralateral to 6-OHDA injection), 9 posterior (5 ipsi-, 2 contralateral to sham injection; 2 contralateral to 6-OHDA injection) from 16 rats (9 sham, 7 6-OHDA)	**Exclusions for this analysis:** -Only data from electrodes implanted in anterior or posterior LO from sham rats or implanted in the contralateral side to the injection in 6-OHDA rats were considered in this analysis.

Note. The experiment started with 36 rats, each implanted with 4 electrodes (bilaterally into NAc and LO). See text for additional details on exclusion criteria. 6-OHDA: amphetamine; LO: orbitofrontal cortex, lateral part; medial part; NAc: nucleus accumbens; OFC: orbitofrontal cortex

Histological assessment of 6-OHDA lesion

The extent of the 6-OHDA lesion was assessed at coronal slices from anterior, central and posterior levels of the NAc (+1.7, +1.2, and +0.7 mm from bregma, respectively). Mean

intensity measurements were taken from regions of interest (ROIs) drawn over the whole of the NAc shell and core, using Image J. Each measurement was normalised by subtracting the background mean intensity from white matter ROIs on the same hemisphere. TH labelling reduction intensity on the lesion side was calculated using the normalised mean intensity values:

$$\%\ reduction\ TH\ labelling\ intensity = \frac{nonlesion\ side - \ lesion\ side}{nonlesion\ side} \times 100$$

We then submitted these values of reduction in TH labelling intensity to a repeated-measures ANOVA, with AP coordinate (+1.7, +1.2, +0.7 mm) and NAc subdivision (core, shell) as within-subjects' factors, brain region (OFC, NAc) and group (sham, 6-OHDA) as between-subjects factor. Multiple-comparison post-hoc analyses were corrected using a Holm-Sidak correction.

To assess the effect of the NAc lesion in tyrosine hydroxylase-expressing (TH^+) neuronal projections to the dorsolateral striatum (DLS) (+1.7, +1.2, and +0.7 mm from bregma) and the lateral OFC (+3.7 mm from bregma), mean intensity measurements were taken from ROIs in these regions. The % reduction TH labelling intensity was then calculated as described for the analysis at the level of the NAc. As there was a general reduction in TH labelling in frontal cortex brain slices, OFC's mean intensity measurements were not normalised to white matter and were instead taken after thresholding the image background. We compared the values for the OFC of sham and 6-OHDA subjects using an independent samples t-test. We submitted the values for the DLS to a repeated-measures ANOVA, with AP coordinate (+1.7, +1.2, +0.7 mm) as within-

subjects' factors, brain region (OFC, NAc) and group (sham, 6-OHDA) as between-subjects factors.

To assess the effect of the NAc lesion in TH^+ neuronal projections to the contralateral NAc, we compared the mean intensity from ROIs in this region across sham and 6-OHDA lesion groups in the lesioned and non-lesioned hemispheres. To make measurements comparable across brains, we divided them by the background mean intensity from white matter ROIs on the same hemisphere.

Lesions were considered successful for subjects in which the average reduction in TH labelling in the coronal slice closest to the NAc electrode placement was greater than 50%. Data from 4 subjects where the 6-OHDA lesion was unsuccessful were excluded from all subsequent analyses. Precise quantification was not possible for 1 subject in the 6-OHDA group and 2 subjects in the sham group due to difficulty obtaining intact brain slices. However, as we observed no changes in any of the other animals in the sham group, these sham group subjects were included in behavioural and amperometric data analyses.

<u>6-OHDA lesion effect on behaviour</u>

We initially analysed the behaviour in all animals that had performed the task except for those in which the 6-OHDA lesion was deemed unsuccessful or could not be assessed, irrespective of whether their O_2 data met our inclusion criteria. We analysed the trial-by-trial lever choices during the 10 days of post-surgery behavioural training. To determine whether the unilateral NAc 6-OHDA lesion had an effect on behaviour, we submitted the % correct choices to a repeated-measures ANOVA with session (8 levels: sessions 1 to 4 and 6 to 9) as within-subjects factor and group (sham, 6-OHDA) as a between-subjects factor. Due to technical reasons, 1/3 of the rats did not have a 5^{th} behavioural session

before the reversal, and another 1/3 of the rats did not have a 5th behavioural session after the reversal. Because of this, we excluded data from sessions 5 and 10 from this analysis.

Additionally, we compared the % total correct responses before and after reversal with a repeated-measures ANOVA with week (pre-, post-reversal) as within-subjects factor and group (sham, 6-OHDA) as between-subjects factor. We also compared the number of trials to reach an 85% correct responses criterion on an exponential moving average (EMA) with a time constant of 6 trials (i.e., only performance on the last 6 trials was considered for this measure), using an independent-samples t-test. Finally, we compared the reaction times between the groups using an independent samples t-test. We repeated these analyses using only the subset of rats that provided useable T_{O2} signals and replicated all findings.

Behavioural modelling

In order to estimate animals' valuation of the two options, trial-by-trial lever choices were modelled using variations of the Rescorla-Wagner model (Rescorla & Wagner, 1972). We started with a model with two free parameters, a learning rate α and inverse temperature β. The value attributed to the lever chosen in each trial (t) was updated by:

(2.1) $$v_{chosen\ lever(t)} = v_{chosen\ lever\ (t-1)} + \alpha(R - v_{chosen\ lever\ (t-1)})$$

Where $v_{chosen\ \mathrm{lever}}$ is the value for the chosen lever, R is the receipt (= 1) or omission (= 0) of reward, and α is the reward learning rate.

During the 1st week of post-surgery training, the lever-reward probability contingencies are the same as in the last pre-surgery training session. Because of this, and since the animals had previously learned the contingencies, we used the true reward probabilities of the levers as the initial value estimates of each lever at the start of the 1st

session. At the beginning of each of the following sessions, the value estimates of each lever were the same as at the end of the preceding session.

In models described as 'single learning rate', all trials used the same learning rate α. To test whether rats were learning faster from rewarded trials than from non-rewarded trials (or vice-versa), we tested models using different learning rates for rewarded (*rew*) and non-rewarded (*no rew*) trials:

(2.2) $$v_{chosen\ lever(t_{rew})} = v_{chosen\ lever\ (t_{rew}-1)} + \alpha_{rew}(R - v_{chosen\ lever\ (t_{rew}-1)})$$

(2.3) $$v_{chosen\ lever(t_{no\ rew})} = v_{chosen\ lever\ (t_{no\ rew}-1)} + \alpha_{no\ rew}(R - v_{chosen\ lever\ (t_{no\ rew}-1)})$$

To test whether the subjects are indeed learning from the outcomes, we also compared our results with those from models in which thelearning rate was set to zero:

(2.4) $$v_{chosen\ lever(t)} = v_{chosen\ lever\ (t-1)}$$

To account for potential lever preference (bias), we tested models including a lever bias parameter. In these models, a bias term T was added to the value estimate of one of the levers (e.g., right lever) so that:

(2.5) $$v_{right\ lever(t)} = v_{right\ lever\ (t-1)} + \alpha\big(R - v_{right\ lever\ (t-1)}\big) + T$$

(2.6) $$v_{left\ lever\ (t)} = v_{left\ lever\ (t-1)} + \alpha\big(R - v_{left\ lever\ (t-1)}\big)$$

thus, decreasing or increasing the likelihood of choosing the biased lever.

To account for lever choice perseveration, i.e., repeating the previous choice regardless of the outcome, we tested models including a perseveration parameter. In these models, a perseveration term P, signed by the choice made on the previous trial, was added to the value estimate calculated in the new trial, so that:

(2.7) $$v_{right\ lever(t)} = v_{right\ lever\ (t-1)} + \alpha\big(R - v_{right\ lever\ (t-1)}\big) + P * a_{t-1}$$

(2.8) $$v_{left\ lever\ (t)} = v_{left\ lever\ (t-1)} + \alpha\big(R - v_{left\ lever\ (t-1)}\big)$$

where a is a constant of negative sign if the lever chosen in the previous trial was the left or positive sign if it was the right one.

To account for multi-trial lever choice perseveration, we tested models including a multi-trial perseveration parameter. In these models, a multi-trial perseveration term *M*, corresponding to the smoothed average of the choices in the preceding trials, was added to the value estimate calculated in the new trial, so that:

(2.9) $$M_t = M_{t-1} + \alpha_{multi\ persv}(a_{t-1} - M_{t-1})$$

(2.10) $$v_{right\ lever(t)} = v_{right\ lever\ (t-1)} + \alpha\big(R - v_{right\ lever\ (t-1)}\big) + M_t$$

(2.11) $$v_{left\ lever\ (t)} = v_{left\ lever\ (t-1)} + \alpha\big(R - v_{left\ lever\ (t-1)}\big)$$

where $\alpha_{multi\ persv}$ is the smoothing factor of the moving average.

Lever choices for each choice trial were predicted using a *softmax* function, which calculates the probability of choosing either option considering their value estimates. These were calculated for each trial *t* as follows:

(2.12) $$P_t(right\ lever) = \frac{1}{1+\exp\left(-\beta(v_{right\ lever} - v_{left\ lever})\right)}$$

(2.13) $$P_t(left\ lever) = 1 - P_t(right\ lever)$$

where β is the inverse temperature parameter.

I then calculated the likelihood of each observation as the Poisson probability to observe the real data given the predicted data. The log-likelihood of a model was calculated as the sum of log-likelihood of choice trials from all rats. We selected the best-fit parameters that maximised the log-likelihood for each training session for each rat within a constrained range (see priors in Table 2.6). To do so, we ran *optimise.minimize*, a function from the Python statistical package Spicy that minimises constrained functions, on the negative log-likelihood, using the Limited-memory Broyden–Fletcher–Goldfarb–Shanno optimisation algorithm ('L-BFGS-B'), with bounded initial parameters. The best-fit parameters out of 15 repeats were chosen. Forced trials were used to update the value of the chosen lever but not for assessing the likelihood of the models.

Table 2.6

Reinforcement learning model-free parameters ranges and distributions from which they are sampled.

Parameters	Range	Sample distribution
α, $\alpha_{multi\ persv}$	0 to 1	β probability distribution, α = 2 and β = 2.
T, P, M	unconstrained	Normal distribution, μ = 2 and σ = 5.
β	≥ 0	γ probability distribution, k = 2 and θ = 2.5.

To compare the models, we used the Bayesian information criterion (BIC), which penalises the likelihood of a model by the number of parameters and the natural logarithm of the number of data points. The model with the lowest BIC score was deemed to give a better fit of the data. We repeated these analyses using only the post-reversal training sessions or using only the subset of rats that provided useable T_{O2} signals and replicated all findings.

In vivo lesion validation

To generate an *in vivo* measure of the lesion on haemodynamic responses and to observe whether there were any changes over the course of the experiment, we investigated the effect of a systemic injection of the dopamine agonist apomorphine. Because of the relatively slow apomorphine pharmacokinetics (Gancher et al., 1996), amperometry data for the *in vivo* lesion validation was downsampled to 1 Hz and normalised to the average signal of a 60 s window prior to injection. To offset potential differences in channels' sensitivity, we standardised the data using z-scoring. Flat or noisy channels were excluded. To detect noisy channels, we calculated the root mean square (RMS) of the processed data (i.e., filtered, down-sampled and normalised, but not z-scored). Channels'

with an average log RMS value < -9.90 or > -4.60 were excluded for being flat or noisy, respectively (see Figure 2.6).

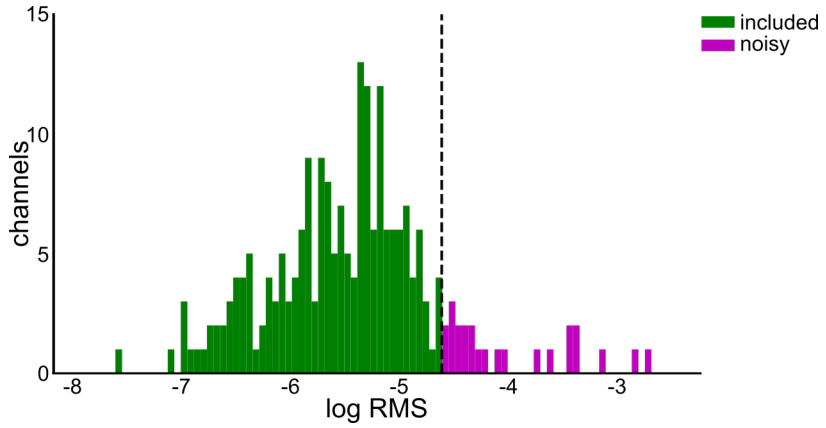


Figure 2.6 Criteria used to exclude noisy channels for *in vivo* validation analysis. Histogram of the root mean square (RMS) of the signals from each channel on sessions with systemic apomorphine administration. Channels with an RMS lower than -9.90 or higher than -4.60 were excluded for being flat or noisy, respectively.

The peak behavioural effect of apomorphine occurs at ~10 min following injection (Gancher et al., 1996). Thus, we calculated the area-under-the-curve (AUC) of T_{O2} signals during the first 15 min of amperometry recordings following injection. Mixed-model ANOVAs were conducted for NAc or OFC signals with factors of group (sham, 6-OHDA) and a repeated factor of apomorphine test (1, 2) and hemisphere (lesion, non-lesion side) using a compound symmetry covariance matrix (best model fit using Schwarz's BIC).

T_{O2} amperometry data analysis

Current (in nanoamperes) recordings from each channel were analysed separately. Linear interpolation was used to replace occasional missing data points, and an anti-aliasing finite impulse response (FIR) filter was applied. Then, the signal was down-sampled to 2 Hz using a decimation algorithm, keeping a single value from each 0.5 s non-overlapping window. Signals were low-pass filtered at 0.1 Hz using a biquad Butterworth

filter. All filtering was performed in both directions to eliminate phase-shifting of the data.

Data from electrodes with noisy, atypical or no meaningful signals were excluded. Our exclusion criteria included: (1) trials where there was an unphysiologically fast change in current (i.e., the difference in current between 2 consecutive data points, 0.5 s, > 1.25 nA) (see Figures 2.7-A and 2.8-A); and (2) trials with a signal variance < 0.2 or > 285 (see Figures 2.7-B and 2.8-B). If > 60% of the trials in a session met either criterion, then the whole session for that channel was excluded. Additionally, any atypical channels where the mean current was > 2.5 SD from the mean of all channels in that region or where the sign differed from the group (n = 2) were excluded.

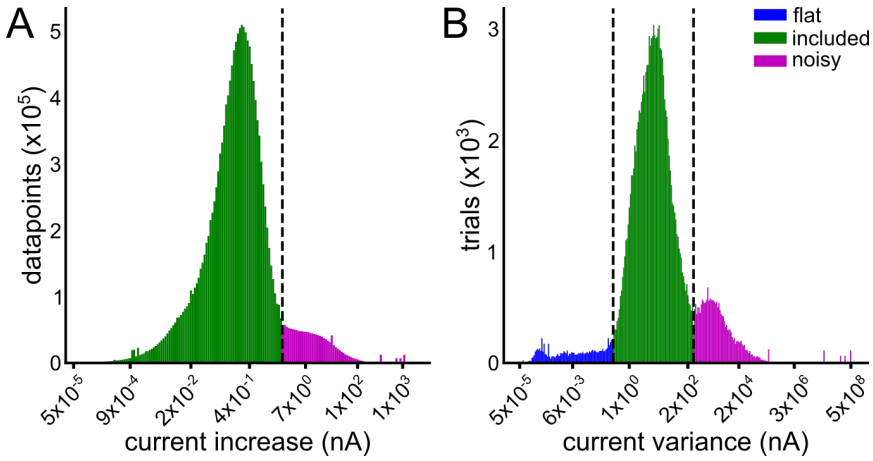


Figure 2.7 Criteria used to exclude T_{O2} signals from noisy or flat trials (Dopamine Lesion Experiment). (**A**) Histogram of the increase in current between every two consecutive data points (0.5 s apart). Trials in which there was at least one unphysiologically fast change in current (> 1.25 nA between two consecutive data points) were excluded for being noisy. (**B**) Histogram of the signal variance within each trial's relevant time window (10 s before trial initiation to 30 s after outcome). Trials with a variance lower than 0.2 or higher than 285 were excluded for being flat or noisy, respectively. T_{O2}: tissue oxygen amperometry.

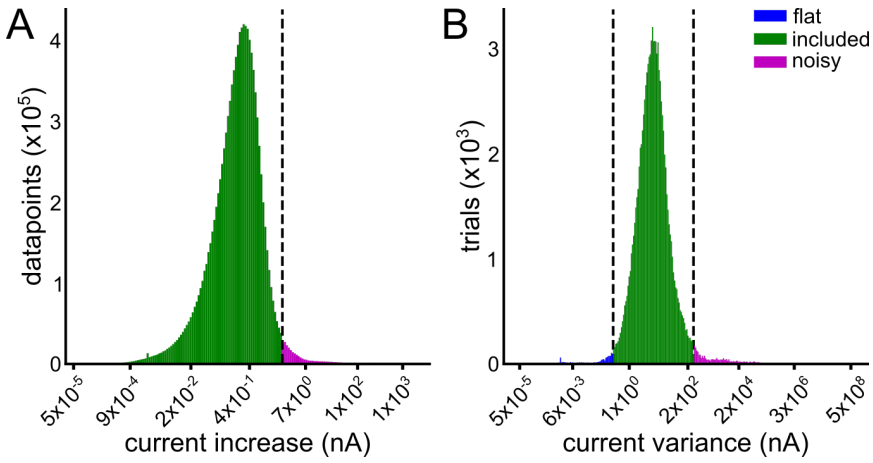


Figure 2.8 Criteria used to exclude T_{O2} signals from noisy or flat trials (Hyperdopamine Experiment). (**A**) Histogram of the increase in current between every two consecutive data points (0.5 s apart). Trials in which there was at least one unphysiologically fast change in current (> 1.25 nA between two consecutive data points) were excluded for being noisy. (**B**) Histogram of the signal variance within each trial's relevant time window (10 s before trial initiation to 30 s after outcome). Trials with a variance lower than 0.2 or higher than 285 were excluded for being flat or noisy, respectively. T_{O2}: tissue oxygen amperometry.

For each behavioural trial, a window of -10 to +45 s of O_2 data was aligned to relevant behavioural events. Because latency to lever press was variable, we time-warped the signals between lever extension and lever press to the mean reaction time across all trials. As amperometry measures relative changes in oxygen concentration and to compensate for different baselines between channels, signals were normalised by subtracting the average signal in the 1 s window before the onset of the behavioural event. Finally, as a *decrease* in the raw data signal reflects an *increase* in oxygen concentration, the sign of all signals was reversed (i.e., positive became negative, negative became positive).

To understand the relationship between the T_{O2} and the subjects' behaviour, we performed two sets of complementary analyses: (i) analyses of the signals dynamics using the average T_{O2} amplitude in NAc and OFC in each trial (taking a 45 s window from cue onset), before or after reversal of the lever-reward probability contingencies; and (ii)

regression of the same signals against estimates from our computational model using the model with the lowest BIC score).

To investigate the dynamics of the T_{O2} signals recorded from NAc or OFC in response to rewarded and unrewarded outcomes after choosing the high- or the low-reward probability levers during the first post-surgery week of behavioural testing, i.e., before reversal of the lever reward contingencies, in sham animals, we compared the T_{O2} amplitude in response to specific behavioural time points by computing the maximum range amplitude (i.e., max – min signal values), zeroed at the time point of interest (i.e., cue onset), for each trial type (i.e., rewarded or unrewarded trials after the subjective high or low high reward probability levers were chosen). Subjective high or low reward probability levers were defined as the preferred or otherwise levers by each rat at each trial using the model-derived trial-by-trial reward estimates. Because we also wanted to test whether there was an effect of lateralisation or of the vehicle injection, we subjected the T_{O2} amplitude values to a 2-way repeated-measures ANOVA, with outcome (rewarded, unrewarded) and lever reward probability (high, low) as within-subjects' factors, brain region (OFC, NAc) and hemisphere (lesion, non-lesion side) as between-subjects factors.

We used similar approaches to investigate the effect of the 6-OHDA lesion on OFC and NAc responses and to compare the effects of the 6-OHDA and vehicle (sham group) injections. In the latter analysis, we used group (sham, 6-OHDA lesion) instead of hemisphere as a between-subjects factor.

In the post-reversal week, we wanted to investigate whether OFC and NAc responses changed as the rats learned the new reward contingencies. Thus, we compared signals from sessions right after the reversal, as the animals are learning new reward contingencies with those from sessions later in training when most animals had learned

the new contingencies. We did this by submitting the T_{O2} amplitude values to a 3-way repeated-measures ANOVA, with outcome (rewarded, unrewarded), lever reward probability (high, low) and sessions period (early and late, i.e., sessions 1-2 and 3-4 post-reversal) as within-subjects' factors, and brain region (OFC, NAc) and hemisphere (lesion, non-lesion side) as between-subjects' factors. We did not use the last fifth session post-reversal in these analyses as, due to technical reasons, 1/3 of the subjects was unavailable for this day. We used a similar approach to investigate the same early vs late sessions effect in the animals that receive the 6-OHDA injection into the NAc. In this analysis, we used group (sham, 6-OHDA lesion) instead of hemisphere as a between-subjects factor.

To investigate whether the 6-OHDA injection affected the responses from the contralateral hemisphere, we compared these signals with those from the contralateral hemisphere to the vehicle injection (sham group). We did this by submitting the T_{O2} amplitude values to a 2-way repeated-measures ANOVA, with outcome (rewarded, unrewarded) and lever reward probability (high, low) as within-subjects' factors, brain region (OFC, NAc) and group (sham, 6-OHDA) as between-subjects factors.

We observed no differences in the signals between the ipsi- and contralateral sides of sham lesioned subjects nor between contralateral sides of sham or 6-OHDA lesioned subjects. Therefore, we repeated the analyses in which we compared T_{O2} signals from sham and 6-OHDA groups by comparing the signals recorded from electrodes placed in the ipsilateral side of 6-OHDA lesioned subjects with those from the ipsi- and contralateral sides of sham lesioned subjects as well as from the contralateral side of 6-OHDA lesioned subjects (as opposed to signals only from the hemisphere ipsilateral to the vehicle injection), replicating all findings.

To investigate potential differences in the trial features encoded by anterior and posterior portions of the OFC, we classified electrodes placed in the coronal section from

+4.7 to + 4.2 mm from bregma as anterior and those contained in the coronal sections from +3.2 to +2.2 mm as posterior (Paxinos & Watson, 1998), based on (Panayi & Killcross, 2018). We then subjected the T_{O2} signals to a 3-way repeated-measures ANOVA, with outcome (rewarded, unrewarded), lever reward probability (high, low) and week (pre-, post-reversal) as within-subjects' factors, electrode placement (anterior, posterior) as between-subjects factors. In this analysis, we used the signals from OFC electrodes from ipsi- and contralateral sides of sham lesioned subjects as well as from the contralateral side of 6-OHDA lesioned subjects.

General linear models: Regression coefficients were estimated for each animal at each time point (i.e., every 0.5 s) within the same time window used for the analyses described above. We fitted a generalised linear model (GLM) with a constant term and a set of regressors. These included outcome (rewarded = 1, unrewarded = -1) and lever value estimate obtained from the RL model. We also included other regressors, including: (i) ITI at the end of the previous trial, (ii) trial type (choice = 1 or forced = -1), (iii) outcome on the previous trial (rewarded = 1, unrewarded = -1), and (iv) previous ITI x previous outcome. In analyses of early vs late sessions in the post-reversal week, we also included a regressor for early/late post-reversal (early = -1; late = 1) and interaction terms between early/late post-reversal and outcome or early/late post-reversal and reward estimate.

All regressors were mean-centred and continuous regressors normalised. Regression coefficients in each animal were averaged, and significance was assessed by comparing these values against a distribution of mean zero, using the function *stats.ttest_1samp* from the Python statistical package *SciPy* (Virtanen et al., 2020). The *p* values were corrected for false discovery rate over the ensemble of timepoints for each

regression variable separately, using the procedure of Benjamini and Hochberg (Benjamini & Hochberg, 1995) using the function *stats.multitest.multipletests* from the Python statistical package *statsmodels* (Seabold & Perktold, 2010). Only *p* values from one in every 8 timepoints (i.e., every 4 s) and corresponding to timepoints post-cue were considered for the multiple-comparisons correction. When testing the regression variable of outcome, only *p* values corresponding to timepoints post-outcome delivery were considered.

To test for differences between the ipsilateral and contralateral hemispheres to the sham injection or for differences between lesion and control groups, we used a 2nd level analysis approach. In this analysis, we fitted a GLM to the regression coefficients of the 1st GLM, using as regressors the hemisphere (lesion side hemisphere = 1, non-lesion side hemisphere = -1) or lesion type (6-OHDA = 1, sham = -1). We used a similar approach to test for differences in signals from electrodes placed at anterior or posterior portions of the OFC (anterior = -1, posterior 1) and the interaction between electrode placement and week (pre-reversal = -1, post-reversal = 1). Significance for each regression coefficient was assessed by comparing these values with the variance of the residuals. The *p* values were corrected for false discovery rate using the same approach used for the 1st level GLM.

<u>6-OHDA lesion effect on trial-by-trial T_{O2} current in OFC and NAc</u>

We assessed whether there was a relationship between the change in T_{O2} current amplitude in the OFC and NAc, and whether these changes are influenced by the NAc dopamine lesion. We examined this in two ways.

First, we compared the trial-by-trial T_{O2} current change amplitude for each trial in each channel in subjects for which all electrodes (i.e., left and right OFC, left and right

NAc) provided data. To do this, we subtracted the maximum and minimum values within each trial time window (i.e., from cue onset to 35 s after reward/no reward delivery). Then we normalised the values obtained on the lesion side by subtracting the amplitude measured on the same trial on the non-lesion side:

$$TO2\ current\ change\ amplitude$$
$$= (\max - \min current\ on\ lesion\ side) - (\max - \min current\ on nonlesion\ side)$$

We then used these values to compute a Pearson's correlation coefficient for the T_{O2} current in the OFC with that in NAc on the same trial.

We also compared the mean trial-by-trial amplitude in each region for each electrode on the lesion side and related it to the extent of lesioning in the NAc following histological assessment, i.e., % TH labelling intensity reduction on lesion side NAc. For this analysis, we used data from all electrodes with useable data implanted on the hemisphere ipsilateral to the sham or 6-OHDA lesions. We then tested for differences between the two lesion groups in each region by applying an independent samples t-test and a k-means clustering algorithm to the mean amplitude values of each electrode.

Hyperdopamine Experiment (data analysis)

The table below summarises the sample sizes and exclusions for the data analyses of the hyperdopamine experiment.

Table 2.7

Summary of sample sizes and exclusions for data analyses of the hyperdopamine experiment.

Analysis	*n*	Reasons for exclusions
Electrode placements	22 rats (11 veh, 11 amph)	**Exclusions for this analysis:** -15 rats were excluded because they died during surgery or during recovery from surgery (*n* = 11), or their implants came off post-surgery (*n* = 4); **Exclusions for all T_{O2} signals analyses:** -13 MO, 6 LO, and 4 NAc electrodes were excluded because their track could not be visualised; -5 MO, 8 LO, and 15 NAc electrodes were excluded due to incorrect anatomical placement.
Behavioural performance and modelling	17 rats (9 veh, 8 amph)	**Exclusions for this analysis and all T_{O2} signals analyses:** -1 amph rat was excluded for having trained on the wrong reward contingencies; -2 rats (1 amph, 1 veh) were excluded due to high proportion of missed trials; -2 amph rats were excluded for failing to reach chance level performance post-reversal on > 2 blocks.
T_{O2} signals during baseline (two first weeks post-surgery)	-3 MO electrodes (2 veh, 1 amph) from 3 rats (2 veh, 1 amph); -6 LO electrodes (3 veh, 3 amph) from 6 rats (3 veh, 3 amph); -23 NAc electrodes (14 veh, 12 amph) from 13 rats (7 veh, 7 amph)	
Effect of amphetamine on T_{O2} signals	-15 NAc electrodes (5 veh, 10 amph) from 8 rats (4 veh, 4 amph)	**Exclusions for this analysis:** -Only NAc electrodes that yielded useable signals for all the key training sessions were considered in this analysis.

Note. The experiment started with 37 rats, each implanted with 4 electrodes (bilaterally into the NAc and unilaterally into lateral and medial OFC). See text for additional details on exclusion

criteria. amph: amphetamine; veh: vehicle; LO: orbitofrontal cortex, lateral part; MO: orbitofrontal cortex, medial part; NAc: nucleus accumbens; OFC: orbitofrontal cortex

Effect of amphetamine administration on behaviour

Data from 15 rats that did not complete all post-surgery training sessions, either because they died during surgery or recovery from surgery (n = 11) or their implants came off post-surgery (n = 4), were excluded. Five additional rats were excluded: 1 rat for having trained on the wrong reward contingencies in one of the post-surgery training sessions; 2 rats for omitting > 50% trials in > 1/3 of the post-surgery behavioural sessions; 2 rats for failing to reach chance level performance (i.e., 50% correct responses) prior to reversal on > 2 blocks during the recordings (note that this was not an effect of the amphetamine administration). All remaining rats were included in the behavioural analyses, irrespective of whether their O_2 data met our inclusion criteria.

To assess the effect of acute and sensitized amphetamine injections, we compared the rats' behaviour on key training sessions after either injection and during baseline 1, i.e., the first week of amperometry recordings, prior to any amphetamine injection and in which rats encountered the same reward contingencies but did not receive an injection. First, to determine whether either amphetamine injection affected one measure of reward-focused responding, we submitted the total number of reward magazine head entries on the first post-reversal session at baseline 1, acute, and sensitisation weeks to a repeated-measures ANOVA with drug challenge (baseline 1, acute, sensitisation) as within-subjects factor and group (veh, amph) as between-subjects factor. We performed a similar analysis for assessing the effects of amphetamine injection on reaction time (i.e., latency to lever press).

To assess whether amphetamine impaired the rats' ability to learn the reward contingencies after reversal, we submitted the % correct choices on key behavioural

sessions to a 2-way repeated-measures ANOVA with drug challenge (baseline 1, acute, sensitisation) and session after reversal (first, second, or last session after reversal) as within-subjects factors and group (veh, amph) as between-subjects factor. Multiple-comparison post-hoc analyses were corrected using a Holm-Sidak correction.

To further explore a potential effect of amphetamine on the animals ability to learn the new reward contingencies after reversal within each relevant session, we submitted the % correct responses across every 28-choice trials block (i.e., first or half of each training session) on the first and second post-reversal training sessions at baseline 1, acute, or sensitised administration weeks. Additionally, we analysed the proportion of win-stay and lose-stay trials, i.e., choice trials in which the animals chose the same lever as in the last non-missed rewarded or unrewarded trial, on the first and second training post-reversal training sessions at baseline 1, acute, and sensitisation weeks. We submitted these measures to a 3-way repeated measures ANOVA with drug challenge (baseline 1, acute, sensitisation) and session after reversal (first, second, or last session after reversal), and trial stay type (win-stay, lose-stay) as within-subjects factors and group (veh, amph) as between-subjects factor.

Behavioural modelling

In order to estimate animals' valuation of the two options, trial-by-trial lever choices were modelled using variations of the Rescorla-Wagner model (Rescorla & Wagner, 1972) as described for the analysis of the dopamine lesion experiment.

T_{O2} amperometry data analysis

Current (in nanoamperes) recordings from each channel were processed as described for the dopamine lesion experiment.

To investigate the trial features encoded in each brain region, we compared the amplitude (i.e., maximum range) of the average responses to specific behavioural time points computed by using the processed oxygen signal, zeroed at the time point of interest (i.e., cue onset), for each trial type (i.e., rewarded or unrewarded trials after the subjective high or low high reward probability levers were chosen). We then subjected T_{O2} amplitude values to a 2-way repeated-measures ANOVA, with outcome (rewarded, unrewarded) and lever reward probability (high, low) as within-subjects' factors, brain region (MO and LO; for OFC analyses only) and group (veh, amph) as between-subjects factors. Due to the large difference in sample size for OFC and NAc electrodes, we ran separate analyses for the two brain regions. Where appropriate, multiple-comparison post-hoc analyses were corrected using a Holm-Sidak correction.

We then investigated the effects of single and repeated doses of systemic amphetamine administration only on haemodynamic signals of the NAc (this analysis was not possible for the OFC due to the small sample size). For this analysis, we focused on electrodes that supplied data throughout all sessions of interest (i.e., first and second sessions after reversal at baseline 2, acute, and sensitisation), resulting in n = 15 (10 amph, 5 veh). We submitted the average amplitude for each trial type (see above) at the first and second post-injection sessions to a 3-way repeated-measures ANOVA with drug challenge/week (baseline, acute, sensitisation), post-reversal session (1, 2), outcome (rewarded, unrewarded) and lever reward probability (high, low) as within-subjects' factors. Because of the large sample size difference between amph and veh groups, we performed this analysis separately for each group.

General linear models: To understand the relationship between the T_{O2} and the subjects' behaviour, we regressed the same signals used in the analyses described above against

estimates from our computational behavioural model using the model with the lowest BIC score, as described for the analysis of the dopamine lesion experiment.

To test for differences in signals from electrodes placed at medial or lateral portions of the OFC (medial = -1, posterior 1), we applied a 2nd level analysis on the GLM regression coefficients from either OFC portion.

2.3 Development of a novel behavioural paradigm for studying non-spatial cue-based serial learning.

Chapter 6 of this thesis describes the development of a novel behavioural paradigm for investigating non-spatial cue-driven decision-making in an operant context, with the intention of being able to investigate generalisation of rules and neuronal representations of cognitive maps (Behrens et al., 2018).

In this generalisation task, animals learn the reward contingencies associated with auditory cues and use these to make choices between a left and a right reward port. After a performance criterion is reached, the auditory cues are replaced by a new set assigned with the same reward contingencies. We wanted to investigate whether rats could acquire a learning set (Harlow, 1948), i.e., there are always a particular set of cues, each associated with a particular reward contingency, and generalise it to new sets of cues.

The development of this behavioural paradigm involved training several cohorts of animals for optimisation of stimuli and other task parameters. Here, we describe in detail the optimised task and training protocol, and in Chapter 6, we explain the rationale for the final task design and present preliminary results and behavioural measures that helped guide its development.

2.3.1 Subjects

A total of 22 male Lister Hooded rats (Envigo, UK), ~275-300 g in weight at the start of the experiment (8-12 weeks old), were used in this study. Cohort 1 (*n* = 4) was used for initial tests of the behavioural boxes and optimising training protocols. Cohort 2 (*n* = 9) was trained on a 3-cue generalisation task. Finally, the third cohort (*n* = 9) was trained on a 2-cue generalisation task with a proportion of control trials (see *Control trials* in *Behavioural experiments* for details). Animals were housed in groups of 3 in polycarbonate cages, with *ad libitum* access to water. Access to food was restricted so that the rats' weight was kept at ~85-90% of their free-feeding body weight. The housing room was maintained at a room temperature of 23 ± 1°C, humidity of 40 ± 10% and on 12-hour light/dark cycles beginning at 7 A.M., with lights during the day.

2.3.2 Behavioural experiments

Apparatus

Training was carried out in custom-built operant chambers (length: 35, width: 30, height: 30 cm) housed within a sound-attenuating box ventilated with a fan and fitted with a houselight. Each chamber was fitted with three nose ports and three speakers for the presentation of auditory stimuli, all located on the same wall (see Figure 2.9–A). Stimulus presentation and responses were controlled and monitored via a custom-made circuit board microcontroller (MicroPython) which communicated via USB with a PC running a custom-written task using pyControl (Akam et al., 2021). Auditory cues will be described in detail below. All behavioural sessions were video-recorded using night-vision cameras (bird box camera, Green Feathers) and a video card (GV800B-16, GeoVision). Liquid sucrose rewards were delivered with a pinch solenoid valve.

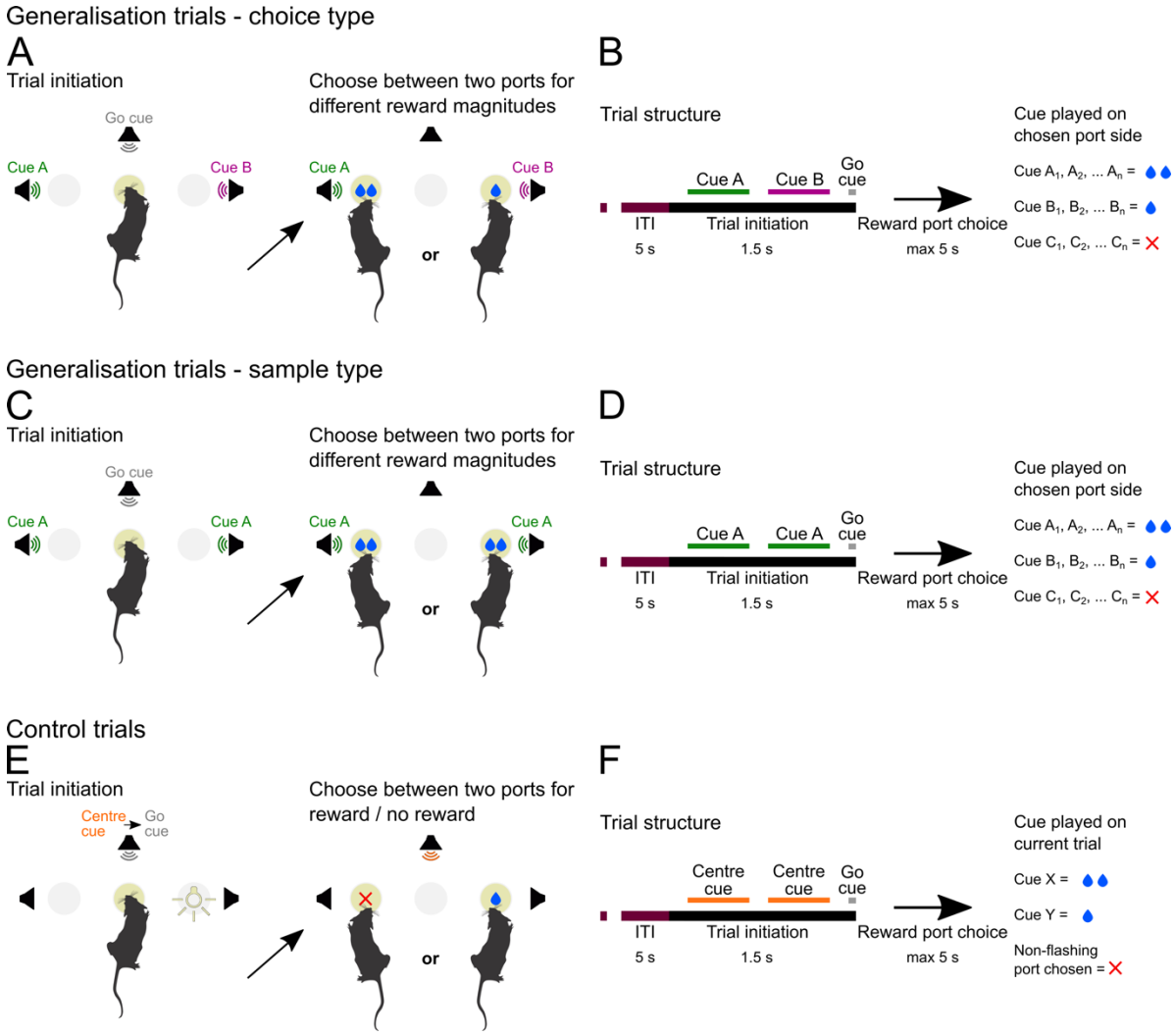


Figure 2.9 Behavioural paradigm. (**A**) In generalisation choice trials, rats poked in a trial initiation port at the centre of the box while two auditory cues played, one from each side speaker, followed by a Go Cue. Then, rats chose between two side ports for receiving rewards of different magnitudes. Two of a set of 2 or 3 auditory cues (in different task versions) played in each choice trial. (**B**) Times of each step of the choice trials. After a performance criterion was reached (80% correct responses in EMA with a time constant of 30 trials), the 3-cue set was replaced by a new one with the same transitive relationship. (**C**) In sample trials, rats poked in a trial initiation port at the centre of the box while one auditory cue played, once from each side speaker, followed by a Go Cue. Then, rats chose between two side ports for receiving a reward. The reward would have the same magnitude regardless of the chosen side port. (**D**) Times of each step of sample trials. (**E**) In control trials, rats poked in the initiation port at the centre of the box while an auditory cue played from the centre speaker followed by a Go Cue. During trial initiation, one of the reward ports light flashed and, if the rat chose this port, it would receive a reward whose magnitude depended on the identity of the auditory cue played; otherwise, it would receive no reward. (**F**) Times of each step of the trial structure of control trials. EMA: exponential moving average.

Generalisation trials

The possibility to initiate a new trial was signalled with illumination of the centre port. To initiate a trial, rats were required to poke and stay in the centre port. After 150 ms, two auditory cues played consecutively, one from each side speaker (500 ms each, with a 150 ms break in between, side and order random; Figure 2.9-A, B). After an additional 150 ms, a white noise burst (50 ms) from the centre speaker signalled the end of holding time. Aborted trials caused by premature centre port exit were signalled by the houselight flashing (10 s). After the holding period, reward side ports illuminated and a nose poke response in either caused delivery of the outcome associated with the auditory cue that had been presented at that location (0.1 ml boluses of 10% liquid sucrose). The cue that had played on the side of the chosen reward port would replay during reward/no reward delivery. If no response was recorded within 5 s, an aborted trial would be recorded and signalled with flashing of the houselight (10 s). After an incorrect choice, i.e., choice of the lowest value option available, a time penalty of 5 s was added to the ITI (5 s). 20% of the generalisation trials were sample trials. In these trials, the same sound would play on one side speaker and then on the other, and either port could be chosen for collecting a reward (Figure 2.9-C, D).

Auditory cue sets were replaced after a behavioural criterion was reached (see *Behavioural Training*). Each training session would terminate after animals had either acquired 475 rewards (~40 mL) or 120 min had elapsed. See Figure 2.9-B for a schematic summary of the times of the events which each trial involves.

Control trials

Trial initiation and the holding time structure were as described for the generalisation trials, with the exception that here the same auditory cue played twice from the *centre*

speaker instead of from one of the side speakers (Figure 2.9-E, F). The sound played was one from a set of two, associated with either a large or small reward that the rat would receive if the correct reward port was subsequently chosen. The correct reward port was cued by its light flashing during holding time. After the holding period, both reward ports now became illuminated during the choice period. The sound would again replay during reward/no reward delivery. 15% of the trials in a session were of the control type; the remaining trials were of the generalisation type.

Auditory cues and reward contingencies

The two control cues were selected from pulsed, stepped, or continuous sinusoidal waves of different frequencies (e.g., pulsed 11 kHz pure tone with pulses at 8 Hz and stepped tones from 3 to 8 kHz with steps at 10 Hz) and remained fixed for an animal throughout training. In the generalisation trials, the cues in each block were randomly drawn from a library of 100 synthesised sounds (based on the gaussian sound synthesis method used by McDermott et al., 2011) and from a library of 4876 bird sounds (*Cornell Guide to Bird Sounds: Essential Set for North America*) and replaced after a behavioural criterion was reached (see *Behavioural Training*). The synthesised sounds were produced using frequencies between 1 and 14 kHz, taking into account the rat audiogram (Heffner & Heffner, 2007). When training Cohort 3, the amplitude of the sounds was normalised so that they had the same root mean square values (i.e., mean power). This allowed correcting for differences in the volume at which sounds were played on the speakers of the behavioural boxes.

To avoid cue presentation order or side bias on the generalisation trials, a bias correction was automatically implemented if one side or cue presentation order was chosen in more than 80% of trials (in an exponential moving average (EMA) with a time

constant of 20 trials, i.e., only performance on the last 20 trials was considered for this measure). For instance, when bias correction was on, if the EMA of choosing the right port was 0.75, then the probability of the larger reward cue playing on the right side on the subsequent trial would be 0.25. This bias correction would turn off when the side/order bias dropped to 60%.

The volume of the auditory cues was set to be perceived at ~80 dB at all times. This means that, for instance, when auditory cues were replayed from side speakers at reward delivery, the volume setting was lower than during trial initiation when the rat was at the centre port.

Rewards were delivered in boluses of 0.1 ml: with large, small and no reward value cues associated with 0.2, 0.1 and 0 ml, respectively.

Behavioural training

Before training commenced, rats were food-restricted and maintained at > 85% of their free-food weights. Each rat was exposed to 5 mL of 20% liquid sucrose on each of the two days before the first training session.

Stage 1 – Reward collection

Initially, rats learned to poke at the illuminated side port to receive a reward. Trials were automatically initiated after a short ITI (3 s) and cued with a white noise burst (50 ms; Go Cue). The randomly assigned, rewarding side port was then cued with a light inside the port, which would turn off after reward delivery. Animals transitioned to the next training stage when two criteria were met: > 250 rewards were received and the proportion of nose poke entries into the illuminated port was > 90%.

Stage 2 – Collect reward at the available reward port after trial initiation

Next, the trial initiation (centre) port was introduced, and animals learned the correct sequence of actions to obtain a reward: poke at the centre port and then at the side port that illuminated. The trial initiation port was illuminated when a trial could be initiated, and its light would turn off after nose poke. Rats then had 10 s after trial initiation to make a response. If no poke was made on a side port during this time, an aborted trial would be recorded, and the houselight would flash for the duration of the abort trial time off (10 s).

A critical step in the behavioural training was teaching the animals to hold at the centre nose poke during a long enough period for playing two sounds on the side speakers and another at the centre speaker. The centre poke holding time necessary for trial initiation increased or decreased by 5 ms after each successful or unsuccessful (i.e., early exit; abort trial) trial initiation. This continued until the rats successfully completed 200 trials with 1000 ms holding time. Aborted trials were signalled with flashing houselight during a time off period (10 s). The Go Cue (50 ms white noise) played from the centre speaker at the end of the holding time.

This training stage usually took several sessions. At the start of a session, the initial and the minimum holding times for that session were set to 80 and 60%, respectively, of the final holding time in the previous session (the minimum holding time was the lowest value the holding time would go even if rats continued to fail to wait for the full holding period).

The reward port that would be available during the current trial would light up during holding time. However, once the holding time was > 600 ms, the reward port would flash during holding time and then switch to steady illumination after the Go Cue. To avoid penalising the rats for small movements during the holding period (e.g., a twist

of the head), there was a nose poke tolerance of 100 ms before an aborted trial was triggered.

Animals automatically transitioned to Stage 3 when the following criteria were met: (i) the holding time for trial initiation was > 1000 ms; (ii) > 500 rewards were obtained; (iii) > 50 rewards were obtained in the current session.

Stage 3 – Collect reward at the correct reward port after trial initiation

In the next stage, *both* rewards ports would light up after trial initiation, and animals learned to choose the port whose light was flashing during the holding period. In this stage, the incorrect choice time out was increased to 5 s and animals kept being trained on increasingly longer holding times as described in Stage 2. Animals automatically transitioned to the next stage when all of the following criteria were met: (i) the holding time for trial initiation was 1500 ms (target holding time); (ii) reward port correct choices rate was > 85% (in EMA, time constant = 50 trials); (iii) > 300 rewards were obtained after reaching correct choices rate threshold with this count would restart if correct choices rate dropped below 50%; (iv) > 50 rewards were obtained in the current session; (v) correct choices rate > 80% at the time of stage transition. Note that in the 3-cue generalisation task version, rats transitioned directly to Stage 5.

Stage 4 – Cued reward magnitude (control cues)

Rats were next trained on the full control trials structure. This was the same as Stage 3 except now, 150 ms after the start of holding time, either the small or large reward control auditory cue played at the centre speaker (see Figure 2.9-F). As in the previous stage, the light of the correct reward port would flash during holding time, and both ports would become illuminated during the choice period. If the rat chose correctly, the sound

would replay during reward delivery. When 500 rewards were obtained in this training stage, and all other criteria indicated in stage 3 were met, the rats automatically transitioned to Stage 5.

Stage 5 – Side cued reward magnitude (generalisation cues) with visual cueing

In this stage, control trials were limited to 15% of the total. On the other 85% of trials, one of the 3 generalisation cue sounds would play from one of the side speakers (500 ms duration, order and left/right side random; same trial initiation structure as in sample trials as schematised in Figure 2.9-D) and the light of the reward port on the same side of the box would flash to indicate the port would be rewarded on that trial. At the end of the holding time, both reward ports would light up and the rat was required to choose the one on the side where the sound played to receive a small reward (i.e., 0.1 ml; see *Auditory cues and reward contingencies* for details). Again, the sound would replay during reward delivery. After a rat had obtained > 100 rewards and all other criteria set out in Stage 3 were met, the rats automatically transitioned to Stage 6.

Stage 6 – Side cued reward magnitude (generalisation cues) without visual cueing

Here, the side port where a reward would be available would not flash during the holding period. Thus, the rats had to rely on tracking the location of the sound (left/right) to receive a reward on these trials. All other aspects of the task were the same as in Stage 5. After a rat had obtained > 1000 rewards and all other criteria set out in Stage 3 were met, the rats automatically transitioned to Stage 7.

Stage 7 – Choice trials

In this stage, rats were required to keep track of the sounds reward contingencies to maximise reward. On generalisation trials, two different sounds would play consecutively during the holding period, one from each side speaker (see Figure 2.9-A, B). The two sounds had different reward contingencies, either one associated with reward and the other with no reward for Cohort 2 trained with 3 cues or one with a small and the other with a large reward for Cohort 3 trained with 2 cues. After a rat had obtained > 500 rewards and all other criteria set out in stage 3 were met, the rats automatically transitioned to the next stage (Stage 8 in the 3-cue version for Cohort 2 or directly to Stage 9 in the 2-cue version for Cohort 3).

Stage 8 – full task structure, free choice trials only

For rats trained on the 3-cue version of the generalisation task: in this stage, three different sounds played from the side speakers. This means that in each trial, two sounds play consecutively, one from each side speaker, and these two sounds were drawn from a set of three. Two of the sounds were those used during training on stage 7. The new sound would have an associated reward that was double the size of the previously rewarded sound. Thus, the three sounds would have reward contingencies of large, small, and no reward. Rats learned to choose the side port with the cue with higher reward volume in each trial. When 300 rewards were obtained in this training stage and all other criteria indicated in stage 3 were met, the rats automatically transitioned to the next stage.

Stage 9 – full task structure, free and choice trials

In this stage, there were sample and choice generalisation trials. On sample trials, the same sound would play on one side speaker and then on the other, and either port could be chosen for collecting a reward (see Figure 2.9-C, D). About 20% of the generalisation trials were sample trials. In this stage, the choice time limit was halved to 5 s and ITI increased to 5 s.

It took on average ~30 behavioural sessions for animals to learn the entire structure of the task, i.e., reach Stage 9 of behavioural training.

Introducing new cue sets

When rats were performing the full task, the generalisation cue sets were replaced with new ones after either (i) rats reached a criterion of 80% "correct" choices (i.e., choices of the highest value option available in an EMA, time constant = 30 trials) plus an additional 30 rewards, or (ii) after 600 rewards with the same cue set even if the rats had not reached this performance criterion.

2.3.3 Data analysis

All data analyses were performed using custom-written Python scripts (Python Software Foundation, 2013).

Performance on generalisation trials

To assess whether rats were getting faster at learning the reward contingencies, we applied a linear regression to the number of trials each rat took to reach the learning criterion for each cue set experienced. Significance was assessed by comparing the slopes of the best fit line for each rat against a distribution of mean zero, using the function

stats.ttest_1samp from the Python statistical package *SciPy* (Virtanen et al., 2020). We removed from this analysis data from the cue set used for learning the initial task. On the 3-cue version of the task, the first two sets of new cues were trained with a larger maximum number of trials limit before triggering a forced transition than subsequent sets (i.e., 3000 on the first two new cue sets and 1500 on subsequent cue sets). Thus, to avoid biasing the analysis to show a reduction in the number of trials taken to trigger a transition after the first cue set, which, because of this difference could have not reflected the animals' learning of the task, we set the number of trials taken to transition on those first two cue sets as 1500 for those rats that took longer than this limit.

To investigate which task components rats used for guiding their trial-by-trial choices, we fitted a GLM with a constant term and a set of regressors to the rats' choices. Thus, to predict the likelihood of the rats choosing a side port (left), we included as regressors the value of the moving average bias towards that port, the order in which the cue on that port played (i.e., first = -1, second = 1), and the reward associated with the cue played on the speaker on that side of the box (smaller reward = -1, larger reward = 1). When analysing the data from the rats trained on the 2-cue set task, we also applied this analysis to the first set of cues. In this analysis, we used only the last 50 choice trials before triggering a transition; otherwise, we used all trials.

To investigate whether rats learned the reward contingencies (2-cue task version) and established a rank order of each cue on a cue set (3-cue version) we also examined response latencies from the end of the holding period to reward port choice on all sample trials. When analysing data from the 3-cue task, we compared reaction times by applying to the data a one-way ANOVA and follow-up pairwise comparisons to investigate any significant effects. When analysing data from the 2-cue task, we used a paired samples t-test for comparing the reaction times. Additionally, on the analysis of 3-cue task version,

we compared the proportion of correct choices on each cue pair (i.e., large vs small, large vs no reward, small vs no reward) after rats had reached performance criteria for triggering a transition, by submitting these values to a repeated measures ANOVA. To assess whether rats were responding above chance level on each of these cue pairs we also compared the proportion of correct responses on each pair against a distribution of mean 50%. One rat was excluded from this last analysis because it did not reach performance criteria for triggering a transition with any of the cue sets it experienced.

To assess whether the sound features for the cues–specifically the mean power for all or just the high frequencies (i.e., > 6 kHz)– affected choice performance for the best performing rat on the 2-cue version, we compared the mean power of each frequency range (all or high frequencies) between cues associated with a small or large reward from pairs that appeared to be discriminable or indiscriminable by this rat using a paired samples t-test.

Performance on control trials

We also wanted to investigate whether the rats learned the reward contingencies on the control light-cued task. Due to the high proportion of aborted control trials, it was not possible to compare the reaction times between the two control cues, so, instead, we compared the proportion of aborted trials in which the cues associated with a small or a large reward played using paired-samples t-tests.

3. Anatomical characterisation of OFC inputs along its axes of organisation

3.1 Introduction

Over recent years there has been an increasing understanding that the rodent orbitofrontal cortex (OFC) is not functionally homogenous, particularly along its medial-lateral axis (see Barreiros, Ishii, et al., 2021; Bradfield & Hart, 2020; Izquierdo, 2017 for reviews). A recent review (Izquierdo, 2017) examined the relationship between the functions reported in rat OFC studies and the anatomical placement of the recording or manipulation sites. This revealed that functional heterogeneity can be mapped, to a certain degree, onto the divisions established by classical OFC parcellation methods. These classical parcellation studies define OFC subregions predominantly along the medial-lateral axis, including medial (MO), ventral (VO), ventrolateral (VLO), lateral (LO), dorsolateral (DLO) and agranular insular (AI) portions (Krettek & Price, 1977b; Price, 2006; Ray & Price, 1992).

While cytoarchitectural, neuroanatomical, and behavioural studies have mostly focused on the medial-lateral axis, recent reports suggest that there may also be important distinctions along the anterior–posterior axis. For example, Panayi & Killcross (2018) found that, while either anterior or posterior LO lesions impaired Pavlovian outcome devaluation, only posterior, but not anterior, LO lesions disrupted Pavlovian reversal learning. Another study revealed that the anterior but not posterior portion of MO is critical for inferring unobservable action-dependent outcomes and for behavioural response adaptation in outcome-devaluation tasks (Bradfield et al., 2018). These findings suggest that, rather than functionally uniform, the currently recognised OFC subregions

might be composed of smaller structural and functional regions along its anterior–posterior axis. However, given the relative lack of clear cytoarchitectonic differences (Van De Werd & Uylings, 2008), it remains an open question whether there are anatomical distinctions that might underpin these functional differences.

Classically, the boundaries between prefrontal cortical regions, including OFC, have been defined by their specific projection patterns with the mediodorsal (MD) thalamus (Rose & Woolsey, 1948). Surprisingly, in comparison to other prefrontal cortical regions, there have been relatively few studies on the anatomical connectivity of the rat OFC, particularly looking at anterior versus posterior differences. Moreover, because there have been few studies systematically characterising the connectivity of different OFC subregions within the same subjects, inferences between subjects have often been necessary to make these comparisons.

Here, we tested whether functional heterogeneity across the anterior–posterior axis of the OFC is underpinned by differences in anatomical connectivity and with a focus on LO. Specifically, we tested the hypothesis that there are distinct anatomical projections to the anterior lateral (ALO) and posterior lateral OFC (PLO) that underlie the functional dissociation reported by Panayi & Killcross (2018). We contrast these regions with the posterior ventral OFC (PVO), an anatomically adjacent portion of the OFC that is thought to be functionally distinct from LO (Balleine et al., 2011; Corwin et al., 1994). Specifically, we used the retrograde tracer cholera toxin subunit B (CTB) to characterise the afferent projections to these regions simultaneously. This approach allowed us to establish whether any differences reflect (1) a gradient of afferent projections that may be organised topographically or (2) a unique pattern of inputs that projects exclusively to a single subregion.

3.2 Methods summary

Please see Chapter 2 of this thesis for a detailed description of the materials and methods. Below is a brief recap of the most important methodological details.

3.2.1 Surgery

14 male rats received intracranial injections of the retrograde tracer CTB coupled to Alexa Fluor 488, 555 or 647 conjugates, performed unilaterally into the right hemisphere. Each animal received injections of one of the three tracers into ALO, PLO, or PVO. Tracer-site combinations were counterbalanced across animals.

3.2.2 Histology

Ten days after tracer injection, animals were perfused transcardially and the brains removed. 60-µm thick coronal sections were obtained with a freezing microtome. Sections were mounted onto slides and coverslipped with mounting medium with DAPI.

3.2.3 Image acquisition and analysis

Images were obtained using a fluorescence slide scanner (Axio Scan.Z1, Zeiss). Levels of cell labelling density were manually categorised offline into four levels: 0, absence of labelling; 1, weak; 2, moderate; 3, strong (see Figure 3.1 for examples). The summary table and figure depicting the average labelling for each injection site are based on the density average across brains, subsequently averaged across slices when a single slice is represented and discretised into four density values: –, absence of labelling; +, weak; ++, moderate; +++, strong (Table 3.2 and Figure 3.14). The weak density level (+) also includes any region that is consistently labelled, i.e., across at least 50% of the brains,

even if only weak labelling was observed. If a brain region contained maximal tracer deposit in a coronal plane in more than 50% of the brains included in the analyses, we represented that area as maximal deposit in the average labelling figure; otherwise, we calculated the average labelling for that region excluding those brains. The classification of brain areas was based on a standard rat brain atlas (Paxinos & Watson, 1998). Depiction of labelling in key areas of interest, including the amygdala, mediodorsal thalamus and submedius nucleus of the thalamus, was obtained by superimposing hand drawings of the labelling in each brain using an opacity value proportional to the density level of labelling observed.

In our analysis, we focused on the quantification of frontal lobe structures, including OFC subdivisions, prelimbic, infralimbic, anterior cingulate cortex and motor cortex; temporal lobe structures, including amygdala, lateral entorhinal cortex and perirhinal cortex; retrosplenial cortex; primary sensory cortices; thalamic nuclei, including paratenial, submedius and mediodorsal. Due to difficulties in obtaining brain slices of consistent quality posterior to -5.30 mm from bregma, we did not quantify labelling in midbrain structures.

In the description and interpretation of our results, we consider anterior and posterior divisions of medial (MO), ventral (VO), and lateral OFC (LO) separately, but not of dorsolateral OFC (DLO). We refer to the portions of LO and VO contained in the coronal slices from +4.70 to + 4.20 mm from bregma as anterior and to those contained in the coronal slices from +3.20 to +2.20 mm as posterior, with the corpus callosum and the claustrum in the coronal section marking the anterior-posterior division as previously established for LO (Panayi & Killcross, 2018). Based on a previously observed functional dissociation in MO (Bradfield et al., 2018), we refer to the areas in coronal slices at +4.70 and +3.70 mm from bregma as anterior and posterior MO, respectively. These

coordinates are based on a standard rat brain atlas (Paxinos & Watson, 1998). While a ventrolateral OFC (VLO) area has been proposed to exist between VO and LO, here, we will only differentiate between areas VO and LO as per the boundaries defined in Paxinos & Watson (1998).

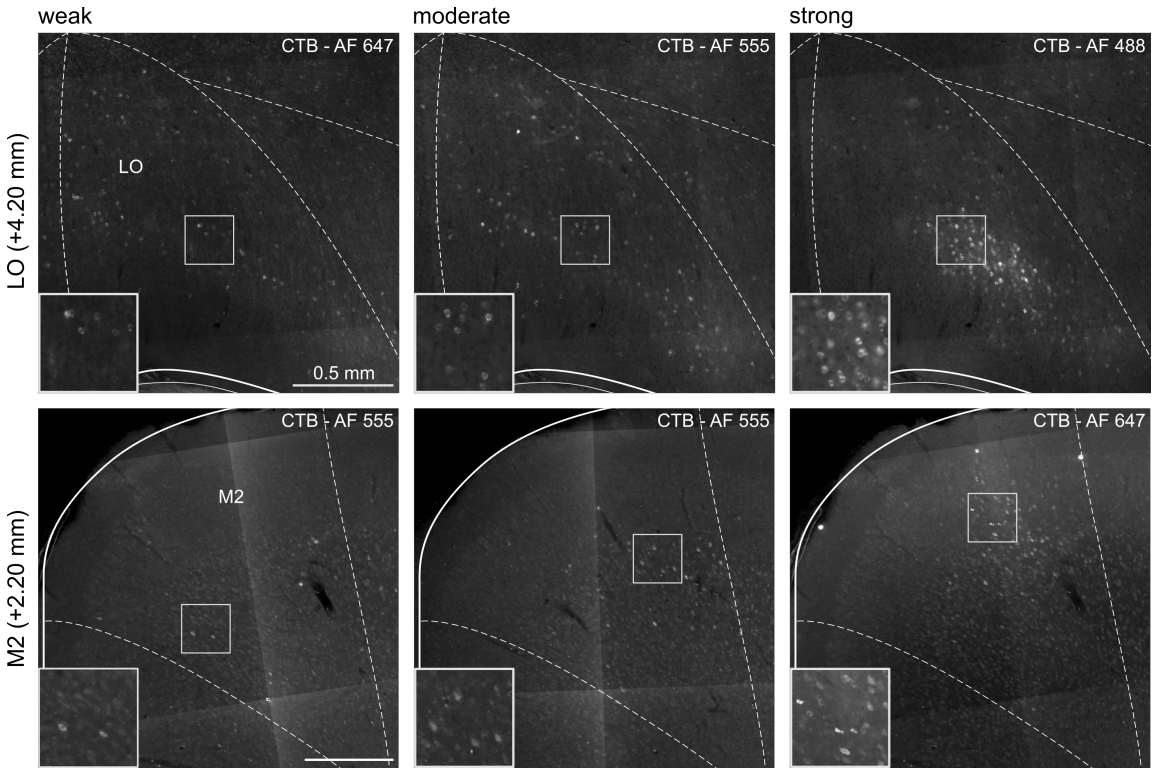


Figure 3.1 Example labelling strength classifications. Depicted are example coronal sections of different brain regions (LO, M2), which labelling strength was classified as weak, moderate, or strong after injection of the retrograde tracer CTB into OFC. Distances shown are distances from bregma in mm. AF: Alexa Fluor; CTB: cholera toxin subunit B; LO: lateral orbitofrontal cortex; M2: secondary motor cortex; OFC: orbitofrontal cortex.

3.3 Results

I compared the projections of the ALO, PLO, and PVO in the same brain while using the same retrograde tracer, coupled to different fluorescent dyes (Conte et al., 2009). Figures 3.2 and 3.3 illustrate the intended injection sites and the core of tracer deposits observed following histological analysis of injected brains included or excluded from further

analyses, respectively. The data presented here were obtained from 7 brains: 4 triple-labelled (n = 12 injection sites), 3 single-labelled (n = 3 injection sites). The location of the core injection deposits included in the analysis was mostly confined to the intended OFC subdivision, with some extension into the ventral agranular insula (AIv) from the PLO injections. Labelling in each brain following tracer deposition into ALO, PLO, and PVO is represented in Figures 3.4, 3.5, and 3.6, respectively. Figure 3.7 depicts the labelling observed following injections in which CTB deposits extended into adjacent AI. The localisation and average density of retrogradely-labelled cells observed are illustrated in Figure 3.8 (see Table 3.1 for exact average density values).

In this study, contralateral labelling was limited to prefrontal afferents (ending at +2.20 mm from bregma; Figure 3.8). The remaining labelling was observed only ipsilaterally. In the triple-labelled brains, the vast majority of cells were single-labelled, with only a small proportion of cells being double- or triple-labelled (see Figure 3.9 for examples). Most often, these were found in the submedius nucleus of the thalamus, where we observed an average of 11% and 1% of labelled cells were double- and triple-labelled, respectively. The density of the retrogradely-labelled cells observed in key areas of interest is summarised in Table 3.2.

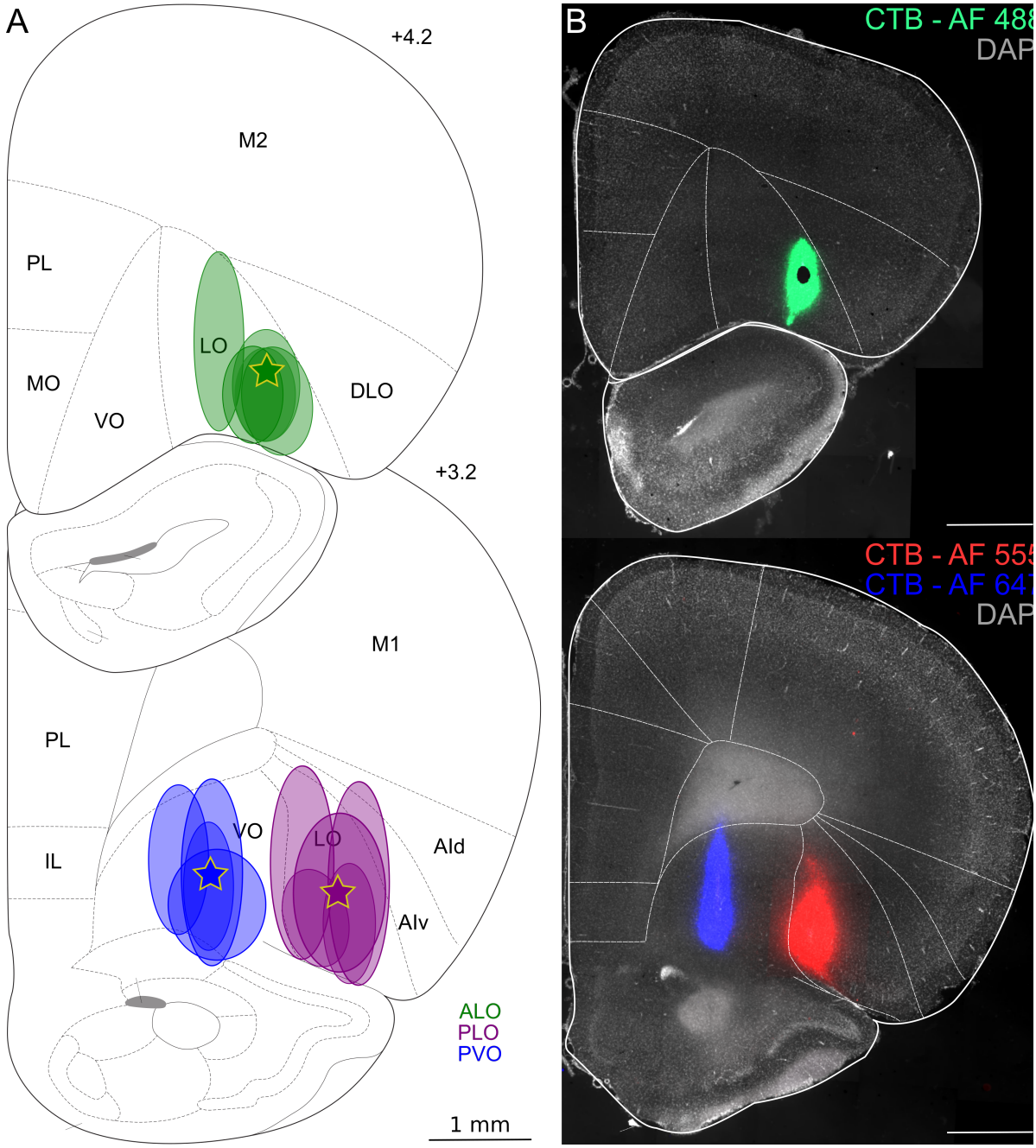


Figure 3.2 Coronal sections depicting core of tracer deposits into ALO, PLO, or PVO, illustrated at the level of maximal tracer deposit. (**A**) Stars and oval shapes represent the intended and the observed injection sites core, respectively. Injections whose core fell outside the orbitofrontal cortex or in which the tracer did not diffuse are not represented here. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. (**B**) Example injection sites of a triple-labelled brain. Distances shown are distances from bregma in mm. AF: Alexa Fluor; AId: agranular insular cortex, dorsal part; AIv: agranular insular cortex, ventral part; ALO: lateral orbitofrontal cortex, anterior part; CTB: cholera toxin subunit B; DLO: orbitofrontal cortex, dorsolateral area; IL: infralimbic cortex; MO: orbitofrontal cortex, medial part; LO: orbitofrontal cortex, lateral part; M1: primary motor cortex; M2: secondary motor cortex; PL: prelimbic cortex; PLO: lateral orbitofrontal cortex, posterior part; PVO: ventral orbitofrontal cortex, posterior part; VO: orbitofrontal cortex, ventral part.

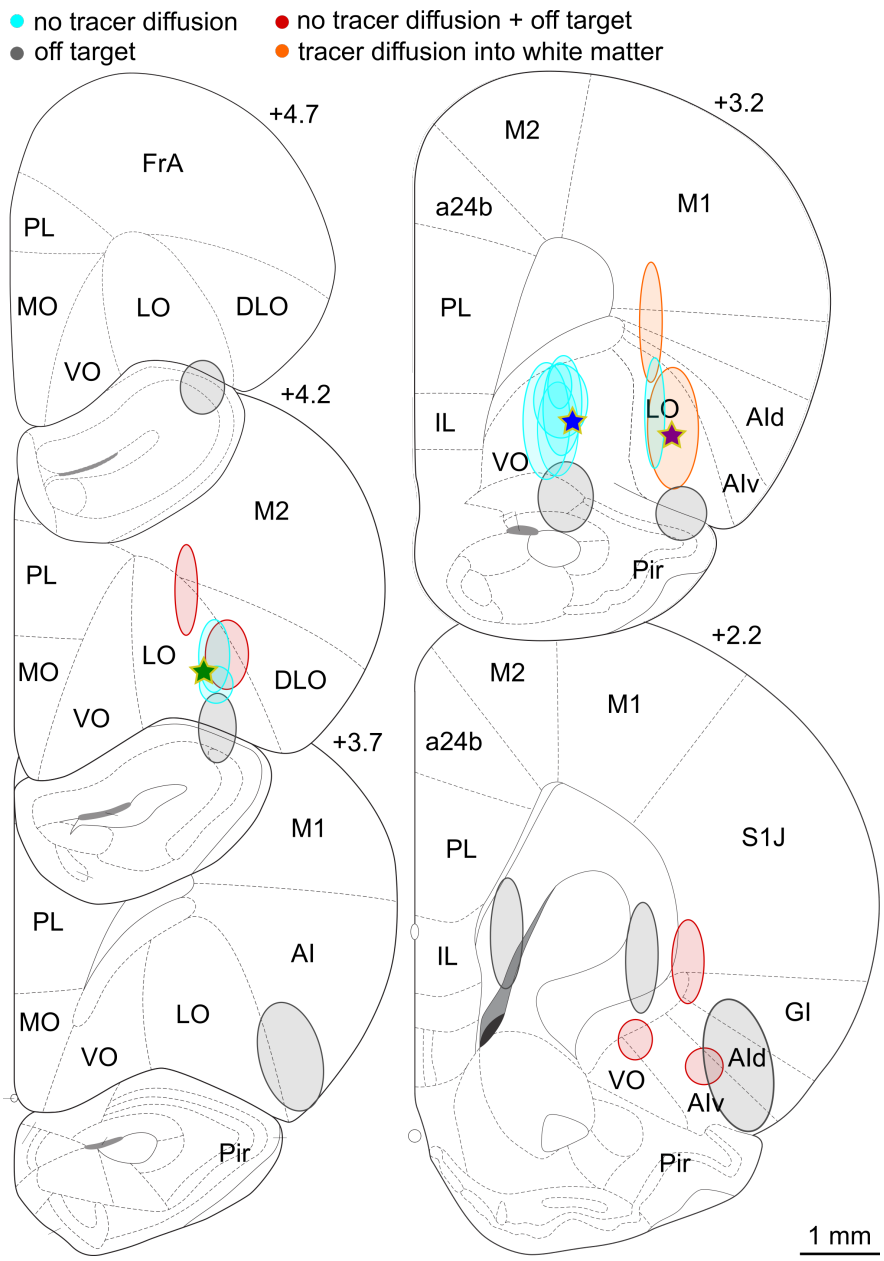


Figure 3.3 Coronal sections depicting core of tracer deposits of excluded injections, illustrated at the level of maximal tracer deposit. Star and oval shapes represent the intended and the observed injection sites core, respectively. Data from injections that were off-target (n = 11), in which tracer did not diffuse (n = 8), were both off target and the tracer did not diffuse (n = 6) or diffused into the white matter (n = 2) were excluded from all analyses. Two of the excluded injections were anterior to +4.7 mm from bregma and are not represented here. Distances shown are distances from bregma in mm. a24b: anterior cingulate cortex, area 24b; AI: agranular insular cortex; AId: agranular insular cortex, dorsal part; AIv: agranular insular cortex, ventral part; DLO: orbitofrontal cortex, dorsolateral part; FrA: frontal association cortex; GI: granular insular cortex; IL: infralimbic cortex; LO: orbitofrontal cortex, lateral part; M1: primary motor cortex; M2: secondary motor cortex; MO: orbitofrontal cortex,

medial part; Pir: piriform cortex; PL: prelimbic cortex; VO: orbitofrontal cortex, ventral part; S1J: primary somatosensory cortex, jaw region.

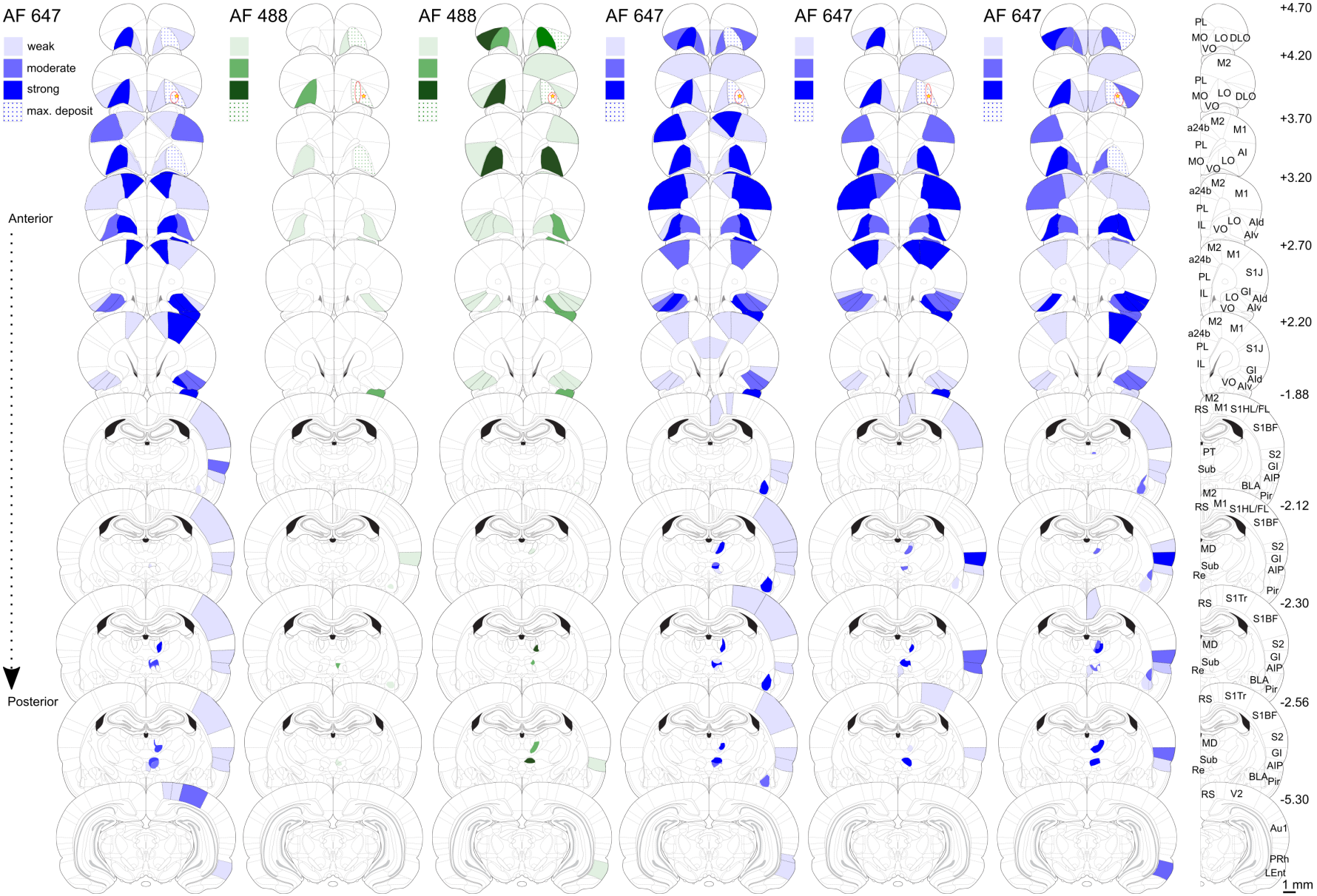

AF 647
weak
moderate
strong
max. deposit
AF 488
AF 488
AF 647
AF 647
AF 647
Anterior
Posterior
+4.70
+4.20
+3.70
+3.20
+2.70
+2.20
-1.88
-2.12
-2.30
-2.56
-5.30
1 mm

Figure 3.4 Density of retrogradely-labelled cells following CTB injection into ALO. Lighter or darker shades represent lower or higher labelling densities, respectively. Red star and oval shapes represent the intended and the observed injection sites core, respectively. Brains are sorted by the location of their observed injection site, from medial to lateral. Distances shown are distances from bregma in mm. AF: Alexa Fluor; CTB: cholera toxin subunit B; a24b: anterior cingulate cortex, area 24b; ALO: lateral orbitofrontal cortex, anterior part; AI, v, d, P: agranular insular cortex, ventral area, dorsal area, posterior area; Au1: primary auditory cortex; BLA: basolateral amygdala; DLO: dorsolateral orbitofrontal cortex; GI: granular insular cortex; IL: infralimbic cortex; LEnt: lateral entorhinal cortex; LO: lateral orbitofrontal cortex; M1: primary motor cortex; M2: secondary motor cortex; MD: mediodorsal nucleus of the thalamus; MO: medial orbitofrontal cortex; Pir: piriform cortex; PL: prelimbic cortex; PRh: perirhinal cortex; PT: paratenial nucleus of the thalamus; Re: nucleus reuniens of the thalamus; RS: retrosplenial cortex; S1BF: primary somatosensory cortex, barrel field; S1J: primary somatosensory cortex, jaw region; S1HL/FL: primary somatosensory cortex, hindlimb and forelimb regions; S1Tr: primary somatosensory cortex, trunk region; S2: secondary somatosensory cortex; Sub: submedius nucleus of the thalamus; V2: secondary visual cortex; VO: ventral orbitofrontal cortex.

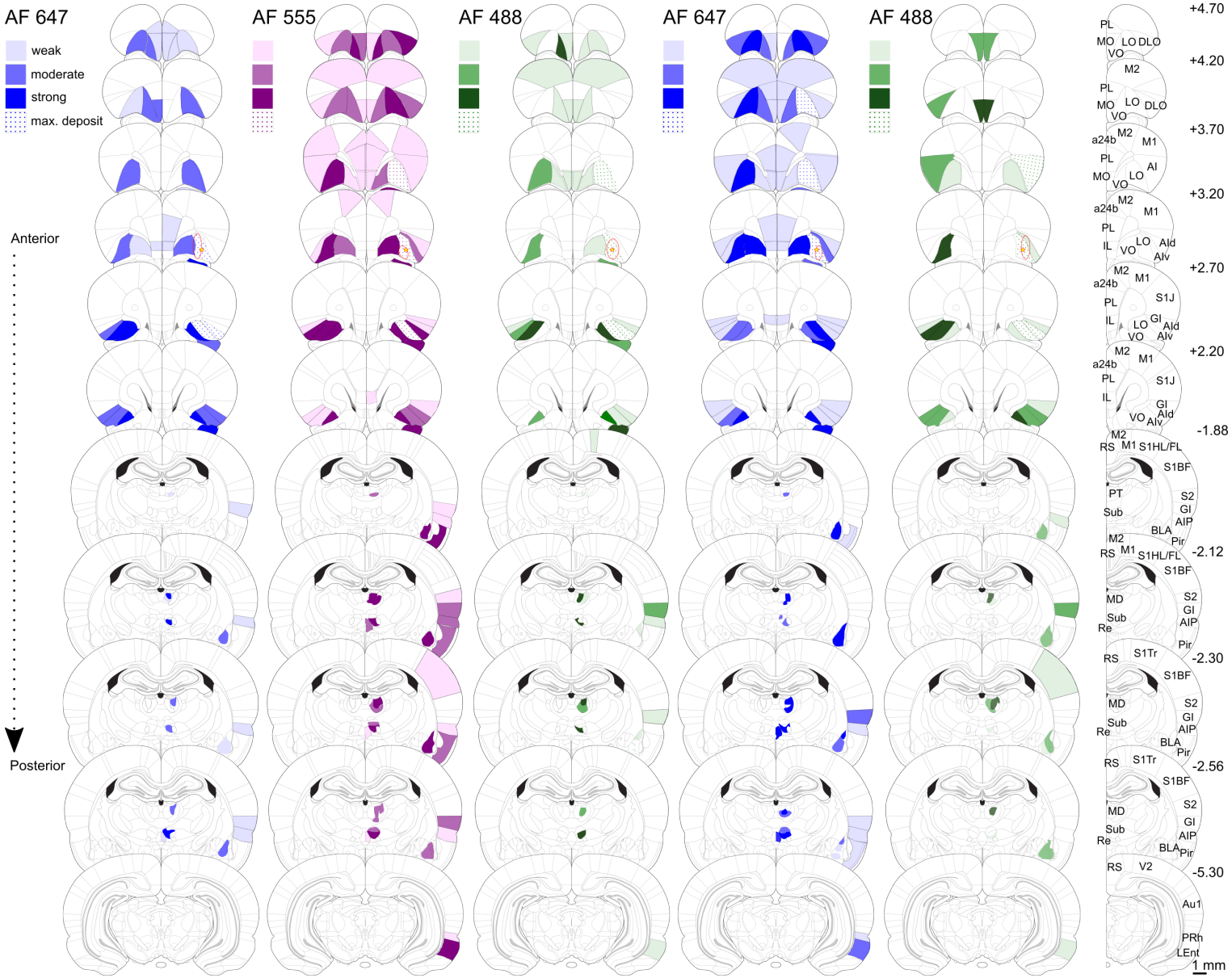

AF 647
weak
moderate
strong
max. deposit
AF 555
AF 488
AF 647
AF 488
Anterior
Posterior
+4.70
+4.20
+3.70
+3.20
+2.70
+2.20
-1.88
-2.12
-2.30
-2.56
-5.30
PL
MO
LO
DLO
VO
M2
M1
a24b
AI
IL
AId
AIv
S1J
GI
S1HL/FL
S1BF
PT
S2
Sub
AIP
BLA
Pir
RS
MD
Re
S1Tr
V2
Au1
PRh
LEnt
1 mm

Figure 3.5 Density of retrogradely-labelled cells following CTB injection into PLO. Lighter or darker shades represent lower or higher labelling densities, respectively. Red star and oval shapes represent the intended and the observed injection sites core, respectively. Brains are sorted by the location of their observed injection site, from medial to lateral. Distances shown are distances from bregma in mm. AF: Alexa Fluor; CTB: cholera toxin subunit B; a24b: anterior cingulate cortex, area 24b; AI, v, d, P: agranular insular cortex, ventral area, dorsal area, posterior area; Au1: primary auditory cortex; BLA: basolateral amygdala; DLO: dorsolateral orbitofrontal cortex; GI: granular insular cortex; IL: infralimbic cortex; LEnt: lateral entorhinal cortex; LO: lateral orbitofrontal cortex; M1: primary motor cortex; M2: secondary motor cortex; MD: mediodorsal nucleus of the thalamus; MO: medial orbitofrontal cortex; Pir: piriform cortex; PL: prelimbic cortex; PLO: lateral orbitofrontal cortex, posterior part; PRh: perirhinal cortex; PT: paratenial nucleus of the thalamus; Re: nucleus reuniens of the thalamus; RS: retrosplenial cortex; S1BF: primary somatosensory cortex, barrel field; S1J: primary somatosensory cortex, jaw region; S1HL/FL: primary somatosensory cortex, hindlimb and forelimb regions; S1Tr: primary somatosensory cortex, trunk region; S2: secondary somatosensory cortex; Sub: submedius nucleus of the thalamus; V2: secondary visual cortex; VO: ventral orbitofrontal cortex.

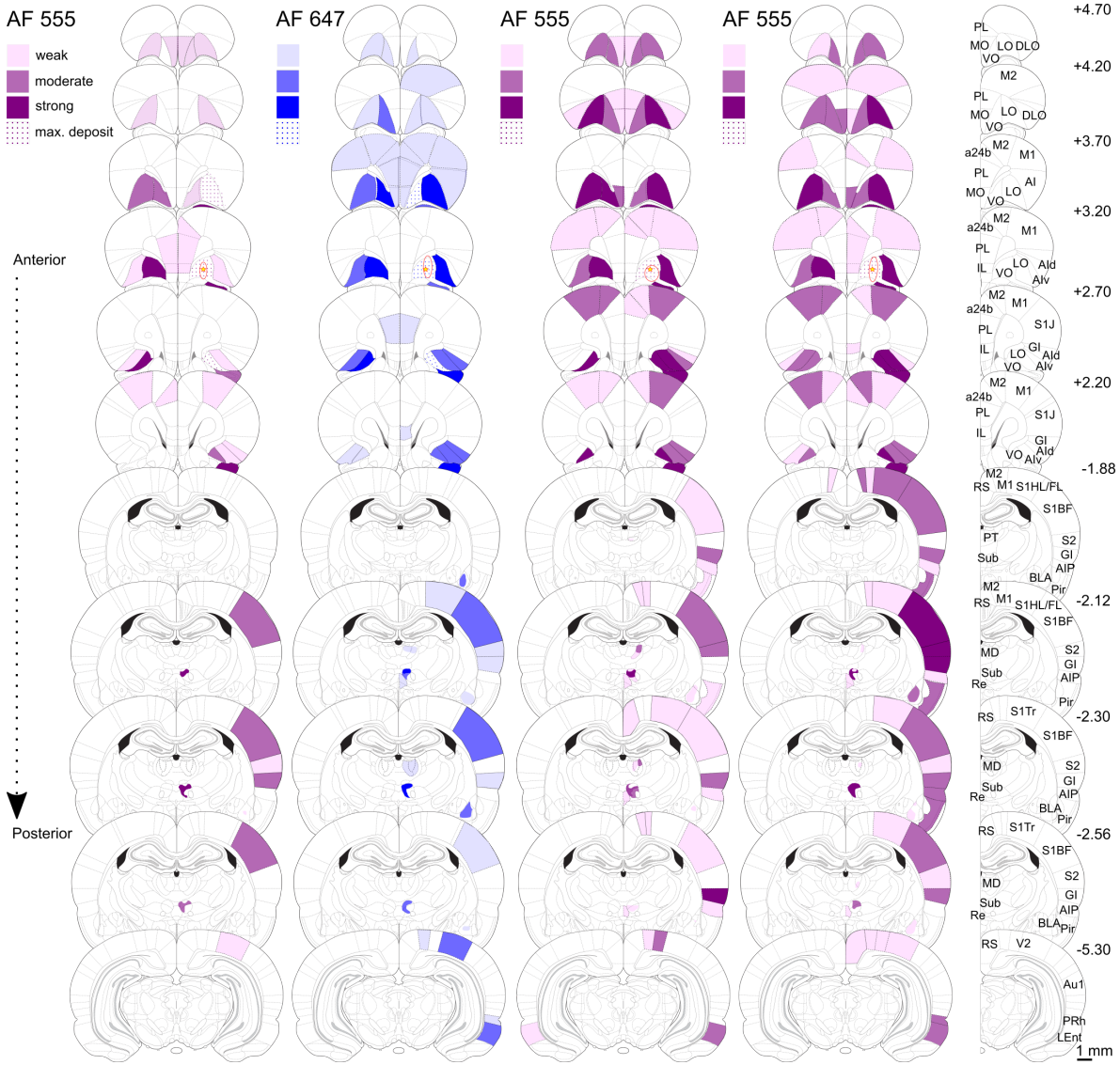


Figure 3.6 Density of retrogradely-labelled cells following CTB injection into PVO. Lighter or darker shades represent lower or higher labelling densities, respectively. Red star and oval shapes represent the intended and the observed injection sites core, respectively. Brains are sorted by the location of their observed injection site, from medial to lateral. Distances shown are distances from bregma in mm. AF: Alexa Fluor; CTB: cholera toxin subunit B; a24b: anterior cingulate cortex, area 24b; AI, v, d, P: agranular insular cortex, ventral area, dorsal area, posterior area; Au1: primary auditory cortex; BLA: basolateral amygdala; DLO: dorsolateral orbitofrontal cortex; GI: granular insular cortex; IL: infralimbic cortex; LEnt: lateral entorhinal cortex; LO: lateral orbitofrontal cortex; M1: primary motor cortex; M2: secondary motor cortex; MD: mediodorsal nucleus of the thalamus; MO: medial orbitofrontal cortex; Pir: piriform cortex; PL: prelimbic cortex; PVO: ventral orbitofrontal cortex, posterior part; PRh: perirhinal cortex; PT: paratenial nucleus of the thalamus; Re: nucleus reuniens of the thalamus; RS: retrosplenial cortex; S1BF: primary somatosensory cortex, barrel field; S1J: primary somatosensory cortex, jaw region; S1HL/FL: primary somatosensory cortex, hindlimb and forelimb regions; S1Tr: primary somatosensory cortex, trunk region; S2: secondary somatosensory cortex; Sub: submedius nucleus of the thalamus; V2: secondary visual cortex; VO: ventral orbitofrontal cortex.

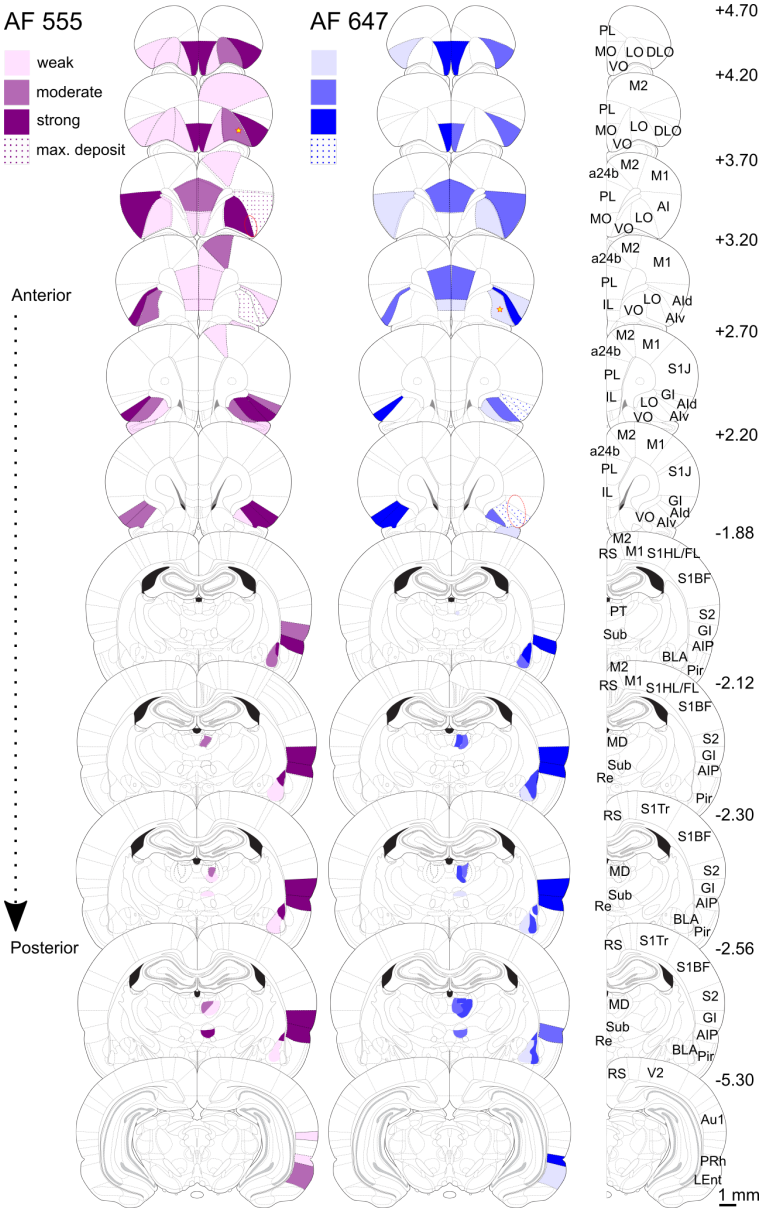


Figure 3.7 Density of retrogradely-labelled cells following injection of CTB into AI. Represented are the labelling patterns of two injections in which the core of the injection extended into AI. Lighter or darker shades represent lower or higher labelling densities, respectively. Red star and oval shapes represent the intended and the observed injection sites core, respectively. Distances shown are distances from bregma in mm. AF: Alexa Fluor; CTB: cholera toxin subunit B; a24b: anterior cingulate cortex, area 24b; AI, v, d, P: agranular insular cortex, ventral area, dorsal area, posterior area; Au1: primary auditory cortex; BLA: basolateral amygdala; DLO: dorsolateral orbitofrontal cortex; GI: granular insular cortex; IL: infralimbic cortex; LEnt: lateral entorhinal cortex; LO: lateral orbitofrontal cortex; M1: primary motor cortex; M2: secondary motor cortex; MD: mediodorsal nucleus of the thalamus; MO: medial orbitofrontal cortex; Pir: piriform cortex; PL: prelimbic cortex; PRh: perirhinal cortex; PT: paratenial nucleus of the thalamus; Re: nucleus reuniens of the thalamus; RS: retrosplenial cortex; S1BF: primary somatosensory cortex, barrel field; S1J: primary somatosensory cortex, jaw region; S1HL/FL: primary somatosensory cortex, hindlimb and forelimb regions; S1Tr: primary somatosensory cortex, trunk region; S2: secondary somatosensory cortex; Sub: submedius nucleus of the thalamus; V2: secondary visual cortex; VO: ventral orbitofrontal cortex.

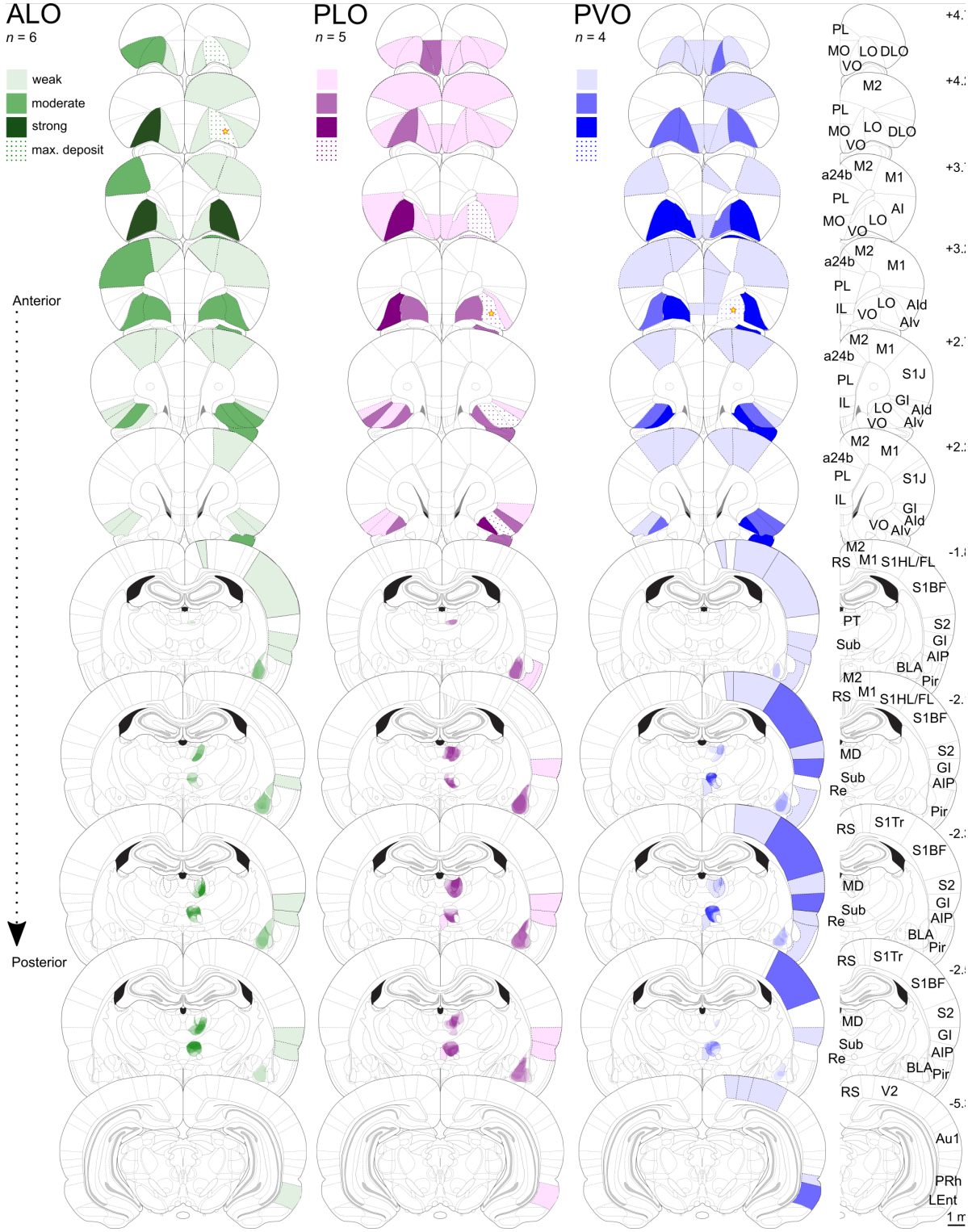


Figure 3.8 Density of retrogradely-labelled cells following CTB injection into ALO, PLO, or PVO. Represented is the average density across brains discretised into four labelling levels (see Table S1 for exact average density values). Lighter or darker shades represent lower or higher labelling densities, respectively. This represents the average density of projections from multiple brains within a given structure but note that this does not indicate the precise location of these projections within the structure. Depiction of labelling in key areas of interest (amygdala, mediodorsal thalamus and submedius nucleus of the thalamus) was obtained by

superimposing hand drawings of the labelling in each brain using an opacity value proportional to the density level of labelling observed. Star shapes represent the intended injection site cores. Distances shown are distances from bregma in mm. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4. a24b: anterior cingulate cortex, area 24b; ALO: lateral orbitofrontal cortex, anterior part; AI, v, d, P: agranular insular cortex, ventral area, dorsal area, posterior area; Au1: primary auditory cortex; BLA: basolateral amygdala; CTB: cholera toxin subunit B; DLO: dorsolateral orbitofrontal cortex; GI: granular insular cortex; IL: infralimbic cortex; LEnt: lateral entorhinal cortex; max. deposit: maximal deposit; LO: lateral orbitofrontal cortex; M1: primary motor cortex; M2: secondary motor cortex; MD: mediodorsal nucleus of the thalamus; MO: medial orbitofrontal cortex; Pir: piriform cortex; PL: prelimbic cortex; PLO: lateral orbitofrontal cortex, posterior part; PVO: ventral orbitofrontal cortex, posterior part; PRh: perirhinal cortex; PT: paratenial nucleus of the thalamus; Re: nucleus reuniens of the thalamus; RS: retrosplenial cortex; S1BF: primary somatosensory cortex, barrel field; S1J: primary somatosensory cortex, jaw region; S1HL/FL: primary somatosensory cortex, hindlimb and forelimb regions; S1Tr: primary somatosensory cortex, trunk region; S2: secondary somatosensory cortex; Sub: submedius nucleus of the thalamus; V2: secondary visual cortex; VO: ventral orbitofrontal cortex.

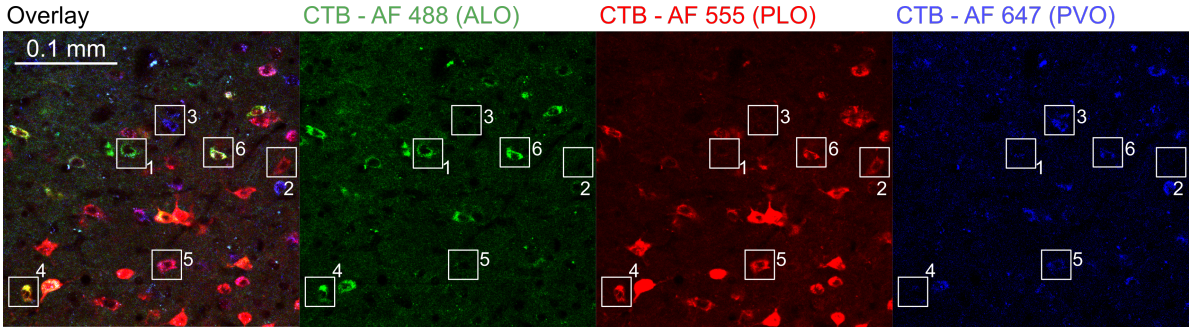


Figure 3.9 Example single-, double-, and triple-labelled cells. Represented is a plane of a z-stack obtained with a confocal microscope of the submedius nucleus of the thalamus in which are present: single-labelled cells after (1) ALO, (2) PLO, or (3) PVO injections; double-labelled cells after (4) ALO and PLO, or (5) after PLO and PVO injections; and (6) triple-labelled cells.

Table 3.1

Average labelling density observed following CTB injection into ALO, PLO, or PVO.

		Ipsilateral			Contralateral		
Structure	AP coordinate (mm)	ALO	PLO	PVO	ALO	PLO	PVO
Prefrontal cortex							
Orbitofrontal cortex							
dorsolateral part	+4.70	1.00	1.00	0	1.50	0.80	0
	+4.20	0.67	0.60	0.25	0.33	0.80	0.25
lateral part	+4.70	m. d.	1.40	1.25	2.17	1.40	1.25
	+4.20	m. d.	1.25	2.00	2.83	2.20	1.75
	+3.70	3.00	1.25	3.00	2.67	2.80	2.50
	+3.20	2.33	m. d.	2.50	2.33	2.80	1.75
	+2.70	2.00	m. d.	2.25	1.67	1.40	1.50
medial part	+4.70	0.33	1.20	0.50	0.17	1.60	0.50
	+4.20	0.33	1.20	0.50	0.33	1.40	0.75
	+3.70	0	1.00	0.75	0	0.80	0.75
ventral part	+4.70	1.00	1.40	1.50	1.00	2.00	1.25
	+4.20	0.83	1.00	1.25	0.67	1.40	1.50
	+3.70	0.83	0.80	1.67	0.83	0.60	2.50

	+3.20	1.83	1.80	m. d.	1.67	1.60	3.00
	+2.70	2.17	2.40	3.00	1.00	1.80	2.75
	+2.20	0	2.80	2.50	0	1.80	1.50
Infralimbic cortex	+3.20	0	0.40	0.50	0	0.40	0.50
	+2.70	0	0.20	0.25	0	0.20	0
	+2.20	0	0.20	0.25	0	0	0
Prelimbic cortex	+4.70	0	0.20	0	0	0.40	0
	+4.20	0	0.20	0.25	0	0.20	0.25
	+3.70	0	0.40	0.25	0	0.40	0.25
	+3.20	0	0.40	0.50	0	0.20	0.25
	+2.70	0.17	0	0.25	0.17	0	0.25
	+2.20	0	0	0	0	0	0.25
Cingulate cortex, area 1	+3.70	0.17	0.20	0.50	0	0.20	0
	+3.20	0	0	0.75	0	0	0
	+2.70	0.17	0	0.25	0	0	0.25
	+2.20	0	0	0.25	0	0	0
Insular cortex							
Agranular insular cortex	+3.70	0	0.50	0.20	0.17	0.80	0
	-1.88	0.67	0.40	0.50	0	0	0
	-2.12	0.83	0.40	0.25	0	0	0
	-2.30	0.67	0.60	0.75	0	0	0
	-2.56	0.50	0.60	0.25	0	0	0
dorsal part	+3.20	0	0.40	0	0.17	0.20	0
	+2.70	1.17	0.80	0.25	0.67	1.20	0
	+2.20	1.33	1.80	1.75	0.83	1.00	0
ventral part	+3.20	0.17	1.00	0	0.33	0.40	0
	+2.70	1.50	m. d.	1.25	1.17	2.20	0.25
	+2.20	1.33	1.80	1.75	0.83	1.20	0.50
Gustatory insula	+2.70	0.17	0.40	0	0.17	0.20	0
	+2.20	0.33	0.60	0	0	0.20	0
	-1.88	0.83	0.40	0.50	0	0	0
	-2.12	0.17	1.20	1.50	0	0	0
	-2.30	0.83	1.00	1.75	0	0	0
	-2.56	0.83	0.80	1.25	0	0	0
Motor cortex							
Primary	+3.70	1.00	0.20	0.50	1.67	0.20	0.50
	+3.20	1.33	0	0.50	1.50	0	1.00
	+2.70	1.17	0	1.00	1.00	0	1.00
	+2.20	1.17	0	1.00	0.33	0	0
	-1.88	0.17	0.20	0.25	0	0	0
	-2.12	0	0	0.50	0	0	0
	-2.30	0	0	0	0	0	0
	-2.56	0	0	0.25	0	0	0
Secondary	+4.20	0.67	0.60	0.50	0.17	0.60	0.25
	+3.70	0.67	0.20	0.25	0.33	0	0.25
	+3.20	1.00	0.20	0.25	1.33	0.20	0.50
	+2.70	1.33	0.20	0.50	0.83	0	1.00
	+2.20	0.33	0	0.75	0.17	0	0.50
	-1.88	1.00	0	0.50	0	0	0.25
	-2.12	0	0	2.00	0	0	1.00
	-2.30	0	0	0.25	0	0	0
	-2.56	0	0	0	0	0	0
Sensory cortices							
Piriform cortex	+3.70	1.50	1.20	2.75	0	0	0
	+3.20	2.17	1.80	2.75	0	0	0
	+2.70	2.17	2.00	3.00	0.14	0	0
	+2.20	2.33	2.20	0	0	0	0
	-1.88	0.17	0.80	0.75	0	0	
	-2.12	0	0.40	0.75	0	0	0
	-2.30	0	0.40	0.50	0	0	0
	-2.56	0.40	0.25	0	0	0	0
Somatosensory cortex							
primary/barrel cortex	-1.88	0.50	0	0.75	0	0	0
	-2.12	0.33	0	2.25	0	0	0
	-2.30	0.33	0.40	1.75	0	0	0
	-2.56	0.33	0	1.50	0	0	0
forelimb and hindlimb regions	-1.88	0	0	0.50	0	0	0
	-2.12	0	0	0.50	0	0	0

trunk region	-2.30	0.17	0	0.50	0	0	0
	-2.56	0.17	0	0.50	0	0	0
secondary	-1.88	0	0.20	0.25	0	0	0
	-2.12	0.40	0.20	1.50	0	0	0
	-2.30	0	0	0.50	0	0	0
	-2.56	0	0	0.25	0	0	0
Auditory cortex							
primary	-5.30	0	0	0	0	0	0
secondary							
dorsal part	-5.30	0	0	0	0	0	0
ventral part	-5.30	0	0	0	0	0	0
Visual cortex							
primary							
binocular area	-5.30	0	0	0	0	0	0
monocular area	-5.30	0	0	1.00	0	0	0
secondary							
lateral area	-5.30	0.33	0	0.75	0	0	0
mediolateral area	-5.30	0.17	0	0.75	0	0	0
mediomedial area	-5.30	0.17	0	0.25	0	0	0
Temporal lobe							
Amygdala							
basolateral	-1.88	1.17	1.60	0.75	0	0	0
anterior part	-1.88	1.17	1.60	0.75	0	0	0
	-2.12	1.33	2.00	1.25	0	0	0
	-2.30	0.67	1.80	1.25	0	0	0
	-2.56	0.33	1.20	0.50	0	0	0
posterior part	-2.30	0	0.40	0	0	0	0
	-2.56	0	0.80	0	0	0	0
lateral							
dorsolateral part	-1.88	0.17	1.20	0.25	0	0	0
	-2.12	0.17	1.40	0	0	0	0
	-2.30	0.50	1.20	0.25	0	0	0
	-2.56	0.17	0.40	0	0	0	0
ventromedial part	-2.56	0	0.60	0	0	0	0
ventrolateral part	-2.56	0	0.20	0	0	0	0
Entorhinal cortex, lateral part	-5.30	0.83	1.20	1.50	0	0	0.33
Perirhinal cortex	-5.30	0.17	0.40	0.75	0	0	0
Parietal lobe							
Retrosplenial cortex	-1.88	0.33	0	0	0	0	0
	-2.12	0	0	0	0	0	0
	-2.30	0.17	0	0.25	0	0	0
	-2.56	0	0	0	0	0	0
Thalamic nuclei							
Mediodorsal	-1.88	0	0	0	0	0	0
central part	-2.30	0.33	2.20	0.75	0	0	0
	-2.56	0.83	2.60	0.25	0	0	0
lateral part	-2.12	1.20	1.80	1.25	0	0	0
	-2.30	2.00	1.40	0.75	0	0	0
	-2.56	2.00	0.80	0	0	0	0
medial part	-2.12	0.20	2.60	0.25	0	0	0
	-2.30	0.17	1.60	0.25	0	0	0
	-2.56	0.50	0.40	0	0	0	0
Paratenial	-1.88	0.40	1.60	0	0	0	0
Reunens	-2.12	0	0.40	0.50	0	0	0
	-2.30	0.33	0.80	0.50	0	0	0
	-2.56	0.17	0.80	0.50	0	0	0
Submedius	-1.88	0	0	0	0	0	0
dorsal part	-2.12	0.60	0.20	2.25	0	0	0
	-2.30	1.50	0.60	3.00	0	0	0
	-2.56	2.33	1.20	1.25	0	0	0
ventral part	-2.12	0.60	2.40	1.25	0	0	0
	-2.30	1.50	2.60	2.50	0	0	0
	-2.56	1.00	2.40	1.25	0	0	0

Note. Average values were obtained following 4-level categorisation (0, absence of labelling; 1, weak; 2, moderate; 3, strong) of labelling density. Areas with maximal tracer deposit, only found around injection sites, were not included in the analyses. m. d.: maximal deposit.

3.3.1 Prefrontal cortex afferents

Intra-OFC projections

DLO

Following injection of CTB into ALO and PLO, we observed weak to moderate labelling in DLO, ipsi- and contralaterally to the injection site (Figure 3.8). In contrast, we detected no labelling in this OFC portion following injection into PVO, both ipsi- and contralaterally.

AI

following ALO injection, we detected only weak labelling in the AI of the same coronal plane as the injection site and weak to moderate labelling in posterior slices. PLO injection resulted in weak or strong ipsilateral labelling in anterior and posterior AI, respectively. Retrograde labelling in the contralateral side was mostly weaker. PVO injection resulted in weak to moderate density of labelled cells in the posterior portions of AI.

LO

ALO injections of CTB resulted in strong or moderate labelling in ALO and PLO, both ipsi- and contralaterally. PLO injections resulted in only weak labelling in ALO, ipsilaterally. Contralaterally, we detected weak to moderate labelling in ALO and strong to weak labelling in PLO along the anterior–posterior axis. Following injection of CTB into PVO, we observed strong to moderate labelling in PLO and only weak to moderate labelling in ALO, ipsilaterally and along the anterior–posterior axis. This was largely mirrored on the contralateral hemisphere.

VO

Following tracer injection into ALO, we detected only weak labelling in anterior ventral OFC (AVO) but moderate labelling in PVO, ipsilaterally. A similar labelling pattern was observed following tracer injection into PLO. This was largely mirrored on the contralateral hemisphere both in ALO and PLO injections. Injection into PVO resulted in weak or strong labelling in ipsilateral AVO or PVO, respectively. Retrograde labelling in the contralateral side largely mirrored that of the ipsilateral cortex, generally with fewer cells present.

MO

MO had almost no labelled cells following ALO injection, with only some very low-density labelling in its most anterior part in both hemispheres in 2 of the 6 brains included in the analysis (see Figure 3.4). By contrast, both PLO and PVO injections resulted in weak but consistent labelling in both the anterior and posterior portions of MO, ipsilaterally. A similar pattern was observed in the contralateral hemisphere, with the exception that the anterior portion of MO contained stronger labelling contralateral to the injection site following PVO injection.

In summary, all three subdivisions receive input from most other parts of the OFC, both ipsi- and contralaterally. One distinction is MO, which sends stronger projections to PLO and PVO than to ALO. By contrast, DLO sends weak projections to both ALO and PLO but not to PVO.

Medial prefrontal projections

Following retrograde tracer injection into ALO, prelimbic (PL) and infralimbic (IL) were mostly devoid of labelled cells, both ipsi- and contralaterally (Figure 3.4). On the other hand, following PLO injection, both IL and PL contained weak labelling in a majority of brains. Similarly, we observed weak but consistent labelling in both IL and PL following PVO injection. No labelled cells were detected consistently in the anterior cingulate cortex (a24b), in either hemisphere following injection into ALO or PLO. By contrast, PVO injection often resulted in weak labelling in this region (Figure 2, Figure 3.6). Thus, IL and PL labelling suggest an anterior–posterior difference within LO, while a24b labelling suggests a medial-lateral distinction between PLO and PVO (Table 3.2).

Motor cortex

Following injections into ALO, we observed weak but widespread labelling in both ipsilateral, primary (M1) and secondary motor (M2) cortex (Figure 3.8). This was largely mirrored in the contralateral side with the difference that, here, ALO resulted in moderate labelling in M1. By contrast, PLO injection resulted only in weak labelling restricted to M2 and to only one coronal plane, both ipsi- and contralaterally. Labelling patterns following PVO injection were similar to those observed following ALO injection, although resulting in weaker labelling in contralateral M1. Overall, input from motor cortex suggests both an anterior–posterior difference and a medial-lateral difference such that PLO receives much weaker motor inputs than ALO and PVO (Table 3.2).

Sensory cortex

Following CTB injection into PVO, but not ALO or PLO, we observed weak labelling present in the lateral, mediolateral, and mediomedial areas of the secondary visual cortex

(V2) and weak to moderate labelling in the barrel cortex and the forelimb, hindlimb, and trunk regions of the primary somatosensory cortex. In contrast, except for weak labelling in the barrel cortex after ALO injection, we observed no labelled cells in the visual or the somatosensory cortices following injection into lateral OFC (either ALO or PLO). Following injection into either ALO, PLO, or PVO, we consistently detected moderate labelling in piriform cortex (Pir) and weak labelling in gustatory insular cortex (GI). However, posterior to -1.88 mm from bregma, Pir exhibited only weak labelling from PVO injections. The auditory cortex (Au1) was devoid of labelled cells following injection into either subdivision. Therefore, there was strong evidence of primary multi-sensory inputs into PVO but not into ALO or PLO (Table 3.2).

3.3.2 Thalamic afferents

Detailed schematics for the thalamus are shown in Figures 3.10 and 3.11. ALO, PLO, and PVO, each exhibit distinct connectivity patterns with the mediodorsal (MD) and submedius (Sub) nuclei of the thalamus.

Mediodorsal nucleus

ALO receives the strongest projections from the lateral part (MDl), particularly its most ventral portion. PLO receives strong projections from both the medial (MDm) and central (MDc) parts of the medial dorsal thalamic nucleus and only weak ones from MDl (Figure 3.10, Figure 3.12-A). In contrast, PVO receives only light projections from either lateral, medial or central MD.

Supporting the observed pattern of labelling statistically (summarised in Figure 3.12-A), OFC subregions received distinct patterns of projections from MD (significant main effect of MD thalamus portion, $F_{(2,\ 30.2)} = 4.32$, $p = .022$; and injection site x MD

thalamus portion interaction, $F_{(4, 30.3)} = 3.80$, $p = .013$). ALO receives stronger projections from MDl than from MDm (ALO: MDl vs MDm, $t_{(36.0)} = 3.40$, $p = .005$) or MDc (ALO: MDl vs MDc, $t_{(36.0)} = 2.95$, $p = .017$), but similar projection strength from MDc and MDm (ALO: MDc vs MDm, $t_{(36.0)} = 0.45$, $p = .956$). In contrast to ALO, there were no differences in the relative strength of projections from MD subregions to PLO and PVO (all $t_{(36.0)} < |1.82|$, $p > .078$). There was a main effect of injection site (ALO, PLO, PVO; $F_{(2, 34.1)} = 16.03$, $p < .001$), with PLO receiving relatively strong projections and PVO receiving relatively weak projections from MD overall (PLO vs PVO: $t_{(32.6)} = 5.60$, $p < .001$; PLO vs ALO: $t_{(35.6)} = 3.47$, $p = .004$; PVO vs ALO: $t_{(34.0)} = -2.27$, $p = .086$).

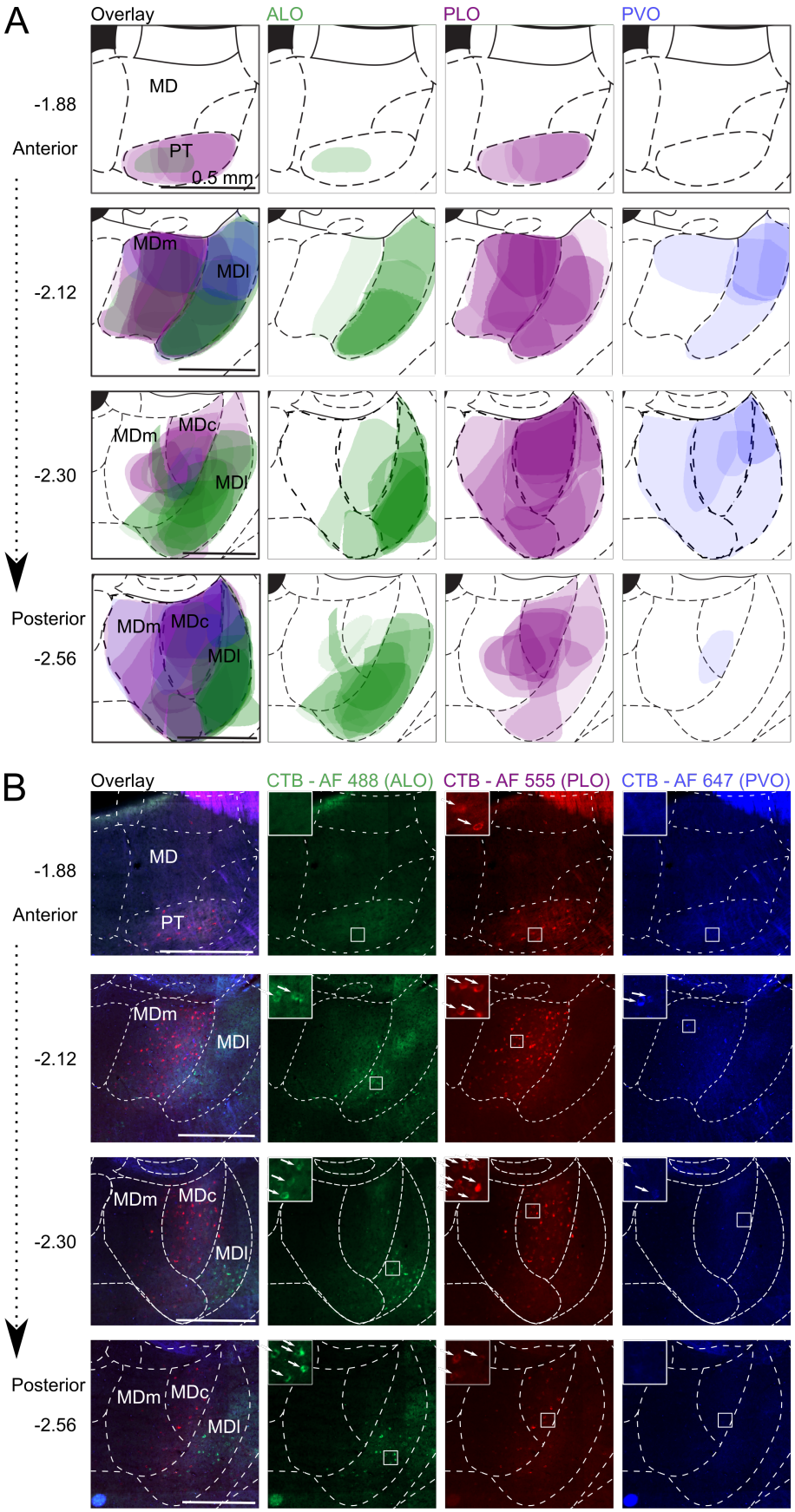


Figure 3.10 Localisation and density of retrogradely-labelled cells in the MD following CTB injection into ALO, PLO, or PVO. (**A**) Shaded areas correspond to superimposed hand

drawings of the labelling in each brain included in the analysis, using an opacity value proportional to the labelling density level. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. (**B**) Micrographs from a representative brain with successful triple labelling. c: central; l: lateral; m: medial; AF: Alexa Fluor; ALO: lateral orbitofrontal cortex, anterior area; CTB: cholera toxin subunit B; MD: mediodorsal nucleus of the thalamus; PLO: lateral orbitofrontal cortex, posterior area; PVO: ventral orbitofrontal cortex, posterior area.

Submedius nucleus

All portions of Sub project to every region of the OFC investigated here, although there is clear spatial segregation in relation to where in Sub the strongest projections originate (Figure 3.11, Figure 3.12-B, C). ALO receives the strongest projections from the posterior part of the dorsal submedius nucleus (SubD). In contrast, most afferents from Sub to PLO come from the ventral portion (SubV). On the other hand, PVO receives most Sub projections from the most anterior half, prominently from the dorsal portion, and slightly less from the ventral part.

First, we quantified the overall pattern of projections from Sub (Figure 3.12-B), focusing on the average labelling strength across coronal sections (averaging across -2.12, -2.38, and -2.56 mm from bregma). This revealed that PLO receives stronger projections from SubV than SubD (PLO: SubV vs SubD: $t_{(24)} = 4.13$, $p = .001$), whereas ALO and PVO (ALO: SubV vs SubD: $t_{(24)} = -0.82$, $p = .804$; PVO: SubV vs SubD: $t_{(24)} = -1.59$, $p = .330$) do not appear to be differentiated (supported by an injection site x Sub portion interaction, $F_{(2,\ 18.2)} = 13.36$, $p < .001$). Furthermore, PVO receives stronger overall projections from Sub than ALO (main effect of injection site $F_{(2,\ 20.2)} = 7.09$, $p = .005$; PVO vs ALO, $t_{(20)} = 3.77$, $p = .004$).

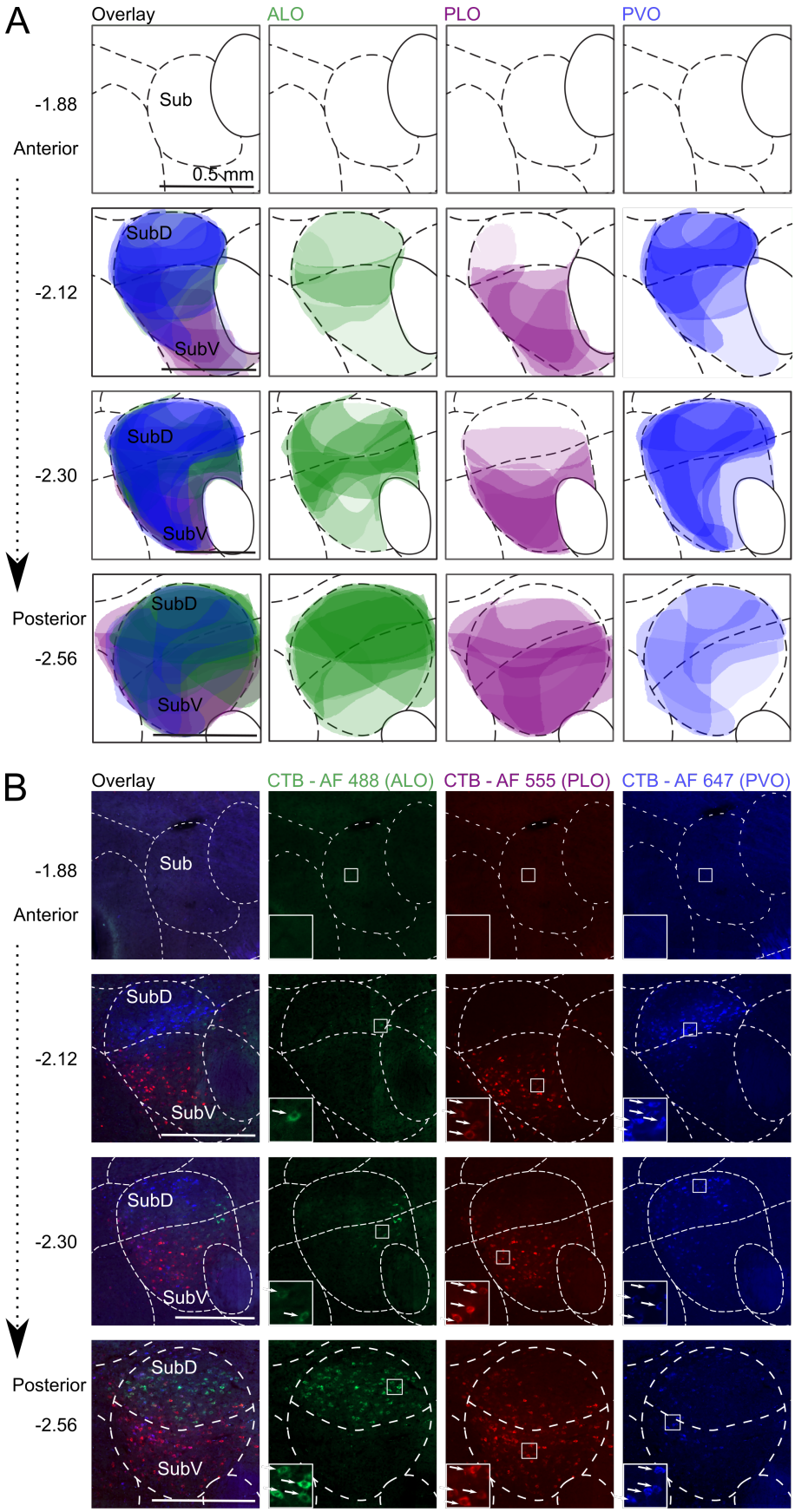


Figure 3.11 Localisation and density of retrogradely-labelled cells in Sub following CTB injection into ALO, PLO, or PVO. (**A**) Shaded areas correspond to superimposed hand

drawings of the labelling in each brain included in the analysis, using an opacity value proportional to the labelling density level. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. (**B**) Micrographs from a representative brain with successful triple labelling. D: dorsal; V: ventral; AF: Alexa Fluor; ALO: lateral orbitofrontal cortex, anterior area; CTB: cholera toxin subunit B; PLO: lateral orbitofrontal cortex, posterior area; PVO: ventral orbitofrontal cortex, posterior area; Sub: submedius nucleus of the thalamus.

Next, we explored the differences in the topographical organisation of these projections along the anterior–posterior axis (Figure 3.12-C). A significant 3-way interaction of injection site x Sub portion x coronal section ($F_{(4,\ 66.1)}$ = 3.27, p = .016) indicated that the distribution and strength of projections from Sub along the anterior–posterior axis differed depending on the orbitofrontal injection site. We explored this 3-way interaction by performing separate coronal section (-2.12, -2.30, -2.56 mm from bregma) x Sub portion (SubV, SubD) ANOVAs for each injection site (ALO, PLO, PVO).

Projections to ALO increased from anterior to posterior Sub (main effect of coronal section, $F_{(2,\ 25)}$ = 4.08, p = 0.029; coronal sections -2.12 vs -2.30 mm: $t_{(25)}$ = -1.98, p = .166; -2.12 vs -2.56 mm: $t_{(25)}$ = -2.77, p = .031; -2.30 vs -2.56 mm: $t_{(25)}$ = -0.79, p = .820), but were not differentiated by dorsal or ventral regions within Sub (main effect of Sub portion: SubV vs SubD, $F_{(1,\ 25)}$ = 0.34, p = .564; coronal section x Sub portion interaction, $F_{(2,\ 25)}$ = 1.88, p = .174). In contrast, Sub projections to PLO were not differentiated along the anterior–posterior axis (main effect of coronal section: $F_{(2,\ 20)}$ = 1.84, p = .185; coronal section x Sub portion interaction: $F_{(2,\ 20)}$ = 2.00, p = .161), but labelling was significantly stronger in SubV than in SubD (main effect of Sub portion, SubV vs SubD: $F_{(1,\ 20)}$ = 74.09, p < .001).

Finally, the relative strength of projections to PVO originating from SubD and SubV changed along the anterior–posterior axis (coronal section x Sub portion interaction: $F_{(2,\ 15)}$ = 5.82, p = .013). Stronger labelling was observed in SubD than in SubV at -2.12 mm from bregma (-2.12 mm: SubD vs SubV, $t_{(18)}$ = 4.88, p < .001), but no differences were

found between labelling in the Sub portions at more posterior coronal sections (-2.30 mm: SubD vs SubV, $t_{(18)}$ = 1.68, p = .298; -2.56 mm: SubD vs SubV: $t_{(18)}$ = 2.46, p = .993). Overall projections from Sub to PVO were strongest at -2.30 mm from bregma (main effect of coronal section, $F_{(2, 15)}$ = 17.76 p < .001; -2.12 vs -2.30 mm, $t_{(15)}$ = -2.81, p = .039; -2.30 vs -2.56 mm, $t_{(15)}$ = 5.96, p < .001; -2.12 vs -2.56 mm, $t_{(15)}$ = 3.15, p = .020). Overall, this pattern of projections suggests an anterior–posterior gradient of projections from Sub to ALO, projections from Sub to PLO originate predominantly in SubD, and a topographically unique pattern of projections to PVO that differs across anterior–posterior and dorsal-ventral axes.

Paratenial nucleus

The paratenial thalamic nucleus (PT; Figure 3.10) consistently exhibited moderate labelling following PLO injection but very light to no labelling with injections into ALO or PVO.

Nucleus reuniens

In the reuniens thalamic (Re) nucleus, we observed light labelling following tracer injections into both lateral and ventral posterior OFC, but mostly absent after injections into ALO (Figure 3.8).

In sum, the thalamus sends distinct and topographically organised projections from MD and Sub into ALO, PLO and PVO (Figure 3.12, Table 3.2). Here, the anterior–posterior distinction between ALO and PLO is as clear and differentiated as the medial-lateral distinction between PLO and PVO. Projections from Re to PLO and PVO, but not to ALO, reinforce the anterior–posterior distinction within LO. Labelling in PT after PLO injection

but not after ALO or PVO also reinforces this anterior–posterior distinction within LO as well as a medial-lateral distinction between PLO and PVO.

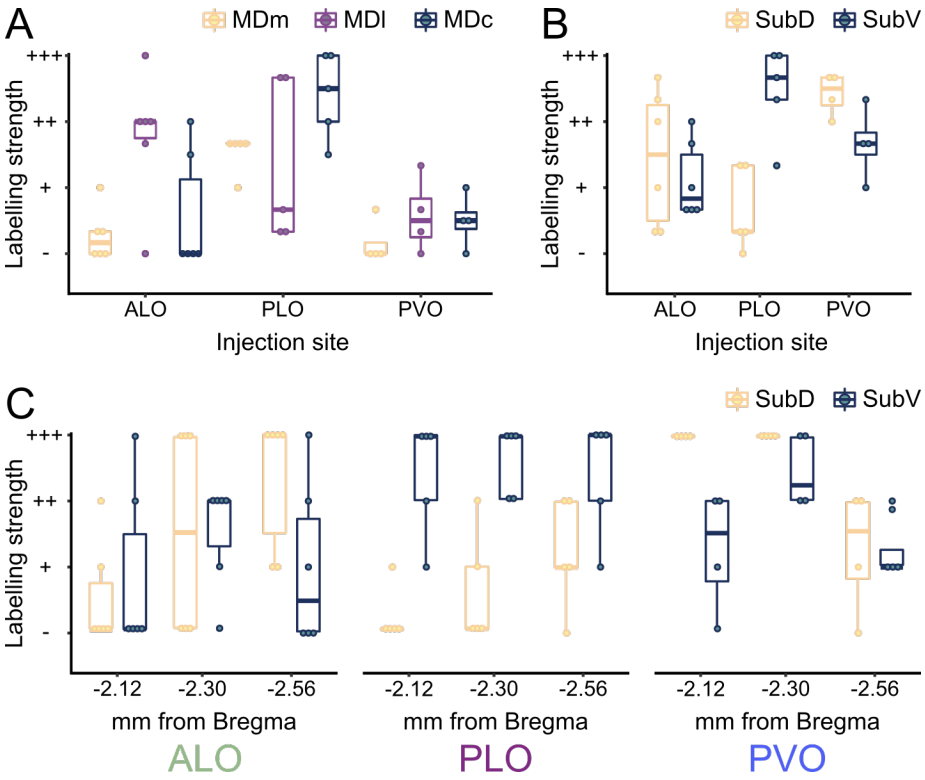


Figure 3.12 Labelling strength in MD and Sub following injection of CTB into ALO, PLO, or PVO. (**A**) Average labelling strength in MD across coronal sections -1.88, -2.12, -2.30, and -2.56 mm from bregma. (**B**) Average labelling strength in Sub across coronal sections -2.12, -2.30, and -2.56 mm from bregma. (**C**) Labelling strength in Sub after ALO, PLO, or PVO injections at coronal sections -2.12, -2.30, or -2.56 mm from bregma. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. c: central; l: lateral; m: medial; D: dorsal; V: ventral; ALO: lateral orbitofrontal cortex, anterior area; CTB: cholera toxin subunit B; MD: mediodorsal nucleus of the thalamus; PLO: lateral orbitofrontal cortex, posterior area; PVO: ventral orbitofrontal cortex, posterior area; Sub: submedius nucleus of the thalamus.

3.3.3 Temporal lobe afferents

Lateral entorhinal cortex exhibited weak labelling following injection into either ALO or PLO but stronger labelling following PVO injection. Labelling was also detected in the perirhinal cortex (PRh) after tracer injection into PVO but was not consistently present following ALO or PLO injections (see Figure 3.4 and 3.5). In summary, although labelling

in the entorhinal and perirhinal was not always consistent, the labelling patterns observed suggest a gradient between their inputs into LO and VO.

Amygdala

A detailed schematic for the amygdala is shown in Figures 3.13. The anterior part of the basolateral amygdala (BLA) exhibited moderate labelling following PLO injection but only weak labelling after either ALO or PVO injections (Figure 3.8). We observed a similar pattern in the dorsolateral part of the lateral amygdala (LaDL), with the difference that it did not contain any labelled cells following PVO injection. We only detected weak labelling in the posterior part of the basolateral amygdala (BLP) and the ventromedial (LaVM) and ventrolateral (LaVL) parts of the lateral amygdala following tracer injection into PLO. There were no labelled cells outside of the basolateral and lateral regions of the amygdala. Overall, amygdala labelled cells were most prevalent in BLA and following injection into PLO (Table 3.2), suggesting both an anterior–posterior gradient within LO and a medial-lateral gradient between PLO and PVO.

Supporting these observations, statistical analysis of the average labelling strength across coronal sections (Figure 3.13-C) revealed a main effect of injection site (ALO, PLO, PVO; $F_{(3, 47.1)} = 8.38$, $p < .001$), with PLO receiving stronger amygdala projections than ALO (PLO vs ALO: $t_{(47.9)} = 3.73$, $p = .002$) or PVO (PLO vs PVO: $t_{(45.8)} = 3.30$, $p = .006$), and ALO and PVO receiving similar strength of projections (ALO vs PVO: $t_{(47.4)} = -0.23$, $p = .994$). There was also a main effect of amygdala portion (BLA, BLP, LaVL, LaVM; $F_{(3, 42.4)} = 15.76$, $p < .001$), with BLA sending the strongest projections to OFC (BLA vs BLP: $t_{(42.4)} = 4.40$, $p < .001$; BLA vs LaVL: $t_{(42.4)} = 4.40$, $p < .001$; BLA vs LaDL: $t_{(42.4)} = 2.68$, $p = .061$), and LaDL sending stronger projections than LaVL (LaDL vs LaDL: $t_{(42.4)} = 3.97$, $p = .002$). However,

the strength of projections from these amygdala portions to OFC subregions did not differ significantly (amygdala portion x injection site interaction, $F_{(6, 42.4)} = 1.84$, $p = .115$).

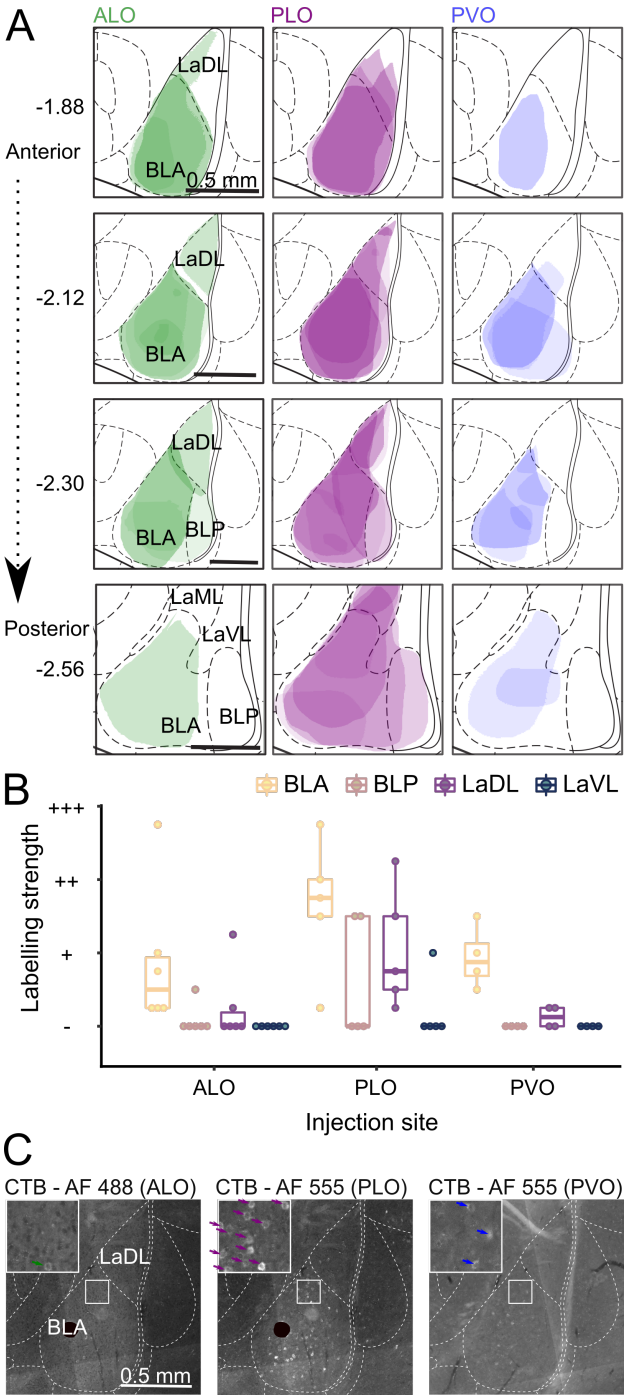


Figure 3.13 Labelling in the amygdala following CTB injection into ALO, PLO, or PVO. (**A**) Shaded areas correspond to superimposed hand drawings of the labelling in each brain included in the analysis, using an opacity value proportional to the labelling density level. *n*

ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. (**B**) Average labelling strength in regions of the amygdala across coronal sections -1.88, -2.12, -2.30 and -2.56 mm from bregma. (**C**) Micrographs from representative labelled brains, at -2.12 mm from bregma. We have masked with a black circle a bright artefact present in the slices representing labelling after ALO and PLO injections. *n* ALO = 6, *n* PLO = 5, *n* PVO = 4 injections. ALO: lateral orbitofrontal cortex, anterior area; CTB: cholera toxin subunit B; BLA: basolateral amygdala, anterior part; BLP: basolateral amygdala, posterior part; LaDL: lateral amygdala, dorsolateral part; LaML: lateral amygdala, mediolateral part; LaVL, lateral amygdala, ventrolateral part; PLO: lateral orbitofrontal cortex, posterior area; PVO: ventral orbitofrontal cortex, posterior area.

Table 3.2

Density of labelling in key areas observed following CTB injection into ALO, PLO, or PVO.

Structures	ALO	PLO	PVO
Prefrontal cortex			
orbitofrontal cortex, medial part	-	+	+
prelimbic cortex	-	+	+
infralimbic cortex	-	+	+
anterior cingulate cortex, area 1	-	-	+
motor cortex	++	-	+
Sensory cortices			
gustatory cortex	+	+	+
piriform cortex	++	++	++
primary somatosensory cortex			
barrel cortex	+	-	++
forelimb, hindlimb & trunk areas	-	-	+
secondary visual cortex			
lateral area, mediolateral & mediomedial areas	-	-	+
auditory cortex	-	-	-
Thalamic nuclei			
mediodorsal			
medial part	+	++	-
central part	+	++	+
lateral part	++	+	+
submedius			
anterior section (-2.12)*			
dorsal part	+	-	++
ventral part	+	++	+
middle section (-2.30)*			
dorsal part	++	+	+++
ventral part	++	+++	+++
posterior section (-2.56)*			
dorsal part	++	+	+
ventral part	+	++	+
paratenial	-	++	-
reuniens	-	+	+

Amygdala			
basolateral			
anterior part	+	++	+
posterior part	-	+	-
lateral	-	+	-

Note. Only regions where main differences and relevant similarities were observed are represented. Because we observed no consistent differences in labelling density between the anterior and posterior portions of MO, here we collapsed its average labelling across all coronal slices. -, absence of labelling; +, weak labelling; ++, moderate labelling; +++, strong labelling; *, distance from bregma in mm; CTB: cholera toxin subunit B.

3.4 Discussion

Here, we tested the hypothesis that OFC functional heterogeneity predicts meaningful differences in connectivity by simultaneously characterising the patterns of inputs of the anterior and posterior portions of LO, which have been previously found to be functional distinct (Panayi & Killcross, 2018). Additionally, we contrast the input patterns into ALO and PLO with inputs into PVO, an anatomically adjacent OFC portion that is thought to be functionally distinct (Balleine et al., 2011; Corwin et al., 1994). Our approach allowed us to assess the extent of overlap, convergence and divergence between the projections of these OFC subdivisions and, thus, define their topographic relationships. By using the same tracer to compare the connectivity across the subdivisions, we minimised potential interpretation problems caused by variability in uptake or transport by the cells, spread in the tissue or extent of local necrosis.

The OFC receives its main inputs from the amygdala, thalamic nuclei, periaqueductal grey, and midbrain dopamine neurons (e.g. Krettek & Price, 1977b; Mcdonald, 1998; Murphy & Deutch, 2018; Ongur & Price, 2000). Our neuroanatomical characterisation revealed substantial differences in cortical and sub-cortical inputs into ALO, PLO, and PVO. Specifically, we identified robust differences in the topographic

organisation and in the gradation of connectivity strength into the OFC subdivisions investigated. Such differences were observed in the inputs from the medial prefrontal cortex, motor cortex, sensory cortices, amygdala, and thalamus (Figure 3.14). We did not assess midbrain inputs to OFC in our study. However, there are moderate dopaminergic projections from the midbrain, including ventral tegmental area, dorsal raphe nucleus and ventral periaqueductal grey to posterior LO and laterally adjacent AI (Murphy & Deutch, 2018), as well as noradrenergic projections from the locus coeruleus to LO, VO and MO (Cerpa et al., 2019) and to AI (Gerfen & Clavier, 1979).

3.4.1 Distinct and topographically organised thalamic inputs to OFC subdivisions

Thalamo-cortical connectivity has historically been one of the key criteria used to segregate cortical regions and define functional circuits (Alexander, 1986; Rose & Woolsey, 1948). Medial, lateral, and central nuclei of the MD thalamus receive partially overlapping but distinct cortical and subcortical afferents (Groenewegen, 1988). In this study, retrogradely-labelled cells in MD thalamus showed a notably separate pattern of projections into ALO, PLO, and PVO. In fact, the distinction between ALO and PLO was just as clear as the medial-lateral distinction between PLO and PVO.

The Sub shares strong reciprocal connections with the OFC, especially with VO/LO (Alcaraz et al., 2015; Coffield et al., 1992; Kuramoto et al., 2017; Reep et al., 1996; Yoshida et al., 1992). In this tracing study, we observed the most pronounced labelling in this area, with all portions of this thalamic nucleus projecting to every region of the OFC that we investigated (Figures 3.11, 3.12). Again, rather than a uniform connectivity pattern, we observed clear spatial segregation in relation to where in the Sub the strongest projections originate. The distinct medial-lateral gradient of OFC inputs from Sub we observed has been previously reported (Alcaraz et al., 2015; Reep et al., 1996). However,

we also identified a novel pattern of anterior–posterior projections from Sub which provides the clearest evidence of a distinction in afferents across the three OFC subregions we investigated in this study: ALO, PLO, and PVO. Our results revealed that PVO receives most Sub projections from the most anterior half of the Sub, especially from the dorsal portion, and slightly less from the ventral part. Intriguingly, results from previous tracing studies (Kuramoto et al., 2017; Reep et al., 1996) revealed that AVO receives inputs from SubV, suggesting a potential anterior–posterior distinction might also be present here. Such distinction within VO is also supported by stronger noradrenergic projections from the locus coeruleus to PVO than to AVO (Cerpa et al., 2019).

The OFC is the main recipient of Sub projections, and, in addition to the subdivisions already mentioned, Sub also sends projections to MO (Kuramoto et al., 2017; Yoshida et al., 1992). Most of the double- or triple-labelled cells we observed were present in the Sub. This means that, unlike other regions studied here, a proportion of information from Sub neurons is broadcast to ALO, PLO, and PVO. It has previously been reported that there are neurons in this region that simultaneously project to lateral and ventral OFC, potentially linking their functions under certain conditions (Kuramoto et al., 2017). Other than OFC inputs, the Sub receives mainly sensory and pain modulation information (Yoshida et al., 1992). The strong VO/LO-Sub reciprocal connection seems to be involved in a pain modulation pathway (Coffield et al., 1992; Kuramoto et al., 2017; Tang et al., 2009). While little is known about Sub function, recent studies exploring the functional significance of its connectivity with the OFC revealed that an intact Sub is necessary for updating both stimulus-outcome (Alcaraz et al., 2015) and action-outcome associations (Fresno et al., 2019).

We also observed distinct labelling patterns in the paratenial thalamic nucleus, which consistently exhibited moderate labelling following PLO injection but very light to no labelling with injections into ALO (Figure 3.8; Table 2). Following injections of anterograde tracer into PT, Vertes & Hoover (2008) observed a similar distinction in its projections to the anterior and posterior portions of LO. Functional studies investigating PT-OFC interaction might help understand the significance of this OFC anterior–posterior distinction.

3.4.2 Multi-sensory nature of OFC inputs

The lateral orbital network, comprised of VO, LO, and AI, is sometimes also referred to as a sensory network because of its strong multi-modal sensory inputs (Ongur & Price, 2000; Price, 2007). Here, we observed evidence of multi-sensory inputs into PVO. This in agreement with the attentional account of VO proposed by Hoover & Vertes (2011) and the connectivity described by Corwin & Reep (1998), showing that VO receives direct visual and somatosensory but not auditory information. By contrast, primary sensory inputs into both ALO and PLO, other than gustatory and olfactory, are sparse or absent. While the OFC is often considered a site of multi-sensory input, these results highlight a medial-lateral distinction in the diversity of primary sensory inputs into LO and VO.

3.4.3 Distinct amygdala input patterns into OFC subdivisions

OFC and BLA share strong reciprocal connections and their projections are topographically organised in both directions (Mcdonald, 1991; Mcdonald, 1998; Reep et al., 1996). In our tracing experiments, amygdala labelled cells were most prevalent following injection into PLO, originating most strongly in anterior BLA. By contrast, we observed only weak labelling after either ALO or PVO injections. The posterior OFC

medial-lateral difference is consistent with the results obtained by Kita & Kitai (1990) following injections of anterograde tracer into anterior BLA: dense labelling in PLO but not in PVO. OFC function has been closely linked to that of the amygdala, in particular its basolateral nucleus. Both OFC and BLA have been implicated in processes that support outcome-guided behaviour through associative learning of sensory-specific representations and predictive-cues, with lesions of either region impairing performance on outcome devaluation and reversal learning tasks (for reviews, see Balleine & Killcross, 2006; Sharpe & Schoenbaum, 2016). The amygdala has been implicated specifically in tracking previous outcomes and comparing them to current outcomes (Izquierdo et al., 2013; Lichtenberg et al., 2017; Schoenbaum, Setlow, Saddoris, et al., 2003; Stalnaker et al., 2007). Interestingly, the stronger inputs from BLA to PLO we observed shed light on previously observed functional differences between ALO and PLO: posterior but not anterior LO lesions disrupt reversal learning, suggesting PLO is critical for updating the value of expected outcomes (Panayi & Killcross, 2018). Moreover, close inspection of a recent paper showing the necessity of OFC inputs from BLA for using cue-generated reward expectations to guide decision-making shows that these mainly terminated in PLO (Lichtenberg et al., 2017).

Paralleling the observations by Panayi & Killcross (2018), a study in non-human primates found that posterior OFC (area 13) is required for updating outcome valuations during a selective satiety devaluation procedure, whereas anterior OFC (area 11) is necessary for translating this knowledge into action during goal selection (Murray et al., 2015). Stronger inputs from BLA to posterior (area 13a) than to anterior OFC (area 11) have been previously observed in primates (Amaral & Price, 1984). This reinforces the idea, which had been previously hinted at by these functional studies (Murray et al., 2015; Panayi & Killcross, 2018), of similar functional organisation principles between rodent

and primate OFC subdivisions when functional differences along the anterior–posterior axis are considered.

3.4.4 Anterior and posterior LO are part of distinct anatomical and functional circuits

While both PVO and PLO receive projections from regions along the medial wall, including IL, PL, and MO (Figure 3.8; Table 3.2), such projections are mostly absent following injection into ALO. This is consistent with previous reports studying inputs from PL and IL to the OFC which also suggest anterior–posterior differences (Takagishi & Chiba, 1991; Vertes, 2004). Even though PVO and PLO injections were in the same anteroposterior plane and PVO is situated more medially, there were comparable degrees of labelling along the medial wall following injections into either of the two subdivisions. This argues against issues with polysynaptic labelling or tracer spreading within the OFC, especially between PVO and PLO injections. Indeed, we observed that CTB deposits that extended into adjacent AI produced a clearly distinct pattern of inputs to the adjacent PLO, with dense innervation from PL, MO and BLA (see Figure 3.7; consistent with Murphy & Deutch, 2018; Reep et al., 1996).

Several studies have proposed a functional dissociation between OFC and PL in supporting Pavlovian and instrumental conditioning, respectively (Balleine & Dickinson, 1998; Corbit & Balleine, 2003; Ostlund & Balleine, 2005; Ostlund & Balleine, 2007). Both PL and OFC share strong reciprocal connections with the BLA (Mcdonald, 1991; Mcdonald, 1998), which in turn has been shown to play a role in both Pavlovian and instrumental outcome-mediated decision making (Balleine & Killcross, 2006; Johnson et al., 2009). Therefore, the similar distinction in the density of projections from PL and BLA to anterior and posterior LO suggest that BLA→PLO and PL→PLO might form a functional

circuit involved in assessing the motivational significance of actions and stimuli, respectively.

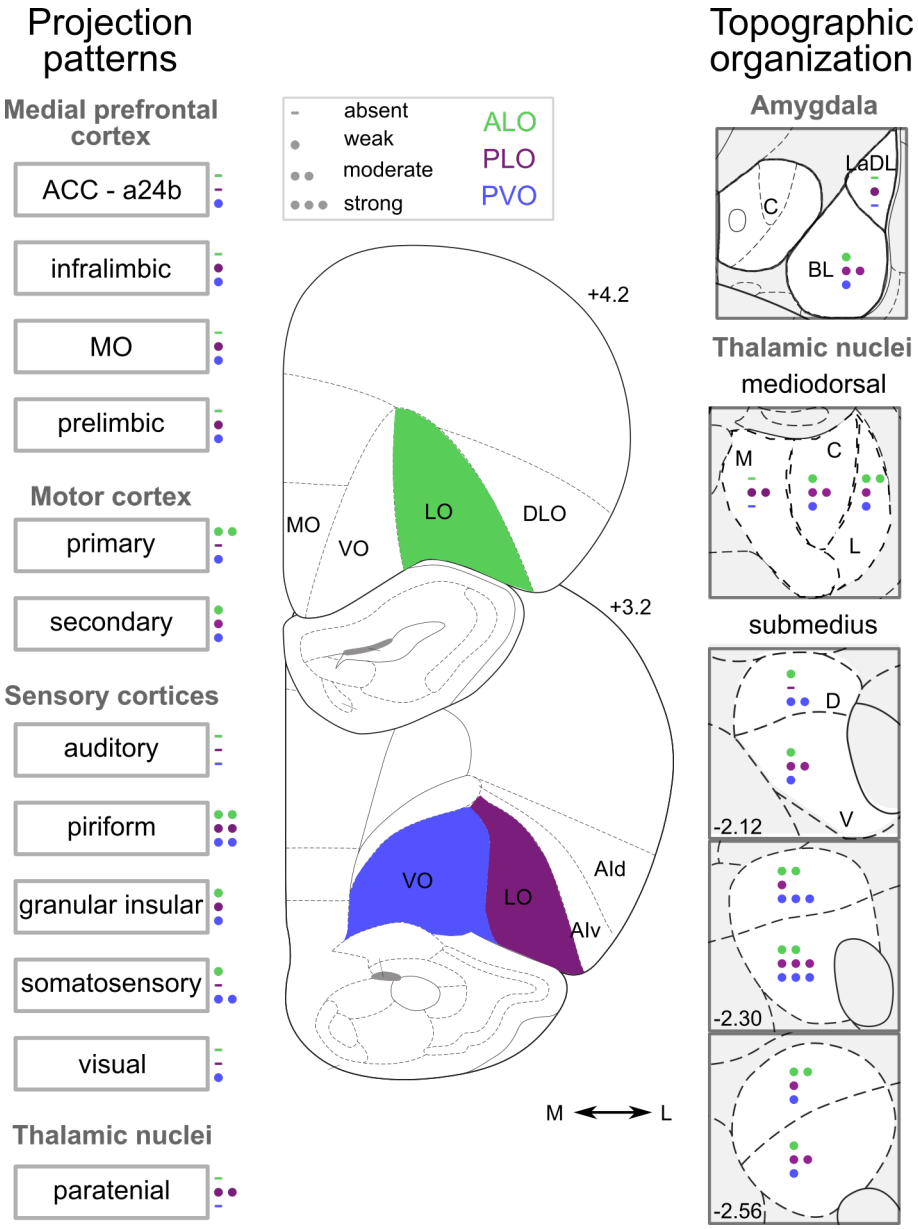


Figure 3.14 Density of key inputs into ALO, PLO and PVO. The different number of circles represent the average density of retrogradely-labelled cells following CTB injection (i.e., absent, weak, moderate or strong) into ALO, PLO, and PVO. Panels depicting labelling in amygdala and MD thalamus represent the average density labelling in all sections quantified and not at that particular coronal section. Distances shown are distances from bregma in mm. a24b: anterior cingulate cortex, area 24b; ACC: anterior cingulate cortex; AId: agranular insular cortex, dorsal part; AIv: agranular insular cortex, ventral part; ALO: lateral orbitofrontal cortex, anterior part; CTB: cholera toxin subunit B; DLO: orbitofrontal cortex, dorsolateral part; IL: infralimbic cortex; LaDL: lateral amygdala, dorsolateral part; LO: orbitofrontal cortex, lateral part; M1: primary motor cortex; M2: secondary motor cortex; MO: orbitofrontal cortex, medial part; PL: prelimbic cortex; PLO: lateral orbitofrontal cortex, posterior part; PVO: ventral

orbitofrontal cortex, posterior part; VO: orbitofrontal cortex, ventral part; A: anterior; P: posterior; BL: basolateral; C: central; M: medial; L: lateral; D: dorsal; V: ventral.

Contributions & acknowledgements

For the studies presented in this chapter, Ines Barreiros (IB), Marios Panayi (MP) and Mark Walton (MEW) planned the experiments; IB collected the data; IB analysed and presented the data with input from MP and MEW.

This work was supported by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship awarded to IB; and by a Wellcome Senior Research Fellowship awarded to MEW (202831/Z/16/Z). We thank Allan Wainman from the Micron Advanced Bioimaging Unit (supported by Wellcome Strategic Awards 091911/B/10/Z and 107457/Z/15/Z) for support and assistance in this work.

4. Causal contribution of dopamine to haemodynamic signatures of reward prediction

4.1 Introduction

In an ever-changing world, individuals need to continuously learn and adapt their behaviour in order to survive. Such decision making is believed to be founded by establishing associations between stimuli or actions and their outcomes, thus creating expectations about what might happen after a particular choice is made. One way to determine when to update these expectations is to detect when these are not met, which elicits reward prediction errors (RPEs) –the difference between expected and actual outcomes. RPEs are believed to support behavioural flexibility by updating stimulus- and action-outcome expectations (Bush & Mosteller, 1951; Montague et al., 1996; Rescorla & Wagner, 1972; Schultz et al., 1997; Sutton & Barto, 1998).

In the brain, the encoding of these teaching signals has long been linked to the activity of midbrain dopamine neurons in the ventral tegmental area (VTA) and substantia nigra pars compacta (Cohen et al., 2012; Mirenowicz & Schultz, 1994; Montague et al., 1996; Roesch et al., 2007; Schultz et al., 1997; Watabe-Uchida et al., 2017) as well as to dopamine release in regions that receive dopaminergic projections, including the nucleus accumbens (NAc) (Blanco-Pozo et al., 2021; Hart et al., 2014). Manipulating dopamine activity or transmission impairs the formation of appropriate stimulus-outcome associations (e.g., Parkinson et al., 2002; Saunders & Robinson, 2012; Schoenbaum & Setlow, 2003) and can also affect reward-guided behavioural flexibility (DenOuden et al., 2013; Hamid et al., 2016; Korn et al., 2021; Parker et al., 2016; Sala-Bayo et al., 2020; Verharen et al., 2018, 2019).

Dopamine release in the NAc in response to reward-predicting cues also signals the subjective benefit associated with them (Berridge & Robinson, 1998; Collins et al., 2016; Flagel et al., 2011; Gan et al., 2010; Papageorgiou et al., 2016; Steinberg et al., 2013). Similarly, NAc neurons also encode value expectations (Averbeck & Costa, 2017; Burton et al., 2015; Day et al., 2011; Roesch et al., 2007; Takahashi et al., 2007; van der Meer et al., 2010), which itself is promoted by dopamine transmission (du Hoffmann & Nicola, 2014; Sugam et al., 2012). Therefore, midbrain dopamine transmission modulates NAc activity, which in turn plays a critical role in reward-guided behavioural flexibility by promoting approach to stimuli or actions with a desirable outcome, particularly in changing or uncertain conditions (as reviewed in Floresco, 2015).

Examining the computations underlying the encoding of reward prediction in the human brain provides a window onto aberrant learning in disorders where dopaminergic signalling dysfunction is a central component (Deserno et al., 2016; Gilmour et al., 2015; Heinz & Schlagenhauf, 2010; Zald & Treadway, 2017). In functional magnetic resonance imaging (fMRI) studies using reward-guided behavioural paradigms, blood oxygenation level-dependent (BOLD) signals in the NAc and ventromedial prefrontal cortex (homologous to the rodent medial frontal and orbitofrontal cortex) were shown to correlate with RPEs (Behrens et al., 2008; Niv et al., 2012; Rutledge et al., 2010). This increase in NAc activity measured in BOLD signals through fMRI has been proposed to be directly modulated by stimulation of dopamine receptors by mesolimbic dopamine release (Knutson & Gibbs, 2007; Lohrenz et al., 2016). However, there has been no direct demonstration of the link between behaviourally derived haemodynamic signatures of RPEs and dopamine. Unlike in rodents, it is not possible to directly measure dopamine release or specifically disrupt its transmission in human subjects. Moreover, rodent fMRI studies are currently largely limited to experiments with anaesthetised or

head restrained animals (e.g., Decot et al., 2017; Lohani et al., 2017). *In vivo* oxygen amperometry, a technique that allows for measuring haemodynamic signals comparable to the BOLD signals of fMRI in behaving animals (Francois et al., 2012), might help bridge the gap between human and rodent studies (Gilmour et al., 2015). Using this technique, Werlen et al. found that haemodynamic signals in the NAc of rats trained on a probabilistic Pavlovian task encode RPEs (Werlen et al., 2019). Here, we attempted to help clarify the link between behaviourally derived haemodynamic signatures of RPEs and dopamine by combining *in vivo* oxygen amperometry recordings with lesioning of NAc dopaminergic inputs.

The orbitofrontal cortex (OFC) is another key node in the anatomical circuits underlying behavioural flexibility. Like the NAc, the OFC encodes the subjective value of expected reward, in addition to information about its probability, identity, magnitude, and the cues predicting it (Bissonette & Roesch, 2015; Padoa-Schioppa & Assad, 2006; Schoenbaum et al., 2009; Stalnaker et al., 2014; van Duuren et al., 2009; Wikenheiser & Schoenbaum, 2016). Similarly to the NAc, OFC signals are scaled by reward expectation during outcome delivery (O'Doherty et al., 2003; Stalnaker et al., 2018; Sul et al., 2010; Tobler et al., 2006; Werlen et al., 2019). However, whereas, for example, responses in the NAc are positively influenced by delivery of reward and negatively influenced by expectation, those in the OFC are scaled by how surprising the outcome is irrespective of valence (i.e., increases in activity to both surprising rewards or reward omissions), potentially reflecting a salience signal (e.g., Ogawa et al., 2013; Werlen et al., 2019). Lesions of the OFC impair behavioural flexibility, particularly under conditions of uncertainty or changing contingencies such as in probabilistic, reversal learning or devaluation tasks (Dalton et al., 2016; Gardner et al., 2017; Kim & Ragozzino, 2005; Parkes et al., 2018; Rudebeck & Murray, 2008; Schoenbaum et al., 2009; Walton et al.,

2010, 2011). The OFC sends glutamatergic projections to VTA and NAc (Heilbronner et al., 2016; Montaron et al., 1996; Vázquez-Borsetti et al., 2009) and might, therefore, directly or indirectly influence mesolimbic dopamine release (Lodge, 2011; Takahashi et al., 2011). Thus, another aim of the study described in this chapter was to investigate how mesolimbic dopamine influences this frontostriatal reward circuit during reward-guided flexible behaviour by measuring OFC and NAc activity simultaneously in rats performing a probabilistic reversal learning task after lesioning of the dopaminergic inputs into NAc.

In sum, the study described here aimed to investigate (1) the causal contribution of dopamine to haemodynamic signatures of RPEs in the NAc (2) and how mesolimbic dopamine influences frontostriatal circuits during reward-guided flexible behaviour. To do this, we used *in vivo* oxygen amperometry to record haemodynamic responses simultaneously in NAc and the lateral OFC in rats performing a probabilistic reversal learning task, half of which had a unilateral lesion of the dopamine terminals in NAc. This allowed us to compare the oxygen signals from rats with and without dopaminergic inputs to NAc as rats used reward expectations to guide choices and received probabilistic rewards, generating RPEs.

4.2 Methods

The experimental procedures described in this chapter were carried out by researchers at Eli Lilly (UK), and the data was analysed by me. Please see Chapter 2 of this thesis for a detailed description of the materials and methods. Below is a brief recap of the most important methodological details.

4.2.1 Probabilistic reversal-learning task

36 male rats were trained on an operant 2-option probabilistic reversal learning task. One option was associated with a 75% probability of reward delivery and the other a 25% reward probability, and this association would reverse each week (Figure 4.1). Each trial consisted of (i) *cue presentation*: 8 s illumination of cue light above either both (choice trials) or just one of the levers (forced trials); (ii) *lever choice*: levers extend, and animals have 10 s to choose one lever; (iii) *outcome anticipation*: levers retract and 5 s magazine light illumination; and (iv) *outcome delivery*: cue and food magazine lights turn off and, on rewarded trials, four 45 mg sucrose pellets were delivered. There was then an inter-trial interval (ITI) of 30-50s. Each session consisted of 112 forced and choice trials, delivered in mini-blocks of 16 trials, consisting of 8 forced trials (4 to each lever, pseudorandom order) followed by 8 choice trials. Rats were trained on the same reward contingencies over a 5-session block, and these were reversed for a subsequent 5-session block.

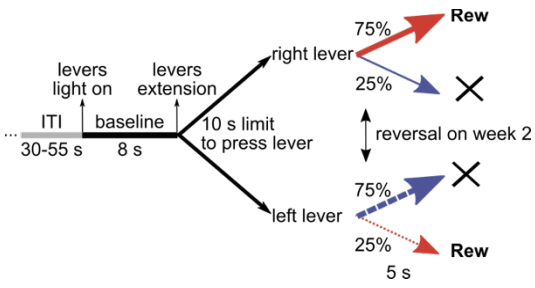


Figure 4.1 Trial structure of probabilistic reversal learning task in choice trials. Reward contingencies were reversed for week 2. In forced trials, only one of the levers was made available. Rew: rewarded.

4.2.2 Surgery

After learning the entire task structure, animals underwent surgical procedures under general anaesthesia. Rats were randomly assigned into two groups and received 6-OHDA

or vehicle (phosphate-buffer saline, PBS) injections into multiple sites in the NAc, unilaterally. Additionally, all rats were implanted with carbon paste electrodes targeted bilaterally at the NAc core and OFC (see Chapter 2 for stereotaxic coordinates details). After a 3-week recovery period, T_{O2} signals were recorded while rats were tested on the behavioural task, first on the same and later on reversed lever-reward contingencies.

4.2.3 O_2 amperometry data recording

O_2 signals were recorded from the NAc and OFC using constant-potential amperometry (−650 mV applied for the duration of the session) as described previously (Francois et al., 2012, 2014).

4.2.4 *In vivo* lesion validation

To provide an assay of the effect of the 6-OHDA lesion *in vivo*, rats received an injection of a non-selective dopamine receptor agonist, apomorphine (0.1 mg/kg, s.c.), on the day before the first and after the last post-surgery testing session.

4.2.5 *Histology*

At the end of experimentation, animals were perfused transcardially and the brains removed. To assess the placement of the electrode tips, brain sections were mounted on glass slides, stained with cresyl violet and viewed microscopically. Signals obtained from any electrode with improper placement or which the track could not be visualised were excluded from all analyses.

To assess the extent of the 6-OHDA lesion, slices were stained for tyrosine hydroxylase (TH). Mean intensity measurements taken from regions of interest on microscope images of these slices were then used to calculate the % reduction in TH

labelling intensity on the lesioned hemisphere. Lesions were considered successful for subjects whose average reduction in TH labelling in the coronal slice closest to the NAc electrode placement was greater than 50%. Subjects with unsuccessful lesions were excluded from subsequent analyses.

4.2.6 *Data Analysis*

Behaviour

To assess the effect of 6-OHDA lesion on behaviour, we analysed the % correct choices from the sham and 6-OHDA lesion groups animals during the post-surgery behavioural training sessions and compared performance before and after reversal of reward contingencies. Additionally, we assessed the number of trials to reach an 85% correct choices criterion (in an exponential moving average (EMA), with a time constant of 6 trials, i.e., only performance on the last 6 trials was considered for this measure). We also compared the reaction times for lever pressing between the two groups.

Behaviour modelling

Trial-by-trial lever choices were modelled using variations of a Rescorla-Wagner model (Rescorla & Wagner, 1972). We started with a model with two free parameters: learning rate α and inverse temperature T. We then compared this model against others that also included free parameters specifying (a) different learning rates for rewarded α_{rew} and unrewarded α_{unrew} choices, (b) lever preference bias T, (c) perseveration P. To capture additional trial-by-trial variance, we also included recency-weighted lever choices rates M. To compare the models, we used the Bayesian information criterion (BIC), which penalises the likelihood of a model by the number of parameters and the natural

logarithm of the number of data points. The model with the lowest BIC score was deemed to give the best fit of the data.

Amperometry

Recordings from each channel were analysed separately. Data were pre-processed to remove noisy or non-meaningful signals and aligned and normalised at relevant behavioural time points (see Chapter 2 for details of exclusions). We then used these signals to investigate the relationship between the T_{O2} and the subjects' behaviour and compare the signal dynamics at key behavioural events (e.g., right after reversal of reward contingencies vs after animals have learned the new contingencies) or between sham and 6-OHDA lesioned subjects. We performed two sets of complementary analyses: (i) analyses of the signals' dynamics using the average T_{O2} amplitude in NAc and OFC in each trial (taking a 45 s window from cue onset), before or after reversal of the lever-reward probability contingencies; and (ii) regression of the same signals against estimates from our computational model using the model with the lowest BIC score (Figure 4.3-F).

4.3 Results

4.3.1 6-OHDA injection in NAc reduces TH labelling in NAc and OFC

We first wanted to assess the average extent of the NAc dopamine lesion by calculating the % reduction in TH labelling intensity over the NAc core and shell and comparing this with another non-targeted region of the striatum, dorsolateral striatum (DLS) (see Chapter 2 for details; Figure 4.2-A, B). We calculated this using mean intensity measurements obtained from regions of interest drawn over the whole of the NAc shell

and core from all subjects for which it was possible to obtain intact brain slices (*n* sham = 17, *n* 6-OHDA = 16). A repeated-measures ANOVA with between-subjects' factors of lesion (sham, 6-OHDA) and within-subjects factors of anterior-posterior (AP) coordinate (+1.7, +1.2, + 0.7 mm) and NAc subdivision (core, shell) showed a reduction in TH labelling on 6-OHDA group (significant effect of lesion with ($F_{(1, 31)}$ = 123.97, p < .001), as well as an AP coordinate x lesion group interaction ($F_{(1, 62)}$ = 25.34, p < .001). Follow-up pairwise comparisons revealed greater lesion damage in the anterior plane on the 6-OHDA lesion group (+1.7 mm vs +1.2 mm, p < .001; +1.7 mm vs +0.7 mm, p < .001; +1.2 mm vs +0.7 mm, p = .036; 4.2-D). The lesion was relatively well confined to the ventromedial striatum as a comparable analysis in DLS did not reveal any differences (all F < 0.39, p > .246). Moreover, there was no evidence of spread to the contralateral hemisphere as a comparison of the mean labelling TH intensity measures from these brain regions across non-lesion hemispheres of sham and 6-OHDA lesion groups revealed no main effects or interaction with group (NAc: all F < 1.34, p > .255; DLS: all F < 0.77, p > .464).

At an individual level, 13 out of the 17 lesions assessed had >50% reduction in the coronal slice closest to the NAc electrode placement (Figure 4.2-C left), so further behavioural and oxygen amperometry analyses focused on these 13 rats.

We used a similar approach to test whether the 6-OHDA lesion also affected projections to the lateral OFC. Unexpectedly, this also revealed a significant reduction in TH labelling in the lesion side OFC of 6-OHDA lesioned animals (6-OHDA vs sham groups, $t_{(31)}$ = 6.91, p < .001, independent samples t-test; Figure 4.2-E, F). A reduction in TH labelling was also noticeable in medial prefrontal cortex structures (e.g., medial OFC, prelimbic cortex; see Figure 4.2-E, not quantified).

Taken together, these results show that the unilateral injection of 6-OHDA into the NAc significantly reduced the dopaminergic projections to this region. This effect did not extend to the DLS nor contralateral NAc but did disrupt cortical dopamine projections.

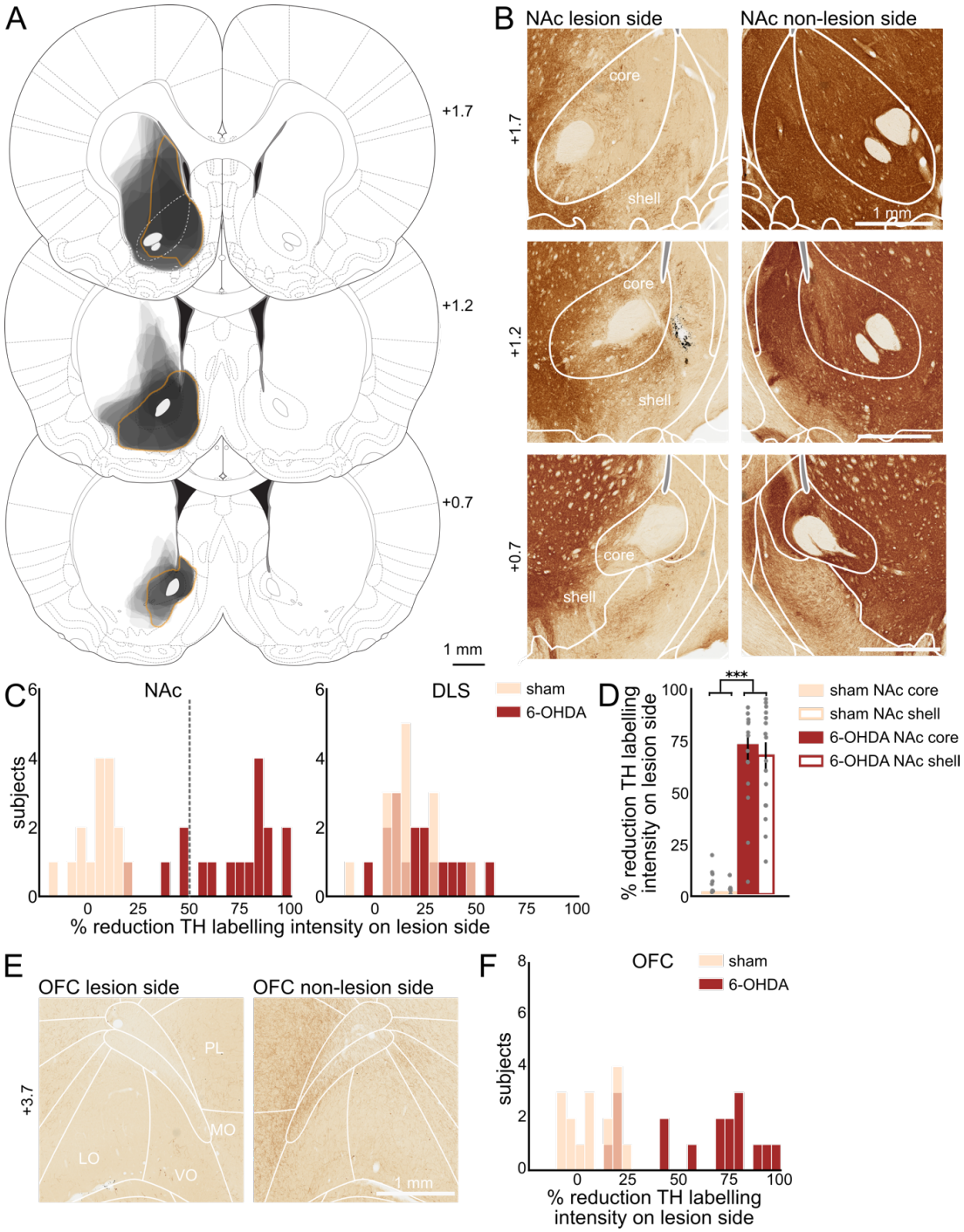


Figure 4.2 Extent of 6-OHDA lesion. 6-OHDA was administered by local intracranial injection targeting the NAc. (**A**) Semi-transparent, grey patches represent lesion damage in a single subject, and darker areas represent overlapping damage across multiple subjects. Orange

contour depicts representative lesion extent at each anterior–posterior coordinate analysed. (**B**) Representative NAc lesion damage in the lesion (left) and non-lesion sides (right). (**C**) Histograms of the % reduction in tyrosine hydroxylase (TH) labelling intensity on the lesion side of the NAc (core and shell) or the DLS in sham and 6-OHDA lesioned subjects in the coronal slice closest to the electrode placement. The vertical dashed line represents the threshold in labelling intensity reduction above which lesions were considered successful (50%). (**D**) % reduction TH labelling intensity in the NAc (core or shell) across coronal planes. (**E**) Representative OFC lesion damage in the lesion (left) and non-lesion sides (right). (**F**) Same as C, for TH labelling in OFC. Error bars depict mean ± SEM. Scatter plots overlaid with bar graphs in panels D represent individual animals averages. n = 33 (sham = 16, 6-OHDA = 17). One asterisk (*) indicates $p < .05$; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests. Coronal sections are identified in mm relative to bregma. 6-OHDA: 6-hydroxydopamine; DLS: dorsolateral striatum; OFC: orbitofrontal cortex; LO: orbitofrontal cortex, lateral part; MO: orbitofrontal cortex, medial part; NAc: nucleus accumbens; PL: prelimbic cortex; TH: tyrosine hydroxylase; VO: orbitofrontal cortex, ventral part.

4.3.2 Unilateral NAc lesion does not affect behavioural performance in probabilistic reversal-learning task

We performed unilateral NAc dopamine lesions to attempt to avoid confounding changes in haemodynamic signals with any changes in behaviour (Burton et al., 2014; Takahashi et al., 2016). To confirm whether this was achieved, we compared the % correct choices in the lesion and sham groups in each post-surgery training session pre- and post-reversal, which revealed no main effects or interaction with group (all $F < 0.23$, $p > .978$, Figure 4.3-A). Comparing the average % correct choices before or after reversal of reward probability contingencies also did not reveal differences between the two groups (all $F < 0.07$, $p > .794$; Figure 4.3-B). Additionally, post-reversal, we observed no significant differences when comparing the number of trials to reach an 85% correct responses criterion (EMA, time constant = 6 trials; $t_{(29)} = 0.54$, $p = .596$, following independent samples t-test; Figure 4.3-C). We also built a simple reinforcement learning model and assessed how well models with different parameters could predict the rats' choices. This showed that the best fit model for both groups was a standard model including a single reward learning rate as the only free parameter other than the inverse temperature

(Figure 4.3-F). This suggests that the rats were learning, updating their subjective value of the chosen lever based on each reward delivery or omission.

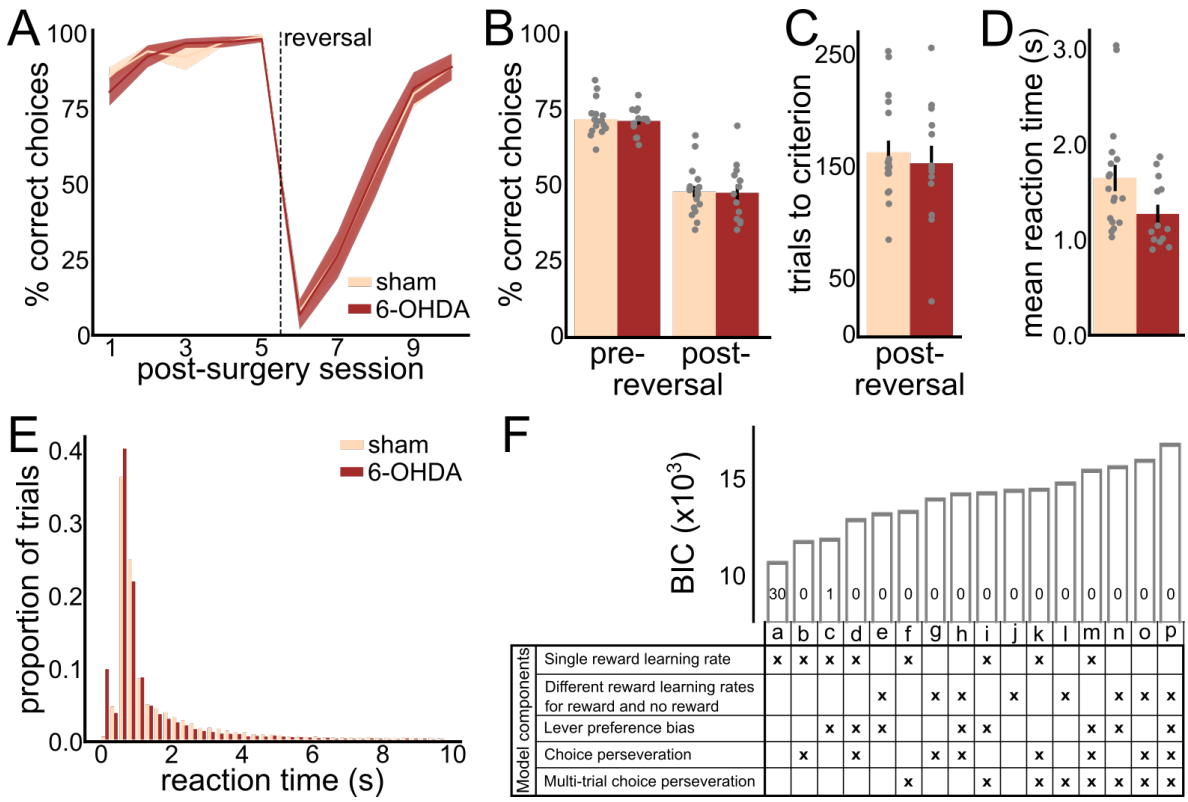


Figure 4.3 Behavioural performance and modelling. (**A**) Mean correct responses (i.e., choice of the lever with the higher probability of reward) of each experimental group across behavioural sessions during the testing phase. (**B**) Mean total correct responses of sham and 6-OHDA lesion groups during all sessions before or after reversal. (**C**) Mean number of trials to reach a criterion of 85% correct responses (in an exponential moving average, time constant = 6). (**D**) Mean reaction time of subjects in each experimental group across all sessions. (**E**) Density histogram depicting the distribution of reaction times of all trials of each experimental group (bin size = 0.25 s). (**F**) Bayesian information criterion (BIC, a measure of the goodness of fit of the model) estimates for different reinforcement learning models predicting the rats' choices on all post-surgery training sessions. BIC penalises the likelihood of a model by the number of parameters and the natural logarithm of the number of data points. The model with the lowest BIC score was deemed to give the best fit of the data. An 'x' in the table indicates the presence of the given component in the model. Numbers within each bar indicate the number of animals for which the given model had the lowest BIC. Represented is the mean ± SEM. Scatter plots overlaid with bar graphs in panels B-D represent individual animals averages. See text for details on statistical tests. *n* = 31 (sham = 18, 6-OHDA = 13). 6-OHDA: 6-hydroxydopamine; BIC: Bayesian information criterion.

Additionally, we compared lever response times in the sham and 6-OHDA lesion groups. This revealed a small difference ($t_{(29)} = 2.06$, $p = .049$, independent samples t-test; Figure 4.3-D), driven primarily by an increase in fast (< 0.25 s) responses in a subset of the 6-OHDA lesioned animals (1 sham and 4 6-OHDA subjects were responsible >97% of these, Figure 4.3-E). Close inspection of the behaviour of these animals suggests that these very fast responses were related to a side bias during anti-bias maintenance training sessions in early stage of the training.

4.3.3 *In vivo* validation of NAc 6-OHDA lesion

To generate an *in vivo* measure of the lesion on haemodynamic responses and to observe whether there were any changes over the course of the experiment, we investigated the effect of a systemic injection of the dopamine agonist apomorphine. Since 6-OHDA lesioning of dopaminergic neurons leads to sensitisation of any remaining dopamine receptors (Schoenfeld & Uretsky, 1972), we expected to observed potentiated responses from ipsilateral NAc after apomorphine injection. There were two apomorphine challenges: one before the first post-surgery session (APO1) and another after the last one (APO2; Figure 4.4-B, top).

We compared the area under the curve during the first 15 min of amperometry recordings following apomorphine injection in sham and 6-OHDA lesioned subjects. Our analysis revealed a group x hemisphere interaction ($F_{(1, 68.76)} = 8.88$, $p = .004$). As can be observed in Figure 4.4, there was a marked increase in T_{O2} signals in the NAc of the 6-OHDA lesioned animals after apomorphine in the lesioned hemisphere ($F_{(1, 64.61)} = 17.98$, $p < .001$) but not contralaterally ($F_{(1, 73.5)} = 0.06$, $p = .801$; see Chapter 2 for details on statistical analysis; Figure 4.4-C). Importantly, while the effects were less sustained in the post-reversal test (APO2), there were no main effects or interaction with time of testing

(NAc: all $F < 1.70$, $p > .197$). By contrast, T_{O2} responses in the OFC did not differ significantly between sham and 6-OHDA lesioned subjects (all $F < 0.95$, $p > .34$). These results validate the functional effect of the 6-OHDA lesion in the NAc and suggest that it could modulate NAc haemodynamic responses throughout behavioural testing.

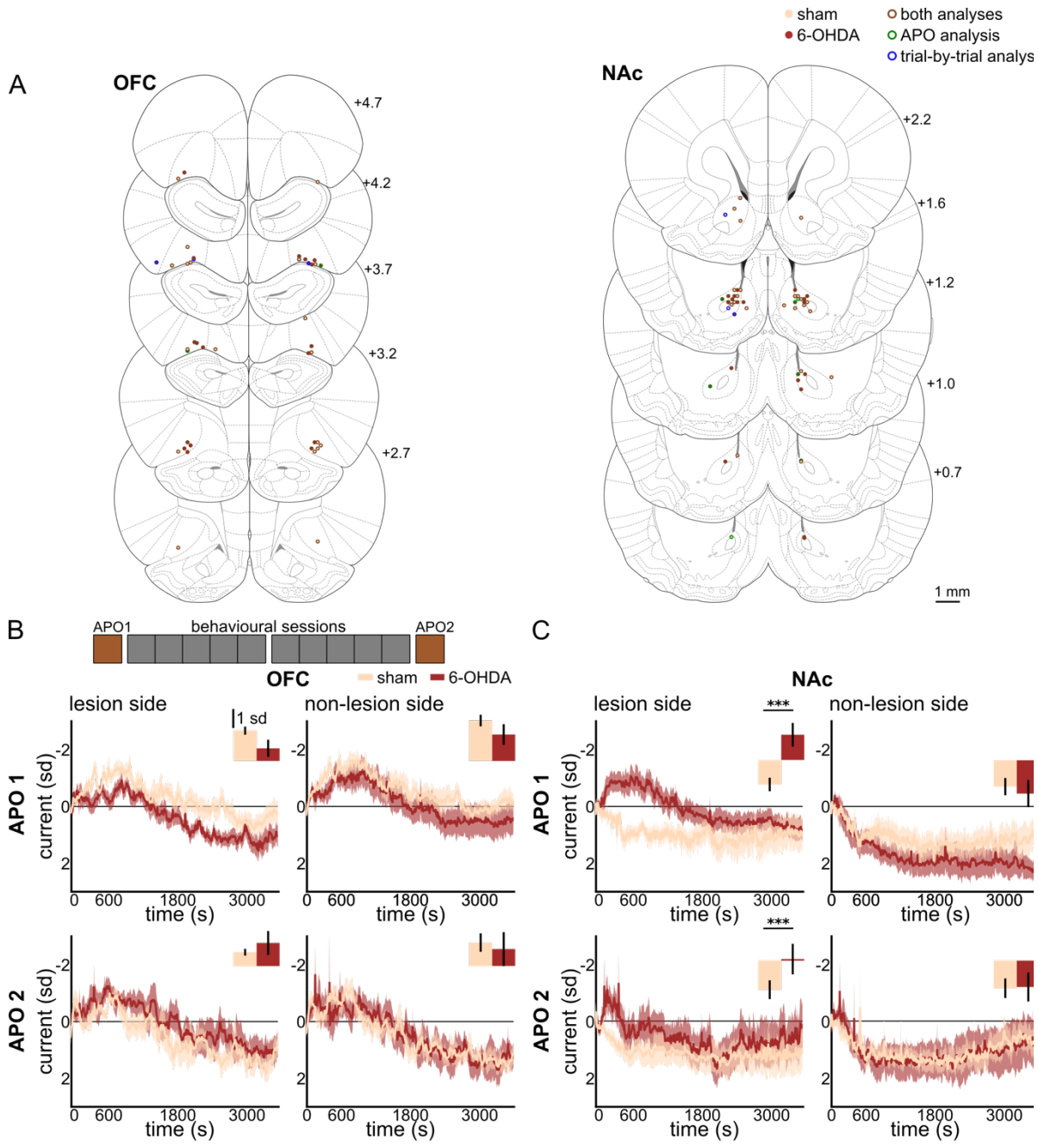


Figure 4.4 Recordings of haemodynamic signals in OFC and NAc and *in vivo* lesion validation with apomorphine. (**A**) Electrode location in OFC (top) and NAc (bottom) of sham

and 6-OHDA groups. Only electrodes from which amperometry or trial-by-trial T_{O2} data were used are represented. OFC electrodes were clustered mainly in the lateral OFC, while NAc electrodes were mainly in the dorsomedial NAc, including dorsomedial core and shell. *n* = 30 subjects. Distances shown are distances from bregma in mm. (**B**) T_{O2} amperometry signals obtained from each channel in the OFC during *in vivo* lesion validation performed before the first (APO 1) and after the last post-surgery behavioural session (APO 2). Apomorphine was administered subcutaneously before the start of amperometry recordings. Bar plots represent the average change in current during the first 15 min after apomorphine injection (t = 0). (**C**) Same as B, but for the NAc signals. Shading or error bars depict mean ± SEM. *n* = 30 (sham = 17, 6-OHDA = 13). One asterisk (*) indicates $p < .05$; otherwise, *m* asterisks indicate $p < 10^{-m}$. See text for details on statistical tests. 6-OHDA: 6-hydroxidopamine; NAc: nucleus accumbens; OFC: orbitofrontal cortex; TH: tyrosine hydroxylase.

4.3.4 NAc and OFC track outcome but not outcome estimates

To determine responses to reward predictions and RPEs, we first characterised the dynamics of the T_{O2} signals recorded from NAc or OFC in response to rewarded and unrewarded outcomes after choosing the high- or the low-reward probability levers during the first post-surgery week of behavioural testing, i.e., before reversal of the lever reward contingencies, in sham animals (Figure 4.5).

As can be observed, NAc signals increased as soon as the cues were presented and during reward expectation (i.e., after lever pressing) and then diverged based on the trial outcome (Figure 4.5-D). By contrast, OFC signals responded primarily to reward expectation and then showed a much stronger prolonged response on rewarded than reward omission trials (Figure 4.5-B). Our analysis of the amplitude of the signals revealed that rewarded trials produced responses with markedly larger amplitudes than unrewarded ones (main effect of outcome: $F_{(1, 41)} = 11.06$, $p = .002$; Figure 4.5-B, D). This effect was particularly prominent in the OFC (outcome x brain region interaction: $F_{(1, 41)} = 23.07$, $p < .001$; $F_{(1, 41)} = 60.88$, $p < .001$ for OFC; $F_{(1, 41)} = 3.35$, $p = .074$ for NAc, reward vs unrewarded trials). One key characteristic of RPE is that the outcome signals should scale with expectation: larger increases after reward on low probability trials and larger

decreases after reward omission on high probability trials. However, this was not apparent in either region. In the OFC, but not the NAc, the reward probability associated with the chosen lever influenced the amplitude of T_{O2} responses, with outcomes on low probability trials generally slightly larger than on high probability trials (reward probability x brain region interaction: $F_{(1, 41)} = 6.18$, $p = .017$; OFC: $p = .013$; NAc: $p = .446$). However, in neither region did this interact with outcome ($F < 0.66$, $p > 0.420$).

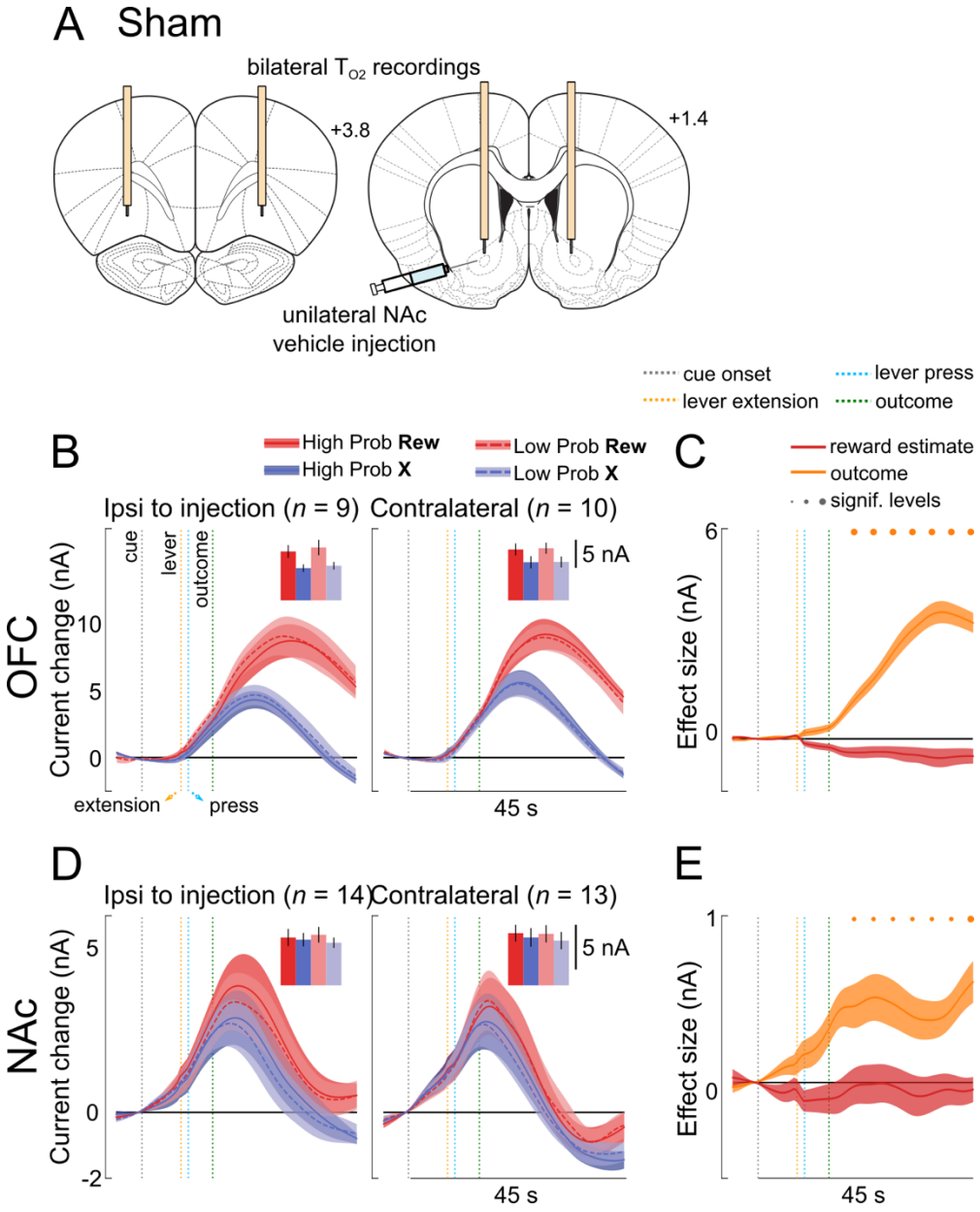


Figure 4.5 Pre-reversal, trial-by-trial haemodynamic responses in sham lesioned rats. (**A**) Electrodes were implanted bilaterally into the OFC and NAc for amperometric tissue

oxygen recordings. Additionally, sham lesioned rats received vehicle injections into the NAc, unilaterally. (**B**, **C**) T_{O2} haemodynamic responses time-locked to cue onset recorded from either OFC (B) or NAc (C) of sham lesioned subjects on training sessions before reversal of lever reward probability contingencies. Inset bar plots represent the maximum amplitude range in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. (**C**, **E**) Average effect size from a general linear model relating T_{O2} responses from both hemispheres from either OFC (C) or NAc (E) to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. Ipsi: ipsilateral; NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

There were no significant differences between the signals in the hemisphere that had the vehicle injection and the contralateral side (all $F < 2.01$, $p > .164$). Therefore, in subsequent analyses, we describe the data collapsed across hemispheres.

We next performed a companion analysis by using a linear regression to unpick the specific influences on the trial-by-trial T_{O2} responses. We included explanatory variables for the value of the chosen option (derived from the reinforcement learning model), trial outcome, as well as other regressors capturing the trial type (i.e., choice or forced), outcome on the previous trial, and inter-time interval at the end of the previous trial, and an interaction term between these last two regressors, as well as a constant term that describes the common activation pattern across all trials (Figure 4.6). For an RPE there should be a *positive* influence of reward expectation when new information is first presented (i.e., cue onset → response), and then at outcome time, both a *negative* influence of reward estimate and a *positive* influence of outcome (Caplin & Dean, 2008; Rutledge et al., 2010). As can be seen in Figure 4.5-C and E, the primary determinant for signal change in both OFC and NAc was whether a trial was rewarded, with a strong positive effect of outcome. By contrast, reward estimate showed only a small negative

influence after outcome delivery in OFC signals whose statistical significance did not survive multiple-comparisons correction.

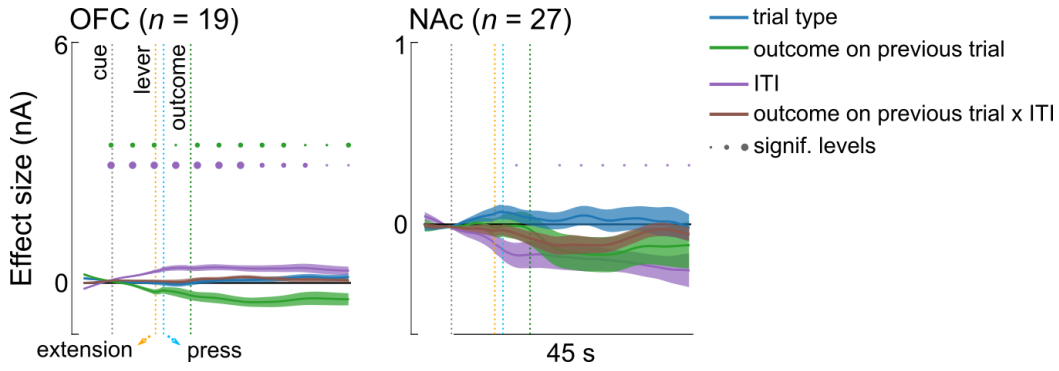


Figure 4.6 Additional explanatory regressors included. Average effect size from a general linear model relating T_{O2} responses from both hemispheres from either OFC (left) or NAc (right) to trial-by-trial trial type (i.e., choice or forced trial), outcome on the previous trial, inter-time interval at the of the previous trial (ITI), and an interaction term between the two. Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

4.3.5 Lesioning of dopaminergic inputs to NAc attenuates responses in the NAc and potentiates those in the OFC

Having established the effects of outcome and reward estimates in NAc and OFC, we next investigated how lesioning of the dopaminergic inputs to NAc influence the encoding of these trial features. We first compared responses recorded ipsilaterally to the 6-OHDA lesion with those recorded ipsilaterally to the sham lesion during pre-reversal sessions (Figure 4.7). Analysis of the T_{O2} responses amplitude revealed a lesion x brain region interaction ($F_{(1, 37)} = 12.05$, $p = .001$). Follow-up pairwise comparisons revealed that while the effect of lesion type in NAc responses was not significant (sham: 5.04 nA [95% CI 3.36–6.73]; 6-OHDA: 3.38 nA [95% CI 1.38–5.39], sham vs 6-OHDA: $p = .210$; Figure 4.7-D), signals in the OFC of 6-OHDA lesioned subjects were significantly larger than those of sham lesioned subjects ($p = .001$, Figure 4.7-B). A linear regression analysis of these

responses also supported these observations. T_{O2} responses in both regions were mainly driven by outcome after 6-OHDA lesion (Figure 4.7-C, E), similarly to what we observed in sham lesioned subjects.

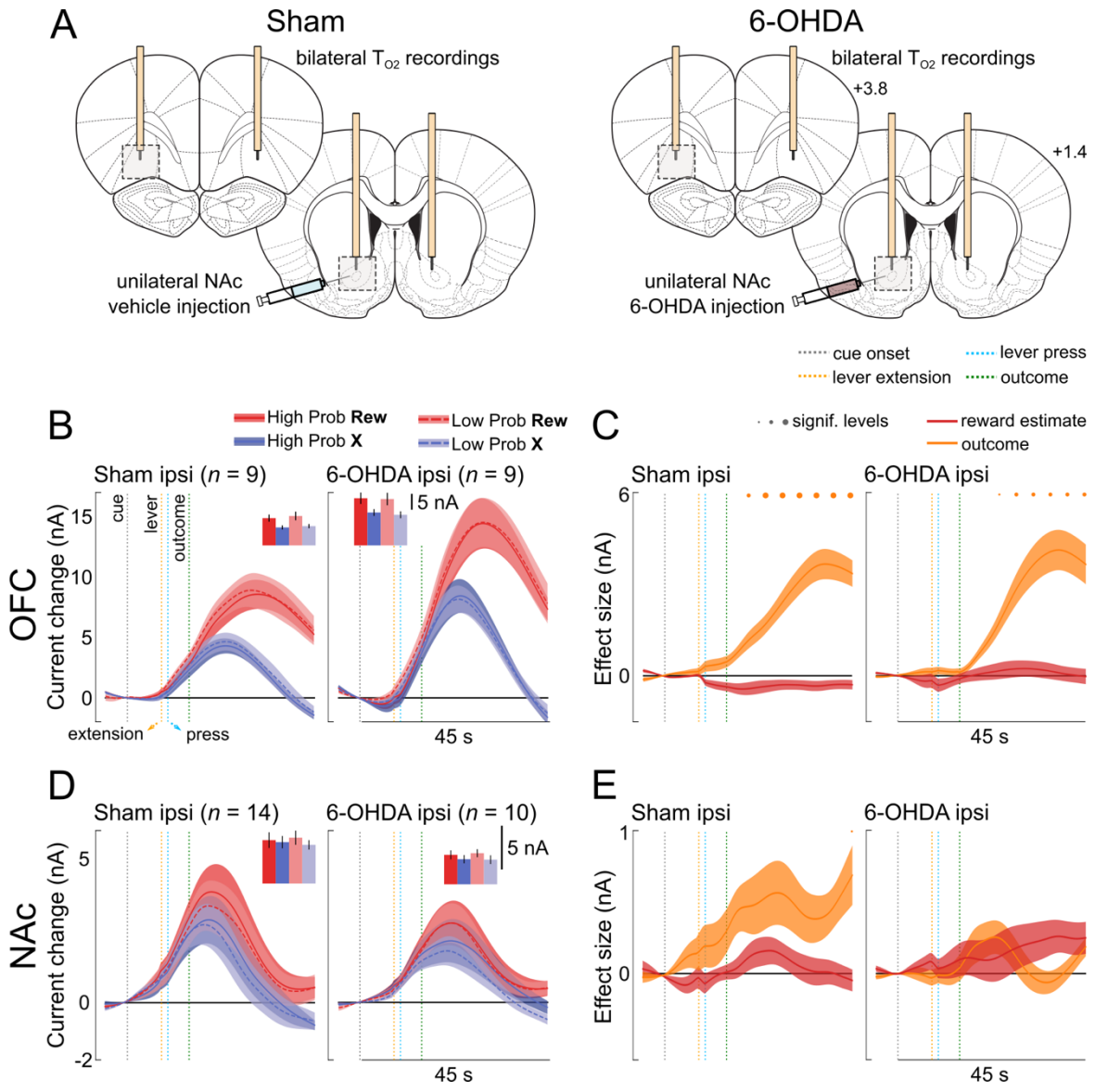


Figure 4.7 Pre-reversal trial-by-trial haemodynamic responses in the ipsilateral hemisphere of sham and 6-OHDA lesioned rats. (**A**) Electrodes were implanted bilaterally into the OFC and NAc for amperometric tissue oxygen recordings. Additionally, rats received either vehicle (sham lesion) or 6-OHDA injections into the NAc, unilaterally. Boxes indicate the regions whose responses are used in these analyses. (**B**, **D**) T_{O2} haemodynamic responses time-locked to cue onset recorded from either ipsilateral OFC (B) or NAc (D) of sham and 6-OHDA lesioned subjects on training sessions before reversal of lever reward probability contingencies. Inset bar plots represent the maximum amplitude range in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were

chosen. (**C**, **E**) Average effect size from a general linear model relating T_{O2} responses from either ipsilateral OFC (C) or NAc (E) of sham or 6-OHDA lesioned animals to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. 6-OHDA: 6-hydroxydopamine; NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

Next, we wanted to confirm that the absence of any other differences was not because the analysis was underpowered. We first confirmed that the 6-OHDA lesion did not extend to the contralateral side, finding no significant interactions between group with brain region, outcome or chosen lever reward probability and their combination on TO_2 responses (all $F < 3.06$, $p > .089$). With this in mind, we then repeated the analysis above but now comparing signals recorded ipsilaterally to the 6-OHDA lesion with those from a control group that included the signals recorded ipsi- and contralaterally to the sham lesion as well as contralaterally to the 6-OHDA lesion. This analysis replicated all the previous results.

Moreover, while a between subjects, 2nd level GLM analysis comparing the control group with those recorded ipsilaterally to the 6-OHDA lesion revealed no significant differences in the encoding of outcome or reward estimate between the two groups, the values of the GLM constant –represents the common effect across all trials– were significantly larger in the OFC of 6-OHDA than of sham lesioned subjects (Figure 4.8). Taken together, these results suggest that rather than affecting the encoding of specific trial features such as outcome or reward estimate, lesioning of dopaminergic NAc inputs primarily affected the overall signals dynamics of OFC responses, potentiating them.

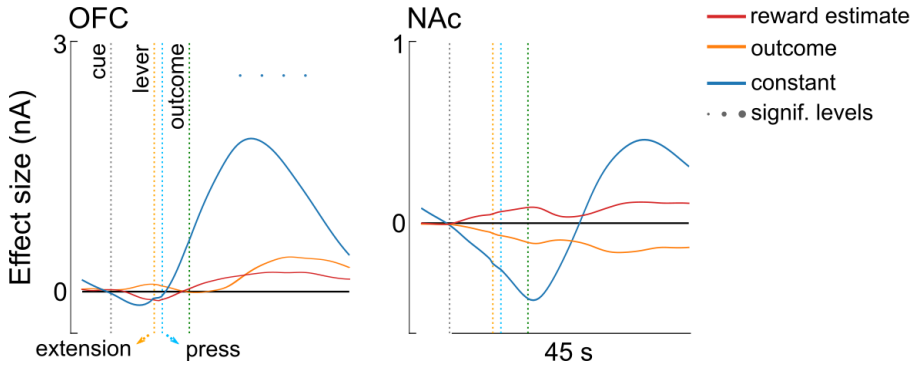


Figure 4.8 Effect of lesion on OFC and NAc haemodynamic responses before reversal. Average effect sizes from 2nd level GLMs relating the effect of trial-by-trial estimates of outcome, reward estimate, or the constant values obtained with the 1st level GLM (pre-reversal week) to lesion type (control (i.e., sham ipsi- and contralateral, and 6-OHDA contralateral) = -1, 6-OHDA ipsilateral = 1). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. *n* OFC = 37 (9 ipsi sham, 10 contra sham, 9 ipsi 6-OHDA, 9 contra 6-OHDA), *n* NAc = 45(14 ipsi sham, 13 contra sham, 10 ipsi 6-OHDA, 8 contra 6-OHDA). NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry

4.3.6 Contrasting effects of expectation in the NAc before and after learning new reward contingencies

During the first week, RPEs were generated due to the probabilistic nature of the reward contingencies. However, as these contingencies were known, there was no need to use these RPEs to update value estimates or change choice policies. Therefore, we next assessed T_{O2} responses after a reversal of the reward contingencies, focusing on signals from sessions 1-2 after the reversal (early) with those from sessions 3-4 (late; Figure 4.9-A, D).

In the OFC, the responses were similar across the post-reversal week, with the exception that post-outcome signals were slightly larger in later sessions (Figure 4.9, A-C). By contrast, in the NAc, T_{O2} patterns changed over the week. In the post-outcome period, there was a significant positive effect of outcome in early post-reversal sessions, which diminished considerably by late post-reversal sessions. Conversely, there was a

shift in the influence of reward expectation, which went from a negative to a positive influence (Figure 4.9, D-F). Therefore, while OFC coding is relatively stable across learning, NAc signals change dynamically as animals update their reward expectations after the reversal of the reward contingencies.

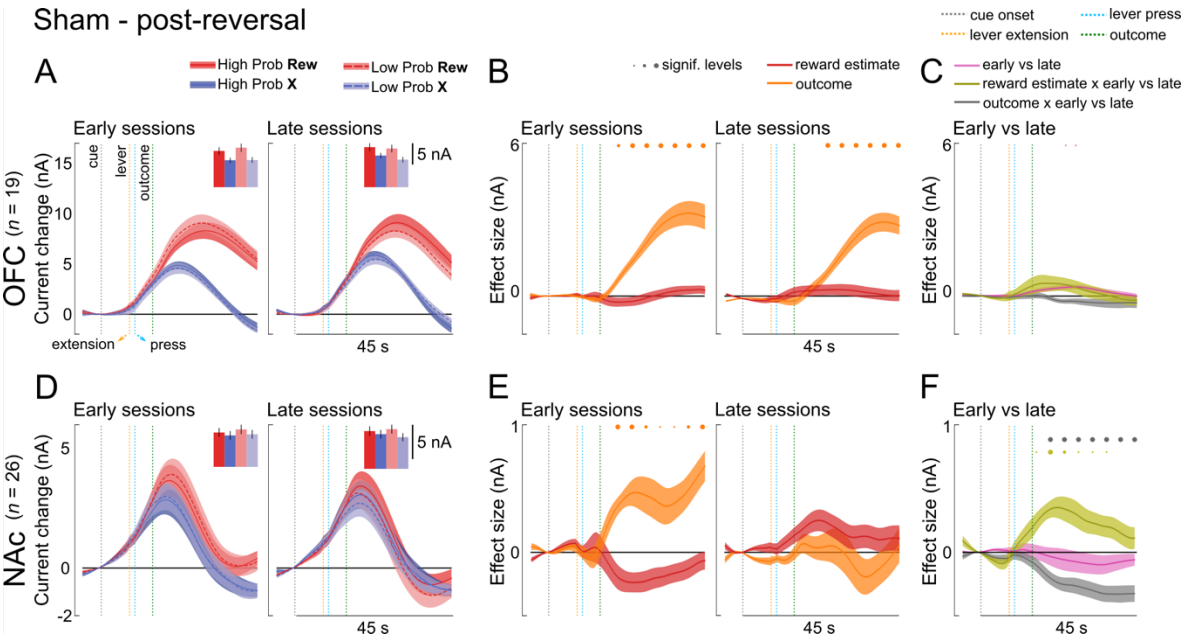


Figure 4.9 Post-reversal, trial-by-trial haemodynamic responses in sham lesioned rats. (**A**, **D**) T_{O2} haemodynamic responses time-locked to cue onset recorded from OFC (A) and NAc (D) on both hemispheres of sham lesioned subjects on the 1st and 2nd (early) or the 3rd and 4th (late) training sessions after reversal of lever reward probability contingencies. Inset bar plots represent the maximum amplitude range in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. (**B**, **E**) Average effect size from a general linear model relating T_{O2} responses from either lesioned side OFC (B) or NAc (E) to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). (**C**, **F**) Average effect size from a general linear model in which a regressor of early (-1) vs late sessions (1) as well as its interaction with reward estimate or outcome were included for predicting T_{O2} responses from OFC (C) and NAc (F). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

4.3.7 Lesioning of dopaminergic inputs to NAc has no effect on the encoding of specific trial features after reversal of reward contingencies

We next wanted to test whether lesioning of dopaminergic inputs to NAc disrupted the encoding of reward estimate or outcome in OFC and NAc after reversal as the rats learned the new contingencies.

We used a general linear model to assess the effects of these trial features on T_{O2} responses in the hemisphere ipsilateral to the lesion of 6-OHDA lesioned animals in early or late post-reversal sessions (Figure 4.10). The results were comparable to what we observed for sham lesioned animals: (1) OFC signals are mainly driven by whether a trial was rewarded or unrewarded, both in early and late post-reversal sessions; (2) reward estimate has only a small, non-significant effect on OFC responses, after outcome delivery, with a minor positive shift in late sessions (Figure 4.10-B, C); (3) the prominent positive effect of outcome present in NAc responses of early sessions is reduced in late sessions; (4) in the NAc, there is also a marked positive shift in the encoding of reward estimate between early and late sessions (Figure 4.10-E, F).

Next, to confirm that there were no significant differences between the encoding of trial features in 6-OHDA and sham lesioned groups, we run a between-subjects, 2nd level GLM analysis. Similar to what we observed in the pre-reversal week, this analysis revealed no significant differences in the encoding of outcome or reward estimate or their interaction with early vs late sessions between the two groups, but the GLM constant values were significantly larger in the OFC ipsilateral to the 6-OHDA lesion (Figure 4.11).

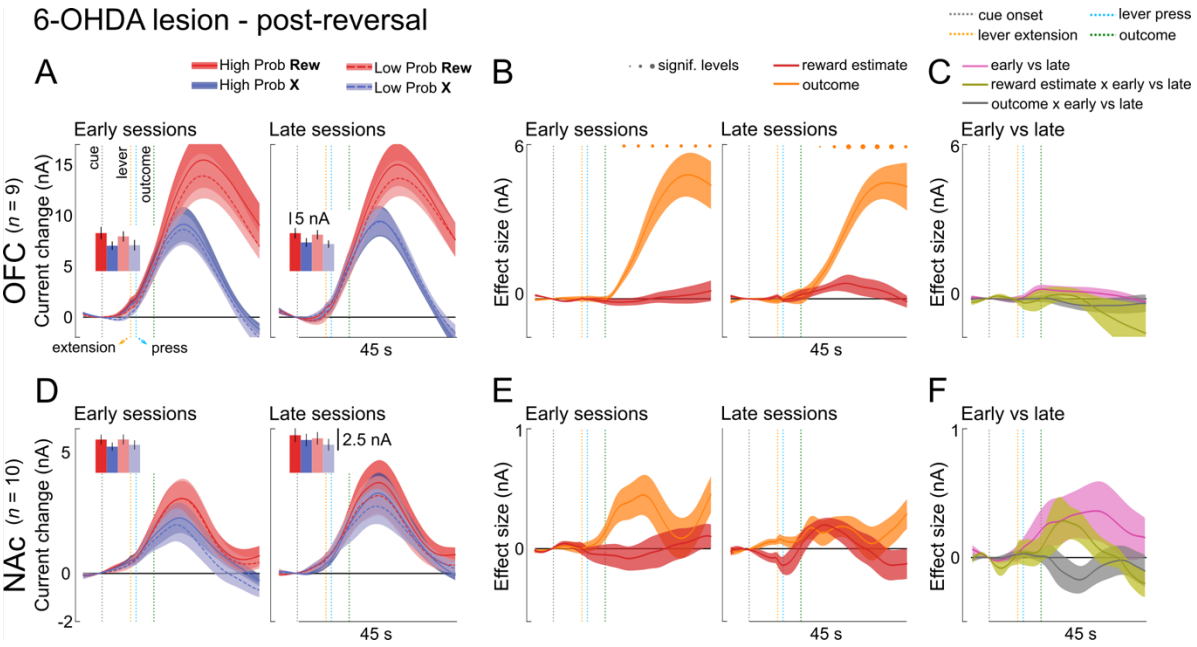


Figure 4.10 Post-reversal, trial-by-trial haemodynamic responses in 6-OHDA lesioned rats. (**A**, **D**) T_{O2} haemodynamic responses time-locked to cue onset recorded from lesioned side of either OFC (B) or NAc (D) of 6-OHDA lesioned subjects on the 1st and 2nd (early) or 3rd and 4th (late) training sessions after reversal of lever reward probability contingencies. Inset bar plots represent the maximum amplitude range in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. (**B**, **E**) Average effect size from a general linear model relating T_{O2} responses from either lesioned side OFC (B) or NAc (E) to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). (**C**, **F**) Average effect size from a general linear model in which a regressor of early (-1) vs late sessions (1) as well as its interaction with reward estimate or outcome were included for predicting T_{O2} responses from OFC (C) and NAc (F). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. 6-OHDA: 6-hydroxydopamine; NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

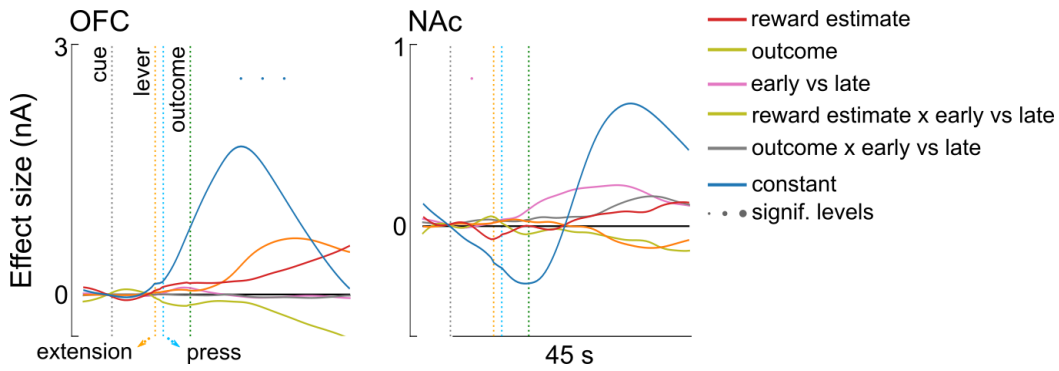


Figure 4.11 Effect of lesion on OFC and NAc haemodynamic responses after reversal. Average effect sizes from 2nd level GLMs relating the effect of trial-by-trial estimates of

outcome, reward estimate, early vs late (post-reversal sessions) and its interaction with reward estimate or outcome, or the constant values obtained with the 1st level GLM (pre-reversal week) to lesion type (control (i.e., sham ipsi- and contralateral, and 6-OHDA contralateral) = -1, 6-OHDA ipsilateral = 1). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. *n* OFC = 35 (9 ipsi sham, 10 contra sham, 9 ipsi 6-OHDA, 7 contra 6-OHDA), *n* NAc = 44 (13 ipsi sham, 13 contra sham, 10 ipsi 6-OHDA, 8 contra 6-OHDA). NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry

4.3.8 Lesioning of dopaminergic inputs to NAc potentiates OFC haemodynamic responses

The analysis of trial-by-trial T_{O2} responses revealed a link between the 6-OHDA lesion in the NAc and a subtle decrease in the amplitude of responses in this region, as well as an unexpected, strong increase in the amplitude of OFC responses (both pre- and post-reversal weeks). This was also clear when we compared the mean trial-by-trial amplitude in each region and related it to the extent of lesioning in the NAc following histological assessment (Figure 4.12; OFC, sham vs 6-OHDA: $t_{(19)} = -2.64$, $p = .017$; NAc, sham vs 6-OHDA: $t_{(23)} = 1.50$, $p = .147$, in independent samples t-test).

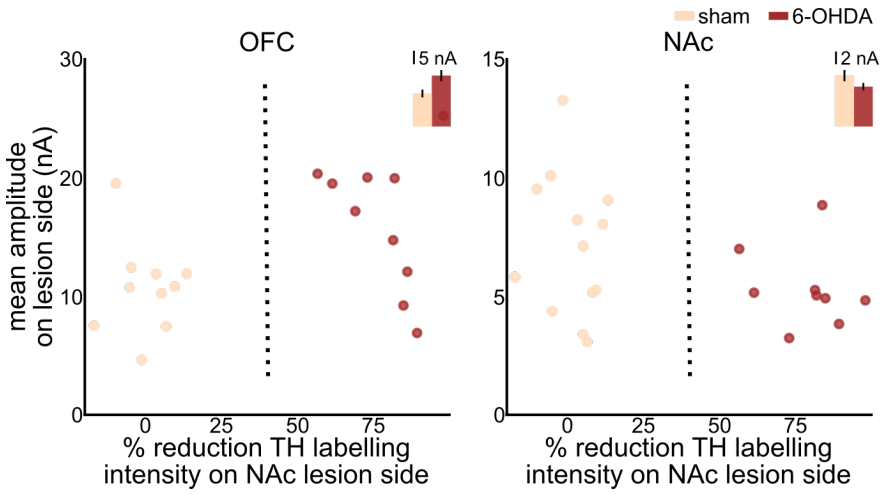


Figure 4.12 Relationship between 6-OHDA lesion and T_{O2} current amplitude changes. Plotted is the mean T_{O2} current trial-by-trial amplitude in the lesion side OFC or NAc against the % reduction TH labelling intensity on lesion side NAc. The dotted line represents the

boundary between the two groups identified through k-means clustering. Inset error bars depict mean ± SEM of average trial-by-trial amplitude for each group. *n* OFC, sham = 10, 6-OHDA = 10; *n* NAc, sham = 14; 6-OHDA = 10. 6-OHDA: 6-hydroxydopamine; OFC: orbitofrontal cortex; NAc: nucleus accumbens; T_{O2}: tissue oxygen amperometry.

We also investigated whether there was a relationship between the trial-by-trial amplitude change in the two brain regions. To do this, we computed the Pearson's correlation coefficient between the average trial-by-trial change in amplitude in the two brain regions for animals that had provided data from both OFC and both NAc hemispheres (see Chapter 2 for details). While the numbers of animals are too low to form strong conclusions, there was no evidence for a correlation between the change in amplitude in the two brain regions in either sham or 6-OHDA lesioned groups (*r* sham: -0.02 [95% CI -0.13–0.08], *r* 6-OHDA: -0.05 [95% CI -0.13–0.04]; Figure 4.13).

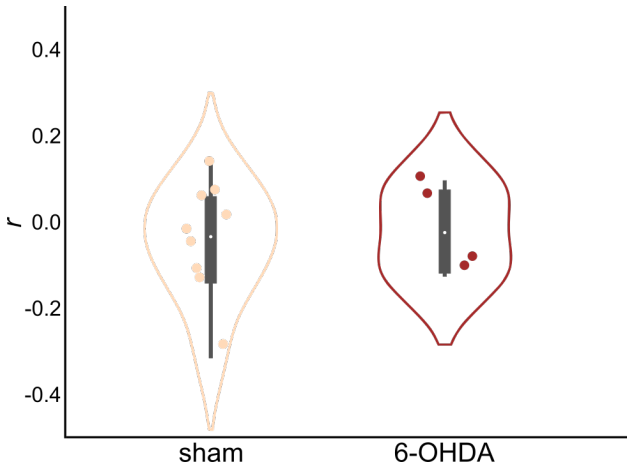


Figure 4.13 Correlation between T_{O2} current amplitude in the lesion side OFC and NAc in sham and 6-OHDA lesioned subjects. Pearson's correlation coefficient between the changes in normalised amplitude in the lesion side OFC and NAc on each trial. Scatter plots overlaid with violin plots represent individual animals' averages. *n* = 13 (sham = 9, 6-OHDA = 4).

4.3.9 Distinct T_{O2} response dynamics in anterior and posterior portions of the OFC

In this study, electrodes placed in the OFC were clustered mainly in its lateral portion (i.e., lateral orbitofrontal cortex, LO; Figure 4.4-A), which has been shown to contain

functionally (Panayi & Killcross, 2018) and anatomical divisions along its anterior–posterior axis (Barreiros, Ishii, et al., 2021; Barreiros, Panayi, et al., 2021). As the electrodes ended up distributed along this axis, we investigated whether the encoding of trial features in T_{O2} responses differed between the anterior and posterior portions of LO.

There were no significant differences between the amplitude of responses from the two LO portions after each trial type (no main effect or interactions with AP placement: all $F < 2.78$, $p > .113$; Figure 4.14-A). However, a between-subjects analysis suggests responses from anterior and posterior LO have different dynamics, with posterior signals exhibiting larger response amplitudes (see the constant effect in Figure 4.14-C).

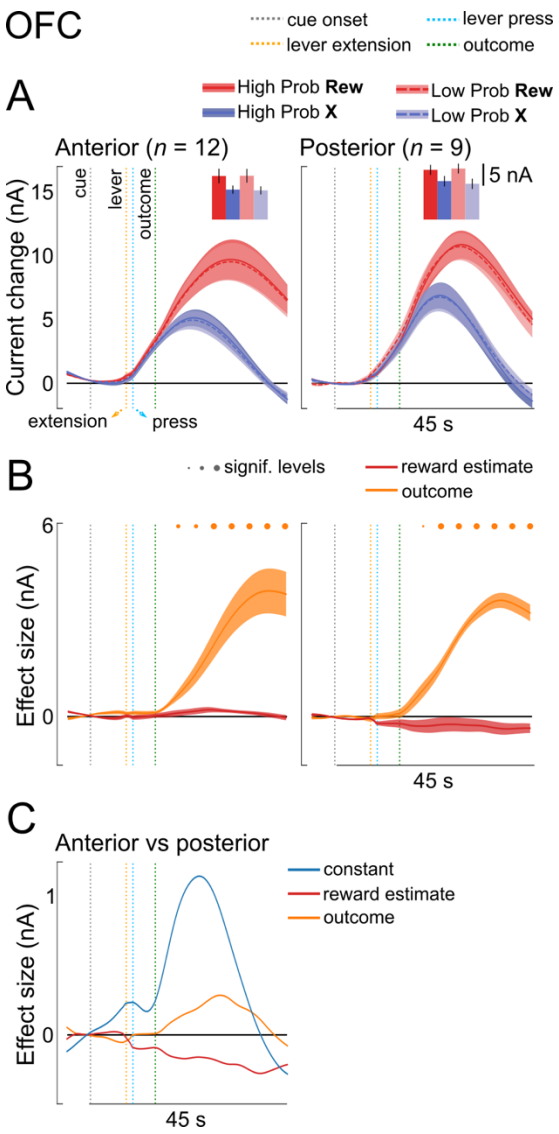


Figure 4.14 Trial-by-trial haemodynamic responses in anterior vs posterior OFC. (**A**) T_{O2} haemodynamic responses time-locked to cue onset recorded from anterior or posterior OFC electrodes of sham lesioned subjects or placed in the contralateral side of 6-OHDA lesioned subjects across all training sessions. Inset bar plots represent the maximum amplitude range in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. (**B**) Average effect size from a general linear model relating T_{O2} responses from either anterior or posterior OFC to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). (**C**) Average effect sizes from 2nd level GLMs relating the effect of trial-by-trial estimates of outcome, reward estimate, or the constant values obtained with the 1st level GLM, as well as interaction terms between these and anterior or posterior electrode placement (anterior = -1, posterior = 1). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. 6-OHDA: 6-hydroxydopamine; GLM: general linear model; NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

4.4 Discussion

Behavioural-derived haemodynamic correlates of RPEs in the NAc are thought to be modulated by mesolimbic dopamine release (as reviewed in Knutson & Gibbs, 2007). The present study investigated this link by combining lesioning of NAc dopaminergic inputs with oxygen amperometry recordings in this region from rats trained on an instrumental probabilistic reversal learning task. Additionally, it examined the role of mesolimbic dopamine in modulating frontostriatal circuits during reward-guided behaviour by combining this approach with simultaneous oxygen amperometry recordings in the OFC.

4.4.1 Lesioning effect of 6-OHDA injection into NAc extends to OFC dopaminergic inputs

Lesioning of the dopaminergic inputs to NAc was achieved through local injection of the catecholaminergic neurotoxin 6-OHDA. The sparing of passing fibres is considered a major advantage of such neurochemical lesions (Lavond & Steinmetz, 2003). However, histological assessment of the lesion through TH^+ staining revealed that, surprisingly, its effect extended to OFC dopaminergic inputs (Figure 4.2-E, F). A similar lesion spread was previously observed in studies that detected a significant reduction in regional dopamine metabolites in the medial prefrontal cortex following injection of 6-OHDA into the NAc (Clarke et al., 1988; Koob et al., 1978; Robbins et al., 1983).

Several potential reasons could account for the spread of the lesion to OFC: (i) spread of 6-OHDA into the ipsilateral ventricle, (ii) degeneration of fibres projecting to the NAc and sending collaterals to OFC, or (iii) neurotoxin uptake by OFC dopaminergic afferents in the vicinity of the NAc.

The first hypothesis is unlikely since injection of 6-OHDA into the ventricle causes widespread dopaminergic neurons degeneration (Quiroga-Varela et al., 2017), and no TH labelling reduction was observed in ipsilateral DLS.

Both the NAc and frontal cortex receive most of their dopaminergic inputs from the VTA (Oades & Halliday, 1987; Salgado & Kaplitt, 2015). Since 6-OHDA can induce cell apoptosis (Gomez-Lazaro et al., 2008), injection into the NAc will cause death of VTA dopamine neurons in addition to degeneration of the axon terminals. If the same neurons projected to the frontal cortex, this could explain the reduction in TH labelling observed. However, neuroanatomy tracing studies suggest that dopaminergic projections to NAc and prefrontal cortex arise from separate subsets of VTA neurons (Deniau et al., 1980; Swanson, 1982). Furthermore, a small proportion of the dopaminergic VTA neurons projecting to NAc also project to ipsilateral DLS (Aransay et al., 2015) and contralateral NAc (Fox et al., 2016; Phillipson & Griffiths, 1985), but there was no TH labelling reduction in either. Thus, degeneration of fibres projecting to the NAc and sending collaterals to OFC is not a likely cause for the spread of the lesion to OFC.

Let us now consider the third hypothesis. During embryonic development, mesocortical dopaminergic neurons pass through the striatum *en route* to the prefrontal cortex (Kalsbeek et al., 1988), with some continuing to grow from the striatum to prefrontal cortex during adolescence (Reynolds et al., 2018). Using single-cell axon tracing analysis, Aransay et al. identified long-range VTA neuronal projections to the prefrontal cortex that pass close to the NAc (Aransay et al., 2015). Therefore, 6-OHDA could enter axons of mesocortical neurons passing near the injection site, causing the degeneration of the OFC dopaminergic terminals observed here.

Despite the substantial degeneration of the dopaminergic terminals in the OFC, systemic injection of the dopamine agonist apomorphine did not potentiate

haemodynamic responses in this region (Figure 4.4-B). This suggests that unlike in the NAc, dopamine release alone might not be sufficient to modulate haemodynamic responses in the OFC. Interestingly, it was previously reported that dopamine release alone is not sufficient for modulating BOLD signals in the medial prefrontal cortex, but instead, these require glutamatergic co-transmission (Helbing et al., 2016). Pharmacological studies combined with oxygen amperometry or fMRI might help understand the relationship between blood oxygenation and dopamine release in the OFC.

4.4.2 Haemodynamic signatures of RPEs in NAc and OFC

Axiomatic approaches to examining RPEs suggest that their neuronal signatures should satisfy three key criteria: (1) a positive influence of reward estimate on reward-predicting cue-elicited signals, with cues that predict larger rewards (or with higher probability) eliciting greater responses; (2) a negative influence of reward estimate at the time of reward delivery, with unexpected rewards producing greater responses; and (3) a positive influence of outcome at the time of delivery, with the delivery of a valued reward producing a greater response than its omission (Caplin & Dean, 2008; Rutledge et al., 2010). In the present study, before reversal of reward contingencies, both OFC and NAc haemodynamic signals were strongly influenced by the delivery of reward at outcome time but not by its expectation (i.e., reward estimate; Figure 4.4). A clear effect of reward estimate was also not clearly detectable during anticipation (i.e., after lever press and before outcome delivery). Thus, these haemodynamic signals do not fulfil the key features of RPEs described above, satisfying only the last condition: rewarded trials produced larger responses than unrewarded.

The absence of a more substantial effect of reward estimate was particularly surprising for the NAc, which has repeatedly been shown to encode value representation (see Averbeck & Costa, 2017 for a review), and NAc haemodynamic signals obtained through oxygen amperometry have been shown to encode RPEs (Werlen et al., 2019). Worth noting, however, that here rats had extensive experience of the same reward contingencies, and previous studies have noticed an attenuation of RPEs encoding for trial features that are less relevant for guiding behaviour (e.g., Klein-Flügge et al., 2011; Werlen et al., 2019). Indeed, a clearer influence of reward estimate on NAc responses emerged after reversal, with signals changing dynamically as animals learned the new reward contingencies (Figure 4.9). Several studies linking dopamine to the encoding of RPEs in the NAc have observed that dopamine release in this region dynamically changes throughout associative learning (Day et al., 2007; Hart et al., 2014, 2015; Saddoris et al., 2015). Moreover, while both the present study and that by Werlen et al. (2019) used oxygen amperometry to record NAc haemodynamic signals from rats trained on a probabilistic reversal learning task, some differences in the experimental approach might account for the differences between their results and ours. For instance, while here rats learned the likelihood of reward after pressing each lever (action-outcome associations), in their study, Werlen et al. trained rats to predict the likelihood of reward after hearing one of two sounds (cue-outcome associations). Additionally, as discussed above, the learning stage affects the encoding of RPEs and while here rats learned the reward contingencies before the start of recordings, in their study, Werlen et al. recorded during task acquisition.

By comparison, haemodynamic signals in the OFC were relatively stable throughout training, with outcome delivery remaining the main driver of signal change across sessions before and after reversal of reward contingencies (Figures 4.4, 4.9). This, again,

contrasts with the encoding of unsigned prediction errors previously observed in this region (Werlen et al., 2019). Worth noting, however, that in the study by Werlen et al. (2019), OFC electrodes were mainly clustered in the posterior portion of ventral OFC (VO), whereas here electrodes were distributed throughout LO. These OFC subdivisions are thought to be functionally distinct (Balleine et al., 2011; Corwin et al., 1994), and we present anatomical evidence supporting this distinction in Chapter 3 of this thesis. Moreover, as discussed above, we trained rats on an instrumental task, whereas Werlen et al. trained rats on a Pavlovian behavioural paradigm, and instrumental and Pavlovian associations might be differently encoded in the rodent OFC (e.g., Ostlund & Balleine, 2007, 2008).

4.4.3 Effect of lesioning of dopaminergic inputs on NAc and OFC haemodynamic signals

Having examined the haemodynamic responses in OFC and NAc across training on the probabilistic reversal learning task, we also compared these responses with those obtained from rats with lesioned NAc dopaminergic inputs. Rather than a clear effect of the lesion in the encoding of specific trial features (e.g., reward estimate or outcome delivery), there was a generalised and sustained shift in the responses of the OFC, which were surprisingly larger than in controls. In contrast, responses in the NAc were only marginally attenuated as a group and individually quite variable, meaning this effect failed to reach statistical significance in any of our analyses.

The lack of a clear effect on NAc responses after lesioning of its dopaminergic inputs highlights the contribution of other inputs from other brain regions in modulating the haemodynamic responses in this region (see Floresco, 2015 for a review on NAc inputs). Even though dopamine is thought to modulate NAc haemodynamic signals

(Knutson & Gibbs, 2007), it has been previously suggested that the link between dopamine release and changes in BOLD signals is not a 1-to-1 relationship (Lohrenz et al., 2016). Thus, the subtle, non-significant attenuation we observed might be caused by a local reduction in neural activity driven by changes in dopamine modulation. Inputs to NAc from other regions involved in reward-guided behaviour (e.g., BLA or prefrontal cortex; Haber, 2011) would account for the remaining haemodynamic responses encoding outcome and reward estimate. Alternatively, these might be modulated by compensatory neuronal mechanisms counteracting the effect of the 6-OHDA lesion. Indeed, several studies have reported that the concentration of striatal extracellular dopamine following a 6-OHDA lesion does not directly correlate with the extent of the lesion assessed histologically, with only dopamine depletion > 95% resulting in substantial loss of function (e.g., Castaneda et al., 1990; Robinson & Whishaw, 1988). In the present study, such an extent of the NAc dopaminergic inputs was achieved in only 2 out of 17 rats (as assessed histologically). Worth noting, however, that systemic injection of the dopamine agonist apomorphine resulted in a pronounced potentiation of signals in the NAc ipsilateral to the 6-OHDA lesion. Further experiments with generalised extensive damage might help clarify whether any remaining dopaminergic inputs are the main driver in modulating the haemodynamic signals observed after the lesion.

OFC haemodynamic responses were significantly potentiated following local injection of 6-OHDA into NAc. As discussed earlier, lesioning of dopaminergic inputs was not restricted to the NAc but instead extended to the OFC and, potentially, to other prefrontal cortex areas (e.g., see TH labelling in prelimbic cortex in Figure 4.2-E). Therefore, the observed change in size of OFC haemodynamic responses could be caused by an unbalanced gating of inputs to this region directly as a consequence of the lesioning of dopaminergic inputs to the OFC or indirectly as a consequence of the lesioning of

dopaminergic inputs to the NAc, or a combination of both. In favour of the change in OFC responses being an indirect downstream effect of the lesioning of NAc dopaminergic inputs, there is a growing body of literature suggesting a potential inverse relationship between cortical and striatal dopamine levels. Decreasing prefrontal cortex dopamine levels, either pharmacologically or through lesioning of cortical dopaminergic terminals, has been shown to result in elevated striatal dopamine (Clarke et al., 2014; Pycock, Carter, et al., 1980; Pycock, Kerwin, et al., 1980; Roberts et al., 1994; Scornaiencki et al., 2009; Thompson & Moss, 1995). Conversely, an increase in prefrontal cortex dopamine levels has been linked to a decrease in striatal dopamine (Doherty & Gratton, 1996; Kayser et al., 2012; Kolachana et al., 1995). Moreover, disruption of the balance between frontal and striatal dopamine has been implicated in neuropsychiatric conditions. Specifically, in schizophrenia, increased mesolimbic dopaminergic signalling coexists with reduced cortical dopamine (Davis et al., 1991; Weinberger, 1987). The regulation of striatal and prefrontal dopamine levels might be mediated by direct or indirect projections from these regions back to VTA (Carr & Sesack, 2000; Nauta et al., 1978; Swanson, 1982; Taber et al., 1995; Zahm & Heimer, 1993). Our results suggest that the previously hypothesised inverse relationship between cortical and striatal modulation might be bidirectionally regulated. Additional experiments targeting only the dopaminergic inputs to OFC or NAc will be necessary to confirm this hypothesis and to assess the relative contribution of the dopaminergic inputs to each region to the modulation of OFC haemodynamic responses.

4.4.4 Haemodynamic responses across anterior–posterior OFC

There has been an increasing understanding that the OFC is functionally heterogeneous along its anterior–posterior and medial–lateral axes (as reviewed in Barreiros, Ishii, et

al., 2021; Izquierdo, 2017). For instance, Panayi & Killcross (2018) found that both anterior and posterior lesions of LO impair cue-outcome devaluation but only posterior LO lesions disrupt reversal learning, suggesting posterior LO is critical for updating the value of expected outcomes while anterior LO is necessary for accessing the current value of an expected outcome (Panayi & Killcross, 2018). In Chapter 3 of this thesis, we present neuroanatomical evidence supporting these functional differences, including that posterior LO receives stronger projections from the basolateral amygdala, a brain region implicated in representing expected value (as reviewed in Averbeck & Costa, 2017) and tracking previous outcomes and comparing them to current outcomes (Izquierdo et al., 2013; Lichtenberg et al., 2017; Schoenbaum, Setlow, Saddoris, et al., 2003; Stalnaker et al., 2007). The current study was not designed to explore functional heterogeneity within the OFC; however, electrodes were mainly clustered in LO and distributed along its anterior–posterior axis. This allowed us to test for differences in the encoding of trial features across anterior and posterior LO. Considering the differences in BLA inputs we observed in Chapter 3 and the functional differences previously reported (Murray et al., 2015; Panayi & Killcross, 2018), we expected to find clearer encoding of unexpected outcomes and, thus, a stronger effect of reward estimate on the haemodynamic signals of posterior LO. However, we observed only subtle differences in the overall signal dynamics, with posterior LO exhibiting larger response amplitudes (Figure 4.14). While such differences might reflect functional differences, they could also potentially be caused by relative proximity to large blood vessels. Moreover, it is possible that the absence of other differences might be due to the small sample size used in this particular analysis. Recording techniques with greater spatial resolution, such as electrophysiology, might be more appropriate to explore differences in the encoding of trial features across regions anatomically close.

Contributions & acknowledgements

For the studies presented in this chapter, Mark Walton (MEW), Francois Gastambide and Gary Gilmour (GG) planned the experiments; Mike Conway (MC) and Julie Foss collected the data; MC performed the surgeries; Ines Barreiros (IB) analysed and presented the data with input from John Huxter, GG and MEW.

This work was supported by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship awarded to IB, by a Lilly Research Award Program grant to MEW and GG, and by a Wellcome Senior Research Fellowship awarded to MEW (202831/Z/16/Z). We thank Marios Panayi, Sebastian Vasquez-Lopez, Marta Blanco Pozo and Yves Wiesenberger for helpful discussion about the data analysis and interpretation; and Greg Daubney for assistance with histological experiments.

5. Effect of hyperdopaminergic states in haemodynamic signatures of reward prediction

5.1 Introduction

In order to maximise reward and avoid punishment, an individual must be able to predict what might happen after a specific action is taken. This ability relies on learning associations between events and their outcomes. In an ever-changing world, the validity of such associations must be continuously monitored and updated by comparing expectations with what actually happens. The mismatch between the two is defined as a prediction error (Bush & Mosteller, 1951; Rescorla & Wagner, 1972; Sutton & Barto, 1998). Brain signals encoding reward prediction errors (RPEs) have long been linked with the activity of dopaminergic neurons in the ventral tegmental area (VTA) and substantia nigra pars compacta (Cohen et al., 2012; Mirenowicz & Schultz, 1994; Montague et al., 1996; Roesch et al., 2007; Schultz et al., 1997; Watabe-Uchida et al., 2017). Neuronal signatures of aspects of prediction errors are present in regions receiving VTA dopaminergic projections, including the nucleus accumbens (NAc) (Averbeck & Costa, 2017; Bissonette & Roesch, 2015; Watabe-Uchida et al., 2017). In this region, dopamine release encodes both positive and negative RPEs (Blanco-Pozo et al., 2021; Hart et al., 2014). Moreover, disrupting striatal transmission impairs the formation of appropriate stimulus-outcome associations (e.g., Parkinson et al., 2002; Saunders & Robinson, 2012; Schoenbaum & Setlow, 2003) and can affect reward-guided behavioural flexibility (DenOuden et al., 2013; Hamid et al., 2016; Korn et al., 2021; Parker et al., 2016; Sala-Bayo et al., 2020; Verharen et al., 2018, 2019).

Dopaminergic signalling dysregulation has been linked with several neuropsychiatric disorders (Deserno et al., 2016; Gilmour et al., 2015; Heinz & Schlagenhauf, 2010; Zald & Treadway, 2017). For example, in schizophrenia, psychotic symptoms such as delusions or hallucinations are thought to be a consequence of a hyperdopaminergic state. Such abnormal dopamine levels are thought to disrupt RPEs encoding, resulting in inappropriate attribution of incentive value or aberrant motivational salience (Deserno et al., 2016; Smith et al., 2007). Indeed, methamphetamine-induced hyperdopaminergic states have been shown to produce schizophrenia-like psychotic symptoms in healthy volunteers (Bernacer et al., 2013). Notably, the appearance of such symptoms was proportional to the disruption of RPEs and incentive value representations in the NAc and prefrontal cortex. The clinical efficacy of dopamine antagonists in ameliorating schizophrenia symptoms also highlights the contribution of dopamine to psychosis (Creese et al., 1976; Seeman et al., 1976). While there is increasing evidence linking dopamine dysregulation and consequent disruption of RPEs encoding to psychotic symptoms, the neuronal mechanisms underlying these processes in neuropsychiatric disorders remain unclear. Combining our understanding of the symptoms in patients with animal research, where more invasive, targeted manipulations are possible, might help shed light on these mechanisms. Even though it is not possible to evaluate delusions in rodents, model-derived estimates of RPEs during associative learning might be used as intermediate endpoints (Gilmour et al., 2015).

The study described in this chapter combined a probabilistic reversal learning task with amphetamine-induced hyperdopaminergic states. Amphetamine-like compounds increase the extracellular concentration of monoamines, including dopamine, entering neurons via their transporters and causing competitive reuptake inhibition. Amphetamine can also enter neurons directly through diffusion across the plasma

membrane. Once inside, it induces changes in molecular signalling that result in the release of dopamine from vesicles into the cytosol, efflux of cytosolic dopamine via reverse transport through dopamine active transporters (DAT), and non-competitive reuptake inhibition through DAT internalization. This results in increased levels of extracellular dopamine, an increase in spontaneous dopamine transients, and an increase or decrease in stimulus-evoked transients for low and high doses of amphetamine, respectively (Amara & Sonders, 1998; Daberkow et al., 2013; Sulzer et al., 2005). In rodents, a single amphetamine injection induces long-term behavioural and neuronal changes, including increased locomotor activity which correlates with increased striatal dopamine release (Vanderschuren et al., 1999). Repeated exposure to amphetamine leads to sensitisation processes that exacerbate its neuronal effects, resulting in dopamine signalling dysregulation similar to that thought to underlie symptoms in schizophrenia. This sensitisation is mediated through cellular mechanisms, including changes in the sensitivity or expression of membrane receptors, and transporters which have been shown to result in an enhanced and reduced ability to increase extracellular dopamine in the NAc and prefrontal cortex, respectively (as reviewed in Featherstone et al., 2007; Pierce & Kalivas, 1997). There is mounting causal evidence that this amphetamine-induced dopaminergic signalling disruption is then reflected in decision-making impairments in probabilistic (St. Onge et al., 2010) and reversal-learning behavioural tasks (Fletcher et al., 2005; Verharen et al., 2018). For example, a recent study in rats found that amphetamine administration disrupted RPEs signalling in NAc and outcome salience signals in OFC during a probabilistic learning task (Werlen et al., 2019), paralleling the disruption observed following methamphetamine administration in humans (Bernacer et al., 2013). These amphetamine-induced neuronal effects were accompanied by an impairment in the behavioural discriminative responses (Werlen et

al., 2019). Another study found that chemogenetic stimulation of VTA-NAc mimicked aspects of the behavioural disruption caused by amphetamine administration (Verharen et al., 2018).

Our current understanding of the role of dopamine in regulating frontostriatal circuits during reward-guided learning comes mainly from a convergence of functional magnetic resonance imaging (fMRI) studies in humans and non-human primates and electrophysiology recordings from animal models. While fMRI studies have the advantage of being non-invasive, they require subjects to remain motionless during scanning, largely limiting fMRI imaging in rodents to non-behavioural studies. This study employed a potential translational technique, *in vivo* oxygen amperometry, which allows for the measurement of haemodynamic signals comparable to the BOLD signals of fMRI in behaving animals (Francois et al., 2012; Lowry et al., 2010). During training on the behavioural task, amperometric tissue oxygen (T_{O2}) signals were recorded simultaneously from NAc and OFC. OFC electrodes targeted both its medial (MO) and lateral (LO) portions, which have been shown to be functionally distinct (as reviewed in Bradfield & Hart, 2020; Izquierdo, 2017).

In sum, the study described here aimed to investigate the neuronal mechanisms underlying maladaptive learning in hyperdopaminergic states similar to those thought to elicit psychotic symptoms in neuropsychiatric conditions. It combined an instrumental probabilistic reversal-learning task with single and repeated injections of amphetamine and simultaneous amperometric recordings from NAc and OFC. We hypothesised that amphetamine administration would disrupt behavioural performance and T_{O2} signals, and any effects would be more pronounced after the repeated injection.

5.2 Methods Summary

The experimental procedures described in this chapter were again carried out by researchers at Eli Lilly (UK) and the data analysed by me. Please see Chapter 2 of this thesis for a more detailed description of the materials and methods. Below is a brief recap of the most important methodological details.

5.2.1 Probabilistic reversal-learning task

37 male rats were trained on the same operant 2-option probabilistic reversal-learning task used in the previous chapter. One option was associated with a 75% probability of reward delivery and the other a 25% reward probability, and this association would reverse each week (Figure 5.1).

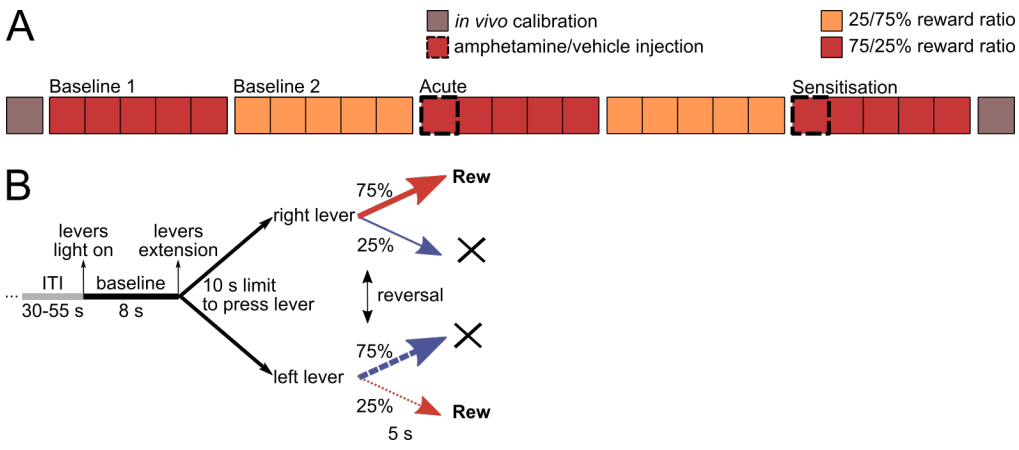


Figure 5.1 Behavioural paradigm. Food restricted rats were trained on a probabilistic reversal learning task before and after undergoing surgery for implantation of electrodes in the OFC and NAc. (**A**) Post-surgery behavioural training stages and key procedures. Reward contingencies reversed at every 5-session block. On the first day of the third and fifth post-surgery 5-session blocks, rats received a systemic injection of amphetamine or vehicle. (**B**) Trial structure of probabilistic reversal learning task in choice trials. In forced trials, only one of the levers was made available. Rew: rewarded.

5.2.2 Surgery

After learning the entire task structure, animals underwent surgical procedures under general anaesthesia. All rats were implanted with carbon paste electrodes targeted at the NAc core (bilateral) and OFC (LO and MO, unilateral) (see Chapter 2 for stereotaxic coordinates details). After a 3-week recovery period, T_{O2} signals were recorded while rats were tested on the behavioural task, with reversal of the reward contingencies every 5-session block.

5.2.3 O_2 amperometry data recording

O_2 signals were recorded from the NAc and OFC using constant-potential amperometry (−650 mV applied for the duration of the session) as described previously (Francois et al., 2012, 2014).

5.2.4 Pharmacological manipulations

D-amphetamine sulphate (Sigma, UK) was dissolved in 5% (w/v) glucose solution, and pH adjusted towards neutral with the dropwise addition of 1 M NaOH as necessary. Amphetamine was dosed at 1 mg/kg (free weight) via the intraperitoneal route.

5.2.5 Histology

At the end of experimentation, animals were perfused transcardially and the brains removed. To assess the placement of the electrode tips, brain sections were mounted on glass slides, stained with cresyl violet and viewed microscopically.

5.2.6 Data analysis

Behaviour

We analysed the % correct choices from the amphetamine and vehicle animals during the post-surgery behavioural training sessions and compared performance at key behavioural events, including the first, the second, and the last training session after reversal at baseline and on the weeks in which there was a systemic injection of amphetamine or vehicle. Additionally, we compared the number of reward magazine head entries and the mean reaction times for lever pressing between the two groups on the training session on the day of injection at acute administration (first amphetamine injection), sensitisation (second amphetamine injection) and a comparable session at baseline. We also investigated a potential effect of amphetamine on the animals' ability to learn the new reward contingencies after reversal within the first and second post-reversal sessions by comparing the % correct responses across every 28-choice trials block (i.e., first or half of each training session) on the first and second post-reversal training sessions at baseline 1, acute, or sensitised administration weeks. Additionally, we analysed the proportion of win-stay and lose-stay trials in these sessions.

Behavioural modelling

Trial-by-trial lever choices were modelled using variations of a Rescorla-Wagner model (Rescorla & Wagner, 1972). We started with a model with two free parameters: learning rate α and inverse temperature T. We then compared this model against others that also included free parameters specifying (a) different learning rates for rewarded α_{rew} and unrewarded α_{unrew} choices, (b) lever preference bias T, (c) perseveration P. To capture additional trial-by-trial variance, we also included recency-weighted lever choices rates M. To compare the models, we used the Bayesian information criterion (BIC), which

penalizes the likelihood of a model by the number of parameters and the natural logarithm of the number of data points. The model with the lowest BIC score was deemed to give the best fit of the data.

Amperometry

Data were pre-processed to remove noisy or non-meaningful signals and aligned and normalized at relevant behavioural time points (see Chapter 2 for details of exclusions). We then used these signals to investigate the relationship between the T_{O2} and the subjects' behaviour and compare the signal dynamics at key behavioural events as well as between veh and amph groups. We performed two sets of complementary analyses: (i) analyses of the signals dynamics using the amplitude (i.e., maximum range) of the T_{O2} responses in NAc and OFC in each trial (taking a 45 s window from cue onset), before or after reversal of the lever-reward probability contingencies; and (ii) regression of the same signals against estimates from our computational model using the model with the lowest BIC score (Figure 5.4).

5.3 Results

5.3.1 Effect of single and repeated systemic amphetamine injections on behaviour

To investigate the effect of single (acute administration) or repeated (sensitisation) systemic injections of amphetamine on the rats' behaviour, we compared their performance at key training sessions after either injection with equivalent sessions in a baseline week in which they encountered the same reward contingencies but received no injection (baseline 1) (Figure 5.2-A).

We first assessed whether amphetamine affected magazine-focused locomotor activity using, as a proxy, the total number of reward magazine head entries on the first session after reversal at baseline 1, acute, and sensitisation session-blocks (Figure 5.2-B), finding no main effect or interaction with group (all $F < 2.15$, $p > .134$). A similar analysis comparing reaction times between veh and amph groups also revealed no differences between the groups (all $F < 0.79$, $p > .463$; Figure 5.2-C).

Next, to assess whether amphetamine injections affected the rats' ability to learn the reward contingencies, we compared the proportion of correct choices on the first, the second, and the last training sessions after reversal at baseline 1, acute, and sensitisation weeks (note that injection test sessions only occurred on the first session in these weeks). This analysis revealed a drug challenge x session x group interaction ($F_{(2, 60)} = 3.87$, $p = .007$). Follow-up comparisons revealed that amphetamine injections changed behavioural performance on the first session after injection and reversal in the sensitisation week (veh vs amph groups at baseline 1, $p = .989$; at acute, $p = .989$; at sensitisation, $p = .009$; Figure 5.2-D). Moreover, while the performance of the vehicle group improved from the first to the second post-reversal session on both the acute and sensitisation weeks (acute: $p = .018$; sensitisation: $p = .013$; Figure 5.2-E), this did not happen in animals that received an injection of amphetamine after either administration (first vs second post-reversal session, at acute: $p = .988$; at sensitisation: $p = .069$; note that the latter reflected a numerical *decrease* in performance from first to second session). However, neither amphetamine injection impacted long-term performance (no significant (veh vs amph groups on the last session after reversal at acute: $p = .147$; at sensitisation, $p = .503$, Figure 5.2-F).

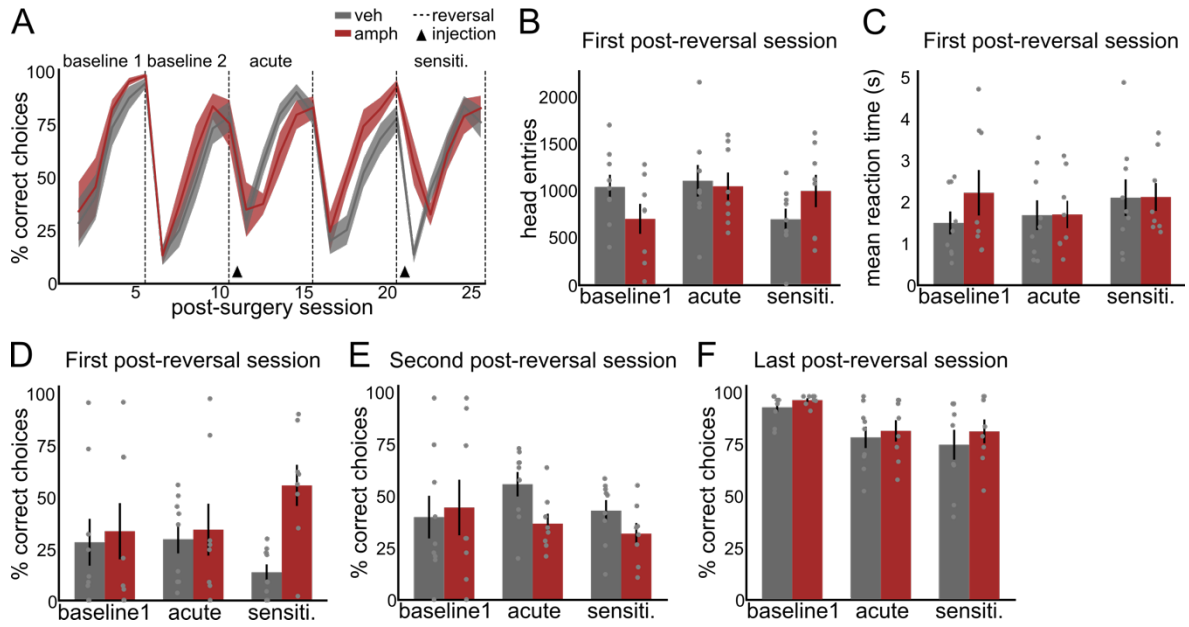


Figure 5.2 Behavioural performance. Subjects received vehicle (veh) or amphetamine (amph) injections in the first session after the second (acute) and after the fifth reversal (sensitisation). (**A**) Mean correct responses (i.e., choice of the lever with the higher probability of reward) of each experimental group across behavioural sessions during the testing phase. (**B**) Mean reward magazine head entries of each experimental group on the first post-reversal session of the baseline 1, acute, and sensitisation session blocks. (**C**) Mean reaction time of subjects in each experimental group on the first post-reversal session of the baseline 1, acute, and sensitisation session blocks. (**D**) Mean correct responses of each experimental group on the first post-reversal session of the baseline 1, acute, and sensitisation session blocks. (**E**) Mean correct responses of each experimental group on the second post-reversal session of the baseline 1, acute, and sensitisation session blocks. (**F**) Mean correct responses of each experimental group on the last post-reversal session of the baseline 1, acute, and sensitisation session blocks. Represented is the mean ± SEM. Scatter plots overlaid with bar graphs in panels B-F represent individual animals averages. *n* = 17 (veh = 9, amph = 8). amph: amphetamine; sensiti. = sensitisation; veh: vehicle.

The results described above suggest a potential delay in learning the new reward contingencies after a reversal. To investigate further this hypothesis, we compared the % correct choices across the first and second half of the first and second-post-reversal training sessions at baseline, acute, and sensitised weeks (Figure 5.3 - A-C). This analysis revealed a group x drug challenge x choice trials block (i.e., first or second halves of the first and second post-reversal training sessions) interaction ($F_{(6, 66)}$ = 4.56, p = .001). Follow up comparisons revealed that this was driven by a difference in the % of correct responses between amph and veh groups on the first post-reversal session at sensitised

administration (amph vs veh at sensitisation block 1: p = .016; block 2: p = .043; Figure 5.3-C). As seen earlier (Figure 5.2-D), the rats that received a vehicle injection exhibited a preference for the previous high reward probability lever, whereas the performance of those that received an injection of amphetamine was around chance level at sensitised administration (Figure 5.3-C). The performance of rats from the amph group was stable within the first post-reversal session (sensitisation blocks 1 vs 2 of amph group: p = .779) but dropped on the following training session (sensitisation blocks 1 vs 3 of amph group: p = .026). Additionally, on the second post-reversal session, rats from amph group appeared to be slower at learning the new reward contingencies (sensitisation blocks 1 vs 3, amph group: p = .427; veh group: p = .013).

To further investigate the potential delay in learning suggested by the results above, we tested whether animals from the two groups differed on their probability to repeat the last chosen lever after a rewarded or an unrewarded trial (i.e., win-stay and lose-stay trials) on the first and second post-reversal sessions at baseline 1, acute, and sensitisation (Figure 5.3 – D-G). Our analysis revealed a drug challenge x session x stay type (i.e., win-stay, lose-stay) x group interaction ($F_{(2, 30)}$ = 10.21, p < .001). Thus, we next examined the proportions of win-stay and lose-stay trials separately. The analysis of the proportion of lose-stay trials across the first and second post-reversal sessions at each drug challenge revealed no main or interaction effects with group (all F < 0.49, p > .617). By contrast, a similar analysis of the proportion of win-stay trials revealed a drug challenge x session x group interaction ($F_{(2, 30)}$ = 6.50, p = .005). Follow up pairwise comparisons revealed that this was driven by a significant change in the proportion of win-stay trials from the first to the second post-reversal session at sensitised administration by rats that received an injection of amphetamine (p = .002; a similar effect was not present for veh rats: p = .828). This means that on the second post-reversal session at sensitisation, rats that received

an injection of amphetamine were less likely to repeat their last lever choice after receiving a reward.

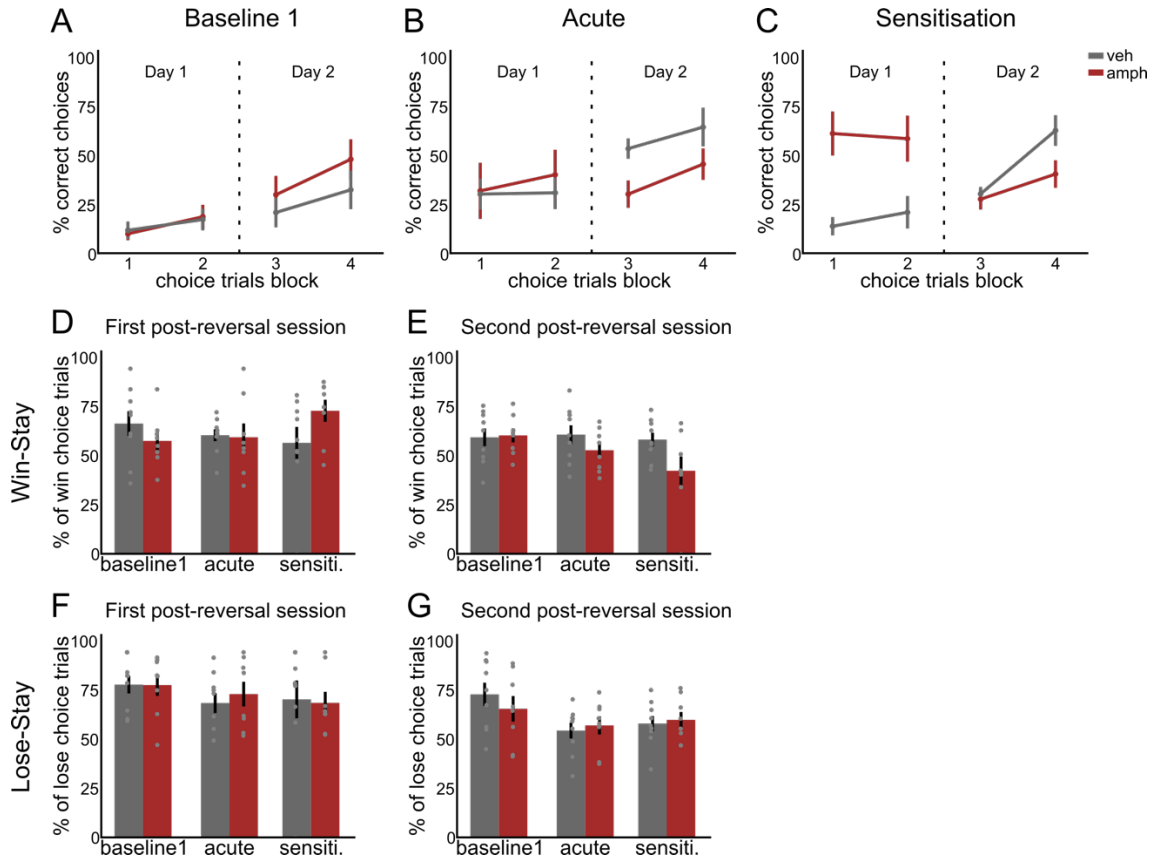


Figure 5.3 Learning within key sessions and stay-switch behaviour. Subjects received vehicle (veh) or amphetamine (amph) injections in the first session after the second (acute) and after the fifth reversal (sensitisation). (**A-C**) Mean correct responses (i.e., choice of the lever with the higher probability of reward) of each experimental group in each block of 28 choice trials in the first or second sessions after reversal (i.e., first or second half of each training session) at baseline 1, acute, and sensitised drug administration. (**D-E**) Proportion of win (i.e., rewarded) trials after which the animal chose the same lever as in the previous trial in the first and second post-reversal sessions at baseline 1, acute, and sensitisation weeks. (**F-G**) Proportion of lose(i.e., unrewarded) trials after which the animal chose the same lever as in the previous trial in the first and second post-reversal sessions at baseline 1, acute, and sensitisation weeks. Scatter plots overlaid with bar graphs in panels D-G represent individual animals averages. Represented is the mean ± SEM. *n* = 17 (veh = 9, amph = 8). amph: amphetamine; sensiti. = sensitisation; veh: vehicle.

Finally, to understand how the distinct lever-reward probability contingencies influenced the choice pattern of the animals, we built a simple reinforcement learning model and assessed how well models with different parameters could predict the rats' choices. The best fit model was a standard model including a single reward learning rate

as the only free parameter other than the inverse temperature (Figure 5.4). This suggests that the amphetamine administration was not changing rats' overall behavioural strategy, but instead how they were updating their subjective value of the chosen lever based on each reward delivery or omission.

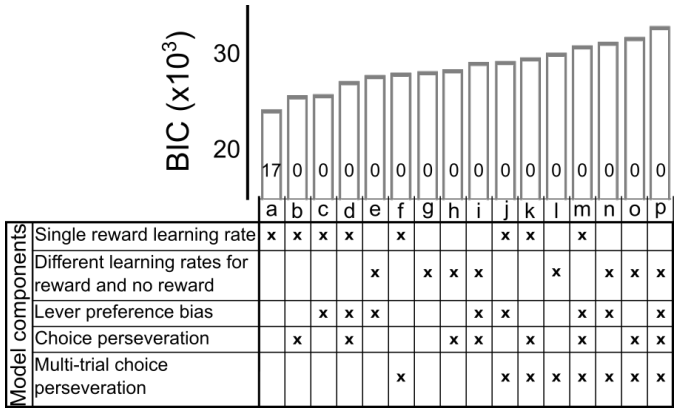


Figure 5.4 Behavioural modelling. Bayesian information criterion (BIC, a measure of the goodness of fit of the model) estimates of different reinforcement learning models. BIC penalizes the likelihood of a model by the number of parameters and the natural logarithm of the number of data points. The model with the lowest BIC score was deemed to give the best fit of the data. An 'x' in the table indicates the presence of the given component in the model. Numbers within each bar indicate the number of animals for which the given model had the lowest BIC. *n* = 17 (veh = 9, amph = 8). BIC: Bayesian information criterion.

5.3.2 OFC haemodynamic signals are mainly driven by outcome, whereas NAc tracks both outcome and reward estimate

To examine how T_{O2} responses in OFC and NAc track the animals' reward expectations and violations of their expectations, signals were recorded from electrodes implanted in these regions during each post-surgery behavioural session. After exclusions (see Chapter 2 for details), 32 electrodes in 15 rats were included in the analysis of trial-by-trial T_{O2} haemodynamic responses (veh: 2 MO electrodes from 2 rats, 3 LO electrodes from 3 rats, 11 NAc electrodes from 7 rats; amph: 1 MO electrode from 1 rat, 3 LO electrodes from 3 rats, 12 NAc electrodes from 6 rats; Figure 5.5).

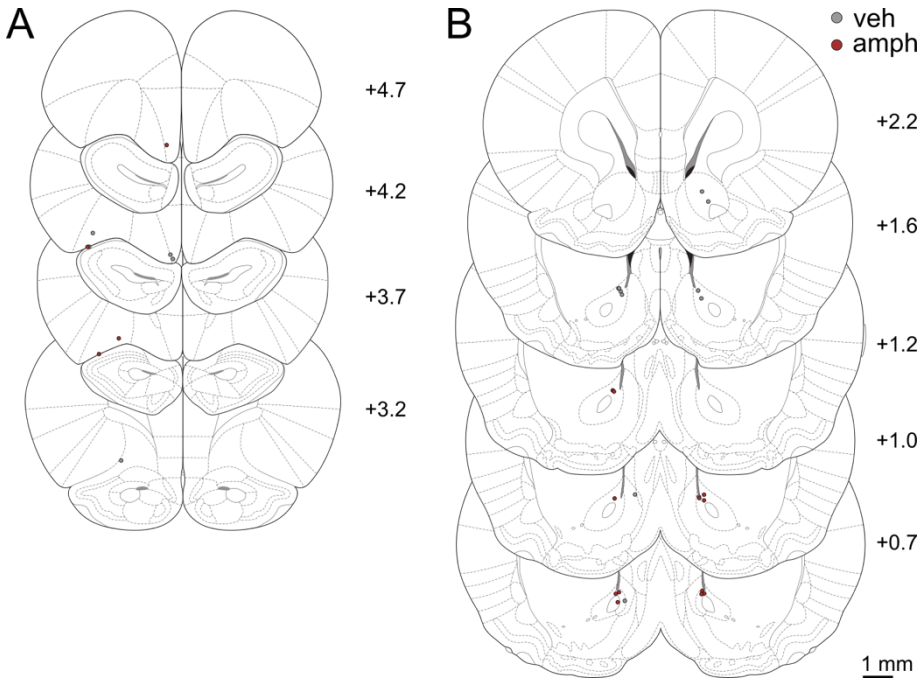


Figure 5.5 Location of electrodes placed in OFC (A) and NAc (B) of veh and amph group subjects whose signals were included in the trial-by-trial analysis. OFC electrodes were clustered in medial and lateral OFC, while NAc electrodes were mainly in the dorsomedial NAc core. Electrodes which were off target, whose signals were noisy or from subjects excluded due to behavioural criteria were not included in the trial-by-trial analysis and are not represented here. *n* = 15 subjects (veh = 8, amph = 7). amph: amphetamine; veh: vehicle; OFC: orbitofrontal cortex; NAc: nucleus accumbens.

As can be observed, MO and LO signals increased primarily after lever pressing. Then, whereas both LO and MO signals diverged on rewarded and unrewarded trials, LO exhibited a much stronger prolonged response on rewarded than reward omission trials (Figure 5.6-B, D). By contrast, similar to what was observed in Chapter 4, NAc signals increased as soon as the cues were presented and during reward expectation (i.e., after lever pressing), and then diverged based on the trial outcome (Figure 5.6-F).

We first analysed the dynamics of the T_{O2} signals in response to rewarded and unrewarded outcomes after choosing the high- (High Prob) or the low-reward (Low Prob) probability levers during the two first weeks of recordings, i.e., before any amphetamine challenge (note that in this experiment, unlike the one reported in Chapter 4, the first post-surgery session involved an immediate reversal in the lever–reward

contingencies that had been trained pre-surgery). As there was no effect or interaction of subsequent group allocation (veh, amph) (all $F < 1.09$, $p > .349$), we show the data collapsed across the two groups (Figure 5.6). As can be observed, in the OFC, rewarded trials produced responses with markedly larger signals than unrewarded trials (main effect of outcome: $F_{(1, 5)} = 14.44$, $p = .013$). This effect was prominent in LO, but not evident in MO (outcome x OFC subdivision interaction: $F_{(1, 5)} = 7.24$, $p = .043$; $p = .534$ for MO; $p = .002$ for LO; reward vs unrewarded trials; Figure 5.6-B, D). There was no clear effect of reward probability associated with the chosen lever on the size of the T_{02} responses in either OFC subdivision (no main or interaction effects with chosen reward probability: all $F < 2.85$, $p > .154$). Similar analyses did not reveal a significant effect of outcome (no main or interaction effects with outcome: all $F < 1.05$, $p > .750$) or chosen lever reward probability (no main or interaction effects with chosen lever reward probability: all $F < 3.87$, $p > .0.63$) on NAc haemodynamic responses (Figure 5.6-F).

Next, we performed a companion analysis, but now asking how animals' subjective reward estimates, calculated from the reinforcement learning model, along with other task and behavioural factors, influenced trial-by-trial haemodynamic signals. We included in this analysis explanatory variables for the value of the chosen option (derived from the reinforcement learning model), trial outcome, as well as other regressors capturing the trial type (i.e., choice or forced), outcome on the previous trial, and inter-time interval at the end of the previous trial, and an interaction term between these last two regressors, as well as a constant term that describes the common activation pattern across all trials. As can be seen in Figure 5.6 (panels C, E and G), the primary determinant for signal change in OFC and NAc was whether a trial was rewarded, with a strong positive effect of outcome. In addition, there was an influence of reward estimate (though this did not survive multiple-comparisons correction) which changed dynamically

throughout the trial for the NAc. As can be seen in Figure 5.6-G, there was a positive influence of reward estimate after cue onset (trial initiation), suggesting that NAc haemodynamic responses encode anticipation of reward at cue onset. This effect then became negative and more pronounced at the time of outcome, which means that more unexpected rewards and/or reward omissions (i.e., low probability rewards or high probability reward omissions) elicit larger changes in response.

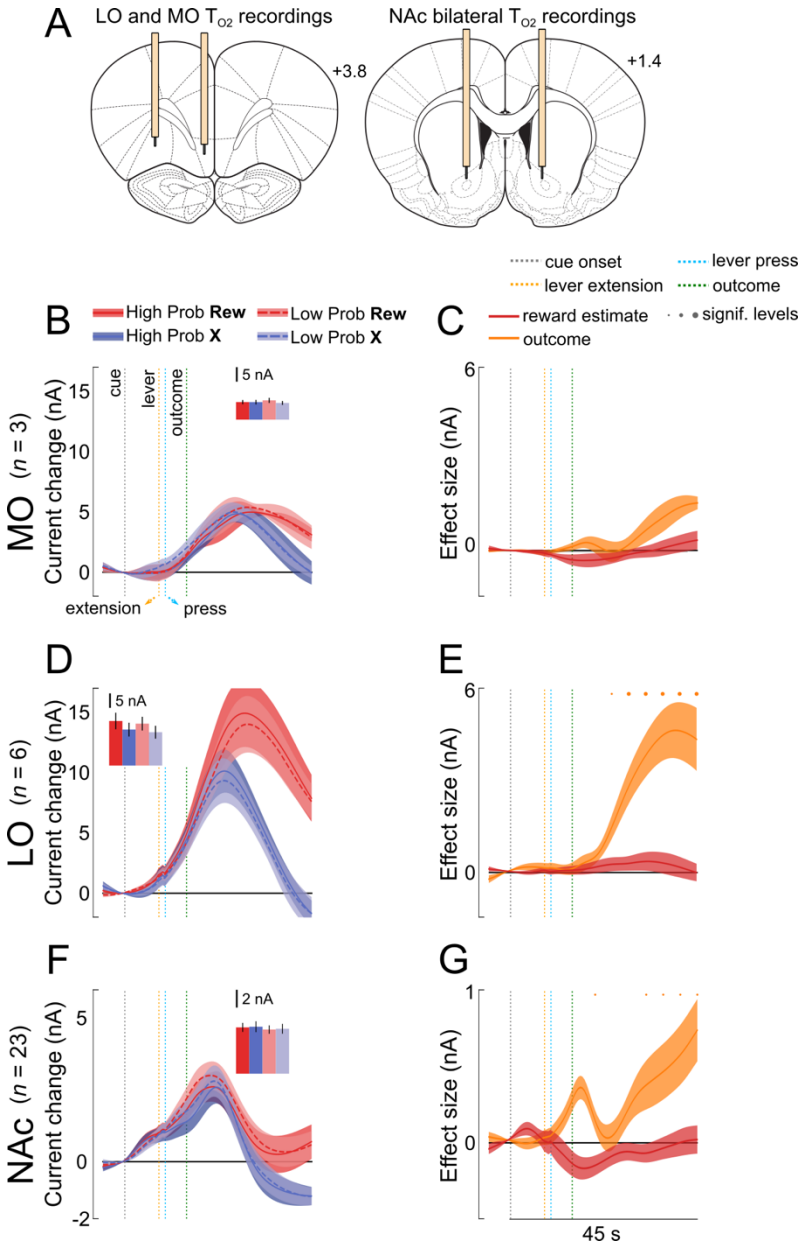


Figure 5.6 Baseline, trial-by-trial haemodynamic responses in all rats. (**A**) Electrodes were implanted into the NAc (bilaterally) and OFC (MO and LO, unilaterally) for amperometric tissue oxygen recordings. (**B**, **D**, **F**) T_{O2} haemodynamic responses time-locked to cue onset recorded from either MO (B), LO (D), or NAc (F) on training sessions on the first two weeks of recordings, i.e., before amphetamine or vehicle administration. Inset bar plots represent the amplitude (i.e., maximum range) of the responses in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. (**C**, **E**, **G**) Average effect size from a general linear model relating T_{O2} responses from both hemispheres from either MO (C), LO (E), or NAc (G) to trial-by-trial estimates of the expected value associated with each lever and trial outcome (reward or no reward). Dots at the top mark timepoints when the corresponding effect is significantly different from zero at $p < .05$ (small dot), $p < .01$ (medium dot), $p < .001$ (large dot). Represented is the mean ± SEM. LO: orbitofrontal cortex, lateral part; MO: orbitofrontal cortex, medial part; NAc: nucleus accumbens; OFC: orbitofrontal cortex; signif. levels: significance levels; T_{O2}: tissue oxygen amperometry.

5.3.3 Distinct TO2 response dynamics in medial and lateral OFC

In this study, OFC electrodes were placed in its medial and lateral portions (Figure 5.5-A), which have been shown to be functionally and anatomically distinct (as discussed in the introduction of this thesis; see also Bradfield & Hart, 2020; Izquierdo, 2017 for reviews). The analysis of the trial-by-trial T_{O2} responses during the first two weeks of recordings described earlier suggests that LO might have a clearer encoding of outcome than MO (Figure 5.6-C, E). To further explore differences in the response dynamics of the two OFC portions, we used a 2^{nd} level, between-subjects approach for comparing their signals. Though not statistically significant (note low sample size of 3 MO and 6 LO electrodes), the results suggest potential differences in response dynamics, with LO signals exhibiting generally larger responses as well as a stronger influence of outcome (see Constant and Outcome effects in Figure 5.7; the GLM constant represents the common effect across all trials).

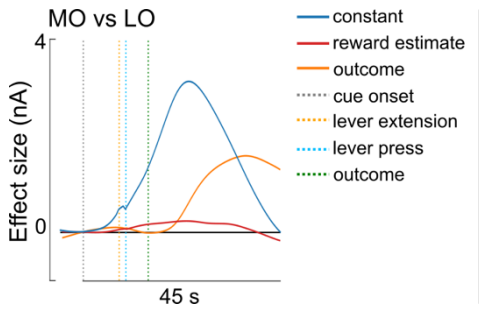


Figure 5.7 Comparison of haemodynamic responses in medial vs lateral OFC. Average effect sizes from 2nd level GLMs relating the effect of trial-by-trial estimates of outcome, reward estimate, or the constant values obtained with the 1st level GLM (baseline weeks; see Figure 5.6-C, E) to medial or lateral electrode placement (medial = -1, lateral = 1). GLM: general linear model; LO: orbitofrontal cortex, lateral part; MO: orbitofrontal cortex, medial part; OFC: orbitofrontal cortex.

5.3.4 Amphetamine disrupts NAc haemodynamic signals

Having established the effects of outcome and reward estimate in the haemodynamic responses during baseline weeks, we next wanted to investigate how single and repeated administration of amphetamine influence the encoding of these trial features. In this analysis, we compared the responses of sessions 1 and 2 of the pre-amphetamine injection week (baseline 2) with those from acute and sensitisation administration weeks. We included in these analyses only data from electrodes that yielded useable data for all sessions of interest. Due to the low sample size for OFC electrodes, we focused on the effect of amphetamine injection on NAc haemodynamic responses.

Visual inspection of the NAc haemodynamic responses on training sessions after drug administration suggested that both single and repeated injections of amphetamine lead to an attenuation of the NAc T_{O2} signals (see Figure 5.8). To test this, we compared the maximum range amplitude for reward or unrewarded trials after choosing the low or high reward probability levers on the first and second post-injection training sessions at baseline, acute and sensitisation, across groups. This analysis revealed a significant difference in the NAc responses across the groups (group x post-injection session x week

x outcome interaction, $F_{(2, 22)} = 6.70$, $p = .005$). Follow-up comparisons revealed that this was driven by amphetamine significantly decreasing the amplitude of the NAc responses to unrewarded trials on the acute and sensitised injection sessions (post-injection sessions 1 vs 2, at acute: $p = .003$; at sensitisation: $p = .035$). By contrast, there was no difference between the encoding of unrewarded trials in those sessions for the vehicle group (post-injection sessions 1 vs 2, at acute: $p = .667$; at sensitisation: $p = .888$). The effect of the acute administration of amphetamine on the amplitude of unrewarded trials was also noticeable when comparing amph and veh groups on the session immediately after injection ($p = .038$). Together, these results suggest a difference in the amplitude of rewarded and unrewarded trials after amphetamine administration, but not on any other sessions or with vehicle injection.

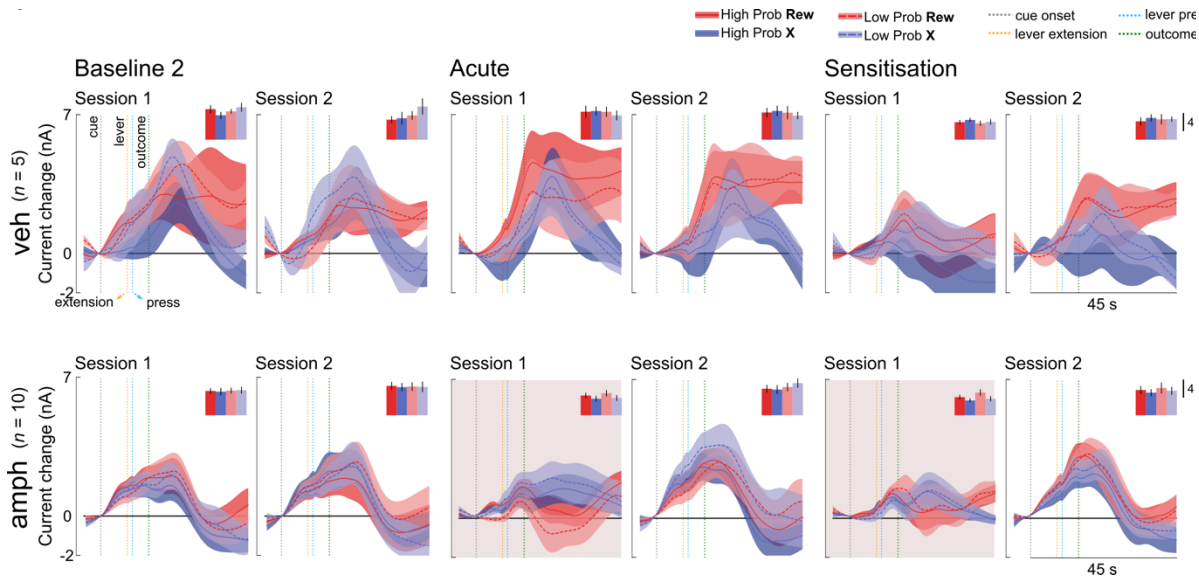


Figure 5.8 Effect of amphetamine administration on NAc haemodynamic signals. Rats received a systemic vehicle or amphetamine injection before the first session of the acute and sensitisation session blocks. The figure shows T_{O2} haemodynamic responses time-locked to cue onset recorded from NAc on training sessions on the first two post-reversal sessions of baseline 2, acute, and sensitisation session-blocks for each experiment group. Inset bar plots represent the average amplitude (i.e., maximum range) responses in each trial type, i.e., rewarded or unrewarded trials after high or low subjective reward probability levers were chosen. Depicted is the mean ± SEM. amph: amphetamine; NAc: nucleus accumbens; veh: vehicle; T_{O2}: tissue oxygen amperometry.

5.4 Discussion

An increasingly large body of evidence implicates dopamine signalling dysfunction in neuropsychiatric conditions (Deserno et al., 2016; Gilmour et al., 2015; Heinz & Schlagenhauf, 2010; Zald & Treadway, 2017). It has been hypothesised that hyperdopaminergic states drive the neuronal mechanisms underlying psychotic symptoms such as delusions through disruption of RPEs encoding in frontostriatal circuits (Deserno et al., 2016; Smith et al., 2007). A better understanding of these phenomena can help build bridges between psychosis rodent models and human disease (Gilmour et al., 2015). To this end, the study described here combined single and repeated administrations of amphetamine with oxygen amperometry recordings in the NAc and OFC during training on a behavioural paradigm which requires animals to learn and update action-outcome associations to maximise reward.

5.4.1 Effect of single and repeated doses of amphetamine on reward-guided behaviour

Rats trained on a probabilistic reversal learning task received a systemic injection of amphetamine at key behavioural sessions when there was a reversal of reward contingencies. On the session of the first drug administration, the proportion of correct choices was similar across the two experimental groups, with both groups of rats displaying a continued preference towards the lever previously associated with higher reward probability (Figure 5.2-D). By contrast, after the second drug administration, while the vehicle group again showed a strong preference for the lever previously associated with higher reward probability, the proportion of correct choices of amphetamine-injected rats was around chance level, without clear preference between

the two levers. An analysis of the behavioural performance on blocks of trials across the session revealed that the higher proportion of correct responses of the amph group was not caused by a faster learning of the reward contingencies as this behaviour was present and stable throughout the session (Figure 5.3-C, left). This loss of discriminative responding after amphetamine administration might be caused by an inability to retrieve or to express their knowledge of the reward contingencies of the two levers. Amphetamine administration has previously been shown to impair behavioural performance in various decision-making tasks (Fletcher et al., 2005; Hashemnia et al., 2020; St. Onge et al., 2010; Verharen et al., 2018; Werlen et al., 2019). The clearer effect after the second drug administration is in line with reports of long term sensitisation following a single amphetamine injection leading to exacerbated effects upon repeated exposure (Featherstone et al., 2007; Vanderschuren et al., 1999). Werlen et al. observed an impairment in discriminative responding following a single amphetamine administration similar to what we saw here after the second injection (Werlen et al., 2019). Perhaps not surprisingly, these observations suggest that the degree of behavioural disruption caused by an amphetamine-induced hyperdopaminergic state might depend on the stage in training and the specific behavioural paradigm being tested. In the study described here, rats experienced reversal of reward contingencies after extensive training on a two-lever probabilistic learning task. By contrast, Werlen et al. (2019) gave rats an amphetamine injection shortly after learning to discriminate between two cues on a Pavlovian probabilistic learning task.

In contrast with rats from the vehicle group, the performance of rats that received an injection of amphetamine did not significantly improve – and was actually worse in the sensitisation week – on the second post-reversal session after either acute or sensitised administration (Figure 5.2-D, E). These results, together with our analysis of

the rats' performance across blocks of trials within the session (Figure 5.3-B and C, right), suggest a delay in learning the new reward contingencies and an impairment after a single amphetamine exposure which was not evident in the proportion of correct responses on the first injection behavioural session. This was accompanied by a decrease in win-stay behaviour, in line with previous observations by Verharen and colleagues (2018). One possibility is that this reflects an inappropriate updating of action-outcome associations as a consequence of amphetamine-induced dopaminergic signalling dysregulation. Bernacer and colleagues (2013) also observed lower learning rates on an instrumental probabilistic learning task in human volunteers after receiving an acute methamphetamine administration (Bernacer et al., 2013). This impairment was accompanied by a disruption of RPE encoding and incentive value representations in the NAc and prefrontal cortex. In our study, the reward contingencies were relatively stable and changed only every 5 sessions. Therefore, rats might have needed a larger amount of evidence that the reward contingencies had changed before starting to modify their behaviour only on the second post-reversal and not on the first post-reversal session (Figure 5.3). It might be for this reason that the delay in learning the new reward contingencies was only noticeable in the second post-reversal session. Further studies in which reward contingencies change more dynamically might help clarify this point.

Nonetheless, the proportion of correct responses was similar across the two experimental groups by the final behavioural session in a week after either single or repeated amphetamine administration. While a possible interpretation of these results is that the learning impairment is a transient effect, a previous study reported long-lasting increased locomotor activity and nucleus accumbens dopamine release after a single amphetamine injection (Vanderschuren et al., 1999). The long term effects of amphetamine administration on task performance might be less pronounced than those

on locomotor activity. However, it is important to note that the present study used a lower drug dose (1 vs 5 mg/kg) which might produce subtler long-term effects. An alternative simpler explanation for the lack of apparent long-term differences in the behaviour of the two experimental groups might be due to a ceiling effect on task performance.

As mentioned above, moderate doses of amphetamine increase locomotor activity (Grilly & Loveland, 2001), with a single exposure being sufficient to induce long-term sensitisation (Vanderschuren et al., 1999). Accordingly, we expected to observe increased locomotor activity, especially after the second amphetamine injection. In our analysis, we found no effect of drug administration when using the number of reward magazine head entries as a proxy for locomotor activity (Figure 5.2-D). However, this might not be an appropriate measure as rats might poke only when expecting delivery of rewards. Rodents' locomotor activity is traditionally assessed by quantifying total distance travelled using video-tracking systems or cage crossovers using photobeams (e.g., Badiani et al., 1995; Vanderschuren et al., 1999). Unfortunately, it was not possible to obtain either of these measures from the dataset provided for this study. In future experiments, video recording of the behavioural sessions combined with machine learning tools for trail tracking (e.g. DeepLabCut, Mathis et al., 2018) could allow for more precise movement quantification. Amphetamine-induced locomotor sensitisation has been shown to correlate with increased dopamine release in the NAc (Vanderschuren et al., 1999). Therefore, a clear readout of locomotor activity could help interpret other effects of amphetamine injection on behaviour as well as any changes in NAc haemodynamic responses.

5.4.2 Haemodynamic signatures of RPEs in NAc and OFC

The cardinal criteria for RPEs encoding include: (i) a positive influence of reward estimate during anticipation, (ii) a positive influence of outcome at the time of delivery, with rewarded trials producing larger responses, and (iii) a negative influence of reward estimate at the time of reward delivery (Caplin & Dean, 2008; Rutledge et al., 2010). Using oxygen amperometry, Werlen and colleagues have previously shown that haemodynamic signals in the NAc of rats trained on a probabilistic Pavlovian task fulfil these criteria (Werlen et al., 2019). The patterns of NAc haemodynamic responses we observed here before amphetamine administration (Figure 5.6-G) are closer to those reported by Werlen and colleagues (2019) than were those observed in the dataset described in Chapter 4. However, the effect of reward estimate on the size of NAc haemodynamic responses was still not statistically reliable. Moreover, an effect of outcome type was observable during periods when the animals did not yet have information on whether the trial would be rewarded on unrewarded (e.g., before reward delivery; see effect of outcome after lever pressing in 5.4-G). This suggests that the outcome effect we measured might be partially influenced by other factors, e.g., a reward is more likely to occur after choosing the high reward probability lever than after choosing the low reward probability lever. As discussed in Chapter 4, differences between our results and those reported by Werlen et al. suggest haemodynamic responses in the NAc might also be influenced by the stage in training (i.e., during learning vs after learning of the reward contingencies) or the type of experimental paradigm (Pavlovian vs instrumental). Indeed, it has previously been shown that ventral striatum lesions specifically impair discrimination of stimuli but not actions (Rothenhoefe et al., 2017).

The medial and lateral portions of the OFC have been shown to be functionally distinct (as reviewed in Bradfield & Hart, 2020; Izquierdo, 2017 and in the main

introduction of this thesis). In brief, while MO appears to be closely linked with decision-making processes in which values have to be integrated and compared, LO is more commonly associated with updating specific stimulus-outcome associations. Since in this study rats were trained to choose between two levers with distinct reward probabilities, we expected to observe a clearer involvement of LO at outcome delivery and of MO during anticipation and around lever press periods, particularly in choice trials. Similar to the responses described in Chapter 4, changes in haemodynamic signals recorded from LO were mainly driven by whether a trial was rewarded with little influence of reward expectation (Figure 5.6-E). Even though to a smaller extent, outcome was also the main driver of signal change in MO (Figure 5.6-C). Direct comparison of the responses in the two regions supports this difference, in addition to an overall larger signal change in LO (Figure 5.7). The stronger encoding of outcome in LO signals is consistent with its role in value updating. MO signals did not have a clear influence of reward estimate around anticipation or lever press. There was also no clear encoding of trial type (i.e., forced vs choice; not shown). Therefore, it is possible that MO's involvement in comparing values between the two levers might be reflected in the activity of distinct neuronal populations whose averaged signals are not discernible in the haemodynamic responses. Haemodynamic responses recorded from NAc were visually more similar to the responses from MO than to those from LO. This observation fits well with the anatomical connectivity between these regions, with the NAc receiving stronger projections from MO than from LO (e.g., Wang et al., 2019). However, due to the small sample size of OFC electrodes, these results must be interpreted with caution.

4.4.3 Effect of single and repeated doses of amphetamine on NAc haemodynamic signals

As discussed earlier, the behavioural analysis revealed a delay in learning the new reward contingencies following amphetamine administration. Therefore, we expected to observe a disruption of trial features encoding in the haemodynamic signals in the NAc. Indeed, we observed an attenuation of the amplitude of the T_{O2} signals after either amphetamine injection. This effect was particularly evident on unrewarded trials. Amphetamine leads to an increase in NAc dopamine release and extracellular dopamine levels (e.g., Daberkow et al., 2013; Sharp et al., 1987; Vanderschuren et al., 1999) as a consequence of increased extracellular dopamine (Amara & Sonders, 1998; Sulzer et al., 2005). Thus, the increased baseline dopamine availability might lead to a reduced signal-to-noise ratio of stimulus-evoked dopamine release. Such a reduction in signal change in this region was previously observed in fMRI studies after methamphetamine or amphetamine was given to human volunteers (Bernacer et al., 2013; Knutson et al., 2004). In another study in rats, Werlen and colleagues observed that, rather than a general amphetamine-induced blunting of the NAc responses, the reduction in haemodynamic signals change was specific to low reward probability trials (Werlen et al., 2019). Because the NAc responses we observed in this study do not clearly fulfil the cardinal criteria for RPE-encoding, it is not possible here to draw conclusions on the effects of amphetamine administration on RPE-encoding in this region. However, the signal changes we observed following injection of amphetamine combined with the delay in learning the new reward contingencies highlights the importance of adequate dopaminergic signalling for encoding key trial features during reward-guided behaviour. Due to a combination of low sample size and high signal variability, it was not possible to clearly establish whether the effect of

amphetamine administration on NAc haemodynamic responses differ between acute and sensitised administrations and these questions remain to be explored in future studies.

Contributions & acknowledgements

For the studies presented in this chapter, Mark Walton (MEW), Francois Gastambide and Gary Gilmour (GG) planned the experiments; Mike Conway (MC) and Julie Foss collected the data; MC performed the surgeries; Ines Barreiros (IB) analysed and presented the data with input from John Huxter, GG and MEW.

This work was supported by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship awarded to IB, by a Lilly Research Award Program grant to MW and GG, and by a Wellcome Senior Research Fellowship awarded to MW (202831/Z/16/Z). We thank Marios Panayi and Sebastian Vasquez-Lopez for helpful discussions about the data analysis and interpretation.

6. Development of a novel behavioural paradigm for non-spatial cue-based serial learning

6.1 Introduction

In the studies described in Chapters 4 and 5, we analysed haemodynamic signals as animals learned through trial and error the likelihood of receiving a reward in a probabilistic reversal learning task. After reversal, the rats learned the new contingencies by gradually updating their value estimates of each lever. However, it has been observed in several species that animals get faster at adapting their behaviour after having experienced more reversals. By learning that one lever is always associated with high reward probability and the other with low, rats would be able to infer the reward likelihood associated with one lever after experiencing the outcome from pressing the other (Izquierdo et al., 2017; Masumi & Sato, 2021). This chapter explores this cognitive process (which will be described as "model-based inference") through which we become better at solving problems as we accumulate experiences or information.

Throughout life, we accumulate large sets of information that we extract from the environment and new experiences, allowing for the formation of "learning sets" –abstract knowledge of the structure of the world (Harlow, 1948), e.g., the relation between lever reward contingencies on the reversal learning task. After an initial period during which problem-solving through trial-and-error is necessary, learning sets allow us to form basic associations between events and to solve problems through inference by extracting a reduced number of key features from the environment (Harlow, 1943, 1948). Consequently, when confronted with new information or choices, memories can guide decision-making, as generalising and integrating information across events enables quick

inferences that go beyond previous experience (Rangel et al., 2008; Wimmer & Shohamy, 2012). This process requires a representation of knowledge that facilitates rapid computations.

It has been suggested that storing knowledge in "cognitive maps" —mental representations of concepts used for orientation within an environment— might support such flexible behaviour (Tolman, 1948; see Behrens et al., 2018 for a review). How information is represented and organised into these relational maps is best understood in the spatial domain (Manns & Eichenbaum, 2009; O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1979). However, the brain can also encode cognitive maps of abstract mental knowledge (Tolman, 1948; Tolman & Brunswik, 1935; Wikenheiser & Schoenbaum, 2016), such as relationships between stimuli, actions and outcomes (i.e., world models), and other non-spatial conceptual information, even in the absence of a continuous organising dimension such as space or time (Constantinescu et al., 2016; Garvert et al., 2017; Tolman & Brunswik, 1935; Wikenheiser & Schoenbaum, 2016).

The orbitofrontal cortex (OFC), a brain region known for its role in decision-making processes (Schoenbaum et al., 2006; Stalnaker et al., 2015), has been proposed to be one critical node for encoding of cognitive maps. One of the core functions of the OFC is representing the future outcomes associated with certain events or choices (Gallagher et al., 1999; Gottfried et al., 2003; Rudebeck & Murray, 2014; Schoenbaum et al., 1998). This allows us to make predictions about the future, based on the knowledge we already have about the rules of the world, by inferring what will happen without having to rely exclusively on previous experience. This process requires the OFC to encode both observable (i.e., cues and outcomes) and implicit variables of the world model and to represent this information in a cognitive map of task space (Jones et al., 2012; Rudebeck & Murray, 2014; Stalnaker et al., 2015; Wilson et al., 2014). Task space includes the

different task states –"latent" representations that help us know which set of rules (implicit or inferred information) are currently in place that go beyond what is available simply from taking into account only the available, observable information. Categorising the world into distinct states and being able to disambiguate between them facilitates decision making processes, maximising reward and avoiding punishments (Takahashi et al., 2013; Wilson et al., 2014). For example, to solve a reversal-learning paradigm, a subject might categorise the task into a state where cue A is rewarded while B is not and another state where cue B is rewarded while A is not. Thus, by categorising the world into these two states, after experiencing the outcome of one of the cues, a subject might be able to assess which of the task states is currently in place and use this information to infer the value of the other cue and to maximise reward.

Even though a number of studies support the involvement of the OFC in encoding aspects of cognitive maps, how these maps are represented is not yet well understood. Cognitive maps might be formed by patterns of abstract relationships independent of specific sensory representations (as reviewed in Behrens et al., 2018). Thus, separate populations of OFC neurons might encode a specific stimulus identity, the outcome associated with the stimulus, and the relative position of the stimulus in the task space. Building upon this hypothesis, here, we aimed to develop a behavioural paradigm that would allow us in the future to investigate the neural mechanisms of task state-dependent representations in the OFC.

6.2 Behavioural paradigm design

To study how the brain encodes cognitive maps it is necessary to establish a behavioural paradigm in which complex representations are formed by linking simpler

representations that share common elements. Furthermore, it must be possible to dissociate representations of cues, choice, outcome, spatial location, and task state. Such dissociation is not possible with paradigms commonly used to investigate behavioural flexibility such as reversal learning or reinforcer devaluation tasks. Thus, we set to establish a novel rodent behavioural paradigm with the characteristics above and in which rats should develop a model of the world to maximise reward.

The intention was that this paradigm could potentially be combined with electrophysiological recordings to allow us to investigate the dynamics of neuronal representations necessary for flexible behaviour, such as the representation of task structure. Thus, when planning the behavioural experiments, we took into consideration previous literature on OFC neuronal responses reporting neurons encoding the subjective value of the expected reward, its magnitude, and the cues predicting it (Bissonette & Roesch, 2015; Padoa-Schioppa & Assad, 2006; Schoenbaum et al., 2009; Wikenheiser & Schoenbaum, 2016). We also introduced several features that would allow for decorrelating the neuronal activity encoding the identity of the cues predicting reward in each trial, the magnitude of the predicted rewards, the order in which cues were presented, the location of the cues presentation, and the relevant task state.

The behavioural paradigm was a non-spatial cue-based serial learning task. In each trial, rats were presented with two auditory cues drawn from a set of three cues associated with distinct reward magnitudes. One cue was presented from a left-sided speaker and the other from a right-sided speaker, with order and side controlled. The rats then approached a left or right reward magazine and received the reward associated with the cue presented on that side.

Once the performance criterion was achieved, the three cues would be replaced by new ones with the same distinct reward magnitudes. Therefore, after experiencing

several sets of cues, animals should be able to abstract the reward structure ("transitive relations") and apply them to the new cues. We hypothesised that, in this task, the neuronal representations of auditory cues associated with large, small, or no reward would be maintained across cue sets and separate from the specific sensory properties (e.g., high or low-frequency tones). This kind of representation would allow for inference about unobserved cues. Thus, when encountering a new set of sounds, if rats experienced the outcomes of cues associated with large and small rewards in the first couple of trials, then they should avoid the novel sound associated with no reward over that associated with a small reward when they first encounter it. This ability to infer the relationship between elements that have not been observed together is known as transitive inference and has been previously observed in rodents (Davis, 1992; Dusek & Eichenbaum, 1997; Roberts & Phelps, 1994).

Below, we summarise the structure of the final behavioural paradigm –please see Chapter 2 for additional details, including the training stages necessary for teaching the rats the entire structure of the task. Then, we present and discuss pilot results and behavioural measures that helped guide its development.

6.3 Methods summary

Please see Chapter 2 of this thesis for a detailed description of the materials and methods. Below is a brief recap of the most important methodological details.

6.3.1 Behavioural task

22 male food-restricted rats were trained on a novel non-spatial cue-based serial learning task with a mix of the trials described below (see Chapter 2 for details on training stages).

The first cohort (n = 4) was used for initial tests of the behavioural boxes and optimising training protocols. The second (n = 9) and the third (n = 9)cohorts were trained on 2-cue and 3-cue versions of the task. See the table below for the main differences between the two versions.

Table 6.1

Summary of main differences between 3- and 2-cue task versions

Task version	Criteria for transition to new cue set	Maximum number of trials before triggering a forced transition	Trial types
3-cue	- 80% performance criterion - Transition would occur 30 trials after the performance criterion was reached *On the first cue set*, a transition was only triggered at this point if performance was still above 80% correct choices; *On subsequent cue sets*, a transition would occur 30 trials after the performance criterion was reached regardless of performance at the time of transition.	-*First cue set*: no limit; -*Second cue set:* 3000 trials -*Subsequent cue sets:* 1500 trials	-only generalisation trials, 20% of which were sample trials; sample trials introduced with third cue set
2-cue	- 80% performance criterion - Transition would occur 30 trials after the above criterion was reached regardless of performance at the time of transition	-*First cue set*: 1000 trials; -*Subsequent cue sets:* 600 trials -*Cue sets after normalisation of cues:* 3000 trials	-75% generalisation trials, 20% of which were sample trials -15% control trials

Note. "Correct" choices were choices of the highest value option available. Trial-by-trial performance was based on an exponential moving average with a time constant of 30 trials.

Task structure

Rats initiated a trial by poking and staying in the centre port. On "generalisation" choice trials, after 150 ms, two auditory cues out of the current set of 3 "generalisation" cues, played consecutively, one from each side speaker (500 ms each, with a 150 ms break in between, side and order pseudorandom; Figure 6.1-A, B). After an additional 150 ms, the left/right side ports illuminated and a white noise burst (50 ms) from the centre speaker signalled the end of holding time, allowing the rat to make a choice. Aborted trials caused by premature centre port exits were signalled by the houselight flashing (10 s). A nose poke response in either reward magazine caused delivery of the outcome associated with the auditory cue that had been presented at that location (0.1 ml boluses of 10% liquid sucrose). If no response was recorded within 5 s, an aborted trial would be recorded and signalled with flashing of the houselight (10 s). The chosen cue was replayed during reward delivery or omission (Figure 6.1-A and C right) to strengthen the association between the auditory cues identity and the distinct reward magnitudes. After an incorrect choice, i.e., choice of the lowest value option available in the trial, a time penalty of 5 s was added to the ITI (5 s).

On sample trials (20% of generalisation trials), the same cue would play on one side speaker and then on the other, and either port could be chosen for collecting a reward (Figure 6.1-C, D).

On control trials (15% of total), trial initiation and the holding time structure were as described, with the exception that here the same auditory cue played twice from the *centre* speaker instead of from one of the side speakers (Figure 6.1-E, F). The cue played was one from a set of two, associated with either a large or small reward, that the rat would receive if the correct reward port was subsequently chosen ("control" cues). The correct reward port was cued by its light flashing during holding time. After the holding

period, both reward ports now became illuminated during the choice period. The sound would again replay during reward/no reward delivery.

The 3 generalisation cues were replaced by 3 new ones using the same reward contingencies after either (i) an 80% performance criterion was reached (based on an exponential moving average (EMA) with a time constant of 30 trials) plus 30 additional rewards, or (ii) if the rats did not reach the performance criterion after 600 rewards with the same cues.

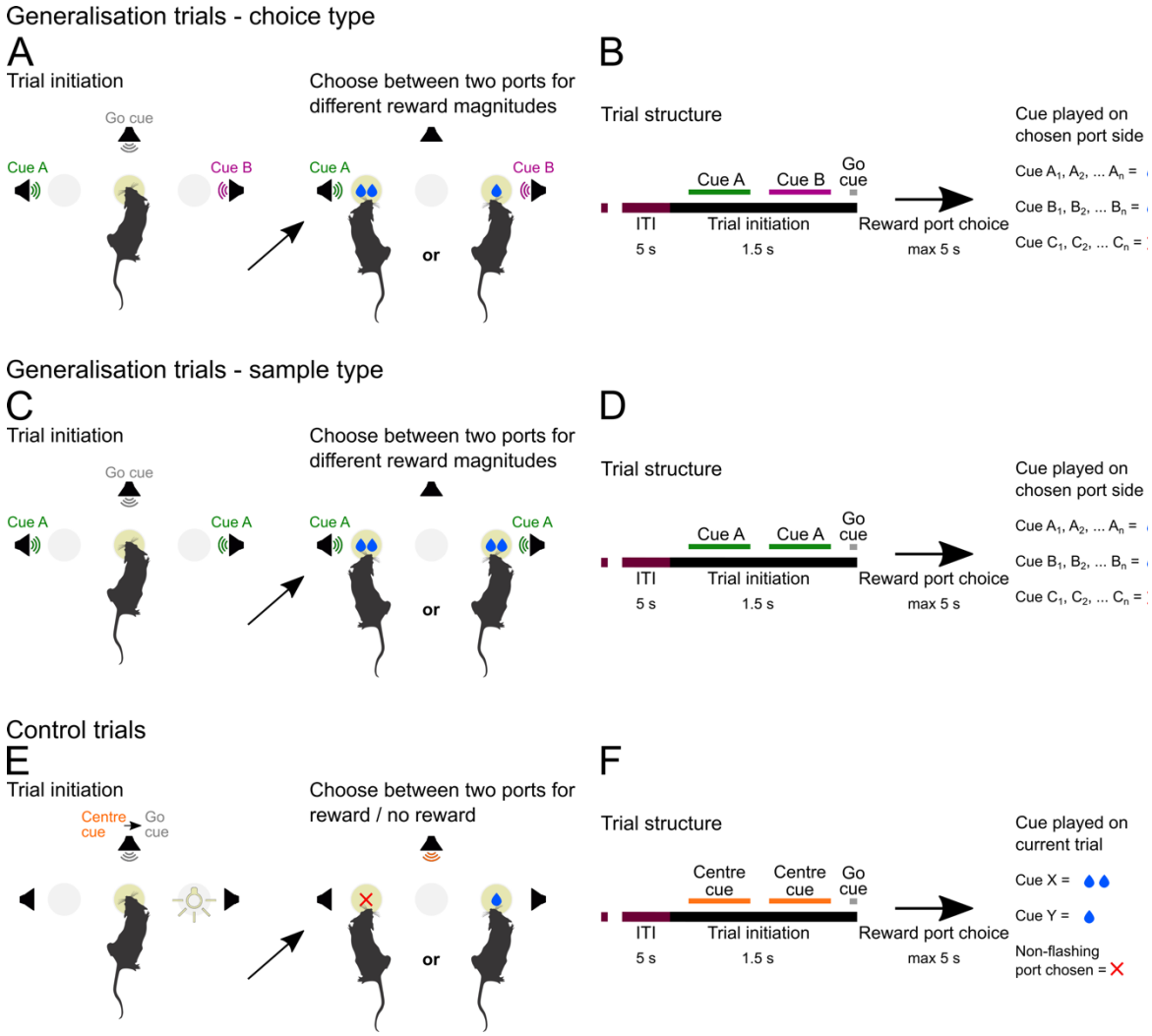


Figure 6.1 Behavioural paradigm. (**A**) In generalisation choice trials, rats poked in a trial initiation port at the centre of the box while two auditory cues played, one from each side speaker, followed by a Go Cue. Then, rats chose between two side ports for receiving rewards of different magnitudes. Two of a set of 2 or 3 auditory cues (in different task versions)

played in each choice trial. The sound would again replay during reward/no reward delivery in all the trial types. (**B**) Times of each step of the choice trials. After a performance criterion was reached (80% correct responses in EMA with a time constant of 30 trials), the 3-cue set was replaced by a new one with the same transitive relationship. (**C**) In sample trials, rats poked in a trial initiation port at the centre of the box while one auditory cue played, once from each side speaker, followed by a Go Cue. Then, rats chose between two side ports for receiving a reward. The reward would have the same magnitude regardless of the chosen side port. The sound would again replay during reward/no reward delivery. (**D**) Times of each step of sample trials. (**E**) In control trials, rats poked in a trial initiation port at the centre of the box while one auditory cue played from the *centre* speaker, follow by a Go Cue. The sound played was one from a set of two, associated with either a large or small reward that the rat would receive if the correct reward port was subsequently chosen. The correct reward port was cued by its light flashing during holding time. After the holding period, both reward ports became illuminated during the choice period. (**F**) Times of each step of control trials. EMA: exponential moving average.

Auditory cues

The two control cues were selected from pulsed, stepped, or continuous sinusoidal waves of different frequencies and remained fixed for an animal throughout training. In the generalisation trials, the cues in each block were randomly drawn from a library of 100 synthesised sounds (based on the gaussian sound synthesis method used by McDermott et al., 2011) and from a library of 4876 bird sounds (*Cornell Guide to Bird Sounds: Essential Set for North America*) and replaced after a behavioural criterion was reached (see *Behavioural Training*).

6.3.2 Data analysis

On choice trials, choices were considered correct when the rats chose the reward port on the side where the sound associated with the larger reward played on that trial. To assess whether rats were getting faster at learning the reward contingencies, we applied a linear regression to the number of trials each rat took for triggering a cue set transition for each cue set experienced. Significance was assessed by comparing the slopes of the best fit line for each rat against a distribution of mean zero. To investigate which task components

rats used for guiding their trial-by-trial choices, we fitted a general linear model (GLM) with a constant term and a set of regressors to the rats' choices, including a left/right choice bias, cue order, and chosen reward size. As an independent test of whether rats had learned the cue-reward associations, we also compared reaction times on sample trials with the different generalisation cues.

We also wanted to investigate whether the rats learned and retained the reward contingencies on the control trials. Due to the high proportion of aborted control trials, it was not possible to compare the reaction times between the two control cues, so, instead, we compared the proportion of aborted trials in which the cues associated with a small or a large reward played using paired samples t-tests.

6.4 Behavioural paradigm development and learnings

6.4.1 Training rats to hold at initiation port during long periods

In this behavioural paradigm, it was important that there was a period where the rats were presented with the two auditory cues consecutively prior to them making their choice. Thus, rats needed to be trained to remain in the trial initiation port for 1500 ms. Therefore, the first step in developing this behavioural paradigm was establishing a protocol for training rats to hold at the trial initiation port for this long. We trained rats in increasingly longer holding times that would increase or decrease by 5 ms after every successful or unsuccessful trial initiation. Most rats reached the target holding time within three training sessions, after a total of 1178 ± 135 trials on average (Figure 6.2).

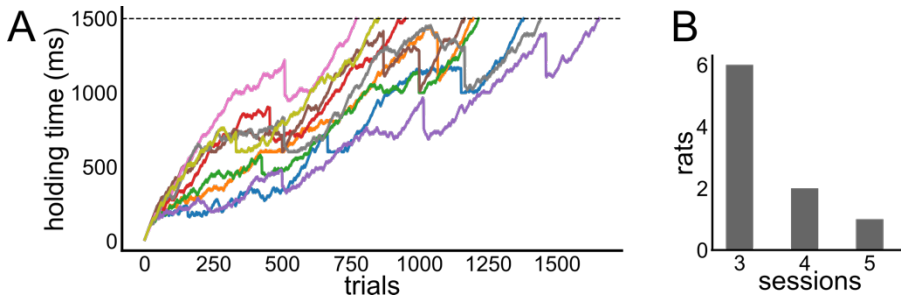


Figure 6.2 Training on increasingly long holding times. (**A**) Rats were trained on increasingly long holding times for initiating a new trial until a target of 1500 ms. After each successful or unsuccessful trial initiation, holding time would increase or decrease by 5 ms. Each coloured line represents the holding times at each trial for each rat in the last cohort trained on the task. (**B**) Number of training sessions it took each rat to reach the target holding time. *n* = 9 rats.

6.4.2 Performance on the 3-cue serial learning task

After learning the associations between cue identity and reward magnitudes on the 3-cue task, rats should choose the cue with the larger reward in each trial, i.e., large over small, small over no reward, and large over no reward. When trained on the first set of cues on the 3-cue task, 7 out of 9 rats did not reach performance criteria for triggering a transition to a new cue set (see Table 6.1 in section 6.3 *Methods summary* for criteria details), so we introduced new cue pairs (forced transition) after 13 behavioural sessions (see Figure 6.3 for the proportion of correct responses in the first 500 trials). The 2 rats that triggered a transition took 92 and 4882 trials.

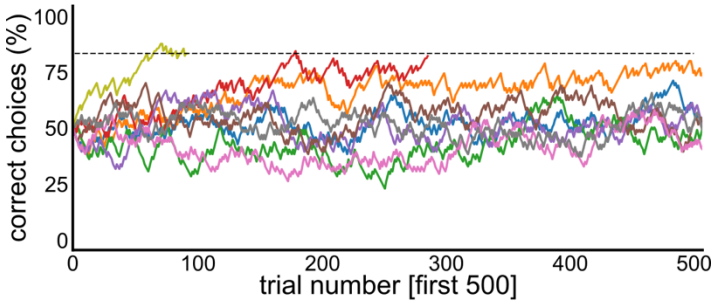


Figure 6.3 Performance on the first cue set on the 3-cue task. Each rat's proportion of correct choices (i.e., choice of reward port on the side where auditory cue associated with larger reward in each trial played) on choice trials with the first cue set, on the 3-cue task. The dashed line represents the performance criterion for replacing the 3-cue set (80% correct

choices in an exponential moving average, time constant = 30 trials). Only the first 500 trials are represented.

We wanted to investigate whether rats were able to learn the task structure and apply it to new cue sets, in which case we would expect to observe a reduction in the number of trials taken to reach performance criteria (Figure 6.4-A). To this end, we applied a linear regression to the number of trials each rat took for triggering a cue set transition and obtained the slope for each rat's learning curve. Subsequent analyses revealed that the distribution of these slopes was significantly different from zero ($t_{(8)}$ = -5.74, p < 0.001; Figure 6.4-B), and their average was negative (-60.32 ± 14.86). This means that the rats took fewer trials to reach criteria as they experienced new sets of cues.

Next, we wanted to understand which strategies rats were using for choosing a reward port on choice trials. Thus, we applied a general linear regression (GLM) to the choices in each choice trial (across all trials), including as regressors (1) the side bias towards a port, (2) whether the first cue played on the side of that port, and (3) whether the cue associated with the largest available reward played on the side of that port. This analysis revealed that reward contingency was the main predictor of choosing a reward port (Figure 6.4-C). In addition, there was also a small effect of cue order, with rats being more likely to choose the reward port on the side where a cue played last on that trial.

Then, to investigate whether rats established a rank order of the cues associated with large, small, and no reward, we compared the reaction times (i.e., trial initiation to reward port choice) on sample trials, revealing a main effect of cue reward contingency ($F_{(2, 16)}$ = 7.98, p = .004; Figure 6.4-D). Follow up pairwise comparisons revealed that rats were significantly faster at collecting rewards in sample trials in which the cue associated with a large reward played than those in which the small reward or the no reward cues

played (large vs small cues: p = .005; large vs no reward cues: p = .003). Surprisingly, reaction times were similar in sample trials with small and no reward cues (p = .961). These results suggest that rats struggled to establish a rank order of the 3 cues: while they learned which cue was associated with the largest reward, they did not develop a strong preference for the small reward cue over that associated with no reward. To confirm whether this was the case, we compared the proportion of correct choices on each cue pair (i.e., large vs small reward, large vs no reward, small vs no reward) after rats had reached performance criteria for triggering a transition, revealing no significant main effect of cue pair ($F_{(2, 14)}$ = 3.45, p = .060; Figure 6.4-E). However, on trials in which the two cues were those associated with small and no reward, rats performance was not significantly different from chance level (large vs small: $t_{(7)}$ = 3.15, p = .016; large vs no reward: $t_{(7)}$ = 4.41, p = .003; small vs no reward: $t_{(7)}$ = 0.09, p = .934), confirming rats did not develop a strong preference for the small over the no reward cue.

Moreover, even the best rat that completed a total of 16 sets of cues over 33 behavioural sessions was still taking on average 662 ±176 trials to reach criterion over its last 10 cue sets (Figure 6.4-F). Therefore, as our planned neuronal recordings for investigating inference representation in the brain would require rats to establish a rank order of the 3 cues and to learn several pairs of cues within a single session, we decided to try a different task design.

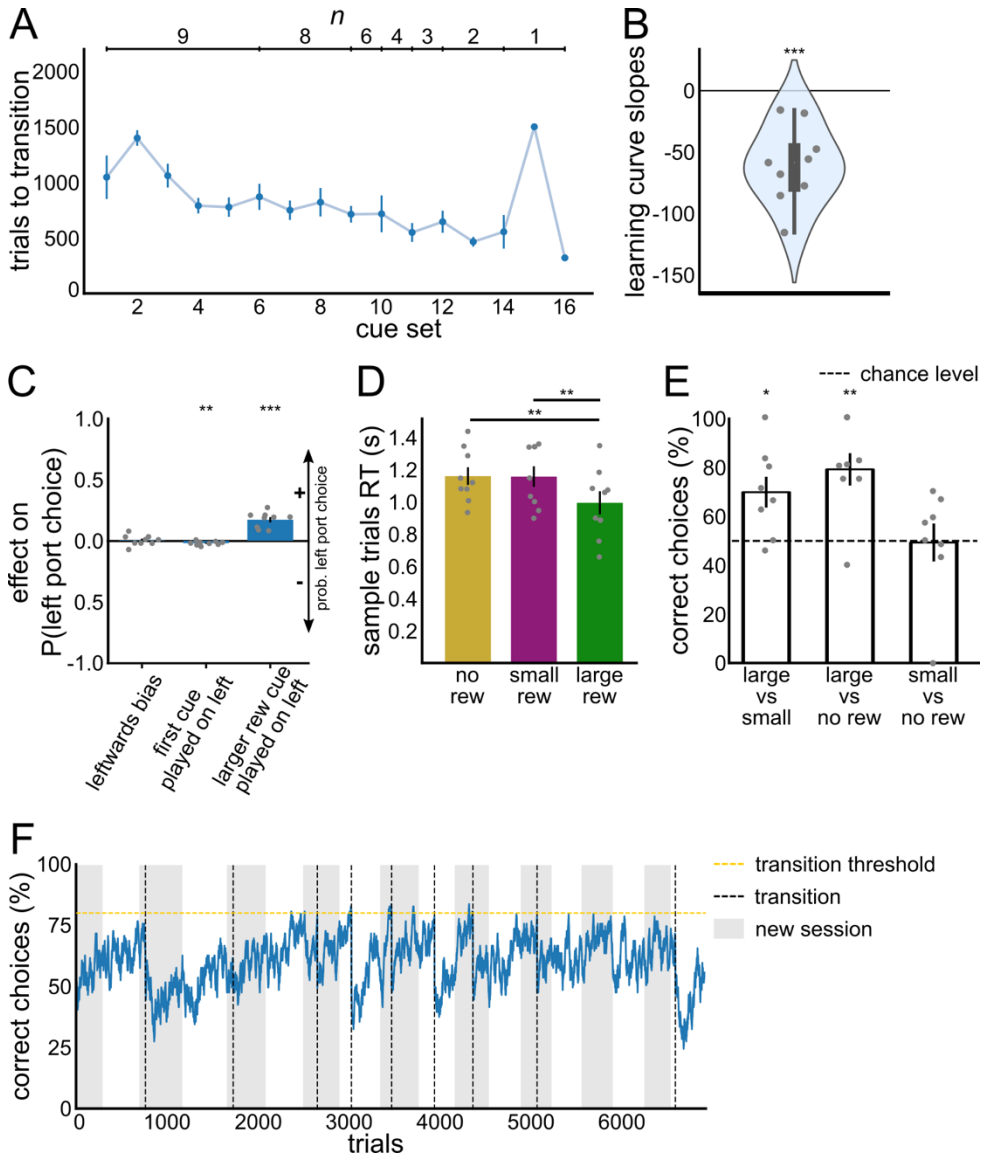


Figure 6.4 Performance after the second transition on the 3-cue task. (**A**) Number of trials taken to transition as a function of cue set after the second transition on the 3-cue task. A forced transition occurred if rats did not reach the performance criterion (80% correct choices in EMA, time constant = 30 trials) after 1500 trials with the same set of cues. The top axis represents the number of animals that experienced each number of cue sets. (**B**) Slopes of the best fit line of each subject number to transition across cue sets, obtained with a linear regression model. (**C**) Average effects size from a general linear model of leftwards (i.e., left side port) bias, first cue in trial playing on the left side, and large reward associated cue playing on the left side on the probability of choosing the left reward port. This analysis was applied to all choice trials. Statistical significance was assessed by comparing the distribution of effects with a random distribution of mean zero. (**D**) Average reaction time (seconds) for reward port entry on all sample trials in which the cue played was the auditory cue associated with no reward, a small or large reward. (**E**) Proportion of correct choices (i.e., choice of reward port on the side where auditory cue associated with larger reward in each trial played) for each choice pair on trials after performance criteria to transition had been reached. (**F**) Behavioural performance on the last 10 cue pairs of the rat with the most cue set transitions. Error bars depict mean ± SEM. Scatter plots overlaid with bar plots represent individual animals'

averages. n = 9 rats. One asterisk (*) indicates $p < .05$; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests. rew: reward; RT: reaction time.

6.4.3 Performance on the 2-cue serial learning task

We next trained a new cohort of 9 rats on a simpler task with only 2-cues associated with large or small rewards. While such task would not allow us to test for transitive inference, it would still allow us to investigate how the brain encodes abstract task structure to facilitate generalisation. When trained on the first set of two cues, rats reached performance criterion (i.e., 80% correct choices) after 337 ± 103 trials on average (Figure 6.5-A, B). However, one of the nine rats did not reach the performance criterion and a new cue set was introduced after 1000 trials (forced transition).

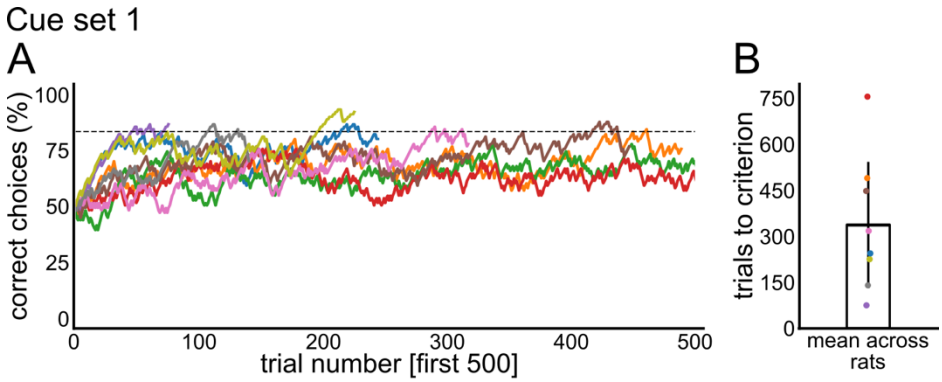


Figure 6.5 Performance on the first cue set on the 2-cue task. (**A**) Each rat's proportion of correct choices (i.e., choice of reward port on the side where auditory cue associated with large reward played) on choice trials with the first cue set, on the 2-cue task. The dashed line represents the performance criterion for replacing the 2-cue set (80% correct choices in an exponential moving average, time constant = 30 trials). Only the first 500 trials are represented. (**B**) Average number of choice trials to criterion. One of the rats (green line in A) is not represented here as it did not reach criterion performance with the first set of cues. Error bars depict mean ± SEM. Scatter plots overlaid with bar plots represent individual animals' averages. n = 9 rats.

A series of analyses similar to those employed for the analysis of the rats' behaviour on the 3-cue task version revealed that, as the rats experienced more sets of cues, they took fewer trials to reach the performance criterion for transition ($t_{(8)} = -2.40$, $p = 0.04$; Figure 6.6-A, B), consistent with a generalisation of the task abstract structure.

Moreover, cue reward contingency was the main factor influencing the choice of a reward port (Figure 6.6-C), and as seen for the first 3-cue set, rats took significantly longer to respond on sample trials when the small reward was available compared to large reward sample trials ($t_{(8)} = 8.88$, $p < .001$; Figure 6.6-D). These results show that most rats were discriminating the auditory cues, learned their reward contingencies and used their identity to guide behaviour and maximise reward.

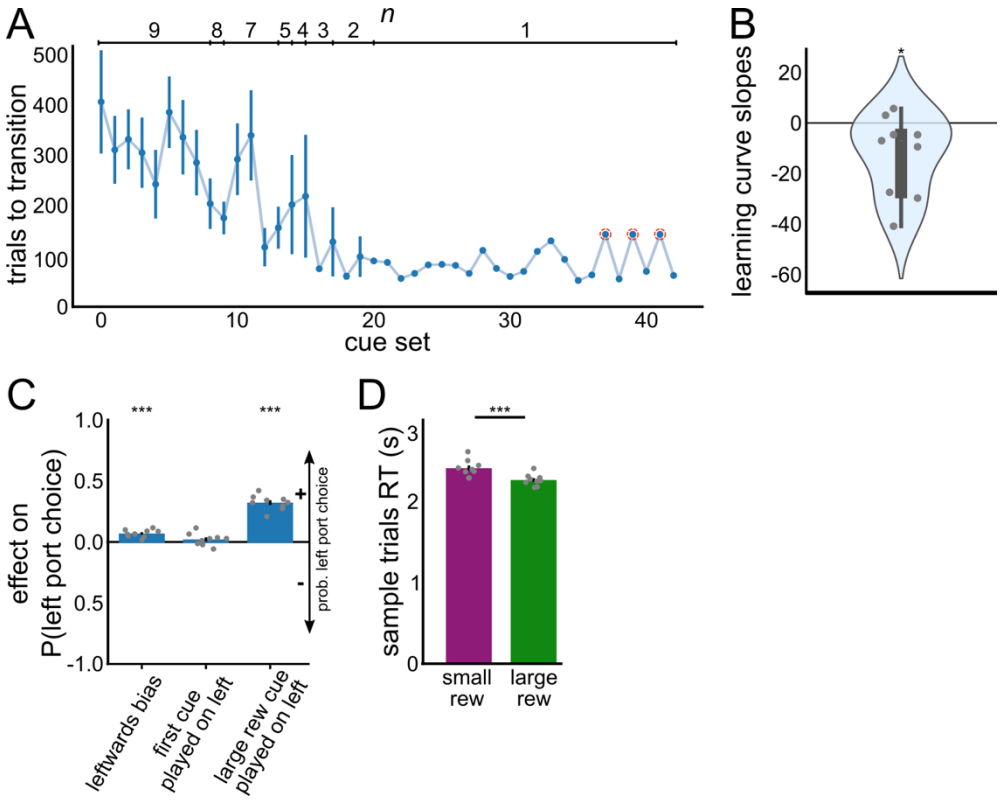


Figure 6.6 Performance after the first transition on the 2-cue task. (**A**) Number of trials taken to transition as a function of cue set after the first transition on the 2-cue task. A forced transition occurred if rats did not reach the performance criterion (80% correct choices in an exponential moving average, time constant = 30 trials) after 600 trials with the same set of cues. Datapoints with red dashed circles were manually introduced forced transitions for one subject that struggled with particular cue sets. The top axis represents the number of animals that experienced each number of cue sets. (**B**) Slopes of the best fit line of each subject number to transition across cue sets, obtained with a linear regression model. (**C**) Average effects size from a general linear model of leftwards (i.e., left side port) bias, first cue in trial playing on the left side, and large reward associated cue playing on the left side on the probability of choosing the left reward port. Statistical significance was assessed by comparing the distribution of effects with a random distribution of mean zero. (**D**) Average reaction time (seconds) for reward port entry on all sample trials in which the cue played was the auditory cue associated with a small or large reward. Error bars depict mean ± SEM. Scatter plots overlaid with bar plots represent

individual animals' averages. n = 9 rats. One asterisk (*) indicates p < .05; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests—rew: reward; RT: reaction time.

While the results above appear to suggest rats learned the abstract task structure and were able to generalise across cue sets, close inspection of the data showed that performance from cue set to cue set could be surprisingly variable. For example, one rat took just 64 ± 15 trials on average to reach the threshold for transition in the last 20 cue sets it experienced. Nonetheless, it occasionally displayed a strong preference for the cue associated with a small reward, requiring a "forced transition" (Figure 6.7-A). To better understand this, we compared the mean spectrograms of the different auditory cues associated with either large or small rewards that this rat experienced.

The mean power across all frequencies was similar for small and large reward cues on both cue pairs that appeared to be discriminable or indiscriminable (discriminable: : $t_{(16)} = 1.37$, $p = 0.188$, Figure 6.7-B, C; indiscriminable: $t_{(2)} = 0.25$, $p = 0.829$, Figure 6.7-D, E). However, a comparison of the mean power across high frequencies (i.e., above 6 kHz) revealed differences between small and large reward cues from discriminable pairs ($t_{(16)} = -5.65$, $p < 0.001$; Figure 6.7-C). Such difference was not present on the pairs the rat struggled with ($t_{(2)} = 1.97$, $p = 0.188$; Figure 6.7-E). A similar analysis applied to the cues that appeared to be discriminable or indiscriminable by the other rats in the cohort replicated the findings for this example rat (mean power across high frequencies of small vs large reward cues for discriminable pairs: $t_{(7)} = -19.80$, $p < 0.002$; for indiscriminable pairs: $t_{(7)} = -1.13$, $p = 0.295$).

These results suggest that rather than learning the specific identity of the cues of each new pair, rats generalised the reward contingencies over low- and high- frequencies —specifically, rats associated small and large rewards with low and high frequencies, respectively.

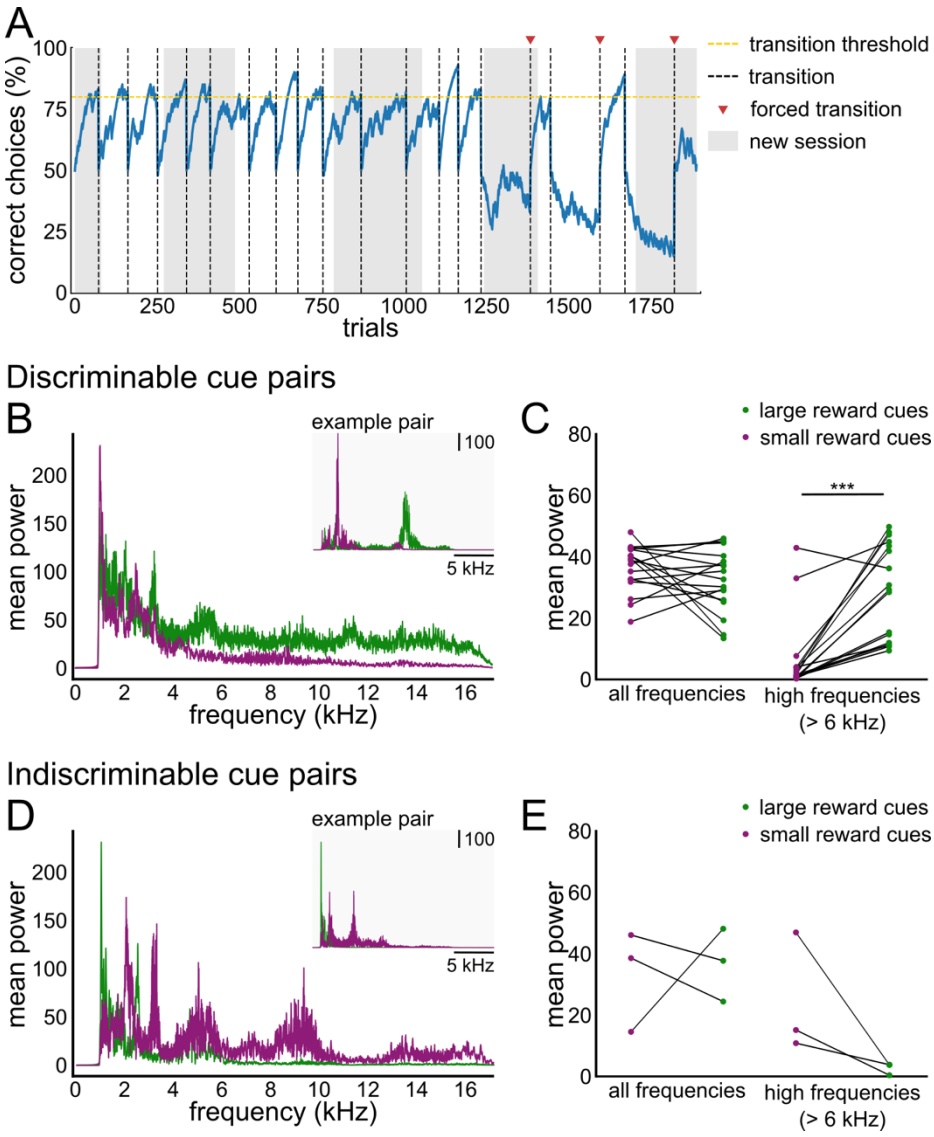


Figure 6.7 Generalisation over high- and low-frequency cues. (**A**) Behavioural performance on the last 20 cue pairs of the rat with the most cue set transitions (see figure 6.6-A). (**B**) Mean spectrogram of all cues associated with small and large rewards which were discriminated by the same rat. Inset plot depicts an example large and small reward cue pair experienced by this rat. (**C**) Mean power over all frequencies or on high frequencies (> 6 kHz) for the cues in B. Black lines link cues associated with a small or large reward that were presented together (i.e., are from the same pair). (**D**) Same as B, but for the cue pairs that this rat was not able to discriminate (see forced transitions in A). (**E**) Mean power over all frequencies or on high frequencies for the cues in D. One asterisk (*) indicates $p < .05$; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests.

6.4.4 Performance on the 2-cue serial learning task after normalisation of auditory cues

OFC neurons have been shown to respond to sounds (Winkowski et al., 2018). Therefore, a correlation between sound frequencies and reward magnitude learned by the rats would make it challenging to determine the neuronal encoding when analysing electrophysiology data. Moreover, the results discussed above highlighted the need to make auditory cues similarly salient.

Therefore, we continued to train the same cohort of rats after modifying the auditory cues so that the peak power of their spectrogram was shifted to 8 kHz, which is the most sensitive frequency of the rat audiogram (Heffner & Heffner, 2007), and then normalised the average power so that it was the same for each stimulus (Figure 6.8). Additionally, to increase the rats' exposure to each stimulus, we increased the duration of each auditory cue from 500 to 750 ms, bringing the total holding time at the nose port for trial initiation to 2 s. We anticipated the rats would take several sessions to develop new strategies for performing the task, i.e., paying attention to specific features of each sound instead of choosing that with the highest frequency. Therefore, we increased the maximum number of trials before triggering a forced transition from 600 to 3000.

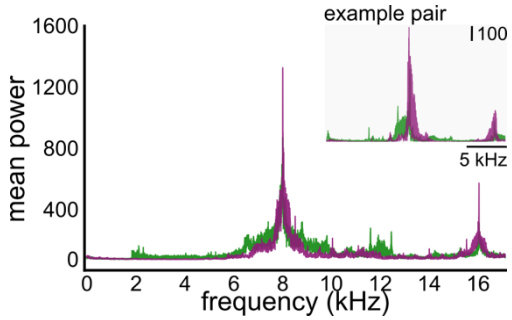


Figure 6.8 Auditory cues with peak power shifted to 8 kHz. Mean spectrogram of a set of 20 cues associated with small and large rewards normalised so that the peak power of their spectrogram was shifted to 8 kHz. Inset plot depicts an example of a large and small reward cue pair.

When we analysed the data from 40 sessions with the new cues, we observed that, unlike with the previous cues, rats were not improving significantly as they experienced more sets of the normalised cues (distribution of learning curve slopes not significantly different from zero: $t_{(8)} = -1.65$, $p = 0.137$; Figure 6.7-A, B). Cue reward contingency was still the main factor influencing the choice of a reward port (Figure 6.9-C), though its influence was reduced compared to before we modified the auditory cues. On sample trials, rats again still took longer to respond when the only available reward was small than when it was large ($t_{(8)} = 5.18$, $p < .001$; Figure 6.9-D). However, this difference was considerably smaller (16.15 $\pm$ 0.45 ms) than with the previous cues (137.45 $\pm$ 5.57 ms). Focusing on the rat described in the previous section (see Figure 6.7), it also correctly discriminated more cue sets with the new, normalised cues (see Figure 6.9-E). However, the average number of trials it took to reach criterion performance for triggering a transition increased from 64 to 526 trials, triggering only 15 transitions in 40 behavioural sessions (9377 trials). Taking these findings into account, we terminated our training on this task.

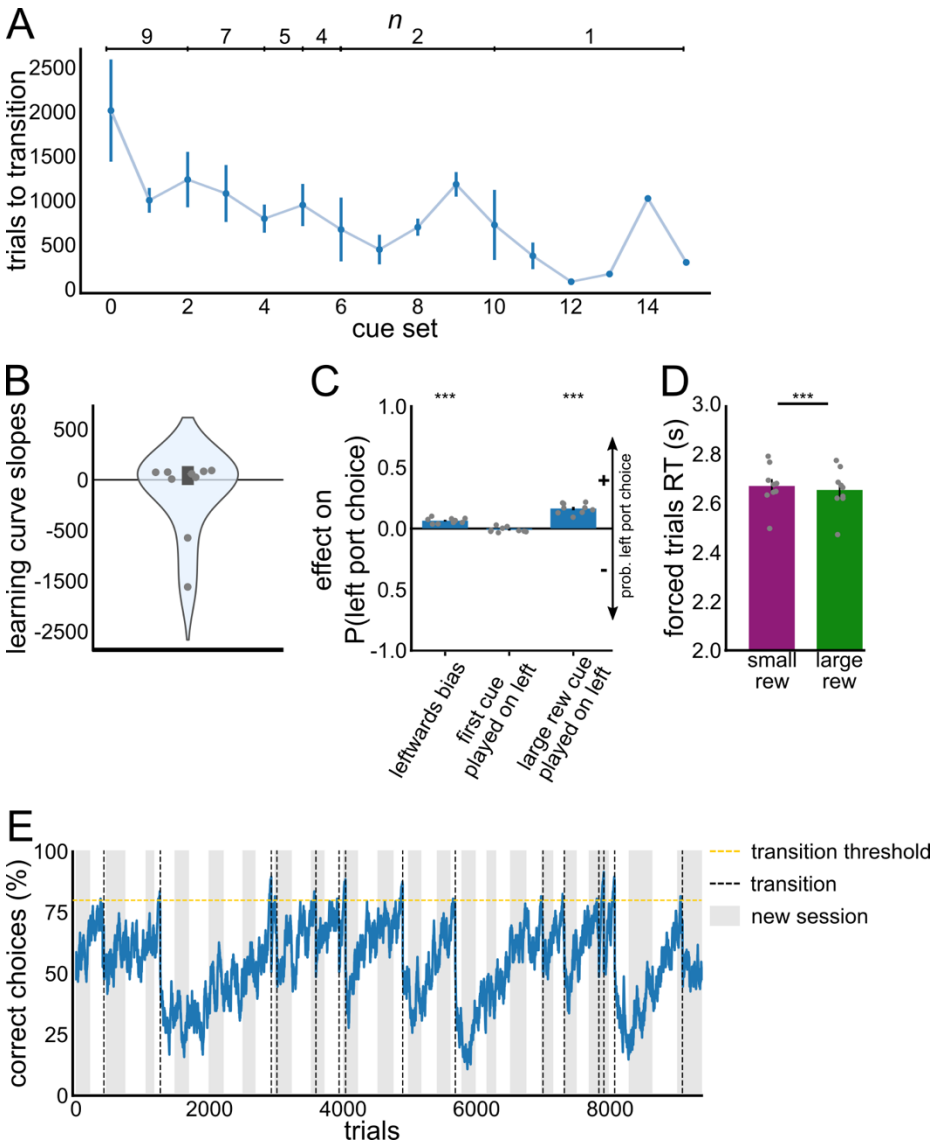


Figure 6.9 Performance on the 2-cue task after normalisation of the cues. (**A**) Number of trials taken to transition as a function of cue set on the 2-cue task. A forced transition occurred if rats did not reach the performance criterion (80% correct choices in EMA, time constant = 30 trials) after 3000 trials with the same set of cues.. The top axis represents the number of animals that experienced each number of cue sets. Represented are the first 40 behavioural sessions with the new cues. (**B**) Slopes of the best fit line of each subject number to transition across cue sets, obtained with a linear regression model. (**C**) Average effects size from a general linear model of leftwards (i.e., left side port) bias, first cue in trial playing on the left side, and large reward associated cue playing on the left side on the probability of choosing the left reward port. Statistical significance was assessed by comparing the distribution of effects with a random distribution of mean zero. (**D**) Average reaction time (seconds) for reward port entry on all sample trials in which the cue played was the auditory cue associated with a small or large reward. Error bars depict mean ± SEM. Scatter plots overlaid with bar plots represent individual animals' averages. (**E**) Behavioural performance of one rat (the same as in Figure 6.7) on the cue pairs experienced after normalisation of the cues. n = 9 rats for panels A-D. One asterisk (*) indicates $p < .05$; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests. rew: reward; RT: reaction time.

6.4.5 Performance on control trials

To be able to investigate task state representation in the task version with 2 cues, we included trials in which different, fixed auditory cues were associated with a large or small reward but within a different task structure. Here, the identity of the auditory stimuli was still informative of the reward magnitude, but now the cues were played from a speaker in a different location (centre vs side speakers), and the location of available reward was signalled with light rather than sound (see Figure 6.1 in section 6.3 *Methods summary*).

While training the animals, we noticed rats seemed to intentionally abort certain trials by exiting the trial initiation port soon after the auditory cue started playing. Due to the high proportion of aborted control trials, it was not possible to compare the reaction times between the two control cues, so, instead, we compared the proportion of aborted trials. This analysis revealed that rats aborted significantly more control trials in which the cue played was that associated with a small reward ($t_{(8)} = 5.79$, $p < .001$; Figure 6.10-A, B).

We tried increasing the cost of aborting trials by temporarily increasing the abort trial time off, initially from 5 to 10 s and then to 15 s, without any noticeable improvement (not shown). Possibly, the rats had become habituated to aborting trials when they heard the small reward cue and increasing the abort trial time off would not change this behaviour. Therefore, we replaced the auditory cues on these trials with new ones. Interestingly, on the first session after the new cues were introduced, rats were significantly more likely to abort trials in which the cue associated with large reward was played ($t_{(8)} = 3.42$, $p = .009$; Figure 6.10-C, D). A possible explanation is that the rats generalised a feature of the auditory cues. When we introduced the new cues, we replaced

the 5 kHz pulsed sine wave cue (cue A) for the 11 kHz pulsed sine wave cue (cue C) and the 12 kHz sine wave cue (cue B) for the 3 to 8 kHz stepped sine wave cue (cue D). Thus, if rats generalised over high- and low-frequency sounds, they would have experienced a reversal of the reward contingencies. Ten sessions after the new cues were introduced, rats were again more likely to abort trials in which the small reward cue was played ($t_{(8)}$ = 3.67, p = .006; Figure 6.10-E, F).

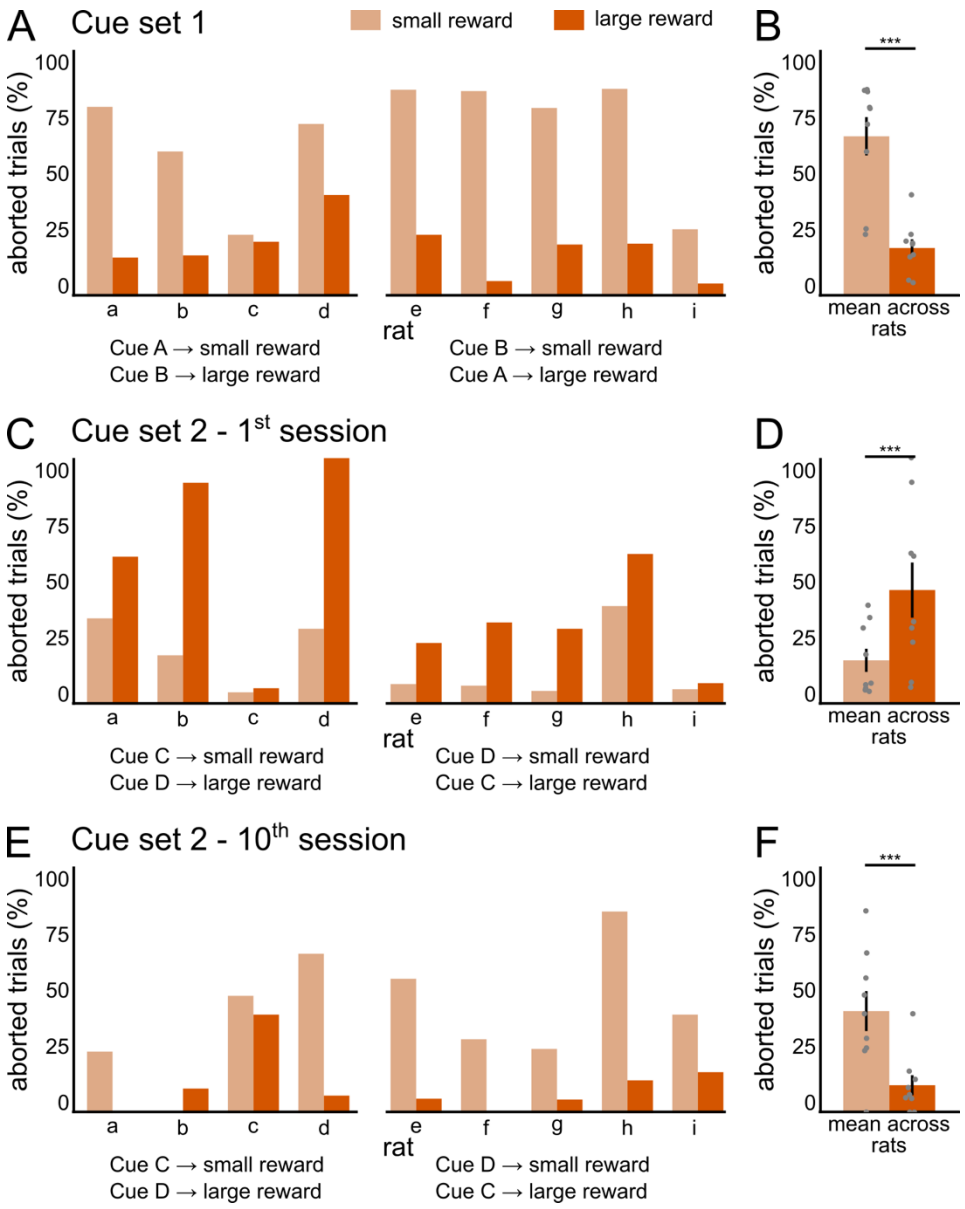


Figure 6.10 Abort rate on control trials. (**A**) Each rat's proportion of aborted trials (i.e., early trial initiation port exit or no choice of either reward port) on control trials in which the

auditory cues associated with small or large rewards were presented, on the last session with cue set 1 (i.e., 5 kHz pulsed sine wave or 12 kHz sine wave, counterbalanced across rats). (**B**) Average proportion of abort trials represented in A, across rats. (**C**) Same as A, but for the first training session after cue set 1 was replaced with cue set 2 (i.e., 11 kHz pulsed sine wave or 3 to 8 kHz stepped sine wave, counterbalanced across rats). (**D**) Average proportion of abort trials represented in C, across rats. (**E**) Same as C, but for the tenth training session after cue set 2 was introduced. (**F**) Average proportion of abort trials represented in E, across rats. Error bars depict mean ± SEM. Scatter plots overlaid with bar plots represent individual animals' averages. n = 9 rats. One asterisk (*) indicates $p < .05$; otherwise, m asterisks indicate $p < 10^{-m}$. See text for details on statistical tests. Cue A: 5 kHz pulsed sine wave with pulses at 12 Hz; cue B: 12 kHz sine wave; cue C: 11 kHz pulsed sine wave with pulses at 8 Hz; cue D: 3 to 8 kHz stepped sine wave with steps at 10 Hz.

6.5 Discussion

An increasingly large number of studies suggest conceptual knowledge is organised in the brain into cognitive maps (as reviewed in Behrens et al., 2018). A clear understanding of the emergence and organisation of these cognitive maps and the relative contribution of different brain regions to these processes is essential to achieving a multi-level understanding of how we represent information and make decisions when confronted with new problems (Kaplan et al., 2017). With this in mind, we aimed to develop a novel behavioural paradigm for investigating non-spatial cue-driven decision-making in an operant context. Initially, we attempted to develop a transitive inference task with sets of 3 auditory cues associated with distinct reward magnitudes; the 3-cue sets would be replaced after a performance criterion was reached. Unfortunately, rats struggled to establish a rank order of the 3 cues, failing to develop a clear preference for the small reward cue over that associated with no reward. Moreover, rats learning of the contingencies was too slow for the planned analysis of neuronal data.

Thus, we then decided to trained rats on a generalisation task with only 2-cues. We anticipated that this simpler task would be easier to learn while still allowing us to investigate abstract task structure representation if we introduced control trials for

decorrelating some of the trial features. While at times behavioural performance was very promising (Figure 6.6), careful inspection of the auditory cues revealed rats were generalising on an unexpected feature of the auditory cues: mean power over low vs high frequencies (Figure 6.7).

OFC neurons have been shown to respond to sounds (Winkowski et al., 2018) and, therefore, a correlation between sound frequencies and reward magnitude learned by the rats would make it challenging to determine the neuronal encoding when analysing electrophysiology data. To counteract this, we modified the cues so that they had similar peak power frequencies by shifting the peak power of their spectrogram to 8 kHz, which is the most sensitive frequency of the rat audiogram (Heffner & Heffner, 2007). However, after this modification, rats took more trials to trigger a cue transition, usually requiring several sessions. Thus, learning was too slow for neuronal recordings as, ideally, rats would need to learn several pairs of cues within a single recording session. Moreover, the rats did not exhibit signs of significant improvement as they experienced new cue sets (Figure 6.9). When reviewing the literature on auditory stimuli discrimination, we learned that rats struggle to discriminate complex sounds if their frequencies are similar. Matching our results, D'Amato & Salmon (1984) observed that rats were able to rapidly discriminate auditory cues with different frequencies, but the performance was around chance if cues were trains of random notes with similar mean frequency: "None of the four rats in the random-notes condition showed the slightest sign of learning the discrimination. […] at the end of 75 sessions (3,600 trials) they were still responding at a chance level, at which point they were terminated." (D'Amato & Salmon, 1984). Taking these findings into account, we terminated our experiments.

Our learnings from these behavioural experiments suggest that stimuli in other dimensions such as the olfactory space might be better suited for investigating

generalisation using rodent tasks requiring a large number of different stimuli. Unfortunately, it was not feasible to adapt the behavioural boxes and set up a new behavioural paradigm using odours within the timeframe of this thesis, but this would certainly be worth exploring in future experiments. In fact, recent studies used a novel odour sequence behavioural paradigm for investigating the representation of cognitive maps in the brain (Zhou, Gardner, et al., 2019; Zhou, Montesinos-Cartagena, et al., 2019).

Even though we did not achieve our goal of establishing a novel behavioural paradigm for task-dependent representations in the brain, we show that rats are able to learn a behavioural task involving a mix of different trial types with complex structures. On the generalisation trials, the identity of the cues indicated the size of the reward on the port on the side where they played, whereas on control trials the identity of the auditory stimuli was still informative of the reward magnitude but now the location of available reward was signalled with light rather than sound. Interestingly, rats also appeared to generalise over the high- and low-frequency cues used on control trials, aborting most trials in which the cue associated with the small reward cue played. While the high proportion of abort trials was not ideal for intended neuronal data analysis, in future experiments, introducing a third cue associated with no reward could help reduce aborting by rescaling the rats' perception of the small reward.

Additionally, we successfully established a protocol for training rats to hold still for long periods (up to 2 s) during the presentation of several stimuli. This protocol might be used in the future to develop behavioural tasks in which long holding periods are necessary, e.g., to reduce impulsivity, increase sunk time cost or present to the animals several stimuli before they make a decision.

Contributions & acknowledgements

For the studies presented in this chapter, Ines Barreiros (IB), Sebastian Vasquez-Lopez (SVL), Mark Walton (MEW), Tim Behrens (TB), and Thomas Akam (TA) planned the experiments; IB and SVL built the behavioural boxes with help from TA; IB and SVL coded the task; IB collected the data; IB analysed and presented the data.

This work was supported by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship awarded to IB; by a Wellcome Senior Research Fellowship awarded to MEW (202831/Z/16/Z); and a Wellcome Collaborator award (214314/Z/18/Z) to TB.

7. General Discussion and Conclusions

7.1 Summary of findings

To thrive in an ever-changing world, individuals must rapidly and flexibly adapt their behaviour according to their needs and environmental constraints. Some problems require a period of learning through trial and error; others can be quickly solved by generalising and integrating information from previous experiences. This thesis aimed to investigate how the brain represents information to facilitate behavioural flexibility during decision making. To this end, we explored anatomical and functional features of the brain systems that support these learning and decision-making processes by focusing on the role of the orbitofrontal cortex (OFC), the nucleus accumbens (NAc) and the dopaminergic reward system. The main goals were: (1) to investigate whether OFC functional heterogeneity is underpinned by differences in anatomical connectivity (Chapter 3); (2) to study the contribution of dopamine to haemodynamic signatures of reward prediction errors (RPEs) during flexible reward-guided decision making (Chapter 4); (3) to investigate the neuronal correlates underlying maladaptive learning in hyperdopaminergic states similar to those thought to elicit psychotic symptoms in neuropsychiatric conditions (Chapter 5); (4) and to develop a novel rodent behavioural paradigm for investigating how the brain encodes information during generalisation (Chapter 6).

My key findings are:

(1) The anterior lateral, posterior lateral, and posterior ventral portions of the rat OFC receive distinct and topographically organised inputs from various thalamic nuclei,

basolateral amygdala, and various cortical areas. Importantly, differences in inputs between anterior and posterior lateral OFC are as striking and clearly defined as already established differences on the mediolateral axis between lateral and ventral OFC.

(2a) RPE-like encoding in NAc haemodynamic signals may depend on environmental demands in a rodent probabilistic reversal-learning task, appearing more pronounced after a reversal of reward contingencies, when animals need to update their value estimates, than when responding to probabilistic but known reward contingencies.

(2b) NAc haemodynamic responses remain largely unchanged after lesioning of NAc dopaminergic inputs, whereas those recorded from the OFC are significantly potentiated.

(3a) Both a single acute injection of a moderate dose of amphetamine and a repeated 2-dose ("sensitised") injection regime cause an impairment in updating reward contingencies after a reversal. This manifested even through into the session after amphetamine administration.

(3b) Both acute and repeated amphetamine cause an attenuation of NAc haemodynamic signals following amphetamine administration.

(4) Rats can learn a complex non-spatial auditory decision-making task. However, their ability to discriminate between pairs of complex auditory cues depends strongly on the peak frequencies of the cues.

Below, I put these findings into a broader context and discuss their implications as well as possibilities for future investigation.

7.2 OFC functional and anatomical heterogeneity

In Chapter 3, we investigated whether OFC functional heterogeneity is underpinned by differences in anatomical connectivity across its subdivisions. The OFC has been anatomically divided into a number of subregions along its medial–lateral axis, which behavioural research suggests have distinct functions (as reviewed in Izquierdo, 2017). Recently, evidence has emerged suggesting functional diversity is also present along the anterior–posterior axis of the rodent OFC (e.g., Bradfield et al., 2018; Panayi & Killcross, 2018). However, the patterns of anatomical connections that underlie these differences have not been well characterised. We studied the input patterns into the anterior lateral (ALO), posterior lateral (PLO), and posterior ventral (PVO) portions of the rat OFC by simultaneously injecting the retrograde tracer cholera toxin subunit B (CTB) coupled to different Alexa Fluor conjugates (488, 594, 648) into each of these subdivisions. Our methodological approach allowed us to simultaneously compare the density and input patterns into these OFC subdivisions. We observed distinct and topographically organised projection patterns into ALO, PLO, and PVO from the mediodorsal and the submedius nuclei of the thalamus. We also observed different levels of connectivity strength into these OFC subdivisions from the amygdala, motor cortex, sensory cortices and medial prefrontal cortical structures, including medial OFC, infralimbic and prelimbic cortices. Interestingly, while labelling in some of these input regions, such as the amygdala, revealed only a gradient in connectivity strength, other regions, e.g. paratenial thalamic nucleus, seem to project almost exclusively to specific OFC subdivisions.

Moreover, differences in input patterns between ALO and PLO on the anterior–posterior axis were as pronounced as those between PLO and PVO, two regions on the medio-lateral axis that have long been treated as distinct.

Even functionally distinct brain regions often show substantial overlap in their inputs and outputs (Passingham et al., 2002). It is, thus, not surprising that different portions of the OFC show the graded connectivity patterns we observed. It might, therefore, be useful to consider the question of what makes two adjacent regions distinct. I believe that integrating neuroanatomical characterisation with functional differences is essential to distinguish brain regions. For instance, ventral (VO) and lateral (LO) OFC are generally considered separate regions (Krettek & Price, 1977; Price, 2006; Ray & Price, 1992). This distinction is supported by differences in efferent (e.g. Schilman et al., 2008) and afferent (see Table 3.2 in Chapter 3) projections and functional distinctions (e.g. Balleine et al., 2011; Izquierdo, 2017). In light of this standard distinction between VO and LO, we ought to consider a division between the anterior and posterior portions of LO. A functional dissociation between ALO and PLO has been previously reported (Panayi & Killcross, 2018) –PLO but not ALOlesions disrupting reversal learning, suggesting PLO is critical for updating the value of expected outcomes– and our neuroanatomical characterisation of their inputs has shed light on the circuitry underlying those differences. We observed a similar degree of distinction between the afferent projections into ALO and PLO and the projections into PLO and PVO. Thus, while further studies might be necessary, the evidence so far firmly supports a distinction along the anterior–posterior axis of LO. We did not explore whether a similar division might exist within VO. However, the projection patterns we observed in the submedius nucleus of the thalamus following injection into PVO differ substantially from those seen by Kuramoto and colleagues (2017) following injection of the same retrograde tracer into anterior VO

(AVO). Such distinction within VO is also supported by stronger noradrenergic projections from the locus coeruleus to PVO than to AVO (Cerpa et al., 2019). Functional differences have also been reported within medial OFC (MO): Bradfield and colleagues (2018) found that lesions of its anterior but not posterior portions disrupt instrumental devaluation. In the same study, the authors show that this functional distinction is supported by anatomical differences, with anterior MO being more strongly connected to the NAc core while posterior MO has a higher density of projections to the BLA (Bradfield et al., 2018). Together, this suggests that distinctions along the anterior–posterior axis may be common across OFC regions.

When designing our neuroanatomical tracing experiments and interpreting its results, we based our categorisation of coronal slices into anterior and posterior on the anatomical coordinates of the studies reporting the functional distinctions described above. However, considering the lack of cytoarchitectonic boundaries for OFC subdivisions, it remains unclear whether these distinctions are better characterised by a separation into anterior and posterior portions or, instead, as a gradient of organisation. Due to the relatively small resolution that the rodent OFC allows for, this question might be particularly challenging to probe through lesion studies alone. Characterisation of OFC connectivity by injecting tracers into afferent and efferent brain regions –as opposed to directly into OFC subdivisions– and subsequent assessment of the resulting patterns along the OFC axes of organisation might lead our understanding on whether clear borders of the various subdivisions indeed exist and where they might lie.

Our results revealed that different portions of the submedius nucleus of the thalamus send distinct, strong and topographically organised projections to OFC subdivisions. Thus, the submedius nucleus presents itself as a key anatomical area within the thalamus whose projections may allow the distinction of orbital subregions more

effectively than those of the mediodorsal thalamus, which have traditionally been used when defining prefrontal circuits (Rose & Woolsey, 1948). The specific pattern and relative density of anatomical projections between brain regions is also one of the criteria used to establish homology between cortical areas of different species (Rose & Woolsey, 1948; Uylings et al., 2003). A detailed understanding of the anatomy of OFC subdivisions will, therefore, allow for connectivity-based inferences on cross species-homologies. Thus, further anatomical and functional characterisation of OFC subdivisions will likely be key to establish clearer homologies between rodent and primate OFC.

Thalamocortical projections from the submedius nucleus to the OFC have recently been shown to be critical for updating action-outcome associations (Fresno et al., 2019), but whether this function is specific to a subset of these projections is not yet known. The function of the various subdivisions of both OFC and submedius nucleus could be further explored using modern genetic tools such as Cre-dependent optogenetics systems (e.g., injection of the retrograde canine adenovirus (CAV) carrying Cre-recombinase into ALO, PLO, or PVO in separate animals and injection of floxed-ArchT (Chow et al., 2010) into the submedius nucleus) in combination with behavioural experiments. Refining anatomical methods to study projections from different portions of a single thalamic nucleus into the same cortical area is also possible by using precise viral injections for fluorescent protein-based labelling (e.g. Furuta et al., 2001; Kuramoto et al., 2017; Vasquez-Lopez et al., 2017).

With this study, we have highlighted the importance of connectivity for understanding brain function. Moving forward, research on the OFC could benefit from detailed anatomical studies to generate testable hypotheses. The employment of tract-tracing techniques that allow for comparison of connectivity of multiple areas within the same brain (e.g. the combination of multiple Alexa Fluor conjugates of CTB that we used

here and that was described in detail by Conte et al., 2009) might be particularly useful for dissociating circuits of anatomically close but functionally distinct subdivisions. In addition to retrograde connectivity, it will also be essential to characterise the anterograde projection patterns to the various OFC subdivisions. Such anatomical studies could employ a combination of anterograde adeno-associated viruses expressing different fluorescent proteins (e.g. yellow, green, and cyan fluorescent proteins). Our understanding of OFC functional heterogeneity might then be built on a detailed description of its anatomical organisation through more confined targeted lesions and recording or manipulation of neuronal activity.

7.3 Link between dopaminergic signalling and haemodynamic signatures of RPEs

In Chapter 4, we investigated the link between behaviourally derived haemodynamic signatures of reward expectation and receipt and dopamine by combining *in vivo* oxygen amperometry recordings in the NAc with unilateral lesions to dopaminergic inputs to this region as animals performed a probabilistic reversal learning task. The encoding of RPE signalling in the brain has long been linked with the activity of midbrain dopamine neurons (Montague et al., 1996; Schultz et al., 1997). The phasic activity of these neurons encodes RPEs both at the time of reward-predicting events and reward delivery. RPEs can also be detected in brain regions that receive input from midbrain dopamine neurons, including the NAc. In this region, RPEs have been observed using various techniques, including measurement of dopamine release with voltammetry or monitoring blood oxygenation level-dependent (BOLD) signals in functional magnetic resonance imaging (fMRI) or comparable haemodynamic responses obtained with oxygen amperometry. For

instance, Werlen and colleagues foundthat haemodynamic signals in the NAc of rats trained on a probabilistic Pavlovian task and measured with oxygen amperometry encode RPEs (Werlen et al., 2019). It is thought that behaviourally-derived haemodynamic correlates of RPEs in the NAc are driven by mesolimbic dopamine release (as reviewed in Knutson & Gibbs, 2007). However, there has been no direct demonstration of this link and it has been suggested that there is not a one-to-one relationship between dopamine release and BOLD signals (Lohrenz et al., 2016). In the study described in Chapter 4, to also assess how mesolimbic dopamine influences frontostriatal circuits during reward-guided flexible behaviour, we recorded haemodynamic signals also from electrodes targeted primarily to the LO region of the OFC during training on a probabilistic reversal-learning behavioural paradigm.

Axiomatic approaches to examining RPEs suggest that their neuronal signatures should satisfy three key criteria: (1) a positive influence of reward estimate on reward-predicting cue-elicited signals, with cues that predict larger rewards (or with higher probability) eliciting greater responses; (2) a negative influence of reward estimate at the time of reward delivery, with unexpected rewards producing greater responses; and (3) a positive influence of outcome at the time of delivery, with the delivery of a valued reward producing a greater response than its omission (Caplin & Dean, 2008; Rutledge et al., 2010). These criteria are important because model-derived RPE estimates correlate strongly with reward. In consequence, general linear models including estimates of RPEs in their set of predictors often exhibit a statistically significant regression coefficient for this predictor that might be simply picking up the reward signalling and without those responses necessarily encoding RPEs. Our results from Chapter 4 show that, before the first reversal of the reward contingencies on this task, both NAc and OFC had a clear encoding of outcome but not of outcome estimates, therefore not fulfilling the criteria for

RPE-like responses. Haemodynamic signals closer to such responses (though failing to reach statistical significance) emerged in the NAc after reversal, with signals changing dynamically as animals learned the new reward contingencies. In Chapter 5, recordings of NAc haemodynamic responses started as the animals experienced a reversal of the reward contingencies. Also here, responses resembled what we would expect for signals encoding RPEs (though the effect of reward estimate was still statistically unreliable). These findings align with previous studies reporting an attenuation of RPEs encoding for trial features that are less relevant for guiding behaviour (e.g., Klein-Flügge et al., 2011; Werlen et al., 2019). It is worth noting, however, that even after reversal of reward contingencies, in our studies, RPE-encoding in NAc haemodynamic responses was not as clear as in the study by Werlen and colleagues (2019). Potential reasons for the contrast between the responses in the studies include differences in the (i) precise location of electrodes; (ii) task design, i.e., Pavlovian vs instrumental; (iii) task context, i.e., during vs after learning of reward contingencies.

Our best fit models of the animals' behaviour for the studies described in Chapters 4 and 5 suggest that their trial-by-trial choices were best predicted by a simple RPE-driven reinforcement learning model in which reward estimates for each lever were updated by comparing the expected and received outcomes, scaled by a learning rate. However, because in our studies, reward contingencies were relatively stable, changing only every five training sessions, and recordings took place only after animals had extensive experience on the behavioural task, RPE might not have been elicited on a trial-by-trial basis depending on whether animals obtained a reward or not. Instead, in our task's context, such updating might have occurred only, or have been particularly prominent, after a reversal of the contingencies, more similarly to what happens in model-based learning. Therefore, future experiments aiming to further our

understanding of the link between dopamine signalling and haemodynamic signatures of RPEs might perform recordings during learning of the behavioural paradigm or in a task in which reward contingencies change more frequently.

Notably, the post-reversal patterns of NAc haemodynamic signals described above were surprisingly similar in rats with intact or lesioned dopaminergic inputs (Chapter 4). Rather than a clear effect of the lesion in the encoding of specific trial features, we observed a general change in the responses in the lesioned hemisphere. While responses from the NAc of lesioned animals appeared attenuated (though this effect failed to reach statistical significance in any of our analyses after correction for multiple comparisons), those produced by OFC were surprisingly larger. The results of this study emphasise that, as previously suggested (Lohrenz et al., 2016), the changes in haemodynamic responses in the NAc are not fully modulated by dopamine. Thus, to understand changes in BOLD signals in this region, it might be necessary to examine the contribution of other NAc inputs for modulating its activity. Moreover, as mentioned in the previous paragraph, it will be important to investigate this link when it is clear subjects are doing RPE-driven reinforcement learning.

OFC signals were driven mainly by outcome delivery and with little influence of reward estimate, thus not fulfilling the key criteria for RPEs encoding (Caplin & Dean, 2008; Rutledge et al., 2010). Like for the NAc, these results contrast with those from the study by Werlen et al. (2019) which found unsigned reward prediction correlates in the OFC using the same technique (oxygen amperometry) to monitor haemodynamic signals in this region. As mentioned earlier, differences in the behavioural paradigms used, including the type of task (i.e., Pavlovian vs instrumental) and the stage in training of the recordings (during vs after acquisition), and subdivisions recorded from (i.e., lateral vs ventral OFC) might account for these discrepancies. Experiments in which recordings are

obtained from the same OFC subdivision during acquisition of the instrumental task used here or after learning the reward contingencies on a comparable Pavlovian task might help understand these differences.

After lesioning of NAc dopaminergic inputs, there was a sustained and generalised potentiation of OFC responses. It has been previously observed that there is an inverse relationship between cortical and striatal dopamine levels. For example, lesioning prefrontal dopamine results in increases in striatal dopamine levels (Clarke et al., 2014; Doherty & Gratton, 1996; Kayser et al., 2012; Pycock, Carter, et al., 1980; Pycock, Kerwin, et al., 1980; Roberts et al., 1994; Scornaiencki et al., 2009; Thompson & Moss, 1995). We did not measure dopamine levels in our study, and dopamine is only one of many neurotransmitters modulating activity in NAc and prefrontal cortex. Nonetheless, the shift in the size of responses we observed in NAc and OFC in opposite directions after lesioning of NAc dopaminergic inputs suggests that the previously hypothesised inverse relationship between cortical and striatal modulation might be bidirectionally regulated. However, since the 6-OHDA lesion unexpectedly extended to prefrontal cortex dopaminergic inputs, it is unclear whether the potentiation of OFC responses was caused by lesioning of dopaminergic inputs to NAc or prefrontal cortex or a combination of both. More refined experiments targeting only the dopaminergic inputs to OFC or NAc will be necessary to clarify the role of mesolimbic dopamine in modulating frontostriatal circuits. Such targeted manipulations might be achieved, for instance, through the combination of transgenic rats expressing the Cre recombinase enzyme under the TH promotor gene (Witten et al., 2011) and Cre-dependent optogenetic systems.

7.4 Dopaminergic signalling dysregulation and maladaptive learning

In Chapter 5, we investigated the neuronal mechanisms underlying maladaptive learning during amphetamine-induced hyperdopaminergic states similar to those thought to elicit psychotic symptoms in neuropsychiatric conditions. Dopaminergic signalling dysregulation has been hypothesised to drive the neuronal mechanisms underlying psychotic symptoms such as delusions through disruption of RPEs encoding in frontostriatal circuits (Deserno et al., 2016; Kapur, 2003; Smith et al., 2007). While there is some evidence suggesting a link between dopamine dysregulation and consequent disruption of RPEs encoding to psychotic symptoms, the neuronal mechanisms underlying these processes in neuropsychiatric disorders remain unclear. We investigated these by combining the same behavioural paradigm (probabilistic reversal-learning task) and brain activity monitoring technique (oxygen amperometry) used in Chapter 4 with acute and sensitised (2 repeated doses two weeks apart) systemic injections of amphetamine. The behavioural performance we observed after both acute and sensitised administration of amphetamine suggests a delay in learning the new, reversed reward contingencies and supports a role of dopaminergic signalling dysregulation in maladaptive learning. Additionally, sensitised but not acute injection of amphetamine led to a transient impairment in discriminative responding, as previously observed by Werlen et al. and colleagues (2019). These behavioural results are consistent with long-term sensitisation following a single exposure to amphetamine administration (Vanderschuren et al., 1999). They also support the idea that amphetamine administration, particularly in its sensitised state, can serve as a model of certain aspects of psychosis (Featherstone et al., 2007).

The behavioural changes discussed above were accompanied by an attenuation of NAc haemodynamic signals on the first training session after amphetamine administration. A similar blunting of haemodynamic responses in this region has been previously observed after methamphetamine or amphetamine administration to humans and rodents (Bernacer et al., 2013; Knutson et al., 2004; Werlen et al., 2019). The delay in learning the new reward contingencies following amphetamine injection combined with attenuation of NAc responses highlights a role of dopamine signalling dysregulation in the inappropriate updating of action-outcome and stimulus-outcome contingencies thought to underlie certain neuropsychiatric conditions (Deserno et al., 2016; Smith et al., 2007).

Interestingly, both reducing (lesioning of dopaminergic inputs, Chapter 4) and boosting dopamine signalling (through systemic amphetamine administration, Chapter 5) had a similar effect on the haemodynamic signals in the NAc, i.e., attenuated coding. However, amphetamine injection appears to have corrupted encoding of reward/no reward differently, whereas there was no evidence of a change in the coding scheme following lesioning of dopaminergic inputs. Haemodynamic responses are an indirect proxy of neural activity, and cerebral veins also contain dopamine receptors (Choi et al., 2006). Therefore, it is important to consider whether the effects of dopaminergic signalling manipulation in the NAc haemodynamic signals are a consequence of the stimulation of receptors in neurons present in this brain region or an effect of neurovascular coupling modulation. Arguing against this, Schwarz and colleagues (2004) have shown that dopamine release following cocaine administration peaks before blood flow increases in this region (Schwarz et al., 2004). Moreover, both our study and that by Werlen et al. (2019) suggest that dopaminergic signalling dysregulation via amphetamine administration might selectively disrupt specific aspects of neural coding

in the NAc. However, it is important to consider that while any changes in behaviour are unlikely to have occurred following lesioning of dopaminergic inputs in Chapter 4 since the lesion was unilateral, this is not controlled for in the amphetamine administration study described in Chapter 5. Therefore, it is not clear whether the changes in haemodynamic responses observed in either study might be a direct reflection of amphetamine leading to dopaminergic signalling dysregulation or an indirect consequence of its effects in behaviour, e.g., increased locomotor activity. Future studies might probe this question by combining haemodynamic recordings with selective unilateral stimulation of NAc dopaminergic terminals by, for example, making use of Cre-dependent optogenetic systems.

Due to the small final sample size in each experimental group, it was not possible to assess the effects of single or repeat amphetamine injections on OFC haemodynamic responses. Additional experiments will be necessary to investigate these as well as how acute and sensitised injections of amphetamine might differ in their disruption of the encoding of RPE components in frontostriatal circuits during reward-guided learning.

7.5 Neuroscience needs histology

One simple yet important lesson to take from the studies presented in this thesis is the importance of comprehensive histological examination. Earlier in this chapter, we discussed the importance of anatomical studies to further our understanding of OFC functional heterogeneity (see section 7.2 on *OFC functional and anatomical heterogeneity*). In addition to studies focusing on the characterisation of connectivity patterns, it is crucial that functional studies clearly report the boundaries of the regions where neuronal activity was recorded or manipulated. In the study described in Chapter

4 (on *Causal contribution of dopamine to haemodynamic signatures of reward prediction*), we saw that local injection of 6-OHDA into the NAc resulted in a reduction in tyrosine hydroxylase-expressing (TH^+) neurons in this region but, unexpectedly, also in the prefrontal cortex. The neurotoxin 6-OHDA is commonly used for lesioning dopaminergic neurons, but studies often fail to assess the extent of the lesion histologically (or at least to report them). A detailed histological assessment is essential for quantifying the extent of the lesion on the targeted brain regions and confirming that its effects were confined to those regions. Such assessments might include (i) a depiction of the lesion boundaries and lesion effects intensity, (ii) the size of the effect of the lesion (e.g., % reduction in TH^+ labelling) in the target vs control regions (e.g., an anatomically close region or the same region on the contralateral side for unilateral lesions), (iii) assessment of regions receiving projections from neurons passing in the vicinity of the target region and that could be affected by the specific lesion type. Such information is crucial for the correct interpretation of any physiological or behavioural effects.

7.6 Bridging rodent and human research with haemodynamic signals recordings

Oxygen amperometry allows for recording haemodynamic signals comparable to BOLD signals of fMRI in behaving animals (Francois et al., 2012; Lowry et al., 2010), thus presenting itself as a promising translational technique to help bridge the gap between rodent and human research (Gilmour et al., 2015). For this reason, it was the technique of choice for recording neuronal activity in the study described in Chapter 5, which aimed to further our understanding of the neuronal mechanisms underlying certain neuropsychiatric conditions. However, because it is a relatively novel technique –in

comparison with, for example, electrophysiology recordings– it can be challenging to reliably obtain stable and meaningful signals (~9% of the electrodes across the studies described in Chapters 4 and 5 were excluded for being too noisy or not yielding meaningful signals in most of the training sessions), and filtering techniques might benefit from further optimisation. These issues will likely be gradually solved as the technique becomes more widely used in translational projects. An alternative approach might be to use functional ultrasound, which also allows measuring haemodynamic changes in response to brain activity in freely moving animals (Urban et al., 2015). Whereas the number of recording sites used during oxygen amperometry recordings is spatially limited by the number of probes that will fit in a rat's skull as well as their combined weight, functional ultrasound has the advantage of allowing for whole-brain imaging. Yet another approach might be to use fMRI in rodents trained on head-fixed behavioural paradigms that allow for testing reward-guided behaviour. This method offers the obvious advantage of being closer to those used in humans, and important steps are being taken towards standardising its application (Mandino et al., 2020) and adapting it to more complex behavioural paradigms (Fonseca et al., 2020). However, it is worth noting that using fMRI might not allow for obtaining certain behavioural measures important for assessing the effect of amphetamine administration, such as changes in locomotor activity (Grilly & Loveland, 2001; Vanderschuren et al., 1999).

A significant setback in our studies that should be addressed in future experiments is the relatively high proportion of incorrectly placed electrodes (~22% across the studies described in Chapters 4 and 5). This was partly caused by inconsistencies between different editions of the standard Paxinos and Watson rat brain atlases around OFC. This considerably contributed to the final small sample size, particularly for the MO region, and made it unfeasible to explore the effect of

amphetamine injections in OFC haemodynamic signals. For future studies, it would be helpful to refine the stereotaxic coordinates for electrode implantation. This could be achieved through simple injection of dyes and subsequent histological assessment. However, for more ethical use of the animals, this process could be combined with tracing studies. Bradfield and colleagues have shown that MO is functionally heterogeneous along its anterior–posterior axis and that these differences are supported by distinct anatomical outputs (Bradfield et al., 2018). However, differences in anterior vs posterior MO afferent projections have not yet been investigated. Thus, initial refinement of the stereotaxic coordinates could simultaneously attempt to characterise MO inputs along its anterior-posterior axis.

7.7 Developing behavioural paradigms for investigating OFC function

The lesion of Phineas Gage hinted at the potential role of the OFC in behavioural flexibility (Harlow, 1848). Since then, careful examination of the behaviour of animals trained on ingeniously designed experimental paradigms combined with neuronal activity recordings or manipulation has guided our understanding of the function of this brain region. We now know that the OFC plays a key role in solving problems that require behavioural flexibility, particularly under changing circumstances. Such problems can often be solved using two strategies (i) gradually updating the value estimates of each available option with new experience, or (ii) by learning the abstract relationships between events and outcomes –cognitive maps– and then use these as orientation and to infer which rules are in place when circumstances change. The OFC seems to be particularly important for the latter strategy. In Chapter 6, we attempted to establish a

novel rodent behavioural paradigm for investigating the contribution of the OFC to the encoding of cognitive maps, specifically during generalisation of abstract task structure. Taking into consideration previous literature on OFC neuronal responses reporting neurons encoding the subjective value of the expected reward, its magnitude, and the cues predicting it (Bissonette & Roesch, 2015; Padoa-Schioppa & Assad, 2006; Schoenbaum et al., 2009; Wikenheiser & Schoenbaum, 2016), we designed a non-spatial cue-based serial learning task in which it would be possible to dissociate representations of cues, choice, outcome, spatial location, and task state. This behavioural paradigm required rats to discriminate and learn the reward magnitudes associated with a large number of different stimuli and to develop a model of the world in order to maximise reward. Our results show that rats are able to learn a behavioural task involving a mix of different trial types with complex structures. Moreover, rats are able to discriminate between pairs of complex auditory cues with different peak frequencies, abstract stimulus features such as mean power over low vs high frequencies, and apply this to new pairs of cues. However, because the OFC also represents sound features (Winkowski et al., 2018) and we wanted to be able to decorrelate sound frequency from reward magnitude, we modified the cues so that they had similar peak frequency power. After this modification, rats struggled to learn the reward contingencies of new cues and learning was too slow for the planned analysis of neuronal data. In short, rats can easily discriminate between pairs of complex auditory cues with different peak frequencies and may generalise across such sets of cues but struggle if their peak power frequencies are similar. These findings replicate those from a previous study by Amato & Salmon (1984) and suggest that stimuli in other sensory modalities might be better suited for developing rodent tasks requiring a large number of different stimuli.

For intended neuronal recordings in our task, rats would need to, ideally, learn several pairs of cues within a single recording session. Therefore, in the development of our behavioural task, it was essential to establish a large stimulus set of discriminable cues. Audition was our sensory modality of choice for generating such a set. Rats have a wide hearing range from 250 Hz to 70 kHz (Heffner et al., 1994). Additionally, there are available large libraries of natural auditory stimuli (e.g., *Cornell Guide to Bird Sounds: Essential Set for North America*), and new cues can be easily computer-generated and specific features manipulated (e.g., see McDermott et al., 2011). Rats are nocturnal animals and spend most of their lives in burrows. It is thought that they do not rely much on visual stimuli in their natural environments and laboratory rats are considered to have generally poor vision (e.g., Burn, 2008). Accordingly, rodent behavioural tasks using visual stimuli often use optical gratings of black and white lines tilted at different angles or with varying luminescence (e.g., Ghodrati et al., 2016). Such simple features would make it challenging to create a sufficiently large set of distinct and discriminable visual cues. Rats rely heavily on somatosensory information when exploring their surroundings, and rats whiskers are thought to be comparable to primate fingertips (Carvell & Simons, 1990). Rodent tasks on this sensory modality typically use as stimuli surfaces of different roughness or whisker stimulation at different frequencies (e.g., Ewert et al., 2008). More than one of these stimuli can be presented sequentially in a single location. However, it would require an additional stimulus, such as a light flickering on the corresponding reward port, to provide information on which reward port would be associated with each somatosensory stimulus in each trial. Rats also rely extensively on odour cues to learn about their environment. Olfactory stimuli allow for a large stimulus space, and it has been shown that rats can acquire a learning set when trained on a series of 2-odour discrimination problems (Slotnick et al., 2000). Importantly, it has

been shown that rats can localise the source of odours to the left or to the right when olfactory stimuli are delivered through an airflow allowing for a difference in the timing at which the odour arrives at each nostril. Therefore, a non-spatial cue-based serial learning task similar to the one we attempted with auditory cues could be developed by sequentially delivering odours through an airflow from one side and then from the other at the beginning of each trial. Such behavioural paradigm might then be used to investigate the representation of abstract task structure in the brain.

7.8 In conclusion

This thesis provided new insights on the neuronal circuits underlying reward-guided flexible behaviour, a complex cognitive process that relies on the adequate contribution of multiple brain regions. This work involved the combination of different experimental approaches, including neuroanatomical characterisation, monitoring and manipulation of neuronal activity, and both the use and development of established and novel behavioural paradigms. A key learning from our work is that a multi-level consideration of these brain processes is essential for the appropriate design, execution and interpretation of neuroscience experiments.

Firstly, we learned that consideration of the anatomical distinctions within the OFC is essential to understand the role of this brain region in promoting behavioural flexibility as well as the functional contribution of its various subdivisions. Such anatomical considerations might also help resolve inconsistencies in the literature on OFC function and clarify outstanding questions regarding homologies between rodents and primates. Moreover, a comprehensive histological assessment is important for clarifying the boundaries of the brain regions targeted during monitoring or

manipulation of neuronal activity and confirming any manipulation effects were confined to the areas intended. Such information is essential for adequate interpretation of any physiological or behavioural effects.

The results of behavioural experiments combined with manipulation of dopaminergic transmission highlight the importance of balanced dopaminergic signalling for appropriate behavioural adaptation. However, in a separate set of experiments, we show that dopamine neuromodulation is not the main driver of haemodynamic signatures of RPEs in the rodent NAc. This emphasises that neural representations within a certain brain area are the result of the integration of multiple inputs. Thus, in order to understand how these representations are formed, it is important to examine the relative contribution of the various input types.

Bibliography

Agster, K. L., & Burwell, R. D. (2009). Cortical efferents of the perirhinal, postrhinal, and entorhinal cortices of the rat. *Hippocampus*, *19*(12), 1159–1186. https://doi.org/10.1002/hipo.20578

Akam, T., Lustig, A., Rowland, J., Kapanaiah, S. K., Esteve-Agraz, J., Panniello, M., Marquez, C., Kohl, M., Kätzel, D., & Costa, R. M. (2021). pyControl: Open source, Python based, hardware and software for controlling behavioural neuroscience experiments. 2. *BioRxiv*, 2021.02.22.432227. https://doi.org/10.1101/2021.02.22.432227

Akam, T., & Walton, M. E. (2021). What is dopamine doing in model-based reinforcement learning? *Current Opinion in Behavioral Sciences*, *38*, 74–82. https://doi.org/10.1016/j.cobeha.2020.10.010

Alcaraz, F., Marchand, A. R., Vidal, E., Guillou, A., Faugère, A., Coutureau, E., & Wolff, M. (2015). Flexible use of predictive cues beyond the orbitofrontal cortex: Role of the submedius thalamic nucleus. *Journal of Neuroscience*, *35*(38), 13183–13193. https://doi.org/10.1523/JNEUROSCI.1237-15.2015

Alexander, G. (1986). Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. *Annual Review of Neuroscience*, *9*(1), 357–381. https://doi.org/10.1146/annurev.neuro.9.1.357

Amara, S. G., & Sonders, M. S. (1998). Neurotransmitter transporters as molecular targets for addictive drugs. *Drug and Alcohol Dependence*, *51*(1–2), 87–96. https://doi.org/10.1016/S0376-8716(98)00068-4

Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (Macaca fascicularis). *The Journal of Comparative Neurology*, *230*(4), 465–496. https://doi.org/10.1002/cne.902300402

Aransay, A., Rodríguez-López, C., García-Amado, M., Clascá, F., & Prensa, L. (2015). Long-range projection neurons of the mouse ventral tegmental area: A single-cell axon tracing analysis. *Frontiers in Neuroanatomy*, *9*(MAY), 1–24. https://doi.org/10.3389/fnana.2015.00059

Arinze, I., & Moorman, D. E. (2020). Selective impact of lateral orbitofrontal cortex inactivation on reinstatement of alcohol seeking in male Long-Evans rats. *Neuropharmacology*, *168*(April 2018), 108007. https://doi.org/10.1016/j.neuropharm.2020.108007

Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*, *20*(4), 505–512. https://doi.org/10.1038/nn.4506

Babayan, B. M., Uchida, N., & Gershman, S. J. (2018). Belief state representation in the dopamine system. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04397-0

Badiani, A., Anagnostaras, S. G., & Robinson, T. E. (1995). The development of sensitization to the psychomotor stimulant effects of amphetamine is enhanced in a novel environment. *Psychopharmacology*, *117*(4). https://doi.org/10.1007/BF02246217

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4–5), 407–419. https://doi.org/10.1016/S0028-3908(98)00033-1

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in Neurosciences*, *29*(5), 272–279. https://doi.org/10.1016/j.tins.2006.03.002

Balleine, B. W., Leung, B. K., & Ostlund, S. B. (2011). The orbitofrontal cortex, predicted value, and choice. *Annals of the New York Academy of Sciences*, *1239*(1), 43–50. https://doi.org/10.1111/j.1749-6632.2011.06270.x

Barreiros, I. V, Ishii, H., Walton, M. E., & Panayi, M. C. (2021). Defining an orbitofrontal compass: Functional and anatomical heterogeneity across anterior–posterior and medial–lateral axes. *Behavioral Neuroscience*, *135*(2), 165–173. https://doi.org/10.1037/bne0000442

Barreiros, I. V, Panayi, M. C., & Walton, M. E. (2021). Organization of Afferents along the Anterior–posterior and Medial–lateral Axes of the Rat Orbitofrontal Cortex. *Neuroscience*,

460, 53–68. https://doi.org/10.1016/j.neuroscience.2021.02.017
Bassareo, V., & Di Chiara, G. (1997). Differential influence of associative and nonassociative learning mechanisms on the responsiveness of prefrontal and accumbal dopamine transmission to food stimuli in rats fed ad libitum. *Journal of Neuroscience*, *17*(2). https://doi.org/10.1523/jneurosci.17-02-00851.1997
Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1). https://doi.org/10.1016/j.neuron.2005.05.020
Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*, *98*(3). https://doi.org/10.1152/jn.01140.2006
Bazzu, G., Puggioni, G. G. M., Dedola, S., Calia, G., Rocchitta, G., Migheli, R., Desole, M. S., Lowry, J. P., O'Neill, R. D., & Serra, P. A. (2009). Real-time monitoring of brain tissue oxygen using a miniaturized biotelemetric device implanted in freely moving rats. *Analytical Chemistry*, *81*(6), 2235–2241. https://doi.org/10.1021/ac802390f
Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. https://doi.org/10.1038/nature07538
Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002
Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
Berendse, H. W., Graaf, Y. G., & Groenewegen, H. J. (1992). Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *Journal of Comparative Neurology*, *316*(3), 314–347. https://doi.org/10.1002/cne.903160305
Berke, J. D. (2018). What does dopamine mean? *Nature Neuroscience*, *21*(6), 787–793. https://doi.org/10.1038/s41593-018-0152-y
Bernacer, J., Corlett, P. R., Ramachandra, P., McFarlane, B., Turner, D. C., Clark, L., Robbins, T. W., Fletcher, P. C., & Murray, G. K. (2013). Methamphetamine-induced disruption of Frontostriatal reward learning signals: Relation to psychotic symptoms. *American Journal of Psychiatry*, *170*(11), 1326–1334. https://doi.org/10.1176/appi.ajp.2013.12070978
Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience*, *21*(8). https://doi.org/10.1523/jneurosci.21-08-02793.2001
Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? In *Brain Research Reviews* (Vol. 28, Issue 3). https://doi.org/10.1016/S0165-0173(98)00019-8
Bissonette, G. B., & Roesch, M. R. (2015). Neurophysiology of Reward-Guided Behavior: Correlates Related to Predictions, Value, Motivation, Errors, Attention, and Action. In *Brain Imaging in Behavioral Neuroscience* (Issue November 2011, pp. 199–230). https://doi.org/10.1007/7854_2015_382
Blaiss, C. A., & Janak, P. H. (2009). The nucleus accumbens core and shell are critical for the expression, but not the consolidation, of Pavlovian conditioned approach. *Behavioural Brain Research*, *200*(1). https://doi.org/10.1016/j.bbr.2008.12.024
Blanco-Pozo, M., Akam, T., & Walton, M. (2021). Dopamine reports reward prediction errors, but does not update policy, during inference- guided choice. *BioRxiv*, 1–44. https://doi.org/https://doi.org/10.1101/2021.06.25.449995
Bonelli, R. M., & Cummings, J. L. (2007). Frontal-subcortical circuitry and behavior. *Dialogues in Clinical Neuroscience*, *9*(2), 141–151. https://doi.org/10.31887/dcns.2007.9.2/rbonelli
Bradfield, L. A., Dezfouli, A., Van Holstein, M., Chieng, B., & Balleine, B. W. (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron*, *88*(6), 1268–1280. https://doi.org/10.1016/j.neuron.2015.10.044

Bradfield, L. A., & Hart, G. (2020). Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. *Neuroscience and Biobehavioral Reviews*, *108*(November 2019), 287–294. https://doi.org/10.1016/j.neubiorev.2019.11.009

Bradfield, L. A., Hart, G., & Balleine, B. W. (2018). Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. *Neurobiology of Learning and Memory*, *155*(August), 463–473. https://doi.org/10.1016/j.nlm.2018.09.008

Brischoux, F., Chakraborty, S., Brierley, D. I., & Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(12). https://doi.org/10.1073/pnas.0811507106

Brodmann, K. (1909). Vergleichende Lokalisationslehre Der Großhirnrinde in Ihren Prinzipien Dargestellt AufGrund Des Zellenbaues. *Leipzig: Barth.*

Brozoski, T. J., Brown, R. M., Rosvold, H. E., & Goldman, P. S. (1979). Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science*, *205*(4409). https://doi.org/10.1126/science.112679

Burke, S. N., Thome, A., Plange, K., Engle, J. R., Trouard, T. P., Gothard, K. M., & Barnes, C. A. (2014). Orbitofrontal cortex volume in area 11/13 predicts reward devaluation, but not reversal learning performance, in young and aged monkeys. *Journal of Neuroscience*, *34*(30), 9905–9916. https://doi.org/10.1523/JNEUROSCI.3918-13.2014

Burn, C. C. (2008). What is it like to be a rat? Rat sensory perception and its implications for experimental design and rat welfare. *Applied Animal Behaviour Science*, *112*(1–2), 1–32. https://doi.org/10.1016/j.applanim.2008.02.007

Burton, A. C., Bissonette, G. B., Lichtenberg, N. T., Kashtelyan, V., & Roesch, M. R. (2014). Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum. *Biological Psychiatry*, *75*(2), 132–139. https://doi.org/10.1016/j.biopsych.2013.05.023

Burton, A. C., Kashtelyan, V., Bryden, D. W., & Roesch, M. R. (2014). Increased firing to cues that predict low-value reward in the medial orbitofrontal cortex. *Cerebral Cortex*, *24*(12), 3310–3321. https://doi.org/10.1093/cercor/bht189

Burton, A. C., Nakamura, K., & Roesch, M. R. (2015). From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. *Neurobiology of Learning and Memory*, *117*, 51–59. https://doi.org/10.1016/j.nlm.2014.05.003

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5). https://doi.org/10.1037/h0054388

Caplin, A., & Dean, M. (2008). Axiomatic methods, dopamine and reward prediction error. *Current Opinion in Neurobiology*, *18*(2), 197–202. https://doi.org/10.1016/j.conb.2008.07.007

Carlén, M. (2017). What constitutes the prefrontal cortex? *Science*, *358*(6362), 478–482. https://doi.org/10.1126/science.aan8868

Carlson, A. (1970). Amphetamine and brain catecholamines. In E. Costa & S. Garattini (Eds.), *Amphetamines and Related Compounds* (pp. 289–300). Raven Press.

Carmichael, S. T., & Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*, *346*(3), 366–402. https://doi.org/10.1002/cne.903460305

Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, *363*(4), 615–641. https://doi.org/10.1002/cne.903630408

Carr, D. B., & Sesack, S. R. (2000). Projections from the rat prefrontal cortex to the ventral tegmental area: Target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *Journal of Neuroscience*, *20*(10), 3864–3873. https://doi.org/10.1523/jneurosci.20-10-03864.2000

Carvell, G. E., & Simons, D. J. (1990). Biometric analyses of vibrissal tactile discrimination in the rat. *Journal of Neuroscience*, *10*(8). https://doi.org/10.1523/jneurosci.10-08-02638.1990

Castaneda, E., Whishaw, I. Q., & Robinson, T. E. (1990). Changes in striatal dopamine neurotransmission assessed with microdialysis following recovery from a bilateral 6-OHDA lesion: Variation as a function of lesion size. *Journal of Neuroscience*, *10*(6), 1847–1854. https://doi.org/10.1523/jneurosci.10-06-01847.1990

Cerpa, J. C., Coutureau, E., & Parkes, S. L. (2021). Dopamine and noradrenaline modulation of goal-directed behaviour in prefrontal areas: Toward a division of labour? *Behavioral Neuroscience*, *135*(2). https://doi.org/10.1037/bne0000426

Cerpa, J. C., Marchand, A. R., & Coutureau, E. (2019). Distinct regional patterns in noradrenergic innervation of the rat prefrontal cortex. *Journal of Chemical Neuroanatomy*, *96*(November 2018), 102–109. https://doi.org/10.1016/j.jchemneu.2019.01.002

Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. P., Bell, A. H., Walton, M. E., & Rushworth, M. F. S. (2015). Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, *87*(5), 1106–1118. https://doi.org/10.1016/j.neuron.2015.08.018

Chen, Y. C. I., Galpern, W. R., Brownell, A. L., Matthews, R. T., Bogdanov, M., Isacson, O., Keltner, J. R., Beal, M. F., Rosen, B. R., & Jenkins, B. G. (1997). Detection of dopaminergic neurotransmitter activity using pharmacologic MRI: Correlation with PET, microdialysis, and behavioral data. *Magnetic Resonance in Medicine*, *38*(3). https://doi.org/10.1002/mrm.1910380306

Choi, J. K., Chen, Y. I., Hamel, E., & Jenkins, B. G. (2006). Brain hemodynamic changes mediated by dopamine receptors: Role of the cerebral microvasculature in dopamine-mediated neurovascular coupling. *NeuroImage*, *30*(3), 700–712. https://doi.org/10.1016/j.neuroimage.2005.10.029

Chow, B. Y., Han, X., Dobry, A. S., Qian, X., Chuong, A. S., Li, M., Henninger, M. A., Belfort, G. M., Lin, Y., Monahan, P. E., & Boyden, E. S. (2010). High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature*, *463*(7277). https://doi.org/10.1038/nature08652

Chudasama, Y., & Robbins, T. W. (2003). Dissociable Contributions of the Orbitofrontal and Infralimbic Cortex to Pavlovian Autoshaping and Discrimination Reversal Learning: Further Evidence for the Functional Heterogeneity of the Rodent Frontal Cortex. *J. Neurosci.*, *23*(25), 8771–8780. https://doi.org/23/25/8771 [pii]

Clarke, H. F., Cardinal, R. N., Rygula, R., Hong, Y. T., Fryer, T. D., Sawiak, S. J., Ferrari, V., Cockcroft, G., Aigbirhio, F. I., Robbins, T. W., & Roberts, A. C. (2014). Orbitofrontal dopamine depletion upregulates caudate dopamine and alters behavior via changes in reinforcement sensitivity. *Journal of Neuroscience*, *34*(22), 7663–7676. https://doi.org/10.1523/JNEUROSCI.0718-14.2014

Clarke, P. B. S., Jakubovic, A., & Fibiger, H. C. (1988). Anatomical analysis of the involvement of mesolimbocortical dopamine in the locomotor stimulant actions of d-amphetamine and apomorphine. *Psychopharmacology*, *96*(4), 511–520. https://doi.org/10.1007/BF02180033

Coffield, J. A., Bowen, K. K., & Miletic, V. (1992). Retrograde tracing of projections between the nucleus submedius, the ventrolateral orbital cortex, and the midbrain in the rat. *Journal of Comparative Neurology*, *321*(3), 488–499. https://doi.org/10.1002/cne.903210314

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, *482*(7383), 85–88. https://doi.org/10.1038/nature10754

Collins, A. L., Greenfield, V. Y., Bye, J. K., Linker, K. E., Wang, A. S., & Wassum, K. M. (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Scientific Reports*, *6*. https://doi.org/10.1038/srep20231

Colwill, R. M., & Rescorla, R. A. (1985). Postconditioning Devaluation of a Reinforcer Affects Instrumental Responding. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(1). https://doi.org/10.1037/0097-7403.11.1.120

Colwill, R. M., & Rescorla, R. A. (1990). Effect of Reinforcer Devaluation on Discriminative Control of Instrumental Behavior. *Journal of Experimental Psychology: Animal Behavior*

Processes, *16*(1). https://doi.org/10.1037/0097-7403.16.1.40
Constantinescu, A. O., OReilly, J. X., Behrens, T. E. J., O'Reilly, J. X., Behrens, T. E. J., Kumaran, D., Summerfield, J. J., Hassabis, D., Maguire, E. A., Barron, H. C., Dolan, R. J., Behrens, T. E. J., Tolman, E. C., Buzsáki, G., Moser, E. I., Eichenbaum, H., Cohen, N. J., Hafting, T., Fyhn, M., … Smith, S. M. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941
Conte, W. L., Kamishina, H., & Reep, R. L. (2009). Multiple neuroanatomical tract-tracing using fluorescent Alexa Fluor conjugates of cholera toxin subunit B in rats. *Nature Protocols*, *4*(8), 1157–1166. https://doi.org/10.1038/nprot.2009.93
Cools, R., & D'Esposito, M. (2010). Dopaminergic Modulation of Flexible Cognitive Control in Humans. In *Dopamine Handbook*. https://doi.org/10.1093/acprof:oso/9780195373035.003.0017
Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, *69*(12), e113–e125. https://doi.org/10.1016/j.biopsych.2011.03.028
Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable electrode mapping study. *Brain Research*, *185*(1). https://doi.org/10.1016/0006-8993(80)90666-6
Corbit, L. H., & Balleine, B. W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behavioural Brain Research*, *146*(1–2), 145–157. https://doi.org/10.1016/j.bbr.2003.09.023
Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.21-09-03251.2001
Cornell Guide to Bird Sounds: Essential Set for North America. (n.d.). Macaulay Library. Retrieved June 7, 2021, from www.macaulaylibrary.org/guide-to-bird-sounds/
Corwin, J. V., Fussinger, M., Meyer, R. C., King, V. R., & Reep, R. L. (1994). Bilateral destruction of the ventrolateral orbital cortex produces allocentric but not egocentric spatial deficits in rats. *Behavioural Brain Research*, *61*(1), 79–86. https://doi.org/10.1016/0166-4328(94)90010-8
Corwin, J. V., & Reep, R. L. (1998). Rodent posterior parietal cortex as a component of a cortical network mediating directed spatial attention. *Psychobiology*, *26*(2), 87–102. https://doi.org/10.3758/BF03330596
Cousins, M. S., Atherton, A., Turner, L., & Salamone, J. D. (1996). Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behavioural Brain Research*, *74*(1–2). https://doi.org/10.1016/0166-4328(95)00151-4
Creese, I., Burt, D. R., & Snyder, S. H. (1976). Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science*, *192*(4238). https://doi.org/10.1126/science.3854
Critchley, H. D., & Rolls, E. T. (1996). Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *Journal of Neurophysiology*, *75*(4), 1673–1686. https://doi.org/10.1152/jn.1996.75.4.1673
Croxson, P. L., Walton, M. E., Boorman, E. D., Rushworth, M. F. S., & Bannerman, D. M. (2014). Unilateral medial frontal cortex lesions cause a cognitive decision-making deficit in rats. *European Journal of Neuroscience*, *40*(12). https://doi.org/10.1111/ejn.12751
Curran, C., Byrappa, N., & McBride, A. (2004). Stimulant psychosis: Systematic review. In *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.185.3.196
D'Amato, M. R., & Salmon, D. P. (1984). Processing and retention of complex auditory stimuli in monkeys (Cebus apella). *Canadian Journal of Psychology*, *38*(2), 237–255. https://doi.org/10.1037/h0080825
D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867). https://doi.org/10.1126/science.1150605
Da Silva, J. A., Tecuapetla, F., Paixão, V., & Costa, R. M. (2018). Dopamine neuron activity before

action initiation gates and invigorates future movements. *Nature*, *554*(7691). https://doi.org/10.1038/nature25457
Daberkow, D. P., Brown, H. D., Bunner, K. D., Kraniotis, S. A., Doellman, M. A., Ragozzino, M. E., Garris, P. A., & Roitman, M. F. (2013). Amphetamine paradoxically augments exocytotic dopamine release and phasic dopamine signals. *Journal of Neuroscience*, *33*(2), 452–463. https://doi.org/10.1523/JNEUROSCI.2136-12.2013
Dahlstrom, A., & Fuxe, K. (1965). Evidence for the Existence of Monoamine-Containing Neurons in the Central Nervous System. *Acta Physiologica Scandinavica*, *65*.
Dalton, G. L., Wang, N. Y., Phillips, A. G., & Floresco, S. B. (2016). Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. *Journal of Neuroscience*, *36*(6), 1996–2006. https://doi.org/10.1523/JNEUROSCI.3366-15.2016
Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science*, *264*(5162). https://doi.org/10.1126/science.8178168
Datiche, F., & Cattarelli, M. (1996). Reciprocal and topographic connections between the piriform and prefrontal cortices in the rat: A tracing study using the B subunit of the cholera toxin. *Brain Research Bulletin*, *41*(6), 391–398. https://doi.org/10.1016/S0361-9230(96)00082-2
Davis, H. (1992). Transitive inference in rats (Rattus norvegicus). *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *106*(4), 342–349. https://doi.org/10.1037/0735-7036.106.4.342
Davis, K. L., Kahn, R. S., Ko, G., & Davidson, M. (1991). Dopamine in schizophrenia: A review and reconceptualization. *American Journal of Psychiatry*, *148*(11), 1474–1486. https://doi.org/10.1176/ajp.148.11.1474
Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560
Daw, N. D., & Tobler, P. N. (2013). Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. *Neuroeconomics: Decision Making and the Brain: Second Edition*, 283–298. https://doi.org/10.1016/B978-0-12-416008-8.00015-2
Day, J. J., Jones, J. L., & Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *European Journal of Neuroscience*, *33*(2). https://doi.org/10.1111/j.1460-9568.2010.07531.x
Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature Neuroscience*, *10*(8), 1020–1028. https://doi.org/10.1038/nn1923
Decot, H. K., Namboodiri, V. M. K., Gao, W., McHenry, J. A., Jennings, J. H., Lee, S. H., Kantak, P. A., Jill Kao, Y. C., Das, M., Witten, I. B., Deisseroth, K., Shih, Y. Y. I., & Stuber, G. D. (2017). Coordination of Brain-Wide Activity Dynamics by Dopaminergic Neurons. *Neuropsychopharmacology*, *42*(3), 615–627. https://doi.org/10.1038/npp.2016.151
Deniau, J. M., Thierry, A. M., & Feger, J. (1980). Electrophysiological identification of mesencephalic ventromedial tegmental (VMT) neurons projecting to the frontal cortex, septum and nucleus accumbens. *Brain Research*, *189*(2), 315–326. https://doi.org/10.1016/0006-8993(80)90093-1
Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, *179*(3). https://doi.org/10.1007/s00213-004-2059-4
DenOuden, H. E. M., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., Franke, B., & Cools, R. (2013). Dissociable Effects of Dopamine and Serotonin on Reversal Learning. *Neuron*, *80*(4). https://doi.org/10.1016/j.neuron.2013.08.030
Descarries, L., Lemay, B., Doucet, G., & Berger, B. (1987). Regional and laminar density of the dopamine innervation in adult rat cerebral cortex. *Neuroscience*, *21*(3), 807–824.

https://doi.org/10.1016/0306-4522(87)90038-8
Deserno, L., Schlagenhauf, F., & Heinz, A. (2016). Striatal dopamine, reward, and decision making in schizophrenia. *Dialogues in Clinical Neuroscience*, *18*(1), 77–89. https://doi.org/10.31887/dcns.2016.18.1/ldeserno
Devlin, J. T., & Poldrack, R. A. (2007). In praise of tedious anatomy. *NeuroImage*, *37*(4), 1033–1041. https://doi.org/10.1016/j.neuroimage.2006.09.055
Di Ciano, P., Robbins, T. W., & Everitt, B. J. (2008). Differential effects of nucleus accumbens core, shell, or dorsal striatal inactivations on the persistence, reacquisition, or reinstatement of responding for a drug-paired conditioned reinforcer. *Neuropsychopharmacology*, *33*(6). https://doi.org/10.1038/sj.npp.1301522
Dodson, P. D., Dreyer, J. K., Jennings, K. A., Syed, E. C. J., Wade-Martins, R., Cragg, S. J., Bolam, J. P., Magill, P. J., & Malenka, R. C. (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(15). https://doi.org/10.1073/pnas.1515941113
Doherty, M. D., & Gratton, A. (1996). Medial prefrontal cortical D1 receptor modulation of the meso-accumbens dopamine response to stress: An electrochemical study in freely-behaving rats. *Brain Research*, *715*(1–2). https://doi.org/10.1016/0006-8993(95)01557-4
du Hoffmann, J., & Nicola, S. M. (2014). Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. *Journal of Neuroscience*, *34*(43), 14349–14364. https://doi.org/10.1523/JNEUROSCI.3492-14.2014
Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 7109–7114. https://doi.org/10.1073/pnas.94.13.7109
Einhorn, L. C., Johansen, P. A., & White, F. J. (1988). Electrophysiological effects of cocaine in the mesoaccumbens dopamine system: Studies in the ventral tegmental area. *Journal of Neuroscience*, *8*(1), 100–112. https://doi.org/10.1523/jneurosci.08-01-00100.1988
Ellwood, I. T., Patel, T., Wadia, V., Lee, A. T., Liptak, A. T., Bender, K. J., & Sohal, V. S. (2017). Tonic or phasic stimulation of dopaminergic projections to prefrontal cortex causes mice to maintain or deviate from previously learned behavioral strategies. *Journal of Neuroscience*, *37*(35), 8315–8329. https://doi.org/10.1523/JNEUROSCI.1221-17.2017
Enomoto, K., Matsumoto, N., Nakai, S., Satoh, T., Sato, T. K., Ueda, Y., Inokawa, H., Haruno, M., & Kimura, M. (2011). Dopamine neurons learn to encode the long-term value of multiple future rewards. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(37). https://doi.org/10.1073/pnas.1014457108
Eshel, N., Bukwich, M., Rao, V., Hemmelder, V., Tian, J., & Uchida, N. (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*, *525*(7568), 243–246. https://doi.org/10.1038/nature14855
Everitt, B. J., Parkinson, J. A., Olmstead, M. C., Arroyo, M., Robledo, P., & Robbins, T. W. (1999). Associative processes in addiction and reward. The role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences*, *877*, 412–438. https://doi.org/10.1111/j.1749-6632.1999.tb09280.x
Ewert, T. A. S., Vahle-Hinz, C., & Engel, A. K. (2008). High-frequency whisker vibration is encoded by phase-locked responses of neurons in the rat's barrel cortex. *Journal of Neuroscience*, *28*(20). https://doi.org/10.1523/JNEUROSCI.0089-08.2008
Featherstone, R. E., Kapur, S., & Fletcher, P. J. (2007). The amphetamine-induced sensitized state as a model of schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *31*(8), 1556–1571. https://doi.org/10.1016/j.pnpbp.2007.08.025
Fibiger, H. C., LePiane, F. G., Jakubovic, A., & Phillips, A. G. (1987). The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *Journal of Neuroscience*, *7*(12). https://doi.org/10.1523/jneurosci.07-12-03888.1987
Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*, *11*(8). https://doi.org/10.1038/nn.2159

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614). https://doi.org/10.1126/science.1077349

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., & Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, *469*(7328). https://doi.org/10.1038/nature09588

Fletcher, P. J., Tenn, C. C., Rizos, Z., Lovic, V., & Kapur, S. (2005). Sensitization to amphetamine, but not PCP, impairs attentional set shifting: Reversal by a D1 receptor agonist injected into the medial prefrontal cortex. *Psychopharmacology*, *183*(2). https://doi.org/10.1007/s00213-005-0157-6

Floresco, S. B. (2015). The nucleus accumbens: An interface between cognition, emotion, and action. *Annual Review of Psychology*, *66*, 25–32. https://doi.org/10.1146/annurev-psych-010213-115159

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. L. (2006). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology*, *31*(2). https://doi.org/10.1038/sj.npp.1300825

Floresco, S. B., & Phillips, A. G. (2001). Delay-dependent modulation of memory retrieval by infusion of a dopamine D1 agonist into the rat medial prefrontal cortex. *Behavioral Neuroscience*, *115*(4), 934–939. https://doi.org/10.1037/0735-7044.115.4.934

Floyd, N. S., Price, J. L., Ferry, A. T., Keay, K. A., & Bandler, R. (2000). Orbitomedial prefrontal cortical projections to distinct longitudinal columns of the periaqueductal gray in the rat. *Journal of Comparative Neurology*, *422*(4), 556–578. https://doi.org/10.1002/1096-9861(20000710)422:4<556::AID-CNE6>3.0.CO;2-U

Fonseca, M., Bergomi, M., Mainen, Z., & Shemesh, N. (2020). *Functional MRI of large scale activity in behaving mice*. 1–98. https://doi.org/10.1101/2020.04.16.044941

Fouriezos, G., & Wise, R. A. (1976). Pimozide-induced extinction of intracranial self-stimulation: response patterns rule out motor or performance deficits. *Brain Research*, *103*(2). https://doi.org/10.1016/0006-8993(76)90809-X

Fox, M. E., Mikhailova, M. A., Bass, C. E., Takmakov, P., Gainetdinov, R. R., Budygin, E. A., & Wightman, R. M. (2016). Cross-hemispheric dopamine projections have functional significance. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(25), 6985–6990. https://doi.org/10.1073/pnas.1603629113

Francois, J., Conway, M. W., Lowry, J. P., Tricklebank, M. D., & Gilmour, G. (2012). Changes in reward-related signals in the rat nucleus accumbens measured by in vivo oxygen amperometry are consistent with fMRI BOLD responses in man. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2012.02.024

Francois, J., Huxter, J., Conway, M. W., Lowry, J. P., Tricklebank, M. D., & Gilmour, G. (2014). Differential Contributions of Infralimbic Prefrontal Cortex and Nucleus Accumbens during Reward-Based Learning and Extinction. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2346-13.2014

Fresno, V., Parkes, S. L., Faugère, A., Coutureau, E., & Wolff, M. (2019). A thalamocortical circuit for updating action-outcome associations. *ELife*, *8*, 1–13. https://doi.org/10.7554/eLife.46187

Fuchs, R. A., Evans, K. A., Parker, M. P., & See, R. E. (2004). Differential involvement of orbitofrontal cortex subregions in conditioned cue-induced and cocaine-primed reinstatement of cocaine seeking in rats. *Journal of Neuroscience*, *24*(29), 6600–6610. https://doi.org/10.1523/JNEUROSCI.1924-04.2004

Furuta, T., Tomioka, R., Taki, K., Nakamura, K., Tamamaki, N., & Kaneko, T. (2001). In vivo transduction of central neurons using recombinant Sindbis virus: Golgi-like labeling of dendrites and axons with membrane-targeted fluorescent proteins. *Journal of Histochemistry and Cytochemistry*, *49*(12), 1497–1507. https://doi.org/10.1177/002215540104901203

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

Comparative Neurology, *492*(2), 145–177. https://doi.org/10.1002/cne.20738
Gallagher, M., McMahan, R. W., & Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(15), 6610–6614. https://doi.org/http://www.jneurosci.org/content/19/15/6610
Gan, J. O., Walton, M. E., & Phillips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience*, *13*(1). https://doi.org/10.1038/nn.2460
Gancher, S., Mayer, A., & Youngman, S. (1996). Changes in apomorphine pharmacodynamics following repeated treatment in 6-hydroxydopamine-lesioned rats. *Brain Research*, *729*(2), 190–196. https://doi.org/10.1016/0006-8993(96)00421-0
Gardner, M. P. H., Conroy, J. C., Sanchez, D. C., Zhou, J., & Schoenbaum, G. (2019). Real-Time Value Integration during Economic Choice Is Regulated by Orbitofrontal Cortex. *Current Biology*, *29*(24), 4315-4322.e4. https://doi.org/10.1016/j.cub.2019.10.058
Gardner, M. P. H., Conroy, J. C., Styer, C. V., Huynh, T., Whitaker, L. R., & Schoenbaum, G. (2018). Medial orbitofrontal inactivation does not affect economic choice. *ELife*, *7*, 1–22. https://doi.org/10.7554/eLife.38963
Gardner, M. P. H., Conroy, J. S., Shaham, M. H., Styer, C. V., & Schoenbaum, G. (2017). Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. *Neuron*, *96*(5), 1192-1203.e4. https://doi.org/10.1016/j.neuron.2017.10.026
Garvert, M. M. M., Dolan, R. J., & Behrens, T. E. J. (2017). A map of abstract relational knowledge in human entorhinal cortex. *ELife*, *April 2016*, 1–21. https://doi.org/10.7554/eLife.17086
Gerfen, C. R., & Clavier, R. M. (1979). Neural inputs to the prefrontal agranular insular cortex in the rat: Horseradish peroxidase study. *Brain Research Bulletin*. https://doi.org/10.1016/S0361-9230(79)80012-X
Gershman, S. J., & Niv, Y. (2010). Learning latent structure: Carving nature at its joints. *Current Opinion in Neurobiology*, *20*(2), 251–256. https://doi.org/10.1016/j.conb.2010.02.008
Ghodrati, M., Alwis, D. S., & Price, N. S. C. (2016). Orientation selectivity in rat primary visual cortex emerges earlier with low-contrast and high-luminance stimuli. *European Journal of Neuroscience*, *44*(10). https://doi.org/10.1111/ejn.13379
Gilmour, G., Gastambide, F., Marston, H. M., & Walton, M. E. (2015). Using intermediate cognitive endpoints to facilitate translational research in psychosis. *Current Opinion in Behavioral Sciences*, *4*(Table 1), 128–135. https://doi.org/10.1016/j.cobeha.2015.04.011
Gomez-Lazaro, M., Galindo, M. F., Concannon, C. G., Segura, M. F., Fernandez-Gomez, F. J., Llecha, N., Comella, J. X., Prehn, J. H. M., & Jordan, J. (2008). 6-Hydroxydopamine activates the mitochondrial apoptosis pathway through p38 MAPK-mediated, p53-independent activation of Bax and PUMA. *Journal of Neurochemistry*, *104*(6), 1599–1612. https://doi.org/10.1111/j.1471-4159.2007.05115.x
Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science (New York, N.Y.)*, *301*(5636), 1104–1107. https://doi.org/10.1126/science.1087919
Gourley, S. L., Zimmermann, K. S., Allen, A. G., & Taylor, J. R. (2016). The medial orbitofrontal cortex regulates sensitivity to outcome value. *Journal of Neuroscience*, *36*(16), 4600–4613. https://doi.org/10.1523/JNEUROSCI.4253-15.2016
Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1). https://doi.org/10.1016/0306-4522(91)90196-U
Grilly, D. M., & Loveland, A. (2001). What is a "low dose" of d-amphetamine for inducing behavioral effects in laboratory rats? In *Psychopharmacology*. https://doi.org/10.1007/s002130000580
Groenewegen, H. J. (1988). Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *Neuroscience*, *24*(2), 379–431. https://doi.org/10.1016/0306-4522(88)90339-9
Groman, S. M., Keistler, C., Keip, A. J., Hammarlund, E., DiLeone, R. J., Pittenger, C., Lee, D., &

Taylor, J. R. (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron*, *103*(4), 734-746.e3. https://doi.org/10.1016/j.neuron.2019.05.042
Haber, S. N. (2011). Frontiers in NeuroscienceNeuroanatomy of Reward: A View from the Ventral Striatum. In *Neurobiology of Sensation and Reward*.
Hall, C. N., Howarth, C., Kurth-Nelson, Z., & Mishra, A. (2016). Interpreting bold: Towards a dialogue between cognitive and cellular neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705). https://doi.org/10.1098/rstb.2015.0348
Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173
Hamilton, D. A., & Brigman, J. L. (2015). Behavioral flexibility in rats and mice: Contributions of distinct frontocortical regions. *Genes, Brain and Behavior*, *14*(1), 4–21. https://doi.org/10.1111/gbb.12191
Harlow, H. F. (1943). Solution by rhesus monkeys of a problem involving the Weigl principle using the matching-from-sample method. *Journal of Comparative Psychology*, *36*(3), 217–227. https://doi.org/10.1037/h0058067
Harlow, H. F. (1948). The Formation of Learning Sets. *Psychological Review*. https://doi.org/10.1037/h0062474
Harlow, J. M. (1848). Passage of an Iron Rod Through the Head. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *11*(2), 281–283. https://doi.org/10.1176/jnp.11.2.281
Hart, A. S., Clark, J. J., & Phillips, P. E. M. (2015). Dynamic shaping of dopamine signals during probabilistic Pavlovian conditioning. *Neurobiology of Learning and Memory*, *117*(2), 84–92. https://doi.org/10.1016/j.nlm.2014.07.010
Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *Journal of Neuroscience*, *34*(3), 698–704. https://doi.org/10.1523/JNEUROSCI.2489-13.2014
Hashemnia, S., Euston, D. R., & Gruber, A. J. (2020). Amphetamine reduces reward encoding and stabilizes neural dynamics in rat anterior cingulate cortex. *ELife*, *9*, 1–22. https://doi.org/10.7554/ELIFE.56755
Heffner, H. E., & Heffner, R. S. (2007). Hearing ranges of laboratory animals. *Journal of the American Association for Laboratory Animal Science*, *46*(1), 20–22.
Heffner, H. E., Heffner, R. S., Contos, C., & Ott, T. (1994). Audiogram of the hooded Norway rat. *Hearing Research*, *73*(2), 244–247. https://doi.org/10.1016/0378-5955(94)90240-2
Heilbronner, S. R., Rodriguez-Romaguera, J., Quirk, G. J., Groenewegen, H. J., & Haber, S. N. (2016). Circuit-Based Corticostriatal Homologies Between Rat and Primate. *Biological Psychiatry*, *80*(7), 509–521. https://doi.org/10.1016/j.biopsych.2016.05.012
Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: Salience attribution revisited. *Schizophrenia Bulletin*, *36*(3), 472–485. https://doi.org/10.1093/schbul/sbq031
Helbing, C., Brocka, M., Scherf, T., Lippert, M. T., & Angenstein, F. (2016). The role of the mesolimbic dopamine system in the formation of blood-oxygen-level dependent responses in the medial prefrontal/anterior cingulate cortex during high-frequency stimulation of the rat perforant pathway. *Journal of Cerebral Blood Flow & Metabolism*, *36*(12), 2177–2193. https://doi.org/10.1177/0271678X15615535
Hervig, M. E., Fiddian, L., Piilgaard, L., Božič, T., Blanco-Pozo, M., Knudsen, C., Olesen, S. F., Alsiö, J., & Robbins, T. W. (2019). Dissociable and Paradoxical Roles of Rat Medial and Lateral Orbitofrontal Cortex in Visual Serial Reversal Learning. *Cerebral Cortex*, *July*, 1–14. https://doi.org/10.1093/cercor/bhz144
Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology*, *80*(2). https://doi.org/10.1152/jn.1998.80.2.947
Hoover, W. B., & Vertes, R. P. (2011). Projections of the medial orbital and ventral orbital cortex in the rat. *Journal of Comparative Neurology*, *519*(18), 3766–3801. https://doi.org/10.1002/cne.22733

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1503550112
Howe, M. W., & Dombeck, D. A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, *535*(7613). https://doi.org/10.1038/nature18942
Hughes, R. N., Bakhurin, K. I., Petter, E. A., Watson, G. D. R., Kim, N., Friedman, A. D., & Yin, H. H. (2020). Ventral Tegmental Dopamine Neurons Control the Impulse Vector during Motivated Behavior. *Current Biology*, *30*(14). https://doi.org/10.1016/j.cub.2020.05.003
Huys, Q. J. M., Cruickshank, A., & Seriès, P. (2014). Reward-Based Learning, Model-Based and Model-Free. In *Encyclopedia of Computational Neuroscience* (pp. 1–10). Springer New York. https://doi.org/10.1007/978-1-4614-7320-6_674-1
Ishii, H., Ohara, S., Tobler, P. N., Tsutsui, K. I., & Iijima, T. (2012). Inactivating anterior insular cortex reduces risk taking. *Journal of Neuroscience*, *32*(45), 16031–16039. https://doi.org/10.1523/JNEUROSCI.2278-12.2012
Ishii, H., Ohara, S., Tobler, P. N., Tsutsui, K. I., & Iijima, T. (2015). Dopaminergic and serotonergic modulation of anterior insular and orbitofrontal cortex function in risky decision making. *Neuroscience Research*, *92*, 53–61. https://doi.org/10.1016/j.neures.2014.11.009
Izquierdo, A. (2017). Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *Journal of Neuroscience*, *37*(44), 10529–10540. https://doi.org/10.1523/JNEUROSCI.1678-17.2017
Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, *345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021
Izquierdo, A., Darling, C., Manos, N., Pozos, H., Kim, C., Ostrander, S., Cazares, V., Stepp, H., & Rudebeck, P. H. (2013). Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *Journal of Neuroscience*, *33*(9), 4105–4109. https://doi.org/10.1523/JNEUROSCI.4942-12.2013
Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology*, *219*(2), 607–620. https://doi.org/10.1007/s00213-011-2579-7
Izquierdo, A., & Murray, E. A. (2004). Combined Unilateral Lesions of the Amygdala and Orbital Prefrontal Cortex Impair Affective Processing in Rhesus Monkeys. *Journal of Neurophysiology*, *91*(5), 2023–2039. https://doi.org/10.1152/jn.00968.2003
Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, *466*(7305). https://doi.org/10.1038/nature09263
Johnson, A. W., Gallagher, M., & Holland, P. C. (2009). The basolateral amygdala is critical to the expression of pavlovian and instrumental outcome-specific reinforcer devaluation effects. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3758-08.2009
Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus-Reinforcement associations. *Experimental Neurology*, *36*(2). https://doi.org/10.1016/0014-4886(72)90030-1
Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science*, *338*(November), 953–956. https://doi.org/10.1126/science.1227489.Orbitofrontal
Kalsbeek, A., Voorn, P., Buijs, R. M., Pool, C. W., & Uylings, H. B. M. (1988). Development of the dopaminergic innervation in the prefrontal cortex of the rat. *Journal of Comparative Neurology*, *269*(1), 58–72. https://doi.org/10.1002/cne.902690105
Kaplan, R., Schuck, N. W., & Doeller, C. F. (2017). The Role of Mental Maps in Decision-Making. *Trends in Neurosciences*, *40*(5), 256–259. https://doi.org/10.1016/j.tins.2017.03.002
Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23. https://doi.org/10.1176/appi.ajp.160.1.13
Kawaguchi, Y., Wilson, C. J., Augood, S. J., & Emson, P. C. (1995). Striatal interneurones: chemical,

physiological and morphological characterization. In *Trends in Neurosciences* (Vol. 18, Issue 12). https://doi.org/10.1016/0166-2236(95)98374-8
Kayser, A. S., Allen, D. C., Navarro-Cebrian, A., Mitchell, J. M., & Fields, H. L. (2012). Dopamine, corticostriatal connectivity, and intertemporal choice. *Journal of Neuroscience*, *32*(27), 9402–9409. https://doi.org/10.1523/JNEUROSCI.1180-12.2012
Kemp, C., & Tenenbaum, J. B. (2009). Structured Statistical Models of Inductive Reasoning. *Psychological Review*, *116*(1). https://doi.org/10.1037/a0014282
Kemp, J. M., & Powell, T. P. (1971). The connexions of the striatum and globus pallidus: synthesis and speculation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *262*(845). https://doi.org/10.1098/rstb.1971.0106
Kennerley, S. W., & Wallis, J. D. (2009). Evaluating choices by single neurons in the frontal lobe: Outcome value encoded across multiple decision variables. *European Journal of Neuroscience*, *29*(10). https://doi.org/10.1111/j.1460-9568.2009.06743.x
Kim, J., & Ragozzino, M. E. (2005). The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiology of Learning and Memory*, *83*(2). https://doi.org/10.1016/j.nlm.2004.10.003
Kim, K. M., Baratta, M. V., Yang, A., Lee, D., Boyden, E. S., & Fiorillo, C. D. (2012). Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS ONE*, *7*(4). https://doi.org/10.1371/journal.pone.0033612
King, V., Corwin, J. V., & Reep, R. L. (1989). Production and characterization of neglect in rats with unilateral lesions of ventrolateral orbital cortex. *Experimental Neurology*, *105*(3), 287–299. https://doi.org/10.1016/0014-4886(89)90132-5
Kishida, K. T., Saez, I., Lohrenz, T., Witcher, M. R., Laxton, A. W., Tatter, S. B., White, J. P., Ellis, T. L., Phillips, P. E. M., & Montague, P. R. (2016). Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(1). https://doi.org/10.1073/pnas.1513619112
Kita, H., & Kitai, S. T. (1990). Amygdaloid projections to the frontal cortex and the striatum in the rat. *Journal of Comparative Neurology*, *298*(1), 40–49. https://doi.org/10.1002/cne.902980104
Kitchin, R. M. (1994). Cognitive maps: What are they and why study them? *Journal of Environmental Psychology*, *14*(1), 1–19. https://doi.org/10.1016/S0272-4944(05)80194-X
Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., & Behrens, T. E. J. (2013). Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex. *Journal of Neuroscience*, *33*(7), 3202–3211. https://doi.org/10.1523/JNEUROSCI.2532-12.2013
Klein-Flügge, M. C., Hunt, L. T., Bach, D. R., Dolan, R. J., & Behrens, T. E. J. (2011). Dissociable Reward and Timing Signals in Human Midbrain and Ventral Striatum. *Neuron*, *72*(4), 654–664. https://doi.org/10.1016/j.neuron.2011.08.024
Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(16). https://doi.org/10.1523/jneurosci.21-16-j0002.2001
Knutson, B., Bjork, J. M., Fong, G. W., Hommer, D., Mattay, V. S., & Weinberger, D. R. (2004). Amphetamine modulates human incentive processing. *Neuron*, *43*(2), 261–269. https://doi.org/10.1016/j.neuron.2004.06.030
Knutson, B., & Gibbs, S. E. B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology*, *191*(3), 813–822. https://doi.org/10.1007/s00213-006-0686-7
Kolachana, B. S., Saunders, R. C., & Weinberger, D. R. (1995). Augmentation of prefrontal cortical monoaminergic activity inhibits dopamine release in the caudate nucleus: An in vivo neurochemical assessment in the rhesus monkey. *Neuroscience*, *69*(3). https://doi.org/10.1016/0306-4522(95)00246-F
Kondo, H., Saleem, K. S., & Price, J. L. (2005). Differential connections of the perirhinal and

parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *Journal of Comparative Neurology*, *493*(4), 479–509. https://doi.org/10.1002/cne.20796
Koob, G. F., Riley, S. J., Smith, S. C., & Robbins, T. W. (1978). Effects of 6-hydroxydopamine lesions of the nucleus accumbens septi and olfactory tubercle on feeding, locomotor activity, and amphetamine anorexia in the rat. *Journal of Comparative and Physiological Psychology*, *92*(5), 917–927. https://doi.org/10.1037/h0077542
Korn, C., Akam, T., Jensen, K. H. R., Vagnoni, C., Huber, A., Tunbridge, E. M., & Walton, M. E. (2021). Distinct roles for dopamine clearance mechanisms in regulating behavioral flexibility. *Molecular Psychiatry*, *June*, 1–12. https://doi.org/10.1038/s41380-021-01194-y
Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., Maciver, M. A., & Poeppel, D. (2017). Neuron Perspective Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041
Krettek, J. E., & Price, J. L. (1977a). Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *Journal of Comparative Neurology*, *172*(4), 687–722. https://doi.org/10.1002/cne.901720408
Krettek, J. E., & Price, J. L. (1977b). The cortical projections of the mediodorsal nucleus and adjacent thalamic nuclei in the rat. *Journal of Comparative Neurology*, *171*(2), 157–191. https://doi.org/10.1002/cne.901710204
Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*(5), 341–372. https://doi.org/10.1016/j.pneurobio.2004.03.006
Kufahl, P. R., Li, Z., Risinger, R. C., Rainey, C. J., Wu, G., Bloom, A. S., & Li, S. J. (2005). Neural responses to acute cocaine administration in the human brain detected by fMRI. *NeuroImage*, *28*(4). https://doi.org/10.1016/j.neuroimage.2005.06.039
Kuramoto, E., Iwai, H., Yamanaka, A., Ohno, S., Seki, H., Tanaka, Y. R., Furuta, T., Hioki, H., & Goto, T. (2017). Dorsal and ventral parts of thalamic nucleus submedius project to different areas of rat orbitofrontal cortex: A single neuron-tracing study using virus vectors. *Journal of Comparative Neurology*, *525*(18), 3821–3839. https://doi.org/10.1002/cne.24306
Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6). https://doi.org/10.1073/pnas.1321596111
Lammel, S., Hetzel, A., Häckel, O., Jones, I., Liss, B., & Roeper, J. (2008). Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System. *Neuron*, *57*(5). https://doi.org/10.1016/j.neuron.2008.01.022
Lammel, S., Ion, D. I., Roeper, J., & Malenka, R. C. (2011). Projection-Specific Modulation of Dopamine Neuron Synapses by Aversive and Rewarding Stimuli. *Neuron*, *70*(5). https://doi.org/10.1016/j.neuron.2011.03.025
Lammel, S., Lim, B. K., & Malenka, R. C. (2014). Reward and aversion in a heterogeneous midbrain dopamine system. In *Neuropharmacology* (Vol. 76, Issue PART B). https://doi.org/10.1016/j.neuropharm.2013.03.019
Lammel, S., Lim, B. K., Ran, C., Huang, K. W., Betley, M. J., Tye, K. M., Deisseroth, K., & Malenka, R. C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, *491*(7423). https://doi.org/10.1038/nature11527
Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, if anything, is rodent prefrontal cortex? *ENeuro*, *5*(5). https://doi.org/10.1523/ENEURO.0315-18.2018
Lavond, D. G., & Steinmetz, J. E. (2003). Lesion Techniques for Behavioral Experiments. In *Handbook of Classical Conditioning*. https://doi.org/10.1007/978-1-4615-0263-0_8
Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). Estimated marginal means, aka least-squares means. In *R package version 1.5.0*.
Lerner, T. N., Holloway, A. L., & Seiler, J. L. (2021). Dopamine, Updated: Reward Prediction Error and Beyond. *Current Opinion in Neurobiology*, *67*, 123–130. https://doi.org/10.1016/j.conb.2020.10.012

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Li, J., Bravo, D. S., Louise Upton, A., Gilmour, G., Tricklebank, M. D., Fillenz, M., Martin, C., Lowry, J. P., Bannerman, D. M., & Mchugh, S. B. (2011). Close temporal coupling of neuronal activity and tissue oxygen responses in rodent whisker barrel cortex. *European Journal of Neuroscience*, *34*(12), 1983–1996. https://doi.org/10.1111/j.1460-9568.2011.07927.x

Lichtenberg, N. T., Pennington, Z. T., Holley, S. M., Greenfield, V. Y., Cepeda, C., Levine, M. S., & Wassum, K. M. (2017). Basolateral amygdala to orbitofrontal cortex projections enable cue-triggered reward expectations. *Journal of Neuroscience*, *37*(35). https://doi.org/10.1523/JNEUROSCI.0486-17.2017

Lipton, P. A., Alvarez, P., & Eichenbaum, H. (1999). Crossmodal associative memory representations in rodent orbitofrontal cortex. *Neuron*, *22*(2). https://doi.org/10.1016/S0896-6273(00)81095-8

Lodge, D. J. (2011). The medial prefrontal and orbitofrontal cortices differentially regulate dopamine system function. *Neuropsychopharmacology*, *36*(6), 1227–1236. https://doi.org/10.1038/npp.2011.7

Lohani, S., Poplawsky, A. J., Kim, S. G., & Moghaddam, B. (2017). Unexpected global impact of VTA dopamine neuron activation as measured by opto-fMRI. *Molecular Psychiatry*, *22*(4), 585–594. https://doi.org/10.1038/mp.2016.102

Lohrenz, T., Kishida, K. T., & Read Montague, P. (2016). Bold and its connection to dopamine release in human striatum: A cross-cohort comparison. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705). https://doi.org/10.1098/rstb.2015.0352

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(22). https://doi.org/10.1073/pnas.0608842104

Lopatina, N., Sadacca, B. F., McDannald, M. A., Styer, C. V., Peterson, J. F., Cheer, J. F., & Schoenbaum, G. (2017). Ensembles in medial and lateral orbitofrontal cortex construct cognitive maps emphasizing different features of the behavioral landscape. *Behavioral Neuroscience*, *131*(3), 201–212. https://doi.org/10.1037/bne0000195

Lowry, J. P., Boutelle, M. G., & Fillenz, M. (1997). Measurement of brain tissue oxygen at a carbon paste electrode can serve as an index of increases in regional cerebral blood flow. *Journal of Neuroscience Methods*. https://doi.org/10.1016/S0165-0270(96)00140-9

Lowry, J. P., Boutelle, M. G., O'Neill, R. D., & Fillenz, M. (1996). Characterization of carbon paste electrodes in vitro for simultaneous amperometric measurement of changes in oxygen and ascorbic acid concentrations in vivo. *Analyst*, *121*(6), 761–766. https://doi.org/10.1039/an9962100761

Lowry, J. P., Griffin, K., McHugh, S. B., Lowe, A. S., Tricklebank, M., & Sibson, N. R. (2010). Real-time electrochemical monitoring of brain tissue oxygen: A surrogate for functional magnetic resonance imaging in rodents. *NeuroImage*, *52*(2), 549–555. https://doi.org/10.1016/j.neuroimage.2010.04.187

Machado, C. J., & Bachevalier, J. (2007). The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *European Journal of Neuroscience*, *25*(9), 2885–2904. https://doi.org/10.1111/j.1460-9568.2007.05525.x

Mandino, F., Cerri, D. H., Garin, C. M., Straathof, M., van Tilborg, G. A. F., Chakravarty, M. M., Dhenain, M., Dijkhuizen, R. M., Gozzi, A., Hess, A., Keilholz, S. D., Lerch, J. P., Shih, Y. Y. I., & Grandjean, J. (2020). Animal Functional Magnetic Resonance Imaging: Trends and Path Toward Standardization. In *Frontiers in Neuroinformatics* (Vol. 13). https://doi.org/10.3389/fninf.2019.00078

Manns, J. R., & Eichenbaum, H. (2009). A cognitive map for object memory in the hippocampus. *Learning & Memory*, *16*(10), 616–624. https://doi.org/10.1101/lm.1484509

Mar, A. C., Walker, A. L. J., Theobald, D. E., Eagle, D. M., & Robbins, T. W. (2011). Dissociable

effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *Journal of Neuroscience*, *31*(17), 6398–6404. https://doi.org/10.1523/JNEUROSCI.6620-10.2011
Marín, O., Anderson, S. A., & Rubenstein, J. L. R. (2000). Origin and molecular specification of striatal interneurons. *Journal of Neuroscience*, *20*(16). https://doi.org/10.1523/jneurosci.20-16-06063.2000
Marota, J. J. A., Mandeville, J. B., Weisskoff, R. M., Moskowitz, M. A., Rosen, B. R., & Kosofsky, B. E. (2000). Cocaine activation discriminates dopaminergic projections by temporal response: An fMRI study in rat. *NeuroImage*, *11*(1). https://doi.org/10.1006/nimg.1999.0520
Marsden, C. A. (2006). Dopamine: the rewarding years. *British Journal of Pharmacology*, *147*(S1), S136–S144. https://doi.org/10.1038/sj.bjp.0706473
Masumi, A., & Sato, T. (2021). Model-based analysis of learning latent structures in probabilistic reversal learning task. *Artificial Life and Robotics*. https://doi.org/10.1007/s10015-020-00674-8
Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9). https://doi.org/10.1038/s41593-018-0209-y
Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248). https://doi.org/10.1038/nature08028
McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research*, *146*(1–2), 97–103. https://doi.org/10.1016/j.bbr.2003.09.019
McDermott, J. H., Wrobleski, D., & Oxenham, A. J. (2011). Recovering sound sources from embedded repetition. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(3), 1188–1193. https://doi.org/10.1073/pnas.1004765108
Mcdonald, A. J. (1991). Organization of amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience*, *44*(1), 1–14. https://doi.org/10.1016/0306-4522(91)90247-L
Mcdonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Progress in Neurobiology*, *55*(3), 257–332. https://doi.org/10.1016/S0301-0082(98)00003-3
McEnaney, K. W., & Butter, C. M. (1969). Perseveration of responding and nonresponding in monkeys with orbital frontal ablations. *Journal of Comparative and Physiological Psychology*, *68*(4), 558–561. https://doi.org/10.1037/h0027639
McHugh, S. B., Fillenz, M., Lowry, J. P., Rawlins, J. N. P., & Bannerman, D. M. (2011). Brain tissue oxygen amperometry in behaving rats demonstrates functional dissociation of dorsal and ventral hippocampus during spatial processing and anxiety. *European Journal of Neuroscience*. https://doi.org/10.1111/j.1460-9568.2010.07497.x
McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, *16*(4), 479–485. https://doi.org/10.1038/nn.3337
Meredith, G. E. (1999). The synaptic framework for chemical signaling in nucleus accumbens. *Annals of the New York Academy of Sciences*, *877*. https://doi.org/10.1111/j.1749-6632.1999.tb09266.x
Miranda, B., Nishantha Malalasekera, W. M., Behrens, T. E., Dayan, P., & Kennerley, S. W. (2020). Combined model-free and model-sensitive reinforcement learning in non-human primates. *PLoS Computational Biology*, *16*(6). https://doi.org/10.1371/journal.pcbi.1007944
Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, *72*(2), 1024–1027. https://doi.org/10.1152/jn.1994.72.2.1024
Mishkin, M. (1964). Perseveration of central sets after frontal lesions in monkeys. In J. Warren & K. Alkbert (Eds.), *The frontal granular cortex and behavior* (pp. 219–241). McGraw-Hill.
Mogenson, G. J., Jones, D. L., & Yim, C. Y. (1980). From motivation to action: Functional interface between the limbic system and the motor system. *Progress in Neurobiology*, *14*(2–3), 69–97. https://doi.org/10.1016/0301-0082(80)90018-0

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/jneurosci.16-05-01936.1996
Montaron, M. F., Deniau, J. M., Menetrey, A., Glowinski, J., & Thierry, A. M. (1996). Prefrontal cortex inputs of the nucleus accumbens-nigro-thalamic circuit. *Neuroscience*, *71*(2), 371–382. https://doi.org/10.1016/0306-4522(95)00455-6
Moore, R. Y., & Bloom, F. E. (1978). Central catecholamine neuron systems: anatomy and physiology of the dopamine systems. *Annual Review of Neuroscience*, *1*. https://doi.org/10.1146/annurev.ne.01.030178.001021
Murphy, M. J. M., & Deutch, A. Y. (2018). Organization of afferents to the orbitofrontal cortex in the rat. *Journal of Comparative Neurology*, *526*(9), 1498–1526. https://doi.org/10.1002/cne.24424
Murray, E. A., Moylan, E. J., Saleem, K. S., Basile, B. M., & Turchi, J. (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *ELife*, *4*(5), 1–18. https://doi.org/10.7554/eLife.11695
Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *Journal of Neuroscience*, *27*(31), 8166–8169. https://doi.org/10.1523/JNEUROSCI.1556-07.2007
Murray, E. A., & Rudebeck, P. H. (2018). Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nature Reviews Neuroscience*, *19*(7), 404–417. https://doi.org/10.1038/s41583-018-0013-4
Nauta, W. J. H., Smith, G. P., Faull, R. L. M., & Domesick, V. B. (1978). Efferent connections and nigral afferents of the nucleus accumbens septi in the rat. *Neuroscience*, *3*(4–5), 385–401. https://doi.org/10.1016/0306-4522(78)90041-6
Nicola, S. M., Surmeier, D. J., & Malenka, R. C. (2000). Dopaminergic Modulation of Neuronal Excitability in the Striatum and Nucleus Accumbens. *Annual Review of Neuroscience*, *23*(1), 185–215. https://doi.org/10.1146/annurev.neuro.23.1.185
Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, *22*(10), 1544–1553. https://doi.org/10.1038/s41593-019-0470-8
Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, *32*(2), 551–562. https://doi.org/10.1523/JNEUROSCI.5498-10.2012
Noonan, M. P., Chau, B. K. H., Rushworth, M. F. S., & Fellows, L. K. (2017). Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision-making in humans. *Journal of Neuroscience*, *37*(29), 7023–7035. https://doi.org/10.1523/JNEUROSCI.0692-17.2017
Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(47), 20547–20552. https://doi.org/10.1073/pnas.1012246107
O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7
O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science*, *304*(5669). https://doi.org/10.1126/science.1094285
O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*(5). https://doi.org/10.1016/S0896-6273(02)00603-7
O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1
O'Keefe, J., & Nadel, L. (1979). The hippocampus as a cognitive map. *Behavioral and Brain Sciences*, *2*(04), 487–494. https://doi.org/10.1017/S0140525X00063949

O'Neill, R. D., Grünewald, R. A., Fillenz, M., & Albery, W. J. (1982). Linear sweep voltammetry with carbon paste electrodes in the rat striatum. *Neuroscience*, *7*(8), 1945–1954. https://doi.org/10.1016/0306-4522(82)90009-4
Oades, R. D., & Halliday, G. M. (1987). Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Research Reviews*, *12*(2), 117–165. https://doi.org/10.1016/0165-0173(87)90011-7
Ogawa, M., van der Meer, M. A. A., Esber, G. R., Cerri, D. H., Stalnaker, T. A., & Schoenbaum, G. (2013). Risk-Responsive Orbitofrontal Neurons Track Acquired Salience. *Neuron*, *77*(2), 251–258. https://doi.org/10.1016/j.neuron.2012.11.006
Olds, J. (1958). Self-Stimulation of the Brain: Its Use To Study Local Effects of Hunger, Sex, and Drugs. *Science*, *127*(3294). https://doi.org/10.1126/science.127.3294.315
Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–427. https://doi.org/10.1037/h0058775
Öngür, D., An, X., & Price, J. L. (1998). Prefrontal cortical projections to the hypothalamus in macaque monkeys. *Journal of Comparative Neurology*, *401*(4), 480–505. https://doi.org/10.1002/(SICI)1096-9861(19981130)401:4<480::AID-CNE4>3.3.CO;2-6
Öngür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology*, *460*(3), 425–449. https://doi.org/10.1002/cne.10609
Ongur, D., & Price, J. L. (2000). The Organization of Networks within the Orbital and Medial Prefrontal Cortex of Rats, Monkeys and Humans. *Cerebral Cortex*, *10*(3), 206–219. https://doi.org/10.1093/cercor/10.3.206
Ono, T., Nishijo, H., & Uwano, T. (1995). Amygdala role in conditioned associative learning. In *Progress in Neurobiology* (Vol. 46, Issue 4). https://doi.org/10.1016/0301-0082(95)00008-J
Ostlund, S. B., & Balleine, B. W. (2005). Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *Journal of Neuroscience*, *25*(34), 7763–7770. https://doi.org/10.1523/JNEUROSCI.1921-05.2005
Ostlund, S. B., & Balleine, B. W. (2007). Orbitofrontal cortex mediates outcome encoding in pavlovian but not instrumental conditioning. *Journal of Neuroscience*, *27*(18), 4819–4825. https://doi.org/10.1523/JNEUROSCI.5443-06.2007
Ostlund, S. B., & Balleine, B. W. (2008). The disunity of Pavlovian and instrumental values. In *Behavioral and Brain Sciences* (Vol. 31, Issue 4). https://doi.org/10.1017/S0140525X08004925
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676
Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95–102. https://doi.org/10.1038/nn2020
Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron*, *96*(4), 736–754. https://doi.org/10.1016/j.neuron.2017.09.031
Panayi, M. C., & 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.37357.001
Papageorgiou, G. K., Baudonnat, M., Cucca, F., & Walton, M. E. (2016). Mesolimbic Dopamine Encodes Prediction Errors in a State-Dependent Manner. *Cell Reports*, *15*(2), 221–228. https://doi.org/10.1016/j.celrep.2016.03.031
Parker, N. F., Cameron, C. M., Taliaferro, J. P., Lee, J., Choi, J. Y., Davidson, T. J., Daw, N. D., & Witten, I. B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience*, *19*(6). https://doi.org/10.1038/nn.4287
Parkes, S. L., Ravassard, P. M., Cerpa, J. C., Wolff, M., Ferreira, G., & Coutureau, E. (2018). Insular and ventrolateral orbitofrontal cortices differentially contribute to goal-directed behavior

in rodents. *Cerebral Cortex*, *28*(7), 2313–2325. https://doi.org/10.1093/cercor/bhx132
Parkinson, J. A., Dalley, J. W., Cardinal, R. N., Bamford, A., Fehnert, B., Lachenal, G., Rudarakanchana, N., Halkerston, K. M., Robbins, T. W., & Everitt, B. J. (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: Implications for mesoaccumbens dopamine function. *Behavioural Brain Research*, *137*(1–2), 149–163. https://doi.org/10.1016/S0166-4328(02)00291-7
Parkinson, John A., Willoughby, P. J., Robbins, T. W., & Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal systems. *Behavioral Neuroscience*, *114*(1). https://doi.org/10.1037/0735-7044.114.1.42
Passingham, R. E. (2017). *The Brain Isn't Porridge*. Departmental Seminar, Department of Experimental Psychology, University of Oxford.
Passingham, R. E., Stephan, K. E., & Kötter, R. (2002). The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience*, *3*(8), 606–616. https://doi.org/10.1038/nrn893
Pavlov, I. P. (1927). Conditioned reflexes. An investigation of the physiological activity of the cerebral cortex. Translated and edited by E. V. ANREP. *Conditioned Reflexes. An Investigation of the Physiological Activity of the Cerebral Cortex. Translated and Edited by E. V. ANREP*.
Paxinos, G., & Watson, C. (1998). The Rat Brain in Stereotaxic Coordinates. In *Elsevier Academic Press* (4th ed.).
Paxinos, G., & Watson, C. (2013). The Rat Brain in Stereotaxic Coordinates Seventh Edition. *Elsevier Academic Press*.
Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The mixed instrumental controller: Using value of information to combine habitual choice and mental simulation. *Frontiers in Psychology*, *4*(MAR). https://doi.org/10.3389/fpsyg.2013.00092
Phillips, A. G., Ahn, S., & Floresco, S. B. (2004). Magnitude of Dopamine Release in Medial Prefrontal Cortex Predicts Accuracy of Memory on a Delayed Response Task. *Journal of Neuroscience*, *24*(2), 547–553. https://doi.org/10.1523/JNEUROSCI.4653-03.2004
Phillips, P. E. M., Stuber, G. D., Helen, M. L. A. V., Wightman, R. M., & Carelli, R. M. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature*, *422*(6932). https://doi.org/10.1038/nature01476
Phillipson, O. T., & Griffiths, A. C. (1985). The topographic order of inputs to nucleus accumbens in the rat. *Neuroscience*, *16*(2), 275–296. https://doi.org/10.1016/0306-4522(85)90002-8
Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behavioral Neuroscience*, *119*(1), 317–322. https://doi.org/10.1037/0735-7044.119.1.317
Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., & Schoenbaum, G. (2003). Different Roles for Orbitofrontal Cortex and Basolateral Amygdala in a Reinforcer Devaluation Task. *Journal of Neuroscience*, *23*(35), 11078–11084. https://doi.org/10.1523/jneurosci.23-35-11078.2003
Pierce, R. C., & Kalivas, P. W. (1997). A circuitry model of the expression of behavioral sensitization to amphetamine-like psychostimulants. *Brain Research Reviews*, *25*(2), 192–216. https://doi.org/10.1016/S0165-0173(97)00021-0
Preuss, T. M. (1995). Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *Journal of Cognitive Neuroscience*, *7*(1), 1–24. https://doi.org/10.1162/jocn.1995.7.1.1
Price, J. L. (2006). Architectonic structure of the orbital and medial prefrontal cortex. In D. H. Zald & S. L. Rauch (Eds.), *The Orbitofrontal Cortex*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198565741.003.0001
Price, J. L. (2007). Definition of the Orbital Cortex in Relation to Specific Connections with Limbic and Visceral Structures and Other Cortical Regions. *Annals of the New York Academy of Sciences*, *1121*(1), 54–71. https://doi.org/10.1196/annals.1401.008

Pycock, C. J., Carter, C. J., & Kerwin, R. W. (1980). Effect of 6-Hydroxydopamine Lesions of the Medial Prefrontal Cortex on Neurotransmitter Systems in Subcortical Sites in the Rat. *Journal of Neurochemistry*, *34*(1). https://doi.org/10.1111/j.1471-4159.1980.tb04625.x

Pycock, C. J., Kerwin, R. W., & Carter, C. J. (1980). Effect of lesion of cortical dopamine terminals on subcortical dopamine receptors in rats. *Nature*, *286*(5768), 74–77. https://doi.org/10.1038/286074a0

Python Software Foundation. (2013). Python Language Reference, version 2.7. In *Python Software Foundation*.

Quiroga-Varela, A., Aguilar, E., Iglesias, E., Obeso, J. A., & Marin, C. (2017). Short- and long-term effects induced by repeated 6-OHDA intraventricular administration: A new progressive and bilateral rodent model of Parkinson's disease. *Neuroscience*, *361*, 144–156. https://doi.org/10.1016/j.neuroscience.2017.08.017

R Core Team (2020). (2020). R: A language and environment for statistical computing. In *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*.

Rahman, S., & McBride, W. J. (2000). Feedback control of mesolimbic somatodendritic dopamine release in rat brain. *Journal of Neurochemistry*, *74*(2), 684–692. https://doi.org/10.1046/j.1471-4159.2000.740684.x

Ramus, S. J., & Eichenbaum, H. (2000). Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. *Journal of Neuroscience*, *20*(21). https://doi.org/10.1523/jneurosci.20-21-08199.2000

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Ray, J. P., & Price, J. L. (1992). The organization of the thalamocortical connections of the mediodorsal thalamic nucleus in the rat, related to the ventral forebrain–prefrontal cortex topography. *Journal of Comparative Neurology*, *323*(2), 167–197. https://doi.org/10.1002/cne.903230204

Ray, J. P., & Price, J. L. (1993). The organization of projections from the mediodorsal nucleus of the thalamus to orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, *337*(1), 1–31. https://doi.org/10.1002/cne.903370102

Reep, R. L., Corwin, J. V., & King, V. (1996). Neuronal connections of orbital cortex in rats: Topography of cortical and thalamic afferents. *Experimental Brain Research*. https://doi.org/10.1007/BF00227299

Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement BT - Clasical conditioning II: current research and theory. *Clasical Conditioning II: Current Research and Theory*.

Reynolds, L. M., Pokinko, M., Torres-Berrío, A., Cuesta, S., Lambert, L. C., Del Cid Pellitero, E., Wodzinski, M., Manitt, C., Krimpenfort, P., Kolb, B., & Flores, C. (2018). DCC Receptors Drive Prefrontal Cortex Maturation by Determining Dopamine Axon Targeting in Adolescence. *Biological Psychiatry*, *83*(2), 181–192. https://doi.org/10.1016/j.biopsych.2017.06.009

Reynolds, S. M., & Zahm, D. S. (2005). Specificity in the projections of prefrontal and insular cortex to ventral striatopallidum and the extended amygdala. *Journal of Neuroscience*, *25*(50), 11757–11767. https://doi.org/10.1523/JNEUROSCI.3432-05.2005

Robbins, T. W., & Roberts, A. C. (2007). Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cerebral Cortex*, *17*(SUPPL. 1), 151–160. https://doi.org/10.1093/cercor/bhm066

Robbins, T. W., Roberts, D. C. S., & Koob, G. F. (1983). Effects of d-amphetamine and apomorphine upon operant behavior and schedule-induced licking in rats with 6-hydroxydopamine-induced lesions of the nucleus accumbens. *Journal of Pharmacology and Experimental Therapeutics*, *224*(3), 662–673.

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., & Robbins, T. W. (1994). 6-Hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin card sort test: Possible interactions with

subcortical dopamine. *Journal of Neuroscience*, *14*(5 I). https://doi.org/10.1523/jneurosci.14-05-02531.1994
Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex*, *10*(3). https://doi.org/10.1093/cercor/10.3.252
Roberts, W. A., & Phelps, M. T. (1994). Transitive Inference in Rats: A Test of the Spatial Coding Hypothesis. *Psychological Science*, *5*(6), 368–374. https://doi.org/10.1111/j.1467-9280.1994.tb00287.x
Robinson, T. E., & Whishaw, I. Q. (1988). Normalization of extracellular dopamine in striatum following recovery from a partial unilateral 6-OHDA lesion of the substantia nigra: a microdialysis study in freely moving rats. *Brain Research*, *450*(1–2), 209–224. https://doi.org/10.1016/0006-8993(88)91560-0
Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*, *10*(12), 1615–1624. https://doi.org/10.1038/nn2013
Roesch, M. R., & Olson, C. R. (2005). Neuronal activity in primate orbitofrontal cortex reflects the value of time. *Journal of Neurophysiology*, *94*(4). https://doi.org/10.1152/jn.00373.2005
Roesch, M. R., Taylor, A. R., & Schoenbaum, G. (2006). Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. *Neuron*, *51*(4), 509–520. https://doi.org/10.1016/j.neuron.2006.06.027
Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery and Psychiatry*, *57*(12). https://doi.org/10.1136/jnnp.57.12.1518
Rose, J. E., & Woolsey, C. N. (1948). The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep and cat. *Research Publications - Association for Research in Nervous and Mental Disease*.
Rothenhoefe, K. M., Costa, V. D., Bartolo, R., Vicario-Feliciano, R., Murray, E. A., & Averbeck, B. B. (2017). Effects of ventral striatum lesions on stimulus-based versus action-based reinforcement learning. *Journal of Neuroscience*, *37*(29), 6902–6914. https://doi.org/10.1523/JNEUROSCI.0631-17.2017
Rudebeck, P. H., & Murray, E. A. (2008). Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(33), 8338–8343. https://doi.org/10.1523/JNEUROSCI.2272-08.2008
Rudebeck, P. H., & Murray, E. A. (2011). Balkanizing the primate orbitofrontal cortex: Distinct subregions for comparing and contrasting values. *Annals of the New York Academy of Sciences*, *1239*(1), 1–13. https://doi.org/10.1111/j.1749-6632.2011.06267.x
Rudebeck, P. H., & Murray, E. A. (2014). The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. In *Neuron* (Vol. 84, Issue 6, pp. 1143–1156). https://doi.org/10.1016/j.neuron.2014.10.049
Rudebeck, P. H., Ripple, J. A., Mitz, A. R., Averbeck, B. B., & Murray, E. A. (2017). *Amygdala contributions to stimulus–reward encoding in the macaque medial and orbital frontal cortex during learning*. *37*(8), 2186–2202. https://doi.org/10.1523/JNEUROSCI.0933-16.2017
Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., & Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nature Neuroscience*, *16*(8), 1140–1145. https://doi.org/10.1038/nn.3440
Russo, S. J., & Nestler, E. J. (2013). The brain reward circuitry in mood disorders. *Nature Reviews Neuroscience*, *14*(9), 609–625. https://doi.org/10.1038/nrn3381
Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *Journal of Neuroscience*, *30*(40), 13525–13536. https://doi.org/10.1523/JNEUROSCI.1747-10.2010
Sadacca, B. F., Jones, J. L., & Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *ELife*, *5*(MARCH2016), 1–13. https://doi.org/10.7554/eLife.13665

Saddoris, M. P., Cacciapaglia, F., Wightman, R. M., & Carelli, R. M. (2015). Differential dopamine release dynamics in the nucleus accumbens core and shell reveal complementary signals for error prediction and incentive motivation. *Journal of Neuroscience*, *35*(33), 11572–11582. https://doi.org/10.1523/JNEUROSCI.2344-15.2015

Sala-Bayo, J., Fiddian, L., Nilsson, S. R. O., Hervig, M. E., McKenzie, C., Mareschi, A., Boulos, M., Zhukovsky, P., Nicholson, J., Dalley, J. W., Alsiö, J., & Robbins, T. W. (2020). Dorsal and ventral striatal dopamine D1 and D2 receptors differentially modulate distinct phases of serial visual reversal learning. *Neuropsychopharmacology*, *45*(5). https://doi.org/10.1038/s41386-020-0612-4

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. In *Psychopharmacology* (Vol. 191, Issue 3). https://doi.org/10.1007/s00213-006-0668-9

Salamone, J. D., Correa, M., Yang, J. H., Rotolo, R., & Presby, R. (2018). Dopamine, effort-based choice, and behavioral economics: Basic and translational research. In *Frontiers in Behavioral Neuroscience* (Vol. 12). https://doi.org/10.3389/fnbeh.2018.00052

Salamone, J. D., Cousins, M. S., & 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. *Behavioural Brain Research*, *65*(2). https://doi.org/10.1016/0166-4328(94)90108-2

Salgado, S., & Kaplitt, M. G. (2015). The nucleus accumbens: A comprehensive review. *Stereotactic and Functional Neurosurgery*, *93*(2), 75–93. https://doi.org/10.1159/000368279

Sallet, J., Noonan, M. P., Thomas, A., O'Reilly, J. X., Anderson, J., Papageorgiou, G. K., Neubert, F. X., Ahmed, B., Smith, J., Bell, A. H., Buckley, M. J., Roumazeilles, L., Cuell, S., Walton, M. E., Krug, K., Mars, R. B., & Rushworth, M. F. S. (2020). Behavioral flexibility is associated with changes in structure and function distributed across a frontal cortical network in macaques. *PLoS Biology*, *18*(5), 1–26. https://doi.org/10.1371/journal.pbio.3000605

Saunders, B. T., & Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of pavlovian-conditioned responses. *European Journal of Neuroscience*, *36*(4), 2521–2532. https://doi.org/10.1111/j.1460-9568.2012.08217.x

Sawaguchi, T., & Goldman-Rakic, P. S. (1991). D1 dopamine receptors in prefrontal cortex: Involvement in working memory. *Science*, *4996*(251). https://doi.org/10.1126/science.1825731

Sawaguchi, T., Matsumura, M., & Kubota, K. (1990). Effects of dopamine antagonists on neuronal activity related to a delayed response task in monkey prefrontal cortex. *Journal of Neurophysiology*, *63*(6). https://doi.org/10.1152/jn.1990.63.6.1401

Schilman, E. A., Uylings, H. B. M., Graaf, Y. G. de, Joel, D., & Groenewegen, H. J. (2008). The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neuroscience Letters*, *432*(1), 40–45. https://doi.org/10.1016/j.neulet.2007.12.024

Schoenbaum, G., Chiba, A., & Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, *1*(2), 155–159. https://doi.org/10.1038/407

Schoenbaum, G., Roesch, M. R., & Stalnaker, T. A. (2006). Orbitofrontal cortex, decision-making and drug addiction. In *Trends in Neurosciences* (Vol. 29, Issue 2, pp. 116–124). https://doi.org/10.1016/j.tins.2005.12.006

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, *10*(12), 885–892. https://doi.org/10.1038/nrn2753

Schoenbaum, G., & Setlow, B. (2003). Lesions of Nucleus Accumbens Disrupt Learning about Aversive Outcomes. *Journal of Neuroscience*, *23*(30), 9833–9841. https://doi.org/10.1523/jneurosci.23-30-09833.2003

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., & Gallagher, M. (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning and Memory*, *10*(2), 129–140.

https://doi.org/10.1101/lm.55203
Schoenbaum, G., Setlow, B., Saddoris, M. P., & Gallagher, M. (2003). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron*, *39*(5), 855–867. https://doi.org/10.1016/S0896-6273(03)00474-4
Schoenfeld, R., & Uretsky, N. (1972). Altered response to apomorphine in 6-hydroxydopamine-treated rats. *European Journal of Pharmacology*, *19*(1), 115–118. https://doi.org/10.1016/0014-2999(72)90085-4
Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., Seidenbecher, C. I., Coenen, H. H., Heinze, H. J., Zilles, K., Düzel, E., & Bauer, A. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *Journal of Neuroscience*, *28*(52), 14311–14319. https://doi.org/10.1523/JNEUROSCI.2058-08.2008
Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27. https://doi.org/10.1152/jn.1998.80.1.1
Schultz, W. (2019). Recent advances in understanding the role of phasic dopamine activity. *F1000Research*, *8*, 1–12. https://doi.org/10.12688/f1000research.19793.1
Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593
Schwarz, A. J., Zocchi, A., Reese, T., Gozzi, A., Garzotti, M., Varnier, G., Curcuruto, O., Sartori, I., Girlanda, E., Biscaro, B., Crestan, V., Bertani, S., Heidbreder, C., & Bifone, A. (2004). Concurrent pharmacological MRI and in situ microdialysis of cocaine reveal a complex relationship between the central hemodynamic response and local dopamine concentration. *NeuroImage*, *23*(1), 296–304. https://doi.org/10.1016/j.neuroimage.2004.05.001
Scornaiencki, R., Cantrup, R., Rushlow, W. J., & Rajakumar, N. (2009). Prefrontal cortical D1 dopamine receptors modulate subcortical D2 dopamine receptor-mediated stress responsiveness. *International Journal of Neuropsychopharmacology*, *12*(9). https://doi.org/10.1017/S1461145709000121
Scott, T. R., Edwards, E. M., Smith, C. A., Hilgert, K. G., Schwartz, G. J., & Pritchard, T. C. (2005). Medial orbitofrontal cortex: Its role in mediating satiety in the macaque. *Chemical Senses*, *30 SUPPL. 1*. https://doi.org/10.1093/chemse/bjh178
Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*. https://doi.org/10.25080/majora-92bf1922-011
Seeman, P., Lee, T., Chau-Wong, M., & Wong, K. (1976). Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature*, *261*(5562). https://doi.org/10.1038/261717a0
Sescousse, G., Caldú, X., Segura, B., & Dreher, J. C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(4), 681–696. https://doi.org/10.1016/j.neubiorev.2013.02.002
Sescousse, G., Redouté, J., & Dreher, J. C. (2010). The architecture of reward value coding in the human orbitofrontal cortex. *Journal of Neuroscience*, *30*(39), 13095–13104. https://doi.org/10.1523/JNEUROSCI.3501-10.2010
Sharp, T., Zetterström, T., Ljungberg, T., & Ungerstedt, U. (1987). A direct comparison of amphetamine-induced behaviours and regional brain dopamine release in the rat using intracerebral dialysis. *Brain Research*. https://doi.org/10.1016/0006-8993(87)91416-8
Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*, *20*(5), 735–742. https://doi.org/10.1038/nn.4538

Sharpe, M. J., & Schoenbaum, G. (2016). Back to basics: Making predictions in the orbitofrontal-amygdala circuit. In *Neurobiology of Learning and Memory* (Vol. 131). https://doi.org/10.1016/j.nlm.2016.04.009

Shi, C., & Cassell, M. D. (1998). Cortical, Thalamic, and Amygdaloid Connections of the Anterior and Posterior Insular Cortices. *Journal of Comparative Neurology*, *468*(May), 440–468. https://doi.org/https://doi.org/10.1002/(SICI)1096-9861(19981005)399:4<440::AID-CNE2>3.0.CO;2-1

Shiflett, M. W., & Balleine, B. W. (2010). At the limbic-motor interface: Disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation. *European Journal of Neuroscience*, *32*(10), 1735–1743. https://doi.org/10.1111/j.1460-9568.2010.07439.x

Skinner, B. F. (1953). *Science and human behaviour* (Macmillan (Ed.)).

Slotnick, B., Hanford, L., & Hodos, W. (2000). Can rats acquire an olfactory learning set? *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(4). https://doi.org/10.1037/0097-7403.26.4.399

Smith, A. J., Li, M., Becker, S., & Kapur, S. (2007). Linking animal models of psychosis to computational models of dopamine function. *Neuropsychopharmacology*, *32*(1). https://doi.org/10.1038/sj.npp.1301086

Smith, D. V., Hayden, B. Y., Truong, T. K., Song, A. W., Platt, M. L., & Huettel, S. A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience*, *30*(7), 2490–2495. https://doi.org/10.1523/JNEUROSCI.3319-09.2010

Sobel, E., & Corbett, D. (1984). Axonal branching of ventral tegmental and raphe projections to the frontal cortex in the rat. *Neuroscience Letters*, *48*(2). https://doi.org/10.1016/0304-3940(84)90006-5

Spanagel, R., & Weiss, F. (1999). The dopamine hypothesis of reward: Past and current status. *Trends in Neurosciences*, *22*(11), 521–527. https://doi.org/10.1016/S0166-2236(99)01447-2

St. Onge, J. R., Chiu, Y. C., & Floresco, S. B. (2010). Differential effects of dopaminergic manipulations on risky choice. *Psychopharmacology*, *211*(2), 209–221. https://doi.org/10.1007/s00213-010-1883-y

St. Onge, J. R., & Floresco, S. B. (2010). Prefrontal cortical contribution to risk-based decision making. *Cerebral Cortex*, *20*(8), 1816–1828. https://doi.org/10.1093/cercor/bhp250

Stalnaker, T. A., Cooch, N. K., McDannald, M. A., Liu, T. L., Wied, H., & Schoenbaum, G. (2014). Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nature Communications*, *5*, 1–13. https://doi.org/10.1038/ncomms4926

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience*, *18*(5), 620–627. https://doi.org/10.1038/nn.3982

Stalnaker, T. A., Franz, T. M., Singh, T., & Schoenbaum, G. (2007). Basolateral Amygdala Lesions Abolish Orbitofrontal-Dependent Reversal Impairments. *Neuron*, *54*(1), 51–58. https://doi.org/10.1016/j.neuron.2007.02.014

Stalnaker, T. A., Howard, J. D., Takahashi, Y. K., Gershman, S. J., Kahnt, T., & Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. *ELife*, *8*. https://doi.org/10.7554/eLife.49315

Stalnaker, T. A., Liu, T. L., Takahashi, Y. K., & Schoenbaum, G. (2018). Orbitofrontal neurons signal reward predictions, not reward prediction errors. *Neurobiology of Learning and Memory*, *153*(January), 137–143. https://doi.org/10.1016/j.nlm.2018.01.013

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, *20*(4). https://doi.org/10.1038/nn.4520

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Stoeckel, K., Schwab, M., & Thoenen, H. (1977). Role of gangliosides in the uptake and retrograde axonal transport of cholera and tetanus toxin as compared to nerve growth

factor and wheat germ agglutinin. *Brain Research*, *132*(2). https://doi.org/10.1016/0006-8993(77)90421-8
Stopper, C. M., Green, E. B., & Floresco, S. B. (2014). Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cerebral Cortex*, *24*(1), 154–162. https://doi.org/10.1093/cercor/bhs297
Sugam, J. A., Day, J. J., Wightman, R. M., & Carelli, R. M. (2012). Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. *Biological Psychiatry*, *71*(3). https://doi.org/10.1016/j.biopsych.2011.09.029
Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033
Sulzer, D., Sonders, M. S., Poulsen, N. W., & Galli, A. (2005). Mechanisms of neurotransmitter release by amphetamines: A review. *Progress in Neurobiology*, *75*(6), 406–433. https://doi.org/10.1016/j.pneurobio.2005.04.003
Surmeier, D. J., Ding, J., Day, M., Wang, Z., & Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in Neurosciences*, *30*(5), 228–235. https://doi.org/10.1016/j.tins.2007.03.008
Sutton, R. S. (1988). Learning to Predict by the Methods of Temporal Differences. *Machine Learning*, *3*(1). https://doi.org/10.1023/A:1022633531479
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning* (Vol. 14, Issue 4). MIT Press.
Swanson, L. W. (1982). The projections of the ventral tegmental area and adjacent regions: A combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Research Bulletin*, *9*(1–6), 321–353. https://doi.org/10.1016/0361-9230(82)90145-9
Swanson, L. W., & Cowan, W. M. (1975). A note on the connections and development of the nucleus accumbens. *Brain Research*, *92*(2), 324–330. https://doi.org/10.1016/0006-8993(75)90278-4
Taber, M. T., Das, S., & Fibiger, H. C. (1995). Cortical Regulation of Subcortical Dopamine Release: Mediation via the Ventral Tegmental Area. *Journal of Neurochemistry*, *65*(3). https://doi.org/10.1046/j.1471-4159.1995.65031407.x
Takagishi, M., & Chiba, T. (1991). Efferent projections of the infralimbic (area 25) region of the medial prefrontal cortex in the rat: an anterograde tracer PHA-L study. *Brain Research*, *566*(1–2), 26–39. https://doi.org/10.1016/0006-8993(91)91677-S
Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., & Schoenbaum, G. (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. *Neuron*, *95*(6). https://doi.org/10.1016/j.neuron.2017.08.025
Takahashi, Y. K., Chang, C. Y., Lucantonio, F., Haney, R. Z., Berg, B. A., Yau, H. J., Bonci, A., & Schoenbaum, G. (2013). Neural Estimates of Imagined Outcomes in the Orbitofrontal Cortex Drive Behavior and Learning. *Neuron*, *80*(2), 507–518. https://doi.org/10.1016/j.neuron.2013.08.008
Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, *91*(1), 182–193. https://doi.org/10.1016/j.neuron.2016.05.015
Takahashi, Y. K., Roesch, M. R., Stalnaker, T. A., & Schoenbaum, G. (2007). Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. *Frontiers in Integrative Neuroscience*, *1*(DEC). https://doi.org/10.3389/neuro.07.011.2007
Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O'Donnell, P., Niv, Y., & Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*, *14*(12), 1590–1597. https://doi.org/10.1038/nn.2957
Tang, J. S., Qu, C. L., & Huo, F. Q. (2009). The thalamic nucleus submedius and ventrolateral orbital cortex are involved in nociceptive modulation: A novel pain modulation pathway. In *Progress in Neurobiology*. https://doi.org/10.1016/j.pneurobio.2009.10.002
Thompson, T. L., & Moss, R. L. (1995). In vivo stimulated dopamine release in the nucleus accumbens: modulation by the prefrontal cortex. *Brain Research*, *686*(1). https://doi.org/10.1016/0006-8993(95)00429-T

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, *2*(4). https://doi.org/10.1037/h0092987
Thorpe, S. J., Rolls, E. T., & Maddison, S. (1983). The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Experimental Brain Research*, *49*(1). https://doi.org/10.1007/BF00235545
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715). https://doi.org/10.1126/science.1105370
Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, *95*(1). https://doi.org/10.1152/jn.00762.2005
Tolman, E. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/h0061626
Tolman, E. C., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, *42*(1), 43–77. https://doi.org/10.1016/j.neuron.2014.1
Tolman, E., & Honzik, C. (1930). Introduction and removal of reward, and maze performance in rats. In *University of California Publications in Psychology* (Vol. 4).
Tolman, E., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning. I. Orientation and the short-cut. *Journal of Experimental Psychology*, *36*(1), 13–24. https://doi.org/10.1037/h0053944
Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*(6729). https://doi.org/10.1038/19525
Tritsch, N. X., & Sabatini, B. L. (2012). Dopaminergic Modulation of Synaptic Transmission in Cortex and Striatum. *Neuron*, *76*(1), 33–50. https://doi.org/10.1016/j.neuron.2012.09.023
Tunbridge, E. M., Bannerman, D. M., Sharp, T., & Harrison, P. J. (2004). Catechol-O-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *Journal of Neuroscience*, *24*(23). https://doi.org/10.1523/JNEUROSCI.1124-04.2004
Urban, A., Dussaux, C., Martel, G., Brunner, C., Mace, E., & Montaldo, G. (2015). Real-time imaging of brain activity in freely moving rats using functional ultrasound. *Nature Methods*, *12*(9), 873–878. https://doi.org/10.1038/nmeth.3482
Uylings, H. B. M., Groenewegen, H. J., & Kolb, B. (2003). Do rats have a prefrontal cortex? *Behavioural Brain Research*, *146*(1–2), 3–17. https://doi.org/10.1016/j.bbr.2003.09.028
Van De Werd, H. J. J. M., & Uylings, H. B. M. (2008). The rat orbital and agranular insular prefrontal cortical areas: A cytoarchitectonic and chemoarchitectonic study. *Brain Structure and Function*, *212*(5), 387–401. https://doi.org/10.1007/s00429-007-0164-y
van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, *67*(1). https://doi.org/10.1016/j.neuron.2010.06.023
van Duuren, E., Escámez, F. A. N., Joosten, R. N. J. M. A., Visser, R., Mulder, A. B., & Pennartz, C. M. A. (2007). Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *14*(6). https://doi.org/10.1101/lm.546207
van Duuren, E., Van Der Plasse, G., Lankelma, J., Joosten, R. N. J. M. A., Feenstra, M. G. P., & Pennartz, C. M. A. (2009). Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. *Journal of Neuroscience*, *29*(28), 8965–8976. https://doi.org/10.1523/JNEUROSCI.0005-09.2009
van Rossum, J. M., van der Schoot, J. B., & Hurkmans, J. A. T. M. (1962). Mechanism of action of cocaine and amphetamine in the brain. *Experientia*, *18*(5). https://doi.org/10.1007/BF02148316
Vanderschuren, L. J. M. J., Donné Schmidt, E., De Vries, T. J., Van Moorsel, C. A. P., Tilders, F. J. H., & Schoffelmeer, A. N. M. (1999). A single exposure to amphetamine is sufficient to induce long-term behavioral, neuroendocrine, and neurochemical sensitization in rats. *Journal of*

Neuroscience, *19*(21), 9579–9586. https://doi.org/10.1523/jneurosci.19-21-09579.1999
Vasquez-Lopez, S. A., Weissenberger, Y., Lohse, M., Keating, P., King, A. J., & Dahmen, J. C. (2017). Thalamic input to auditory cortex is locally heterogeneous but globally tonotopic. *ELife*, *6*, 1–18. https://doi.org/10.7554/eLife.25141
Vázquez-Borsetti, P., Cortés, R., & Artigas, F. (2009). Pyramidal neurons in rat prefrontal cortex projecting to ventral tegmental area and dorsal raphe nucleus express 5-HT2A receptors. *Cerebral Cortex*, *19*(7). https://doi.org/10.1093/cercor/bhn204
Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology*, *44*(13), 2195–2204. https://doi.org/10.1038/s41386-019-0454-0
Verharen, J. P. H., De Jong, J. W., Roelofs, T. J. M., Huffels, C. F. M., Van Zessen, R., Luijendijk, M. C. M., Hamelink, R., Willuhn, I., Den Ouden, H. E. M., Van Der Plasse, G., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2018). A neuronal mechanism underlying decision-making deficits during hyperdopaminergic states. *Nature Communications*, *9*(1), 1–15. https://doi.org/10.1038/s41467-018-03087-1
Verharen, J. P. H., den Ouden, H. E. M., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2020). Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology*. https://doi.org/10.1007/s00213-020-05454-7
Vertes, R. P. (2004). Differential Projections of the Infralimbic and Prelimbic Cortex in the Rat. *Synapse*, *51*(1), 32–58. https://doi.org/10.1002/syn.10279
Vertes, R. P., & Hoover, W. B. (2008). Projections of the paraventricular and paratenial nuclei of the dorsal midline thalamus in the rat. *Journal of Comparative Neurology*, *508*(2), 212–237. https://doi.org/10.1002/cne.21679
Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. https://doi.org/10.1038/s41592-019-0686-2
Völlm, B. A., De Araujo, I. E., Cowen, P. J., Rolls, E. T., Kringelbach, M. L., Smith, K. A., Jezzard, P., Heal, R. J., & Matthews, P. M. (2004). Methamphetamine activates reward circuitry in drug naïve human subjects. *Neuropsychopharmacology*, *29*(9). https://doi.org/10.1038/sj.npp.1300481
Wallis, G. G., Mcharg, J. F., & Scott, O. C. A. (1949). Acute Psychosis Caused by Dextro-amphetamine. *British Medical Journal*. https://doi.org/10.1136/bmj.2.4641.1394
Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, *15*(1), 13–19. https://doi.org/10.1038/nn.2956
Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. *Neuron*, *65*(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027
Walton, M. E., Behrens, T. E. J., Noonan, M. P., & Rushworth, M. F. S. (2011). Giving credit where credit is due: Orbitofrontal cortex and valuation in an uncertain world. *Annals of the New York Academy of Sciences*, *1239*(1), 14–24. https://doi.org/10.1111/j.1749-6632.2011.06257.x
Walton, M. E., Chau, B. K. H., & Kennerley, S. W. (2015). Prioritising the relevant information for learning and decision making within orbital and ventromedial prefrontal cortex. *Current Opinion in Behavioral Sciences*, *1*, 78–85. https://doi.org/10.1016/j.cobeha.2014.10.005
Wang, Z., Yue, L., Cui, C., Liu, S., Wang, X., Li, Y., & Ma, L. (2019). Top - down control of the medial orbitofrontal cortex to nucleus accumbens core pathway in decisional impulsivity. *Brain Structure and Function*, *224*(7), 2437–2452. https://doi.org/10.1007/s00429-019-01913-w
Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience*, *40*(1), 373–394. https://doi.org/10.1146/annurev-neuro-

072116-031109
Weinberger, D. R. (1987). Implications of Normal Brain Development for the Pathogenesis of Schizophrenia. *Archives of General Psychiatry*, *44*(7). https://doi.org/10.1001/archpsyc.1987.01800190080012
Werlen, E., Shin, S. L., Gastambide, F., Francois, J., Tricklebank, M. D., Marston, H. M., Huxter, J. R., Gilmour, G., & Walton, M. E. (2019). Amphetamine disrupts haemodynamic correlates of prediction errors in nucleus accumbens and orbitofrontal cortex. *Neuropsychopharmacology*, *45*(5), 793–803. https://doi.org/10.1038/s41386-019-0564-8
Wernick, N. L. B., Chinnapen, D. J. F., Cho, J. A., & Lencer, W. I. (2010). Cholera toxin: An intracellular journey into the cytosol by way of the endoplasmic reticulum. *Toxins*, *2*(3), 310–325. https://doi.org/10.3390/toxins2030310
Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat Rev Neurosci*, *17*(8), 513–523. https://doi.org/10.1038/nrn.2016.56
Wilson, C., Nomikos, G. G., Collu, M., & Fibiger, H. C. (1995). Dopaminergic correlates of motivated behavior: Importance of drive. *Journal of Neuroscience*, *15*(7 II). https://doi.org/10.1523/jneurosci.15-07-05169.1995
Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, *81*(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005
Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science*, *338*(6104), 270–273. https://doi.org/10.1126/science.1223252
Winkowski, D. E., Nagode, D. A., Donaldson, K. J., Yin, P., Shamma, S. A., Fritz, J. B., & Kanold, P. O. (2018). Orbitofrontal Cortex Neurons Respond to Sound and Activate Primary Auditory Cortex Neurons. *Cerebral Cortex*, *28*(3), 868–879. https://doi.org/10.1093/cercor/bhw409
Wise, R. A., & Rompre, P. P. (1989). Brain dopamine and reward. *Annual Review of Psychology*, *40*, 191–225. https://doi.org/10.1146/annurev.ps.40.020189.001203
Wise, Roy A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494. https://doi.org/10.1038/nrn1406
Wise, Roy A., Spindler, J., Dewit, H., & Gerber, G. J. (1978). Neuroleptic-induced "anhedonia" in rats: Pimozide blocks reward quality of food. *Science*, *201*(4352). https://doi.org/10.1126/science.566469
Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends in Neurosciences*, *31*(12), 599–608. https://doi.org/10.1016/j.tins.2008.08.008
Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., Yizhar, O., Cho, S. L., Gong, S., Ramakrishnan, C., Stuber, G. D., Tye, K. M., Janak, P. H., & Deisseroth, K. (2011). Recombinase-driver rat lines: Tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, *72*(5). https://doi.org/10.1016/j.neuron.2011.10.028
Wittmann, M. K., Fouragnan, E., Folloni, D., Klein-Flügge, M. C., Chau, B. K. H., Khamassi, M., & Rushworth, M. F. S. (2020). Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17343-w
Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/1978942.1978963
Xie, J., & Padoa-Schioppa, C. (2016). Neuronal remapping and circuit persistence in economic decisions. *Nature Neuroscience*, *19*(6). https://doi.org/10.1038/nn.4300
Yoshida, A., Dostrovsky, J. O., & Chiang, C. Y. (1992). The afferent and efferent connections of the nucleus submedius in the rat. *Journal of Comparative Neurology*. https://doi.org/10.1002/cne.903240109
Yun, I. A., Nicola, S. M., & Fields, H. L. (2004). Contrasting effects of dopamine and glutamate

receptor antagonist injection in the nucleus accumbens suggest a neural mechanism underlying cue-evoked goal-directed behavior. *European Journal of Neuroscience*, *20*(1). https://doi.org/10.1111/j.1460-9568.2004.03476.x

Zahm, D. S., & Brog, J. S. (1992). On the significance of subterritories in the "accumbens" part of the rat ventral striatum. *Neuroscience*, *50*(4). https://doi.org/10.1016/0306-4522(92)90202-D

Zahm, D. S., & Heimer, L. (1993). Specificity in the efferent projections of the nucleus accumbens in the rat: Comparison of the rostral pole projection patterns with those of the core and shell. *Journal of Comparative Neurology*, *327*(2). https://doi.org/10.1002/cne.903270205

Zahrt, J., Taylor, J. R., Mathew, R. G., & Arnsten, A. F. T. (1997). Supranormal stimulation of D1 dopamine receptors in the rodent prefrontal cortex impairs spatial working memory performance. *Journal of Neuroscience*, *17*(21). https://doi.org/10.1523/jneurosci.17-21-08528.1997

Zald, D. H., & Treadway, M. T. (2017). Reward Processing, Neuroeconomics, and Psychopathology. *Annual Review of Clinical Psychology*, *13*, 471–495. https://doi.org/10.1146/annurev-clinpsy-032816-044957

Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2019). Rat Orbitofrontal Ensemble Activity Contains Multiplexed but Dissociable Representations of Value and Task Structure in an Odor Sequence Task. *Current Biology*, *29*(6), 897-907.e3. https://doi.org/10.1016/j.cub.2019.01.048

Zhou, J., Montesinos-Cartagena, M., Wikenheiser, A. M., Gardner, M. P. H., Niv, Y., & Schoenbaum, G. (2019). Complementary Task Structure Representations in Hippocampus and Orbitofrontal Cortex during an Odor Sequence Task. *Current Biology*, *29*(20), 3402-3409.e3. https://doi.org/10.1016/j.cub.2019.08.040

Zweifel, L. S., Fadok, J. P., Argilli, E., Garelick, M. G., Jones, G. L., Dickerson, T. M. K., Allen, J. M., Mizumori, S. J. Y., Bonci, A., & Palmiter, R. D. (2011). Activation of dopamine neurons is critical for aversive conditioning and prevention of generalized anxiety. *Nature Neuroscience*, *14*(5). https://doi.org/10.1038/nn.2808