

Understanding compulsive behaviour in psychiatric disorders with a touchscreen rodent model of reversal learning

Mark Daniel Rafter, BSc.

Thesis submitted to the University of Nottingham for
the degree of Doctor of Philosophy.

SEPTEMBER 2016



**The University of
Nottingham**

UNITED KINGDOM • CHINA • MALAYSIA

*Man can do what he wills but he cannot will what he wills.
– Arthur Schopenhauer*

*The diminutive chains of habit are seldom heavy enough to be
felt, till they are too strong to be broken.
– Samuel Johnson*

*There is no fate unless there is someone or something to be
fated. There is no trap without someone to be caught. There is,
indeed, no compulsion unless there is also freedom of choice, for
the sensation of behaving involuntarily is known only by
contrast with that of behaving voluntarily.
– Alan Watts*

*Reality is what we take to be true. What we take to be true is
what we believe. What we believe is based upon our perceptions.
What we perceive depends on what we look for. What we look
for depends on what we think. What we think depends on what
we perceive. What we perceive determines what we believe.
What we believe determines what we take to be true. What we
take to be true is our reality.
– David Bohm*

Abstract

Behaviour is considered to be compulsive when it is performed automatically regardless of whether it results in deleterious consequences. Although most prominently associated with obsessive-compulsive disorder, compulsivity is also present in a host of other psychiatric disorders and likely represents a trans-diagnostic trait with shared dysfunctional circuitry (Gillan et al., 2016a). Despite this, no single treatment shows anti-compulsive efficacy across disorders, and disorder-specific treatments are not particularly efficacious either (Grant et al., 2016). This may be because different circuitry parameters are disrupted in different disorders, but result in similar behavioural outcomes, therefore a treatment targeting one parameter will not alleviate dysfunction caused by alterations in a different parameter. This thesis investigates the circuitry of compulsivity by administering drugs that differentially target these parameters to rats undergoing associative learning tasks shown to be dependent on this neural circuitry.

We found that the acute administration of phencyclidine – a drug which models the psychotic state (by blocking NMDA receptors; Rafter et al., 2016) – promoted compulsive approach towards a formerly rewarded stimulus but not compulsive avoidance to a formerly unrewarded stimulus. We also found that administration of this drug to neonatal pups in combination with adolescent social isolation led to the opposite effect, that is, reduced compulsive approach towards a formerly rewarded stimulus once it became unrewarded. Administration of a serotonin 5-HT_{2c} receptor antagonist (a putative anti-compulsive agent) had no effect on choices but accelerated the

speed of responding. Meanwhile intra-orbitofrontal cortex infusion of the dopamine neurotoxin 6-hydroxydopamine or the serotonin neurotoxin 5,7-dihydroxytryptamine failed to reliably induce neurotransmitter depletion, and subsequently had no effect on any behavioural measure.

These findings suggest that targeting glutamate systems upstream of dopamine and serotonin systems may result in better treatment outcomes for compulsivity driven by formerly reinforced associations, e.g. in delusions, behavioural addictions, and drug addiction.

Table of Contents

Chapter 1 – General Introduction	1
1.1 Introduction	1
1.2 Decision making and reinforcement learning.....	2
1.2.1 Modelling reinforcement learning	2
1.2.2 Multiple reinforcement learning systems	4
1.3 Dual system decision making in animals	7
1.3.1 Behavioural evidence for a dual system	7
1.3.2 Neural evidence for a dual system	11
1.4 Disruption of dual system balance in psychiatric disorders	22
1.4.1 Obsessive-compulsive disorder	22
1.4.2 Schizophrenia.....	26
1.4.3 Drug addiction	31
1.4.4 Parkinson’s disease.....	33
1.4.5 Summary	34
1.5 Studying dual system function with rodent reversal learning	34
1.5.1 Reversal learning	34
1.5.2 Role of OFC	36
1.5.3 Optimal conditions and measures.....	41
1.6 Aims.....	45
Chapter 2 – The effect of phencyclidine on the expression of learned associations	49
2.1 Introduction	49
2.2 Methods	51
2.2.1 Animals	51
2.2.2 Apparatus.....	52
2.2.3 Touchscreen task protocol.....	53
2.2.4 Measures and statistics	56
2.3 Results	56
2.3.1 Pre-training and acquisition phase.....	56
2.3.2 Testing Phase.....	57
2.4 Discussion	59
Chapter 3 – The effect of phencyclidine on learning after reversal of learned associations	62
3.1 Introduction	62
3.2 Methods	63
3.2.1 Animals	63

3.2.2 Apparatus.....	63
3.2.3 Touchscreen task protocol.....	64
3.2.4 Locomotor activity protocol.....	67
3.2.5 Measures and statistics	67
3.3 Results	68
3.3.1 Acquisition phases.....	68
3.3.2 Pre-Reversal phase.....	68
3.3.3 Reversal phase	69
3.3.4 Novel Acquisition phase.....	74
3.3.5 Locomotor activity	77
3.4 Discussion	80
Chapter 4 – The effect of phencyclidine on the reversal of positive and negative associations	84
4.1 Introduction	84
4.2 Methods	85
4.2.1 Animals	85
4.2.2 Apparatus.....	85
4.2.3 Touchscreen task protocol.....	86
4.2.4 Locomotor activity protocol.....	88
4.2.5 Measures and statistics	88
4.3 Results	88
4.3.1 Acquisition phases.....	88
4.3.2 Perseverance test.....	88
4.3.2 Learned Irrelevance test.....	91
4.3.3 Novelty Bias	93
4.3.4 Extinction.....	95
4.3.5 Latency Measures	96
4.3.6 Locomotor activity	99
4.4 Discussion	101
Chapter 5 – The effect of SB-242084 on the reversal of positive and negative associations	105
5.1 Introduction	105
5.2 Methods	108
5.2.1 Animals	108
5.2.2 Apparatus.....	108
5.2.3 Touchscreen task protocol.....	108
5.2.4 Measures and statistics	108

5.3 Results	109
5.3.1 Acquisition phases.....	109
5.3.2 Perseverance test.....	109
5.3.3 Learned Irrelevance test	111
5.3.4 Novelty Bias	113
5.3.5 Latency Measures	114
5.4 Discussion	118
Chapter 6 – The effect of dual-hit neonatal PCP/post-weaning social isolation on the reversal of positive and negative associations	120
6.1 Introduction	120
6.2 Methods	123
6.2.1 Animals	123
6.2.2 Apparatus.....	124
6.2.3 Locomotor activity protocol	124
6.2.4 Novel object discrimination protocol	124
6.2.5 Social interaction protocol	125
6.2.6 Touchscreen task protocol.....	125
6.2.7 Measures and statistics	126
6.3 Results	126
6.3.1 Locomotor activity	126
6.3.2 Novel object discrimination.....	127
6.3.3 Social interaction	129
6.3.4 Reversal learning task.....	130
6.4 Discussion	139
Chapter 7 – The effect of intra-OFC 6-hydroxydopamine on the reversal of positive and negative associations	142
7.1 Introduction	142
7.2 Methods	144
7.2.1 Animals	144
7.2.2 Apparatus.....	144
7.2.3 Touchscreen task protocol.....	144
7.2.4 Surgery protocol	145
7.2.5 Brain monoamine quantification.....	146
7.2.6 Measures and statistics	148
7.3 Results	148
7.3.1 Acquisition phases.....	148
7.3.2 Perseverance test.....	149

7.3.3 Learned irrelevance test	151
7.3.4 Novelty bias test.....	153
7.3.5 Extinction.....	154
7.3.6 Latency measures	155
7.3.7 Locomotor activity	157
7.3.8 Post-mortem neurochemistry	158
7.4 Discussion	162
Chapter 8 – The effect of intra-OFC 5,7-dihydroxytryptamine on the reversal of positive and negative associations	166
8.1 Introduction	166
8.2 Methods	167
8.2.1 Animals	167
8.2.2 Apparatus.....	167
8.2.3 Touchscreen task protocol.....	167
8.2.4 Surgery protocol	168
8.2.5 Brain monoamine quantification.....	168
8.2.6 Measures and statistics	168
8.3 Results	168
8.3.1 Acquisition phases.....	168
8.3.2 Perseverance test.....	170
8.3.3 Learned irrelevance test	171
8.3.4 Extinction.....	173
8.3.5 Latency measures	174
8.3.6 Post-mortem neurochemistry	176
8.4 Discussion	179
Chapter 9 – General Discussion	183
9.1 Summary of findings	183
9.2 Interpretations for compulsivity.....	184
9.2.1 Split task interpretation	184
9.2.2 Task difficulty.....	187
9.2.3 Stimulus associability.....	187
9.2.4 Solving problems of interpretation	190
9.3 Future questions and treatment implications	191
9.3.1 Serotonin.....	191
9.3.2 Glutamate.....	192
9.3.3 Reversal learning for modelling compulsivity	195
9.4 Concluding remarks	197

References.....	199
Appendix.....	225

List of Figures

Figure 1.1 Hypothetical performance of four agents in a two-choice task where only one option is rewarded (correct).....	4
Figure 1.2 Model-free and model-based learning.....	6
Figure 1.3 The organisation of feed-forward spiralling cortico-striato-thalamo-cortical loops.....	13
Figure 1.4 Action selection via corticostriatal plasticity.....	17
Figure 1.5 Prediction error dynamics and learning.....	20
Figure 1.6 Theoretical performance and controller system engagement during a two-choice multi-trial reversal learning task.....	36
Figure 1.7 Credit assignment by orbitofrontal cortex.....	38
Figure 1.8 Effects of pharmacological manipulations within corticostriatal circuitry on goal-directed behaviour.....	47
Figure 2.1 A rat touchscreen operant box displaying two stimuli.	53
Figure 2.2 A flowchart showing the different stages of a trial for the final discrimination phase of training.	55
Figure 2.3 Measures during acquisition training	57
Figure 2.4 The effect of acute PCP administration (0 – 5mg/kg, n=12 for each condition) on various measures of visual discrimination performance.....	59
Figure 3.1 Responses made during a simple reversal test for PCP or vehicle treated animals.....	70
Figure 3.2 Sessions to criterion on a simple reversal test for PCP or vehicle treated animals.....	71
Figure 3.3 Latency measures during a simple reversal test for PCP or vehicle treated animals.....	74
Figure 3.4 Responses made during a novel acquisition test for PCP or vehicle treated animals.....	75
Figure 3.5 Sessions to criterion on a novel acquisition test for PCP or vehicle treated animals.....	76
Figure 3.6 Responses made split by image assignment during a novel acquisition test for PCP or vehicle treated animals.	77
Figure 3.7 The effect of PCP on locomotor activity after receiving PCP during a simple reversal task.	79
Figure 4.1 Stimulus-reward contingency changes for each phase of the split reversal task.	87
Figure 4.2 Responses made during the perseverance test for PCP or vehicle treated animals.....	90
Figure 4.3 Sessions to criterion during the perseverance test for PCP or vehicle treated animals.	91
Figure 4.4 Responses made during the learned irrelevance test for PCP or vehicle treated animals.	92

Figure 4.5 Sessions to criterion during the learned irrelevance test for PCP or vehicle treated animals.....	93
Figure 4.6 Responses made during the novelty bias test for PCP or vehicle treated animals.....	94
Figure 4.7 Responses made during the novelty bias test for PCP or vehicle treated animals, split by prior reversal contingency.....	95
Figure 4.8 Responses made during the extinction for PCP or vehicle treated animals.	96
Figure 4.9 Latency measures during the perseverance and learned irrelevance test (collated) for PCP or vehicle treated animals.	99
Figure 4.10 The effect of PCP on locomotor activity after receiving PCP during a perseverance or learned irrelevance task.	100
Figure 5.1 Responses made during the perseverance test for SB-242084 or vehicle treated animals.	110
Figure 5.2 Sessions to criterion during the perseverance test for SB-242084 or vehicle treated animals.....	111
Figure 5.3 Responses made during the learned irrelevance test for SB-242084 or vehicle treated animals.....	112
Figure 5.4 Sessions to criterion during the learned irrelevance test for SB-242084 or vehicle treated animals.....	113
Figure 5.5 Responses made during the novelty bias test for SB-242084 or vehicle treated animals.	114
Figure 5.6 Latency measures during the perseverance and learned irrelevance test (collated) for SB-242084 or vehicle treated animals.	117
Figure 6.1 The effects of PCP-isolation or vehicle group-housing on (A) ambulatory activity and (B) rears during the locomotor activity test (PND 59).	127
Figure 6.2 Novel object discrimination in PCP-isolates and group-housed animals.	128
Figure 6.3 Responses made during the perseverance test for PCP-isolates or controls.....	131
Figure 6.4 Sessions to criterion during the perseverance test for PCP-isolates or controls.	132
Figure 6.5 Responses made during the learned irrelevance test for PCP-isolates or controls.....	133
Figure 6.6 Sessions to criterion during the learned irrelevance test for PCP-isolates or controls.....	134
Figure 6.7 Responses made during the novelty bias test for PCP-isolates or controls.....	135
Figure 6.8 Responses made during the extinction test for PCP-isolates or controls.....	136

Figure 6.9 Latency measures during the perseverance and learned irrelevance test (collated) for PCP-isolates or controls.	138
Figure 7.1 Approximate slice locations for monoamine analysis after 6- OHDA or vehicle infusion.	147
Figure 7.2 Response and latency measures before and after surgery in 6- OHDA or vehicle infused animals.....	149
Figure 7.3 Responses made during the perseverance test for 6-OHDA or vehicle infused animals.....	150
Figure 7.4 Sessions to criterion during the perseverance test for 6-OHDA or vehicle infused animals.....	151
Figure 7.5 Responses made during the learned irrelevance test for 6-OHDA or vehicle infused animals.	152
Figure 7.6 Sessions to criterion during the learned irrelevance test for 6- OHDA or vehicle infused animals.....	153
Figure 7.7 Responses made during the novelty bias test for 6-OHDA or vehicle infused animals.....	154
Figure 7.8 Responses made during the extinction test for 6-OHDA or vehicle infused animals.	155
Figure 7.9 Latency measures during the perseverance and learned irrelevance test (collated) for 6-OHDA or vehicle infused animals.	157
Figure 7.10 Locomotor activity for 6-OHDA or vehicle infused animals.	158
Figure 7.11 Neurochemistry for 6-OHDA or vehicle infused animals.	160
Figure 7.12 Neurochemical-behaviour correlates for 6-OHDA or vehicle infused animals.	162
Figure 8.1 Response and latency measures before and after surgery in 5,7- DHT or vehicle infused animals.	169
Figure 8.2 Responses made during the perseverance test for 5,7-DHT or vehicle infused animals.....	170
Figure 8.3 Sessions to criterion during the perseverance test for 5,7-DHT or vehicle infused animals.....	171
Figure 8.4 Responses made during the learned irrelevance test for 5,7-DHT or vehicle infused animals.	172
Figure 8.5 Sessions to criterion during the perseverance test for 5,7-DHT or vehicle infused animals.....	173
Figure 8.6 Responses made during the extinction test for 5,7-DHT or vehicle infused animals.	174
Figure 8.7 Latency measures during the perseverance and learned irrelevance test (collated) for 5,7-DHT or vehicle infused animals.....	176
Figure 8.8 Neurochemistry for 5,7-DHT or vehicle infused animals.....	178
Figure 8.9 Neurochemical-behaviour correlates for 5,7-DHT or vehicle infused animals.	179

List of Tables

Table 1.1 Comparison of the model-based and model-free decision making systems.....	11
Table 1.2 Tasks affected by corticostriatal disruption.	14
Table 1.3 A non-exhaustive list of reversal learning task parameters that affect behavioural strategy and therefore the circuitry utilised.....	43
Table 5.1 Effects of the selective 5-HT _{2c} R antagonist SB-242084 in rodent reversal learning tasks.	107
Table 6.1 A timeline of behavioural procedures in PCP-isolates and group-housed animals.....	123
Table 6.2. Discrimination measures for PCP-isolates and group-housed animals.	129
Table 7.1 Previously published levels of monoamines for OFC, mPFC and dorsal striatum in post-mortem rat brain tissue.	161
Table 9.1 A summary of the experimental findings of this thesis.	184
Table 9.2 Hypothetical parameters for two animals in the split reversal task.	190

Acknowledgements

I'd firstly like to thank my supervisors Kevin Fone and Paula Moran for giving me the opportunity to do a PhD and for their kind support throughout the process, along with the University of Nottingham for funding my position.

I'd also like to thank Ian Topham, Dave Watson, Karen Swift, and Clare Spicer for their valuable technical assistance.

I must also thank all the family and friends who've supported me throughout the years.

Finally, I'd like to thank all the people who've inspired me, and all of the giants whose shoulders we stand on today in order to better understand – and improve the circumstances of – the world.

Publications

Book chapters

Rafter M, Fone KCF, Moran PM. (2016). Chapter 10 – Glutamate Pharmacological Models Relevant to Schizophrenia and Psychosis: Can a Receptor Occupancy Normalization Approach Reduce the Gap between Animal and Human Experiments? In Pletnikov MV & Waddington JL (Ed.), Handbook of Behavioral Neuroscience (Vol. 23, pp. 139–174). Elsevier.

Conference abstracts

Rafter MD, Moran PM, Fone KCF. (2015). Selectively altered associative learning in the dual-hit post-natal PCP/social isolation rat model of schizophrenia. *Journal of Psychopharmacology* S29 A72.

Rafter MD, Moran PM, Fone KCF. (2014). Effect of phencyclidine on perseverance and learned irrelevance in a touchscreen reversal learning task in rats. *Journal of Psychopharmacology* S28 A113.

Abbreviations

5-HIAA	5-hydroxyindoleacetic acid
5-HT	5-hydroxytryptamine (serotonin)
5,7-DHT	5,7-dihydroxytryptamine
6-OHDA	6-hydroxydopamine
ANOVA	Analysis of variance
A-O	Action-outcome (association)
CS+	Conditioned reinforced (rewarded) stimulus
CS-	Conditioned non-reinforced (unrewarded) stimulus
DOPAC	3,4-dihydroxyphenylacetic acid
DMS / DLS	Dorsomedial / dorsolateral striatum
GABA	Gamma-aminobutyric acid
HPLC	High performance liquid chromatography
i.p.	Intraperitoneal
NMDA	N-methyl-D-aspartate
OCD	Obsessive-compulsive disorder
OFC	Orbitofrontal cortex
p(x)	Probability of x
PCP	Phencyclidine
PFC	Prefrontal cortex
RM-ANOVA	Repeated measures analysis of variance
RPE / PE	Reward prediction error
S+	Current reinforced (rewarded) stimulus
S-	Current non-reinforced (unrewarded) stimulus
S-R	Stimulus-response (association)
s.c.	Subcutaneous
SMC	Sensorimotor cortex
SSRIs	Selective serotonin reuptake inhibitors
VTA	Ventral tegmental area

Chapter 1 – General

Introduction

1.1 Introduction

In every moment of every day, people can – and frequently do – make decisions that produce undesirable outcomes for their future selves. Among the common everyday examples are: cigarette smoking, eating junk food, and procrastination. For a large proportion of the population, these behaviours are in check to the point that they can live meaningful lives; however for some, these behaviours become disruptive enough (“compulsive”) that these people are unable to live in a way that they wish, and thus they seek treatment. Despite differences in diagnoses for these people (e.g. addiction, obsessive-compulsive disorder, schizophrenia), compulsive behaviours are a common trans-diagnostic trait (Gillan et al., 2016a; Robbins et al., 2012) and are often resistant to existing drug treatment (Grant et al., 2016).

This thesis aims to explore the neural substrates of mammalian decision making in a way that may aid in the discovery of treatments for compulsivity. In this introductory chapter, the essential parameters that shape decision making in any agent capable of learning will be introduced using simplified models. Then the corresponding Supraprimate (rodent and primate) neural substrates of these parameters will be described, followed by how these parameters and substrates are dysfunctional in psychiatric

disorders. I will then describe how reversal learning in rodents can be used to specifically assess the functioning of this circuitry. The subsequent chapters report experiments using rodent reversal learning tasks with various manipulations considered to affect this circuitry, followed by a discussion and synthesis of the thesis.

1.2 Decision making and reinforcement learning

In order to understand how decision making goes awry in psychiatric disorders, one must first understand and define what “normal” or optimal decision making is. In the contemporary neuroscience literature, it is generally accepted that information is received by the central nervous system (CNS), neural computations occur, a motor action (or inaction) is performed, and a change in the state of the world is the result. Meanwhile, the neural computations that occur will differ between individuals based on their current brain state, which in turn is affected by prior conditioning, genetics, and a host of other environmental factors. Therefore, when attempting to understand aberrant decision making, we must be careful when assigning causation of the aberration to a single process or factor ($X_1 \rightarrow Y$), and instead try to factor in as many causes of variance as possible ($X_{1-n} \rightarrow Y$), which can be achieved via modelling.

1.2.1 Modelling reinforcement learning

Models can represent processes in a simplified way by including the factors that are thought to produce the most variance, whilst disregarding the factors that produce minor variance (deemed “noise”). In the case of

decision making, we are most interested in what factors most heavily influence the decision.

In the chain of events described earlier (information – computation – action), the neural computation is the hidden variable; this can be inferred by applying different models/parameters to a dataset where all the other variables known.

With decision making, the simplest models can be applied to datasets obtained in a two-choice task with a binary outcome (e.g. reward/no reward). An ethological example of such a task would be an animal deciding each day which tree to climb in order to obtain fruit. Intuitively, the probability that an option (tree) is chosen will be proportional to the estimated amount of reward value (fruit) it produces compared to the other available option. In turn, the estimated value of the choice will be updated after each choice if there is a deviation between the actual value and prior expected value, i.e. when reward prediction errors (RPEs) occur. The degree to which the animal alters its perceived value of the option in response to the prediction error is known as the learning rate (α), and the degree of certainty in those valuations when making a decision (versus choosing randomly) is known as the exploration rate (β). These three parameters are vital components of reinforcement learning and will be referred to throughout the thesis. Succinctly, higher learning rates and larger prediction errors lead to more sensitive learning, whilst higher exploration rates lead to less exploitation of learned valuations (see Figure 1.1).

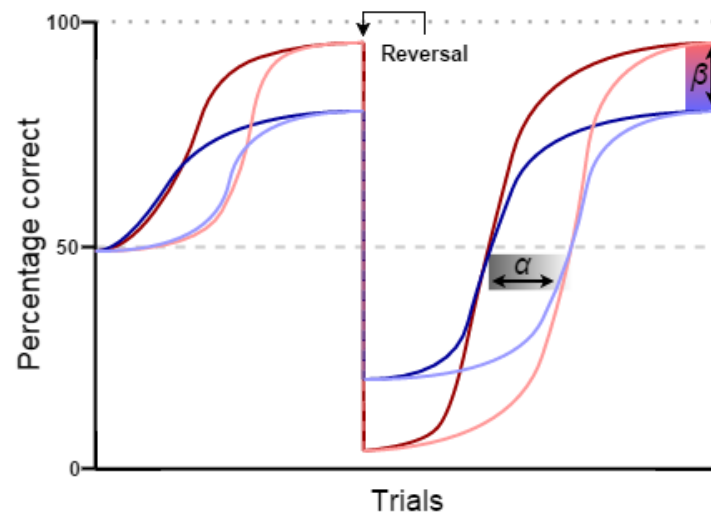


Figure 1.1 Hypothetical performance of four agents in a two-choice task where only one option is rewarded (correct).

In early trials, all agents perform at roughly 50% correct (grey dashed line) as they sample the two options randomly. After numerous trials, they select the correct option more often. Then a reversal occurs, whereby the reward contingencies of the two options are switched. Initially, agents continue choosing based on the learned values of the options. An accumulation of prediction errors leads to re-evaluation of the choices and the agents begin to choose the now-rewarded option more. The four agents differ in their choices based on the learning rate (α) and exploration rate (β) parameters. Darker lines represent agents with a higher learning rate, whilst bluer lines represent agents with a higher exploration rate. An ethological example of this abstract task would be an animal choosing between two fruit trees. Initially one tree bears fruit whilst the other doesn't. After extended pillaging, the "correct" tree becomes barren, at which point, the other tree is now bearing fruit.

1.2.2 Multiple reinforcement learning systems

Simple reinforcement learning equations using the three parameters above follow a rule where learning can only occur based on real, externally experienced prediction errors. Thus, simple, linear associations between choices and outcome values are learned. This is often referred to as "model-

free” learning. Of course animals are able to infer value without having to externally experience an outcome, and these inferred values can override learned values to control behaviour (termed “model-based” learning). An obvious example of this is when a human is explicitly told that an association has been changed (Biele et al., 2011). Other animals can also infer, e.g. by witnessing another animal being rewarded or punished for a choice, an animal can infer that it will receive a similar outcome. Similarly, in a two-choice task, an animal may develop a model of the task which states that the reward contingencies of the two choices are anticorrelated (Bromberg-Martin et al., 2010). For example, by experiencing punishment or reward omission with one choice, it is automatically inferred that the other choice produces reward (see Figure 1.2). This model would accelerate learning in environments where reward contingencies reverse (such as in Figure 1.1).

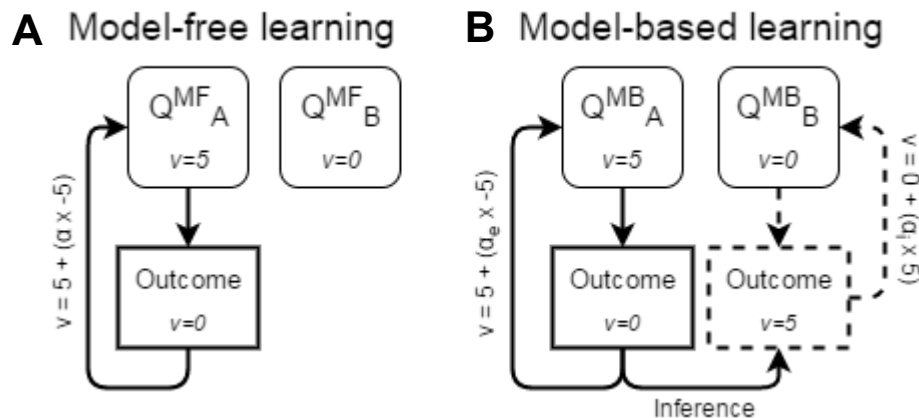


Figure 1.2 Model-free and model-based learning.

(A) A model-free learning system that updates values based on experienced outcome values. Choice A (Q^{MF}_A) has a prior estimated value of 5 and choice B (Q^{MF}_B) a prior value of 0. Q^{MF}_A is chosen and receives an outcome with value 0. The value of Q^{MF}_A is updated in response to this prediction error. (B) A model-based learning system that updates values based on experience and inference. In addition to updating Q^{MB}_A after a prediction error, it is inferred that choice of Q^{MB}_B would have led to a prediction error (dashed lines), and so the value of Q^{MB}_B is updated. This model-based system has two learning rates, α_e which is for experienced (real) outcomes, and α_i for inferred outcomes.

The situation in Figure 1.2 shows separate model-based and model-free valuation systems. However most Supraprimate brains are capable of both model-free and model-based learning (discussed later), which provides a problem as the choice selection system can only use one set of values. This can be solved by having a meta-valuation system that decides which system's values to use by averaging each system's values according to a confidence weight (Lee et al., 2014), and introduces another parameter: the balance between model-based and model-free control (ω).

In summary, decision making in a multi-trial two-choice task can be adequately modelled by a generic agent containing: a model-based and a model-free option valuation system, an option selection system, and a meta-valuation system that weighs the confidence of the valuation systems. In the following section, I will discuss the evidence for these systems in Supraprimate brains, and how specific disruptions within them are proposed to cause compulsive behaviour.

1.3 Dual system decision making in animals

1.3.1 Behavioural evidence for a dual system

Initial studies of reinforcement learning showed that animals placed into testing apparatus (e.g. mazes, operant boxes, traps) were capable of finding a reward or escape, and in subsequent exposures, would perform this feat more quickly, indicating learning (reviewed in Dolan and Dayan, 2013). Early theorists proposed that the animals solved these tasks by performing chains of rewarded stimulus-response (S-R) associations, e.g. learning that location Z was associated with reward, and then that location Y was associated with Z, location X with Y, etc. However later studies using more complex designs showed forms of latent learning that disproved this, e.g. rats that learned to avoid pressing a lever that delivered an aversive salt solution would later press that lever when in a salt-deprived state (Kriekhaus and Wolf, 1968). This latter finding implies that the animals had learned action-outcome (A-O) and outcome-value (O-V) associations, and could flexibly use these associations to achieve goals in novel contexts/states.

Since then, it has been shown that the type of learning (S-R or A-O) that predominates is dependent on task conditions (discussed later, see Yin and Knowlton, 2006). These two forms of learning are termed habitual and goal-directed, respectively, and are proposed to be mediated by separate – but connected – neural systems. Succinctly, the goal-directed system chooses actions based on the belief that they will result in a currently desirable (valuable) outcome. The habitual system chooses actions based on past reinforcement, regardless of the *current* value of the paired outcome.

The goal-directed system is active during early learning and novel situations, and deliberately holds A-O associations online in working memory, choosing responses based on those which are calculated to produce the most valuable outcome. The habitual system dominates behaviour later in learning, and can perform responses rapidly based on previously reinforced S-R associations, regardless of currently desirable outcomes (Yin and Knowlton, 2006). The goal-directed system is always model-based, as current internal states have to be monitored to calculate the most desirable outcome (goal). Meanwhile the habitual system is suggested to be a model-free system that is trained by the model-based system, i.e. a model-based decision tree which calculates action values on-the-fly is converted into a pruned tree containing cached action values based on previous reinforcement (Dolan and Dayan, 2013). As a result, the learning rate of the habitual system is slow, whereas the goal-directed system can contain multiple models and therefore flexible learning rates (see Figure 1.2).

What determines which system dominates behaviour, and how? Task experience and parameters such as reinforcement schedules predict how quickly the transition to model-free behaviour occurs (Yin and Knowlton, 2006; also see Table 1.3). As mentioned in the previous section, a meta-valuation system must arbitrate between the two systems. This system likely arbitrates based on:

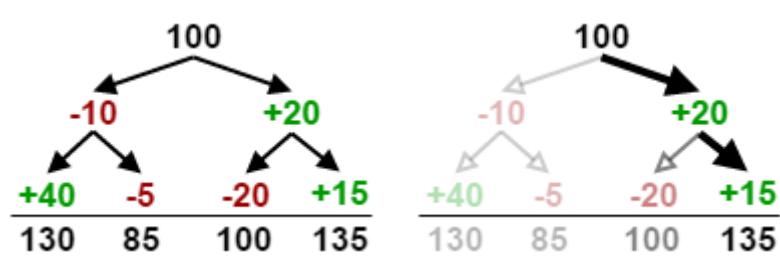
- The difference in uncertainty between the two systems
 - If the model-free system is able to make a high certainty choice in a short period of time (which often occurs in familiar situations), that choice will be taken. Conversely, if the model-free system's choice is uncertain, the model-based system is afforded more time to compute a more certain choice.
- The average reward rate of the current state the individual is in
 - Higher tonic reward favours model-free behaviour (Niv et al., 2007), because in these states the opportunity cost of engaging the (computationally slow) model-based system is higher.
- The “controllability” of the state (the difference in value between activity and passivity, see Boureau et al., 2015 for an elaborated explanation).

It should be noted that another theory proposes that rather than there being two distinct systems, a gradient of model-based systems nest model-free systems, and the value of a choice can be modulated given information received from higher systems (see Koechlin and Summerfield, 2007). For

example, the lowest level system chooses an action given the sensory/motor information available, akin to stimulus-response. The next level chooses an action given sensory information and the context, i.e. context-response. Higher levels then add in factors from working memory such as the current episode, imagination of future outcomes, etc. In this framework, the meta-valuation system engages increasingly higher systems based on the arbitrators listed above.

In summary (see Table 1.1), a model-based (goal-directed) system is active during early learning, building and simulating a model of the environment to select choices that will achieve desired outcomes given the agent's current goals. This model-based system trains a model-free (habitual) system to contain a pruned cognitive map of the most valued (reinforced) choices. This model-free system dominates behaviour in later learning and inflexibly chooses the historically reinforced choices without calculating the current value of that choice's outcome.

Table 1.1 Comparison of the model-based and model-free decision making systems.

	Model-based system	Model-free system
Behaviour	Goal-directed	Habitual
Learning rate	Rapid	Slow
Exploration rate	Flexible	Low
Tree-search of task space – arrows represent choices; numbers represent outcome value of the choice (Huys et al., 2012).		
Computational cost	Expensive	Tractable
- Temporally	Slow	Fast
- Metabolically	High	?
Associations formed	Action-outcome	Stimulus-response
Association awareness	Explicit / conscious	Implicit / unconscious
Outcome valuations	Computed online	Cached
Contexts used in	Novel, unfamiliar, early learning	Habitual, familiar, late learning
Circuitry	Associative loop (PFC-DMS)	Sensorimotor loop (SMC-DLS)

1.3.2 Neural evidence for a dual system

Dopamine-dependent prediction errors

For decades, it has been known that dopamine-releasing neurons in the midbrain (ventral tegmental area, VTA; substantia nigra pars compacta, SNc) change their firing rate in response to prediction errors. Specifically, they increase their firing rate when an unexpected reward is received, and decrease their firing rate when an expected reward is omitted, signalling positive and negative reward prediction errors (RPEs), respectively (Schultz

et al., 1997). Moreover, recent evidence has shown that these neurons are directly responsible for the RPE calculation (Eshel et al., 2015) and are causally involved in RPE-dependent learning (Chang et al., 2016; Steinberg et al., 2013).

Cortico-striatal loops

Midbrain dopamine neurons project to multiple brain regions, but most prominently to the striatum (Björklund and Dunnett, 2007). The striatum sits in between the cortex and basal ganglia to form cortico-striatal-thalamo-cortical loops, with different regions of striatum connecting to different cortical areas (see Haber, 2003). The dorsal striatum is particularly important for instrumental learning (such as the exemplar two-choice task of the previous section; Yin et al., 2008). In rodents, the dorsomedial striatum (DMS) connects to regions of prefrontal cortex as part of the associative loop (aka cognitive or prefrontal loop), whereas the dorsolateral striatum (DLS) connects to sensorimotor cortex (SMC) as part of the sensorimotor loop (Yin and Knowlton, 2006). In primates, the caudate nucleus corresponds to the DMS, and the putamen corresponds to the DLS (Haber and Knutson, 2009). In addition, there is a limbic corticostriatal loop (aka affective loop) which involves the orbitofrontal cortex (OFC), amygdala and ventral striatum.

One proposal is that – via these feed-forward loops (see Figure 1.3) – the ventral striatum selects sets of actions (e.g. eat food or work on thesis) based on global motivational goals (e.g. satisfy hunger or graduate); the DMS selects abstract actions (e.g. type sentence A or B) based on current external

cues (e.g. what sentence precedes it); and the DLS selects concrete chains of motor actions (e.g. press “k-e-y” on keyboard) based on sensorimotor feedback (Ito and Doya, 2011).

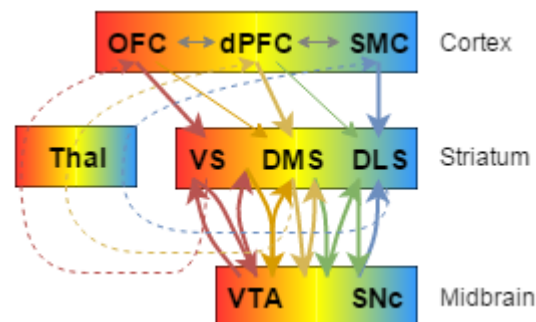


Figure 1.3 The organisation of feed-forward spiralling cortico-striato-thalamo-cortical loops.

As described by (Haber and Knutson, 2009), the limbic loop (red) is able to engage the associative loop (yellow) to invigorate goal-driven behaviour. The associative loop computes optimal abstract choices and engages the sensorimotor loop (blue) to perform concrete motor actions. DLS: dorsolateral striatum. DMS: dorsomedial striatum. dPFC: dorsal prefrontal cortex. OFC: orbitofrontal cortex. SMC: sensorimotor cortex. SNc: substantia nigra pars compacta. Thal: thalamus. VS: ventral striatum. VTA: ventral tegmental area.

In terms of decision making, the limbic loop is responsible for invigorating action selection and determining the urgency at which actions are performed. The associative loop is responsible for goal-directed (model-based) learning, whilst the sensorimotor loop is responsible for habitual (model-free) learning. Consequently, disruption to nuclei in the associative loop leads to an impairment in goal-directed behaviour and disruption to nuclei in the sensorimotor loop leads to an impairment in habitual behaviour (see Table 1.2 for examples). Likewise, greater anatomical and

functional connectivity between regions of the associative loop leads to a greater bias towards goal-directed behaviour, and greater connectivity between sensorimotor loop regions leads to a greater bias (ω parameter) towards habitual behaviour (Morris et al., 2015; Voon et al., 2014; Wit et al., 2012).

Table 1.2 Tasks affected by corticostriatal disruption.

A summary of tasks that differentiate between goal-directed and habitual learning and reviews of the evidence that cortico-striatal disruption affects them (see also Figure 1.8).

Task	Description
Outcome devaluation	The agent learns two associations, e.g. A_1-O_1 and A_2-O_2 , where outcomes 1 and 2 are of equal value and so actions 1 and 2 are chosen equally. One of the outcomes is then devalued either via selective satiation or pairing with illness. The agent is then re-exposed to the context where A_1 and A_2 are available to choose (with no outcome delivered to avoid new learning). The agent's bias towards the action associated with the non-devalued outcome represents the balance of goal-directed learning.
Contingency degradation	As above, the agent learns two associations. One of the A-O contingencies is then degraded by also presenting the outcome in the absence of the action, thus the action is no longer necessary. The rate at which the agent stops choosing the devalued action represents the balance of goal-directed learning.
Reversal learning	As above, the agent learns two associations, but outcomes 1 and 2 are not of equal value, so one action is chosen more frequently than the other. The associations are then reversed, i.e. A_1-O_2 and A_2-O_1 . The rate at which the agent stops choosing the formerly rewarded action represents the balance of goal-directed learning.

Table 1.2 Continued

Probabilistic n-step	The agent learns that choices produce different outcomes depending on task state, e.g. in state A, left = reward and right = punishment; in state B, right = reward, left = punishment. After each choice, a change of state occurs which is probabilistic, e.g. $p(A) = 90\%$, $p(B) = 10\%$. The frequency of choice errors whilst in the rare state represents the balance of habitual learning.
Review	Findings
(Rudebeck and Murray, 2014)	A review of dozens of primate studies looking at disruption of the orbitofrontal cortex using all of the tasks described above. They conclude that OFC influences choice valuation by integrating the sensory qualities of a predicted outcome with internal state information, e.g. the value of salty or sweet popcorn is influenced by internal salt or sugar levels.
(Yin et al., 2008)	A review of rodent and primate studies concluding that outcome devaluation and contingency degradation are impaired by lesions to the dorsomedial striatum, enhanced by lesions to the dorsolateral striatum, and not affected by lesions to the ventral striatum.
(Hamilton and Brigman, 2015)	A review of over 30 studies in rodents looking at the evidence that lesions/deactivation of prefrontal and orbitofrontal regions slows reversal learning and impairs devaluation.
(Fellows, 2011)	A review of studies in humans showing that damage to ventromedial PFC/medial OFC leads to impairments of model-based behaviour such as reversal learning and transitive preference (preferring A over C when $A > B$ and $B > C$).

Action valuation plasticity

What computations do these nuclei perform during decision making?

Most theories propose a “tripartite” reinforcement model where cortical glutamatergic projections (signalling actions to take) and midbrain dopaminergic projections (signalling RPEs) synapse onto striatal GABAergic medium spiny neurons that are part of the direct and indirect pathways of

the basal ganglia (also known as “go” and “nogo” pathways, see Figure 1.4; Koehlin and Summerfield, 2007; Maia and Frank, 2011; Schultz, 2015; Shiflett and Balleine, 2011). *Go* cells express dopamine D₁ receptors (D₁R) and facilitate the execution of an action, whilst *nogo* cells express dopamine D₂ receptors (D₂R) and suppress the execution of an action. Thus the likelihood of an action being executed depends on the balance of *go/nogo* cell firing that action elicits. This balance is determined by the synaptic strength of cortical projections onto these cells, which is malleable by dopamine. Briefly, dopamine spikes (induced by positive RPEs) in conjunction with glutamate release (by cortical projections) leads to long-term potentiation (LTP) of the direct pathway via NMDA-R/D₁R co-stimulation, and long-term depression (LTD) of the indirect pathway via mGluR/D₂R stimulation. Dopamine dips (induced by negative RPEs) cause an opposite pattern of plasticity (Gurney et al., 2015; Shen et al., 2008).

In accordance with this, neural activity in the striatum has been shown to correspond to action values, and manipulation of this activity via stimulation of D₁R-expressing (*go*) or D₂R-expressing (*nogo*) cells causes behavioural changes in line with increases or decreases in action value, respectively (Tai et al., 2012). It should be noted that the activity of *go/nogo* cells is also affected by other cell types in the striatum, and manipulation of these cells is sufficient to change action values, e.g. fast-spiking GABAergic interneurons (Burguière et al., 2013; Gage et al., 2010) and cholinergic interneurons (Bradfield et al., 2013; Okada et al., 2014).

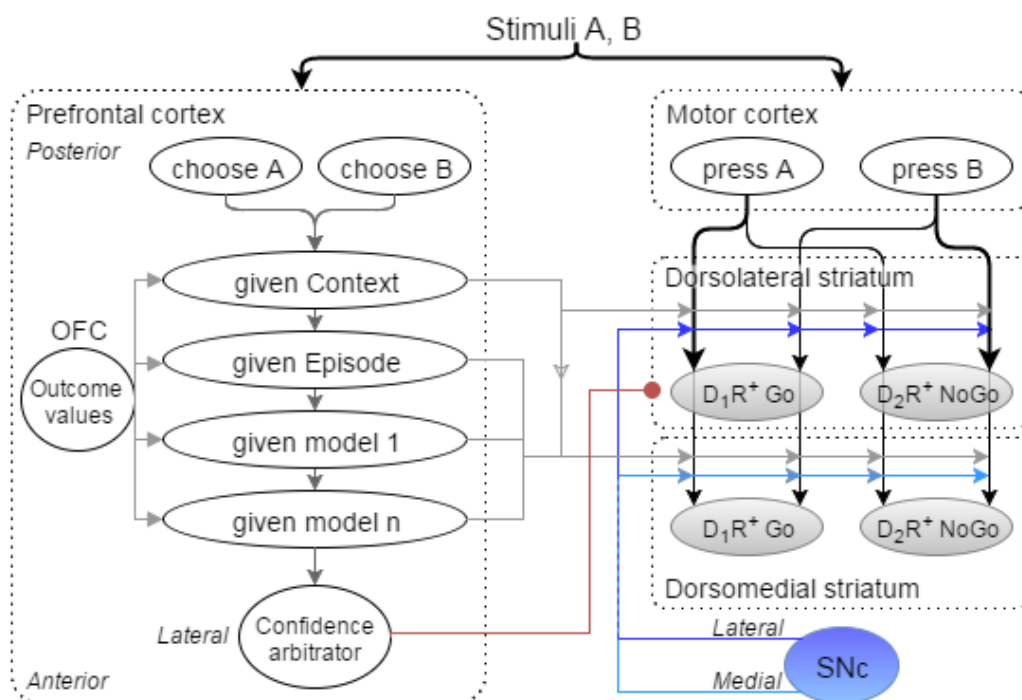


Figure 1.4 Action selection via corticostriatal plasticity.

A simplified representation of the integration of associative and sensorimotor corticostriatal circuitry at the dorsal striatum. Briefly, motor cortical ensembles representing potential motor actions project to dorsal striatum “go” and “nogo” cells, which can respectively facilitate or inhibit execution of a particular action via the basal ganglia direct or indirect pathways (see text). Positive prediction errors reported by nigrostriatal dopamine neurons (blue) reinforce selected actions by D₁R-mediated LTP of the direct pathway synapses and D₂R-mediated LTD of the indirect pathway synapses. In the DLS, this leads to long-term “cached” valuations of actions based on reinforcement history. Meanwhile prefrontal circuitry provides top-down modulation of sensorimotor corticostriatal synapses using action valuations calculated to factor in higher level representations of the environment/task. Posterior regions factor in contextual cues such as location and time; more anterior regions factor in episode (e.g. recent outcomes) and increasingly complex inferred models of the environment (Koechlin and Summerfield, 2007). Orbitofrontal cortex (OFC) appears to be particularly important for updating expected outcome values given the current internal state and inferred environmental state (Stalnaker et al., 2015). Dorsolateral PFC plays an important modulatory role on the OFC to factor in higher, long-term goals e.g. dieting (Rangel and Hare, 2010). Inferior lateral PFC can also inhibit the influence of the DLS (red) if there is

Figure 1.4 Continued

conflict in the valuations of each system (Lee et al., 2014). In this particular example, the action “press A” has been previously reinforced in this context and so has a high cached valuation by the model-free system, reflected by stronger synaptic weight onto DLS *go* cells; whereas “press B” has been negatively reinforced and so has a lower valuation, reflected by stronger synaptic weight onto DLS *nogo* cells.

Striatal learning rates

The tripartite model also adequately explains how the different corticostriatal loops can value actions differently despite the same reinforcement history by acknowledging functional heterogeneity in the striatum.

Firstly, it is suspected that the learning rate of DLS cells is lower than the DMS, resulting in this system being slower to learn the value of actions but also slower to devalue actions (i.e. less flexible; Kimchi and Laubach, 2009; Smith and Graybiel, 2013; Thorn et al., 2010). The DMS is capable of sustaining flexible action-outcome representations via persistent firing (Bissonette and Roesch, 2015; Histed et al., 2009; Kim et al., 2013; Watanabe and Hikosaka, 2005; Yanike and Ferrera, 2014), which allows for multi-trial summation. On the other hand, DLS neuronal firing rapidly decays post-outcome (Kim et al., 2013). Although the neural substrates for this are not clear, there is modest evidence for a dorsolateral-ventromedial gradient of dopamine transporter expression in the striatum – and correspondingly – RPE decay kinetics (Wickens et al., 2007; Zhang et al., 2009), since the dopamine transporter is responsible for rapid reuptake of dopamine from the synapse (Schmitz et al., 2003). Rapid decay of RPE signals in the DLS

would mean that only action signals (from sensorimotor cortex) immediately preceding the RPE are reinforced/extinguished. Whereas slower decay of RPE signals in the DMS/ventral striatum would lead to RPE accumulation (see Figure 1.5), and thus a higher – more recency-biased – learning rate. It is also known that D₂R expression is higher in the DLS and that DLS-dependent (habitual) actions are disrupted by D₂R antagonism but not D₁R antagonism (Yin et al., 2009). However the relevance of this for learning rate is unclear, since D₂Rs are suspected to be more important for learning from negative RPEs (Frank et al., 2007).

Besides dopamine, differences in the flexibility of firing between DLS and DMS cells may be due to cortical inputs, in particular the orbitofrontal cortex, which will be discussed in more detail later (see section 1.5.2 Role of OFC).

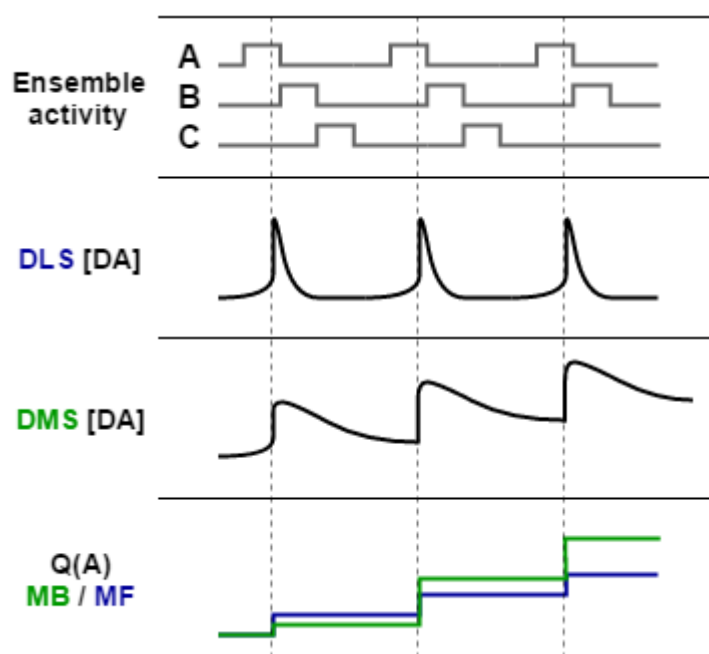


Figure 1.5 Prediction error dynamics and learning.

Hypothetical role of dopamine release kinetics in model-based and model-free valuation of an action (Wickens et al., 2007; see text). Cortical ensembles representing actions A, B, and C lead to execution of those actions at different time points. When action A is performed, a reward is delivered (dashed line), and a positive prediction error occurs inducing dopamine release in the DLS and DMS. Rapid-decaying dopamine spikes in the DLS lead to fixed increments in the valuation (Q) of action A by the model-free system (blue, reflecting SMC-DLS synaptic weights). Slow-decaying dopamine spikes in the DMS lead to cumulative increments in the valuation of action A by the model-based system (green, reflecting PFC-DMS synaptic weights).

Summary

Overall, this simplified model suggests that:

- Motor cortex firing reflects action availability/affordance.
- The summed synaptic strength between motor cortical ensembles and DLS *go/nogo* cells reflects an action's valuation by the model-free system.

- The firing of prefrontal cortical ensembles projecting to DMS *go/nogo* cells reflects an action's valuation by the model-based system.
- The balance or confidence between DLS and DMS activity determines the weighting of each system's valuations towards the final valuation (ω).
- The action with the highest valuation (striatal *go/nogo* firing ratio) wins the action selection arbitration.
- Phasic dopamine changes reflect reward prediction error.
- Tonic striatal dopamine levels (inversely) reflect the exploration parameter (β ; Humphries et al., 2012).
- Dopamine receptor expression and reuptake kinetics contribute to learning rate (α ; caveat below).

What is unclear in this model is how the model-based and model-free valuations are integrated, especially during times of conflict. The synaptic weights reflect valuations but not confidence of those valuations, which may be determined by cholinergic interneuron modulation of *go/nogo* population firing entropy (see Franklin and Frank, 2015; Thorn et al., 2010). It has been proposed that lateral inferior PFC arbitrates based on confidence and correspondingly inhibits DLS to favour model-based valuations in the DMS (i.e. calculates the ω parameter, see Lee et al., 2014) although how the lateral PFC receives confidence estimates from both systems is unclear.

In addition, this model doesn't provide a clear substrate for learning rate, since although multiple interventions affect learning rate

behaviourally, it is difficult to disentangle learning rate from prediction error amplitude. It is tempting to assume that learning rate is determined solely by synaptic dopamine factors, and that DMS and DLS learning rates differ because of this. However this is not the case since firstly, the RPE amplitude reported by dopamine neurons is not fixed (Eshel et al., 2016); secondly, a host of other neurotransmitters modulate dopamine-dependent striatal plasticity, e.g. histamine (Bolam and Ellender, 2016), serotonin (Mathur et al., 2011), adenosine (Shen et al., 2008), endocannabinoids (Gerdeman et al., 2002), and opioids (Atwood et al., 2014); and finally, behavioural markers of learning often arise too rapidly to be accounted for purely by synaptic plasticity (discussed in Gurney et al., 2015); therefore learning rate is likely complex and variable depending on a multitude of conditions and computations from other regions.

1.4 Disruption of dual system balance in psychiatric disorders

In this section, I will introduce psychiatric disorders that feature compulsive symptoms and propose how the decision making systems discussed so far are affected in them.

1.4.1 Obsessive-compulsive disorder

Aetiology and treatment

Obsessive-compulsive disorder (OCD) affects over 2% of the population and is characterised by intrusive thoughts (obsessions), anxiety, and compulsions aimed at alleviating these obsessions (Pauls et al., 2014). Obsession-compulsions commonly fall into one of four pairings: symmetry-ordering, taboo thoughts-checking, contamination-cleaning, or hoarding.

The current first-line treatment for OCD is behavioural therapy and selective serotonin reuptake inhibitors (SSRIs). SSRIs alone substantially improve symptoms in ~50% of patients (Fineberg et al., 2012), whilst therapy tends to improve symptoms in ~60% of patients, and combined treatment has ~70-80% success rate (Romanelli et al., 2014). In treatment-resistant patients, deep-brain stimulation of striatal or thalamic regions has been trialled with mixed success (Kohl et al., 2014).

Although the causes of OCD have not been established, it is known to have a strong genetic component. Family studies show that first-degree relatives of someone with OCD are five times more likely to have OCD, meanwhile twin studies show that genetic variance accounts for 40% of phenotypic variance, whilst shared (family) environmental factors had no effect (reviewed in Pauls et al., 2014). However genetic linkage studies have failed to find a single gene that significantly increase the risk of OCD, which suggests that multiple genes of low effect size contribute towards the development of OCD. A few candidate gene studies (Pauls et al., 2014; Zohar et al., 2012) implicate genes involved in monoamine and glutamate signalling e.g. COMT (catechol-O-methyltransferase) and 5-HTTLPR (5-HT transporter linked polymorphic region) however these genes have also been implicated in a host of other disorders/traits and have a notoriously low replication rate (Flint and Munafò, 2013; Green et al., 2008). Meanwhile genome-wide association studies implicate genes involved in synaptic development, scaffolding and glutamate signalling (Pauls et al., 2014),

although these studies were not sufficiently powered to detect genome-wide significance.

Mechanisms

Given that SSRIs improve symptoms in half of all OCD patients, it has been suggested that dysfunctional serotonergic circuitry contributes to compulsions. Several studies have shown that agonists of the 5-HT_{1B} receptor (such as *meta*-chlorophenylpiperazine and sumatriptan) are able to exacerbate compulsive and anxious symptoms in OCD (although other studies found no effect, see Zohar et al., 2012). Additionally, dopamine releasing agents (such as methylphenidate and amphetamine) can exacerbate OCD symptoms, but also improve symptoms in others (reviewed in Zohar et al., 2012).

OCD patients show impairments on numerous tests of cognitive flexibility, with a predominant consensus that inhibition of habitual actions is impaired (reviewed in Gillan et al., 2016b; Gruner et al., 2015). Likewise, OCD patients display abnormal anatomy and connectivity of corticostriatal loops (reviewed in Burguière et al., 2015; Menzies et al., 2008). Some have proposed that OCD symptoms can be explained by hyperactivity in OFC-DMS circuitry (Burguière et al., 2015), which may be a result of abnormal engagement by the lateral PFC meta-decision making system (Gruner et al., 2015).

The evidence for this is wide-spanning. For example, OFC and striatal hyperactivity is reduced after behavioural and pharmacological treatment, and this reduction strongly correlates with symptom improvement

(reviewed in Pauls et al., 2014). Secondly, OFC-dependent goal-directed behaviours such as outcome devaluation (Gillan et al., 2011, 2014a) and inference of future regret (Gillan et al., 2014b) are impaired, despite the later actual experience of regret being enhanced. Thirdly, DMS hyperactivity is associated with the development of – and urge to perform – avoidance habits that are no longer linked to their outcome (Gillan et al., 2014c). Fourthly, OCD patients show enhanced engagement of OFC circuitry during implicit learning tasks which rely on model-free circuitry in neurotypical participants (reviewed in Gruner et al., 2015), whereas in a task that measures both model-based and model-free behaviour, people with OCD (or other compulsive disorders) display a bias towards model-free behaviour (smaller ω parameter; Voon et al., 2014). In healthy individuals, bias towards model-based learning positively correlates with left medial OFC volume and mOFC-DMS white matter connectivity (Wit et al., 2012). Finally, OCD patients display lesser ventral striatal BOLD activity during reward anticipation, which can be ameliorated with ventral striatal deep-brain stimulation (Figeet et al., 2013), a treatment that reduces compulsive symptoms by ~50% in approximately half of treatment-resistant patients (Kohl et al., 2014).

These findings fit with theories proposing that the OFC assists goal-directed learning by integrating internal and external information to infer new states in the world, and adjust action selection accordingly (detailed in 1.5.2 Role of OFC). Succinctly, non-OCD sufferers are able to learn that during certain states (S), an action (A) leads to the avoidance of an outcome

(O) and an unobservable change of state. For example, after handling raw chicken (S_1), hand washing (A_1) leads to the avoidance of *Salmonella* infection (O_1) and the unobservable state of cleanliness (S_2), where O_1 -avoidance is not contingent upon A_1 . Medial OFC is proposed to generate the representation of S_2 , in which the inhibitory A_1 - O_1 association is absent (Bradfield et al., 2015). In OCD, OFC dysfunction may lower the confidence of this representation, or may disrupt shifting from S_1 to S_2 (Maia and Cano-Colino, 2015; Sharpe et al., 2015), and therefore if the A_1 - O_1 association has been habitised, it will dominate action selection (become compulsive).

In summary, OCD patients are impaired at updating associations that have become habitual, especially if those associations are linked to unobservable, inferred outcomes. This is likely due to impaired inference of the new state in which the habitual associations no longer apply, which is OFC-dependent. As stated earlier, genetic and pharmacological studies suggest that abnormal glutamate and/or 5-HT signalling could be a cause of OFC-striatal dysfunction (Pauls et al., 2014). Both of these neurotransmitters are important for flexible maintenance of information in the OFC (Maia and Cano-Colino, 2015), which is discussed in more detail later.

1.4.2 Schizophrenia

Aetiology and treatment

Schizophrenia affects ~0.7% of the population and is characterised by: psychotic (positive) symptoms, such as hallucinations and delusions; negative symptoms, such as lack of motivation and social withdrawal; and

cognitive symptoms, such as impaired memory and attention (van Os and Kapur, 2009). Compulsive behaviour presents in numerous forms in schizophrenia, from the stubbornness of delusions, to repetitive behaviour, to motivational anhedonia. Additionally, 12% of people diagnosed with schizophrenia fulfil the criteria for OCD (Swets et al., 2014).

The current first-line treatments for schizophrenia are first-generation antipsychotics, which primarily antagonise dopamine D₂ receptors, and second-generation antipsychotics, which antagonise D₂Rs and serotonin 5-HT_{2A} receptors. Antipsychotics improve positive symptoms in roughly two thirds of patients (Barnes, 2011), but produce minimal improvements in negative and cognitive symptoms. Drugs targeting the glutamate and GABA systems have been trialled for schizophrenia on the premise that the disorder arises due to an imbalance of their signalling (discussed later, Stephan et al., 2009), but these have all failed so far (Beck et al., 2016).

There is growing evidence that schizophrenia is caused by a combination of genetic and environmental factors. Twin studies yield heritability estimates of ~80% (Cardno and Gottesman, 2000), and a recent large-scale study including ~38,000 patients found 108 genes associated with schizophrenia at genome-wide significance (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Many of these genes are linked to synapse development and glutamate signalling (Friston et al., 2016). Environmental factors are also associated with schizophrenia, such as prenatal stressors, urbanicity, immigration, and cannabis use (van Os and Kapur, 2009). The causal role of these factors is unclear since many of them

correlate, so causality in one factor confounds all the others. For example, dozens of studies have linked vitamin D levels with schizophrenia, but vitamin D is also associated with diet, ethnicity, and sun exposure, which in turn is associated with sedentary lifestyle, urban living, social withdrawal, and socioeconomic status. Aptly, a recent Mendelian randomization analysis in over 34,000 patients showed that vitamin D is not causally related to schizophrenia (Taylor et al., 2016).

Mechanisms

One of the most established findings in schizophrenia is that of enhanced presynaptic striatal dopamine function (Howes and Kapur, 2009), which leads to aberrant RPE signalling. This aberrant RPE signalling causes surprise at predictable outcomes, which instils salience/meaning to irrelevant or coincidental but unrelated events, and a state of psychosis (Kapur, 2003). Delusions may result in order to explain these aberrant associations, and have been shown to correlate with aberrant RPE signalling (Corlett and Fletcher, 2015). Increased tonic striatal dopamine (Abi-Dargham et al., 2000) may also lead to paradoxical backfire effects where contradictory evidence actually strengthens a delusion, due to intrinsic recall of the delusion in conjunction with dopamine hyperactivity (leading to enhanced consolidation; Fineberg and Corlett, 2016). In accordance with this, D₂R antagonists are effective at reducing psychosis and the salience of delusions by restoring *nogo* cell signalling, which weakens irrelevant associations (Morrison and Murray, 2009) and prevents irrelevant stimuli from being gated into working memory and conscious attention systems (Frank, 2008).

Related to delusionality is a jumping-to-conclusions bias in schizophrenia. This bias is commonly tested with sampling tasks, where coloured beads in an urn are sampled and a decision made as to which colour is most predominant. People with schizophrenia are much more likely to guess the predominant colour early on, sometimes after one sample (Averbeck et al., 2011). Jumping to conclusions may result from overweighting of current sensory evidence compared to prior evidence (Speechley et al., 2010) or a reduced decision threshold (Averbeck et al., 2011), and there is evidence for and against both.

Related to this is the prominent finding in schizophrenia of impaired cognitive flexibility, which is often attributed to increased perseveration. Early studies on the Wisconsin Card Sorting Test (WCST) mostly analysed attentional set shifting errors, which correlate with negative but not positive symptoms (Nieuwenstein et al., 2001). However later analyses show that schizophrenics also make more non-perseverative errors (Li, 2004), and more errors prior to contingency shifts (Prentice et al., 2008; Vogel et al., 2013). The reasons for these errors may be due to either: overconfidence of valuations despite negative feedback (in line with the jumping-to-conclusions bias), an inability to maintain feedback in working memory (Collins et al., 2014; Hartman et al., 2003), decreased uncertainty-induced exploration (Strauss et al., 2011), or lack of motivation to perform the task effortfully (discussed below). Given that schizophrenia is a heterogeneous disorder, all of these factors likely play a role, and this may explain why no

single reinforcement learning parameter correlates with cognitive inflexibility in schizophrenia (Reddy et al., 2016).

These discrepant findings can also be partially explained by the different neuropathologies occurring during state vs. trait symptoms in schizophrenia. State symptoms such as psychosis are acute and respond to antipsychotics. Trait symptoms such as the negative and cognitive symptoms are chronic and do not respond to antipsychotics. One account proposes that trait symptoms result from a decrease in the confidence of prior evidence (i.e. internal models of the world, including option valuations), which may be due to high-level cortical noise via NMDA receptor hypofunction. Meanwhile state symptoms result from an acute compensatory decrease in the confidence of current sensory evidence, which leads to aberrant prediction error signalling such that the dopamine response to surprising outcomes is relatively decreased whilst the dopamine response to unsurprising outcomes is increased (Adams et al., 2013). This fits with pharmacological evidence that pro-dopaminergic drugs given to healthy individuals sometimes produce psychotic-like behaviours (Steeds et al., 2015) – possibly precipitating as an interaction with stress, sleep deprivation, and/or genetic factors. Meanwhile NMDA-R antagonists consistently produce positive, negative, and cognitive symptoms in healthy individuals and exacerbate symptoms in patients (reviewed in Rafter et al., 2016).

Schizophrenia patients often display avolition and motivational anhedonia, that is, they put in less effort to pursue rewards compared to

healthy people, despite rating the experienced value of those rewards similarly (Strauss et al., 2014). A collection of evidence suggests this is due to dysfunction of model-based inference (reviewed in Strauss et al., 2014). Specifically, patients are: impaired at learning from positive RPEs; unable to mentally compare the value of options beyond historical reinforcement (i.e. lack transitive inference; Gold et al., 2012); and are impaired at effort-cost computations; all of which lead to a bias in decision making such that “risky” options with stochastic reinforcement (e.g. trying new activities, socialising) are less likely to be chosen compared to “safe” historically-rewarded options (e.g. eating junk food, smoking). Consistent with this, reduced activation of regions in the limbic and associative loop during reinforcement learning tasks correlates with many negative symptoms, suggesting that overall, people with schizophrenia lack motivation due to an impairment in the circuitry that invigorate goal-directed behaviour (Strauss et al., 2014).

In summary, numerous alterations in reinforcement learning are presented in schizophrenia, which may be incongruent depending on psychotic state and symptom severity. Genetic and pharmacological evidence suggests that trait (negative and cognitive) symptoms may be a result of disrupted excitatory/inhibitory balance (secondary to glutamate and/or other synaptic deficits; Pavão et al., 2014), whilst state (positive) symptoms arise from an interaction of the above plus stressors or other factors that alter dopamine system function (Howes et al., 2015).

1.4.3 Drug addiction

Drug addiction is commonly characterised by compulsive drug use despite its profound negative consequences. The most prevalent addictive drugs (e.g. cocaine, heroin, nicotine) are all capable of artificially increasing striatal dopamine release (Redish, 2004), and this is particularly important for reinforcing their consumption since it interferes with action valuation. Specifically, with natural rewards the value of the action associated with them only increases as a function of the difference between expected and actual reward. However with drug use the value can theoretically rise without limit since positive RPEs are reported regardless of whether the actual reward matches the expectation (Redish, 2004).

As well as being highly reinforcing, addictive drugs induce plastic changes in the brain that bias the balance of corticostriatal circuitry from goal-directed to habitual (reviewed in Everitt and Robbins, 2016; Lucantonio et al., 2014a). This leads to an over-reliance on model-free behaviour (Voon et al., 2015), which may be a result of impaired confidence valuations in associative striatal regions (DMS, ventral striatum) and enhanced confidence valuations in sensorimotor striatum (DLS; Lucantonio et al., 2014a). In behavioural terms, this manifests as doubtfulness towards a drug's negative effects and an assurance in its positive effects. Neuroadaptations in the OFC are also particularly prominent (Lucantonio et al., 2012), with decreased grey matter correlating with compulsive drug use. Concordantly, drug addiction (in humans, non-human primates, and rodents) leads to the typical behavioural impairments in OFC-dependent (model-based) tasks such as reversal learning, outcome devaluation,

Pavlovian overexpectation, and delay discounting (reviewed in Lucantonio et al., 2012). Likewise, in rats, task-related neural activity in the OFC is disrupted – even months after drug withdrawal (Stalnaker et al., 2006) – and a recent study showed that optogenetic rescue of this activity reversed OFC-dependent task impairments (Lucantonio et al., 2014b).

Thus, drugs that cause long-term disruption of OFC function promote their compulsive intake by impairing goal-directed (model-based) inference of the negative outcomes of their use; which in turn facilitates cue-evoked drug seeking by the habitual system via strongly cached drug-reward associations (Lucantonio et al., 2014a).

1.4.4 Parkinson's disease

Parkinson's disease is characterised by major cell loss of dopamine neurons in the substantia nigra, which is accompanied by low dopamine levels in the dorsolateral striatum (Kish et al., 1988). Accordingly, Parkinson's patients off medication are quicker to learn from negative RPEs dependent on decreased dopamine receptor signalling, compared to positive RPE learning dependent on dopamine receptor stimulation (Frank et al., 2004). Additionally, when on medication (L-dopa and D₂R agonists), this bias reverses, since phasic dopamine spikes are restored by L-dopa, allowing D₁R-mediated positive RPEs, but phasic dopamine dips are not followed by disengagement of D₂R signalling, causing an impairment at learning from negative RPEs (Cools et al., 2006; Frank et al., 2004). Additionally, dopamine agonists can lead to a bias in learning from gains vs losses that eventually leads to behavioural addictions such as compulsive

gambling and shopping (Voon et al., 2010), due to the valuation of these actions increasing after unexpected gains but not decreasing after unexpected losses.

1.4.5 Summary

In summary, the compulsive behaviours displayed in these disorders all relate to an imbalance between the goal-directed (model-based) and habitual (model-free) decision making systems. These imbalances can arise from different mechanisms, e.g. aberrant RPE signalling, or impaired inference of alternate outcomes, however they all lead to compulsivity. Thus compulsivity represents a symptom dimension that spans disorders and is mediated by common neural circuitry (Gillan et al., 2016a; Voon et al., 2015).

1.5 Studying dual system function with rodent reversal learning

Since compulsive behaviours arise due to dysfunction in evolutionarily conserved Supraprimate corticostriatal circuitry, it follows that primate compulsivity can be modelled in rodents. This section will discuss the evidence that rodent reversal learning can be used to assess compulsivity arising from OFC-striatal dysfunction.

1.5.1 Reversal learning

Reversal learning refers to situations where associations between actions and outcomes are switched, and often applies to two-choice tasks where the A-O associations are anti-correlated (see Figure 1.6). Reversal learning tasks are often split into two phases: the acquisition phase – where

initial A-O associations are formed, and the reversal phase – where A-O associations are changed. It is specifically during the reversal stage that OFC dysfunction causes learning impairments (reviewed in Hamilton and Brigman, 2015, discussed in detail shortly).

These phases can be further split into early and late, and putatively engage the model-based and model-free systems to differing degrees (see Figure 1.6). One hypothesis is that OFC dysfunction impairs reversal learning by slowing the transition from model-free to model-based control during the early-mid reversal phases, therefore tasks more reliant on model-free control (e.g. more habitised) will more likely be impaired by OFC dysfunction (although this depends on a host of variables discussed later, see Table 1.3).

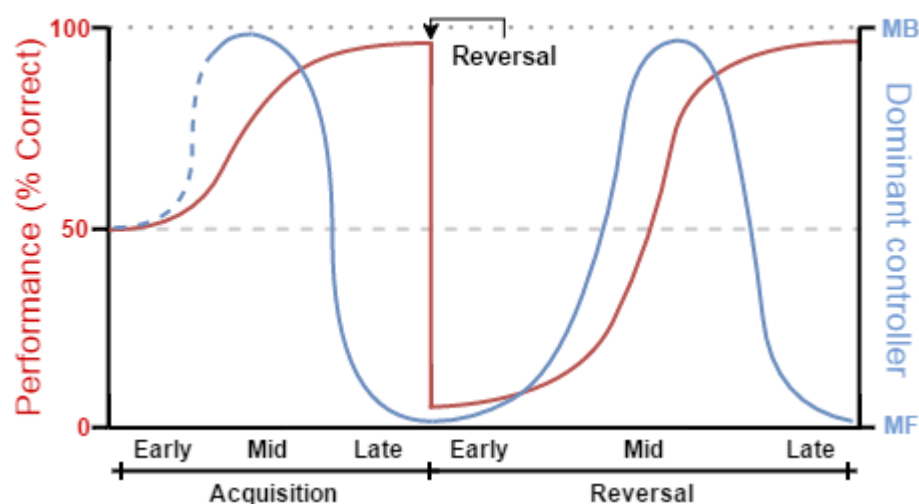


Figure 1.6 Theoretical performance and controller system engagement during a two-choice multi-trial reversal learning task.

For initial trials (early acquisition), performance is roughly 50% correct as the agent chooses actions randomly to determine their outcomes. With reinforcement of one action but not the other, the agent begins to bias selection of this action (mid acquisition). After repeated reinforcement, the agent habitises selection of this action (late acquisition). After the reversal, the agent initially continues to perform the previously rewarded action due to habit (early reversal), but eventually overcomes the habit and starts selecting the previously unrewarded – but now rewarded – action (mid reversal). After repeated reinforcement of the new A-O association, the agent begins to habitise performance of this action. The model-based (MB, associative loop) system is most heavily engaged during the mid-acquisition and mid-reversal phases, whilst the model-free (MF, sensorimotor loop) system is most engaged during late acquisition, early reversal, and late reversal.

1.5.2 Role of OFC

How does OFC facilitate the transition from model-free to model-based behaviour? Currently the most adequate explanation is that OFC integrates information from numerous regions to form a salient associative map of the environment/task, which can then be traced to infer states in the environment (reviewed in Stalnaker et al., 2015).

The expected action-outcome values given the currently inferred state are signalled to midbrain dopamine neurons (and possibly serotonin neurons; Vázquez-Borsetti et al., 2011), which compare this information with the actual received value to signal prediction error (Takahashi et al., 2011). Therefore the stronger (and longer) the OFC signals the expected (pre-reversal) value, the larger the RPEs produced (Schoenbaum et al., 2007). These larger RPEs may facilitate system transition by decreasing the confidence of the old valuation by DLS neurons, or by signalling a new environmental state.

OFC-dependent inference is also important for accurate action-outcome assignment (Walton et al., 2010). Tasks in which different choices with different outcomes are chosen in a small space of time (e.g. post-reversal) are particularly affected by OFC dysfunction, since animals are unable to infer that the most recent outcome should only be assigned to the most recent action (see Figure 1.7).

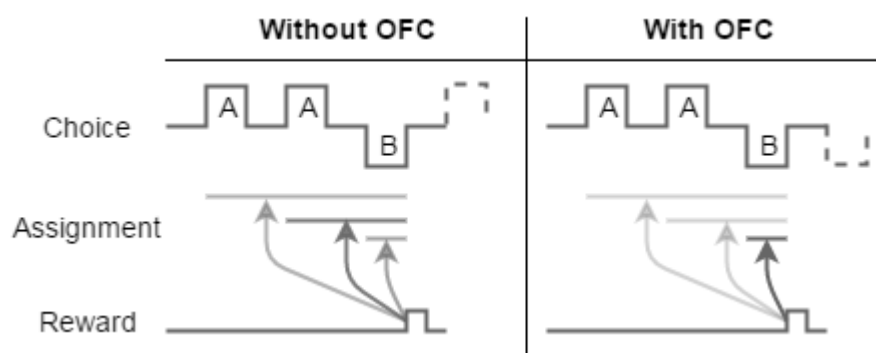


Figure 1.7 Credit assignment by orbitofrontal cortex.

How does an agent know if an outcome was a result of the immediate prior action or a combination of prior actions? The OFC – by representing a learned map of the task structure – may infer that the reward was a result of the immediate prior action due to its inference that rewards are only affected by choices within that trial, and are independent of previous trials. In the diagram above, the agent without OFC is uncertain of whether the reward was a result of choice B, or the combination of AB or AAB. The agent with OFC confidently infers that outcomes are only affected by the choice made immediately prior, and so assigns the reward to B.

What neural substrates allow the OFC to form – and flexibly trace – a map of associations (states) in the environment?

Besides the striatum, OFC is innervated by a host of brain regions including the insula, amygdala, hippocampus, basal forebrain, and other cortical regions spanning prefrontal, motor, and sensory areas (Kringelbach and Rolls, 2004). Therefore OFC is in a prime position to receive information across sensory modalities and memory systems that allow it to form associations between stimuli, actions and outcomes that vary according to context and internal states.

The maintenance of this associative information is likely reliant on NMDA receptors, since intra-OFC infusion of NMDA-R antagonists abolishes cue-outcome selective firing (van Wingerden et al., 2012), and consequently impairs reversal learning (Bohn et al., 2003; Brigman et al., 2013; Homayoun and Moghaddam, 2008). This is akin to the impairments caused by cocaine withdrawal mentioned previously (Stalnaker et al., 2006), which may be due to disrupted OFC glutamate signalling (Lucantonio et al., 2014b).

OFC is also a target of mesocortical dopamine afferents. These afferents signal aversive sensory prediction errors rather than RPEs (Lammel et al., 2011), and this signal has low spatial and temporal fidelity due to lack of dopamine transporter expression in the prefrontal cortex (Sesack et al., 1998). Since OFC dopamine depletion has not been shown to impair reversal learning (Clarke et al., 2007), the importance of this signal is unclear. However prefrontal dopamine is known to play a critical role in stabilising representations in working memory (Durstewitz and Seamans, 2008), and the study by Clarke et al., 2007 had very little working memory load (trial intervals of 3s). Likewise, other studies suggest a subtle but measurable role for OFC dopamine in reinforcement learning (Clarke et al., 2014; Krugel et al., 2009; Walker et al., 2009).

Information about RPEs may come indirectly from the DMS neurons that rapidly encode A-O associations (see section 1.3.2). However conventional theories of basal ganglia loops state that information can only ascend – not descend – the spirals (see Figure 1.3), therefore OFC may be receiving this information cortico-cortically.

OFC is also richly innervated by 5-HT neurons from the dorsal raphe nucleus (reviewed in Roberts, 2011). These neurons convey a multiplexing mix of reward and punishment signals (Dayan and Huys, 2015) that may be particularly important for the OFC. This fits with numerous studies showing changes in 5-HT signalling post-reversal, and that alteration of this signalling (especially within OFC) bidirectionally affects reversal learning (reviewed in Roberts, 2011). Some have proposed that deficient OFC 5-HT signalling impairs reversal learning specifically by increasing perseverance towards formerly valuable choices (i.e. impaired negative RPE processing Clarke et al., 2007), and that this is due to perseverative neural ensemble activity (Maia and Cano-Colino, 2015). However OFC 5-HT lesioned animals decrease responding during extinction (albeit are still biased towards the formerly valuable option), suggesting intact negative RPE signalling but impaired negative RPE-driven exploration (Roberts, 2011). Alternately then, increased exploration of a previously unrewarded option may be dependent on inference of a new state by the OFC, which occurs in response to negative RPEs signalled by 5-HT.

In summary, OFC may accelerate reversal learning via multiple mechanisms. Firstly, by signalling pre-reversal expected outcomes to the VTA – which amplifies RPEs. Secondly, by facilitating accurate assignment of outcomes to the choices that caused them (by maintaining associations of the learned task structure). And thirdly, by inferring new states or rules of the task, e.g. anti-correlated choice-reward contingencies. In the next section

I will discuss the ideal conditions in which OFC manipulations affect reversal learning.

1.5.3 Optimal conditions and measures

Despite the plethora of studies finding specific reversal learning deficits with OFC lesions (reviewed in Hamilton and Brigman, 2015), there are some contradictory findings. Because studies vary in the numerous experimental parameters that are critical in determining strategies (and thus circuitry) used during the tasks, it is not feasible to pinpoint the exact cause of these discrepancies. However the general trend seems to be that the more consolidated or habitual (i.e. confidently valued by the model-free system) an association is, the more OFC-dependent its reversal becomes (see Table 1.3 for more detail).

Based on these suggestions, a reversal learning task optimally designed for assessing OFC function would be designed such that:

- The acquisition phase continues until stringent, stable criteria are met, so that behaviour is under habitual (model-free) control.
- Interventions suspected to target OFC function are administered after this point to check for intact behaviour
 - If the intervention alters non-choice behaviour, it may be targeting other areas.
 - If it alters choice behaviour, it may be cueing a novel context (Overton, 1966) or disrupting sensorimotor corticostriatal circuitry.
- The reversal stage commences (with continued intervention)

- If the intervention selectively alters choice behaviour at this stage, it may be targeting OFC.
- The reversal stage continues until stringent, stable criteria are met.
 - If the intervention continues to alter choice behaviour at this stage, it may be targeting other corticostriatal circuitry.

With a task of this design, measures such as choice behaviour can shed light on RPE-dependent learning in the model-based and model-free systems. In addition, latency measures may act as surrogate markers of system engagement, since habitual actions require less deliberation than goal-directed actions and so are performed more rapidly (see Keramati et al., 2011). Indeed the dual-controller engagement timelines proposed in Figure 1.6 approximately fit the response latency measures demonstrated by (Keramati et al., 2011). Based on these principles, this thesis will test reversal learning in non-albino (Lister-hooded) rats using a well-validated visual touchscreen task (Mar et al., 2013).

Table 1.3 A non-exhaustive list of reversal learning task parameters that affect behavioural strategy and therefore the circuitry utilised.

Parameter	Description
Reinforcement contingency	Probabilistic reinforcement destabilises consistent choices, relies on OFC-dependent credit assignment (Walton et al., 2010). Acquisition learning of probabilistic discriminations is often OFC-dependent (Dalton et al., 2016).
Quantity of training	More training during acquisition of the association promotes habitual responding (Yin and Knowlton, 2006).
Inter-trial interval (ITI)	Longer ITIs perturb use of working memory, preventing associations from being maintained in model-based circuitry (Collins and Frank, 2012).
Trials per session	More trials (and longer sessions) may lead to satiation, fatigue, or decline of treatment effects (Gastambide et al., 2013).
Discovery / practice trials	Trials in which animals can sample all choices freely without it counting towards their performance. These accelerate learning and render the task independent of counterfactual inference (Gastambide et al., 2013).
Forced choice trials	Similar to above except animals are forced to choose an option on a trial to accelerate learning (Keiflin et al., 2013).
Number of choices available	More choices increase deliberation (Keramati et al., 2011), and thus the likelihood that responding reflects actual associative learning rather than chance responding. More choices slow learning since more associations must be learned, and counterfactual inference is more difficult (Walton et al., 2010).
Number of “task-irrelevant” cues	More cues may facilitate learning post-reversal. For example, the (task-relevant) left and right reaches of a T-maze are associated with reward/non-rewarded during acquisition. During reversal, fresh learning to intra-maze cues that differentiate the left and right reaches facilitates learning as it only requires the old associations to be ignored in favour of the new associations, rather than the old associations being unlearned and overwritten (Yin et al., 2008).
Choice modality	Unconditioned (instinctive) affinity or aversion for an option (stimulus or action) affects choice behaviour. Inhibiting these prepotent biases requires prefrontal circuitry (Walker et al., 2006).

Table 1.3 Continued

Treatment timing	State-dependent learning may occur during treatment administration/withdrawal. Different circuits are sensitive to manipulations depending on the stage of learning (West et al., 2011), and whether they are temporary or permanent (Keiflin et al., 2013; Otchy et al., 2015).
Reversal rate / criteria	Similar to quantity of training. Associations that are more stringent (consistently chosen) and stable (for a long period of time) are more likely to be habitual (Riceberg and Shapiro, 2012).
Number of reversals	Later reversals do not rely on model-based inference as the reversal rule (e.g. lose-shift (Jang et al., 2015; Rayburn-Reeves et al., 2013) becomes cached in model-free circuitry (Boulougouris et al., 2007a).
Outcome measures	Data averaging may mask true or nuanced changes in choice behaviour (Gallistel et al., 2004). Conversely, over-flexible post-hoc data binning may yield false positives by misinterpreting noise as signal (Kriegeskorte et al., 2009; Smith and Ebrahim, 2002).

1.6 Aims

Thus far it has been established that compulsivity in psychiatric disorders can result from disrupted function within corticostriatal loops, which leads to an over-reliance on habitual (model-free) decision strategies. Reversal learning tasks provide measures that assess these learning and strategy abnormalities. Therefore this thesis aims to:

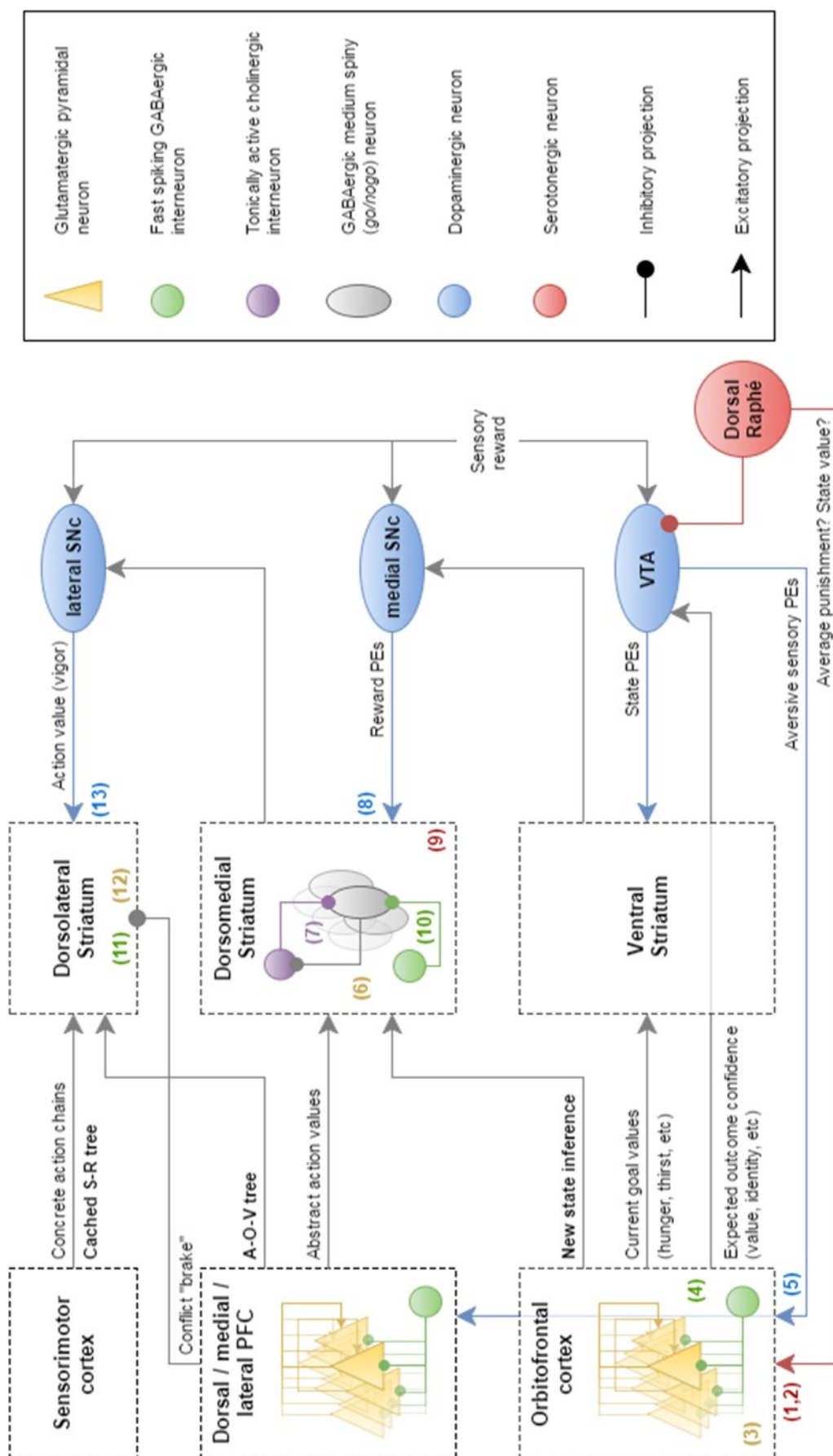
- Develop a rodent reversal learning task that relies on goal-directed and habitual learning (Chapters 2 and 3).
- Develop arms of this task that differentiate between learning from positive RPEs and negative RPEs (Chapter 4 onwards).
- Test the effects of different pharmacological, environmental, and depletion treatments on performance of these tasks.
 - Pharmacological – NMDA receptor antagonism (Chapter 4), 5-HT_{2C} receptor antagonism (Chapter 5).
 - Environmental – Post-weaning social isolation after neonatal NMDA receptor antagonism (Chapter 6).
 - Lesioning – OFC depletion of dopamine (Chapter 7) and 5-HT (Chapter 8).

Based on the circuitry described in Figure 1.8, the general hypothesis of the thesis is that interventions targeting different monoamine and glutamate signalling pathways will alter corticostriatal circuitry in different ways, resulting in specific behavioural alterations on each test. Specifically, low dose systemic NMDA receptor antagonism is expected to preferentially disturb OFC glutamate signalling, which behaviourally manifests as

disrupted early but not late reversal learning (see Chapter 3). 5-HT_{2c}R antagonism is expected to enhance motivation and impulsive responding by disinhibiting the limbic cortico-striatal loop, and enhance learning in response to negative feedback by increasing the signal-to-noise ratio of OFC 5-HT signalling (see Chapter 5). Post-natal NMDA receptor antagonism combined with post-weaning social isolation is expected to enhance learning from positive RPEs and motivation for reward via VTA sensitisation but impair reversal learning via cortical NMDA receptor hypofunction (see Chapter 6). OFC dopamine depletion is expected to impair early reversal learning in response to positive feedback by disrupting PFC-dependent rule stability in working memory (see Chapter 7), whilst OFC 5-HT depletion is expected to impair early reversal learning by disrupting new state inference within OFC circuitry (see Chapter 8).

Figure 1.8 Effects of pharmacological manipulations within corticostriatal circuitry on goal-directed behaviour.

A schematic expanding upon the corticostriatal circuitry shown in Figure 1.4 linking to evidence that regional manipulations specifically affect goal-directed learning. The colour of the numbered locations indicates the neurotransmitter system targeted by the manipulation. **OFC:** (1) 5-HT_{2C} receptor antagonists affecting reversal learning (see Table 5.1). (2) 5-HT depletions affecting reversal learning (Clarke et al., 2007; Rygula et al., 2014); prepotent responding (Walker et al., 2006); reinforcement sensitivity (Rygula et al., 2014); conditioned responding, and extinction (Walker et al., 2009). (3) NMDA-R antagonists affecting reversal learning (Bohn et al., 2003; Brigman et al., 2013) but not acquisition (McKee et al., 2010). (4) GABA agonists affecting reversal learning (Alsiö et al., 2015; Burke et al., 2009; Keiflin et al., 2013); outcome devaluation (West et al., 2011); extinction (Panayi and Killcross, 2014); overexpectation (Takahashi et al., 2009); sensory preconditioning, and blocking (Jones et al., 2012). (5) DA depletions affecting reinforcement sensitivity and exploitation (Clarke et al., 2014); exploitation during extinction (Walker et al., 2009); not affecting reversal learning (Clarke et al., 2007). **DMS:** (6) NMDA-R antagonists affecting outcome devaluation (Yin et al., 2005a), acquisition (McKee et al., 2010), and reversal learning (Palencia and Ragozzino, 2004). (7) Muscarinic M₂/M₄ receptor agonists affecting reversal, state formation and retrieval (Bradfield et al., 2013). (8) DA depletions affecting reversal learning (Clarke et al., 2011; Tait et al., 2016) and contingency degradation, but not outcome devaluation (Lex and Hauber, 2010). (9) 5-HT depletions not affecting reversal learning (Clarke et al., 2011). (10) GABA agonists affecting outcome devaluation (Corbit and Janak, 2010) and contingency degradation (Yin et al., 2005b). **DLS:** (11) GABA agonists affecting contingency degradation (Yin et al., 2006). (12) NMDA-R antagonists affecting late reversal learning (Brigman et al., 2013); not affecting outcome devaluation (Yin et al., 2005a) or early reversal learning (Brigman et al., 2013; Palencia and Ragozzino, 2005). (13) DA depletion affecting development of outcome-insensitivity (Faure et al., 2010).



Chapter 2 – The effect of phencyclidine on the expression of learned associations

2.1 Introduction

An impaired ability to flexibly update learned associations (rules) is a common finding in psychiatric disorders, especially in schizophrenia (see section 1.4.2 Schizophrenia). It has also been observed that acute infusion of the *N*-methyl-*D*-aspartate receptor (NMDA-R) antagonist ketamine to healthy humans impairs rule shifting – but not rule acquisition – on the Wisconsin Card Sorting Test (Krystal et al., 2000), similar to the phenotype seen in schizophrenia (Nieuwenstein et al., 2001). Acute NMDA-R antagonism can also induce a range of other cognitive and behavioural changes in humans that resemble a psychotic state, and consequently these drugs are frequently used to model schizophrenia in primates and rodents (see Rafter et al., 2016). This chapter sought to develop a rodent task which assesses discrimination ability and would be sensitive to low doses of an NMDA-R antagonist capable of disrupting learning about changes in associations (rule shifting). The objective of the first experiment was to optimise a two-choice visual discrimination task for rats and determine a non-sedating, choice-impairing dose of the NMDA-R antagonist, phencyclidine (PCP). PCP was used since it has a more stable pharmacokinetic profile than ketamine (Gastambide et al., 2013) and is

widely characterised in rodent behavioural research (Jentsch and Roth, 1999; Rafter et al., 2016).

Rodent instrumental learning studies looking at the effect of non-motor impairing doses of NMDA-R antagonists have been equivocal, with some finding impairments on rule acquisition (de Bruin et al., 2013; Fellini et al., 2014; Pallarés et al., 1995), others on rule expression (Talpos et al., 2012), but others solely on rule-reversal and extinction (van der Meulen et al., 2003).

Studies using regional infusion of NMDA-R antagonists in rodents have produced findings that corroborate with the dual-controller theories of decision making described in Chapter 1. Namely, infusion into regions of the associative corticostriatal loop during the goal-directed (rule acquisition) stage of learning impairs the rate of acquisition (McKee et al., 2010), but infusion after the rule is acquired does not affect rule expression (McKee et al., 2010; Palencia and Ragozzino, 2004). Infusion into the sensorimotor loop during the habitual (rule expression) stage impairs rule expression (Palencia and Ragozzino, 2005), although Brigman et al., 2013 found no late stage impairments with DLS infusion, but mid-stage “consolidation phase” impairments. Infusion into the orbitofrontal cortex (OFC) during acquisition does not impair acquisition rate (McKee et al., 2010), but infusion after rule reversal does impair re-learning (Brigman et al., 2013). These findings fit with the timeline of corticostriatal circuitry engagement proposed in Figure 1.6.

Given these findings, an experiment was conducted to determine a dose-response relationship between PCP and expression of a strongly acquired visual discrimination rule (performed at $\geq 90\%$ correct for ≥ 2 sessions) in a touchscreen apparatus. Doses up to 5mg/kg i.p. were used as this dose has been widely reported to induce hyperactivity, impair working memory, and modify mPFC function for at least 80 minutes (Adams and Moghaddam, 1998; Suzuki et al., 2002), whilst doses as low as 1.5mg/kg have been shown to impair contextual discrimination performance for at least 90 minutes (Idris et al., 2005).

2.2 Methods

2.2.1 Animals

Twelve male Lister-hooded rats (delivery weight: 150-200g, CRUK) were housed in groups of four in ventilated cages with sawdust, shredded paper, a wooden block and a cardboard cylinder. The cages were maintained at $21 \pm 2^\circ\text{C}$ and $50 \pm 5\%$ humidity, on a 12/12 hour dark/light cycle (lights on at 07:00). Seven days after arriving at the facility, animals began food restriction (16g/rat/day) with ad libitum water intake. Animals were maintained at a body weight of greater than 80% of their free-feeding equivalent. After seven days of food restriction, animals began behavioural training for 5 days per week and remained food restricted for the duration of training and testing. All experiments were conducted during the light phase and in compliance with the UK Animals (Scientific Procedures) Act, 1986; under project PPL: 40/3283, with ethical approval of the University of Nottingham. Where applicable, all experimental methods and results have

been reported in accordance with the twenty item checklist of the ARRIVE guidelines (Kilkenny et al., 2010).

2.2.2 Apparatus

Touchscreens

The automated touchscreen operant boxes each contained a four-bay modular chamber (internal dimensions: 30.5 × 31.8 × 29.2 cm; ENV-009A, Med Associates Inc., Vermont, USA) and a 24V DC fan to mask external noise. On one side of each chamber was a 12" LCD screen situated 2mm behind an infra-red beam break detector (collectively: touchscreen). Approximately 1cm in front of the touchscreen was a black aluminium mask (29 × 19 cm) containing two evenly spaced, centred square apertures (9 × 9 cm), and in front of the mask was a raised spring-hinged shelf to encourage rats to deliberate before touching the screen. Opposite the touchscreen were a sugar pellet magazine (with infra-red beam break detector and an internal 3W light), a tone generator and a 3W house light (see Figure 2.1). The chambers were controlled by K-Limbic software (SOF-740, Med Associates Inc., Vermont, USA).

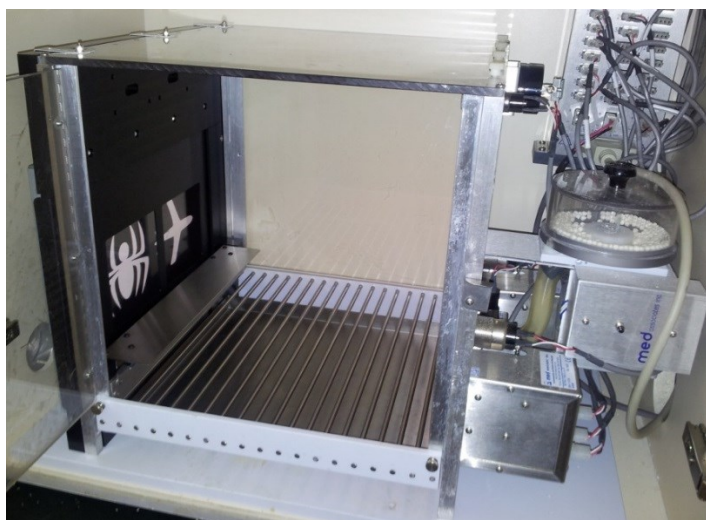


Figure 2.1 A rat touchscreen operant box displaying two stimuli.

Compounds

Phencyclidine hydrochloride (PCP) (Sigma-Aldrich, UK) was dissolved in 0.9% saline to produce solutions for intraperitoneal (i.p., 1ml/kg) injection. All solutions were made within 24h of administration and kept refrigerated.

2.2.3 Touchscreen task protocol

Habituation phase

Rats were placed into the operant chambers, with fans running and non-responding blue screens, for 20 minutes. The chambers were primed with 3 sugar pellets (45mg 5TUL tablets, TestDiet) in front of each screen aperture, and 3 pellets in the pellet receptacle. Halfway through the habituation session, chambers were re-primed with pellets in the same position. After two sessions of habituation, animals were progressed to the pre-training phase.

Pre-training phase

Rats were placed into the operant chambers, and upon a nose poke into the pellet receptacle, an image (the “S+”) was displayed in one of the screen apertures. Each rat was randomised to be shown either a spider or a plane as their S+ (n=6 for each), as these stimuli can be reliably discriminated (Bussey et al., 2008). Upon touching the S+, the screen would blackout, a tone would sound, the pellet receptacle would illuminate and a pellet would be dispensed. Upon receipt of the pellet, the same S+ would again display in one of the touchscreen apertures (pseudorandomised between trials) and this process continued for a maximum of 50 trials or 50 minutes. For the duration of the session, the chamber house light remained on. Animals were progressed to the acquisition phase upon reaching criteria (50 trials in less than 50 minutes).

Acquisition phase

This phase added some additional stages to pre-training. Firstly, rats were now shown two images (spider and plane) simultaneously, and only a nose poke to the S+ was rewarded with a pellet. A nose poke to the S- led to the punishment stage, characterised by the house light turning off for 5s. A timeout period was also added, whereby if the animal failed to nose poke within 25s of image display, the screen would blackout and the animal would have to initiate the next trial. At the end of each trial (regardless of outcome), a 20 second inter-trial interval would occur (see Figure 2.2). Image locations were pseudorandomised between trials, with the exception of correction trials (which occurred after an incorrect response) – this was to prevent adoption of suboptimal strategies such as always going to the same aperture in the knowledge that it will lead to reward roughly 50% of

the time. The entire process was repeated for 60 trials (not including correction trials) or 45 min.

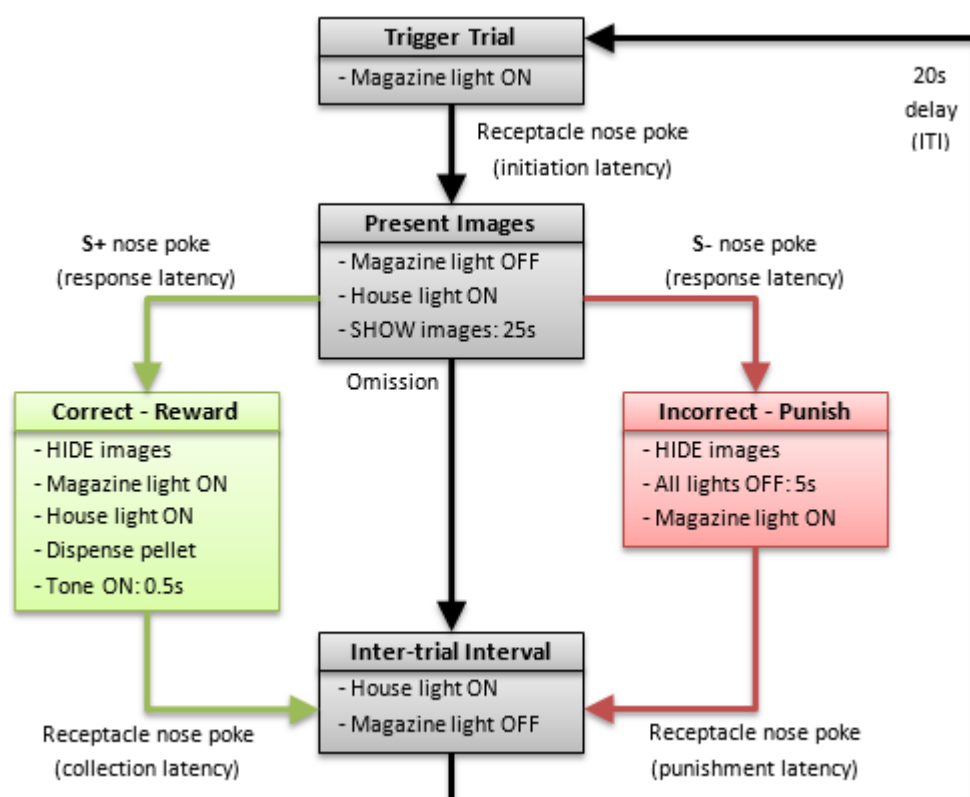


Figure 2.2 A flowchart showing the different stages of a trial for the final discrimination phase of training.

Testing phase

Once all animals had reached criteria performance ($\geq 90\%$ trials completed and $\geq 90\%$ trials correct), the testing phase commenced. The chamber protocol was identical to the acquisition phase, with the exclusion of correction trials. This phase aimed to establish a dose-response relationship between PCP and visual discrimination performance. Thirty minutes prior to entering the touchscreen chamber, rats were administered either vehicle (saline; 1ml/kg, i.p.), 0.5, 1, 2.5, or 5mg/kg PCP hydrochloride

(all in vehicle), under a modified Latin square design with all animals receiving the 5mg/kg dose on the last session (due to this dose not being planned in the original schedule). Dosed testing occurred on Tuesdays and Fridays, while drug-free testing occurred on Mondays and Thursdays. Wednesdays, Saturdays and Sundays were washout, non-test days.

2.2.4 Measures and statistics

Measures taken included: number of trials completed, percentage of correct trials (correct trials / trials completed excluding correction trials multiplied by 100), trial initiation latency, image response latency and pellet collection latency. Measures were analysed by repeated measures ANOVA with Dunnett's post-hoc multiple comparison test in order to obtain multiplicity adjusted P values for each dose compared to the control (vehicle). Where appropriate, measures were transformed for statistical analyses to produce normality (Log₁₀ transformation applies to all latency measures due to positive skew; Whelan, 2008). All analyses were performed using GraphPad Prism 6.

2.3 Results

2.3.1 Pre-training and acquisition phase

During the pre-training phase, rats took a median of 3 sessions (range 2–10) to reach criteria and be progressed to the acquisition phase. For the acquisition phase, rats took a median of 11.5 sessions (range 5–17) to perform to criteria (see Figure 2.3).

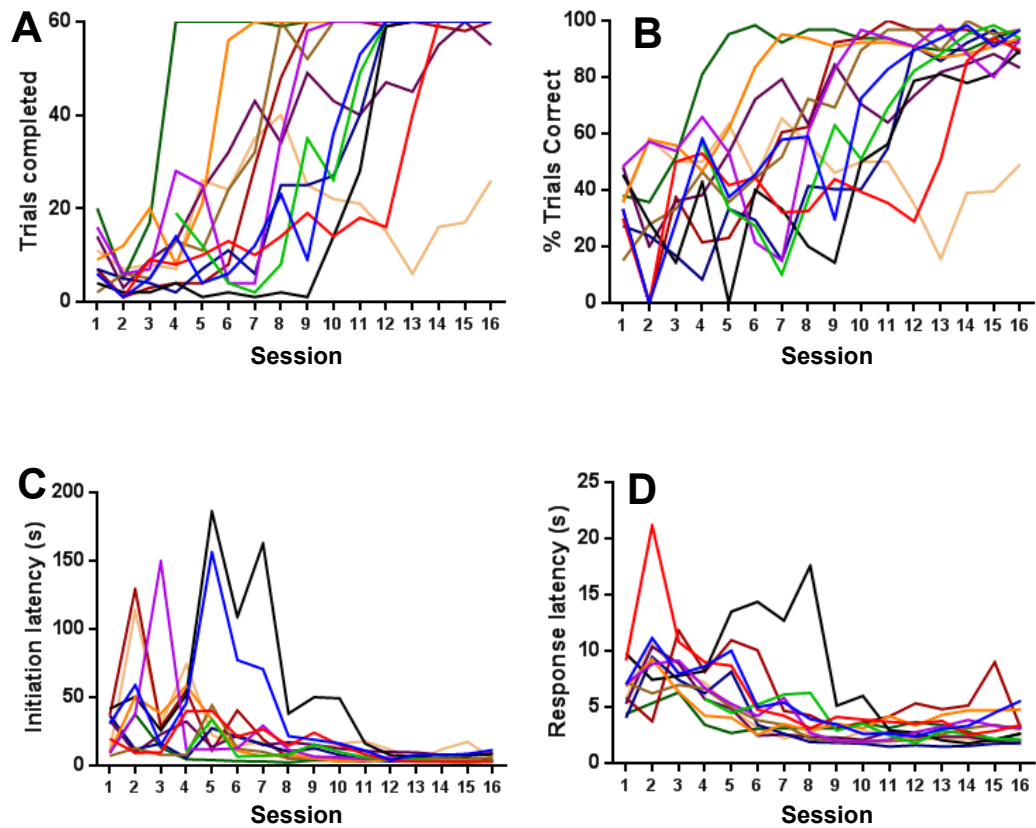


Figure 2.3 Measures during acquisition training.

Data showing visual discrimination performance measures during the acquisition phase for each **male Lister-hooded rat (n=12)** across sessions. Figures show: Number of trials completed per session (A); percentage of correct responses (B); mean session initiation latency (C); and mean session response latency (D).

2.3.2 Testing Phase

There was no effect of PCP treatment on the number of trials completed (>98% throughout the phase) or percentage of correct responses ($F_{(4, 44)} = 0.22$, $p = 0.93$; see Figure 2.4A), nor on trial initiation latency ($F_{(4, 44)} = 1.53$, $p = 0.21$; see Figure 2.4B). There was a significant effect of treatment on trial response latency ($F_{(4, 44)} = 4.1$, $p = 0.007$; see Figure 2.4C) and pellet collection latency ($F_{(4, 44)} = 10.72$, $p < 0.0001$; see Figure 2.4D), with post-hoc tests showing a significant decrease in collection latency with the 1mg/kg dose (mean difference: -0.212s, $p = 0.003$), and a significant

increase in collection latency with the 5mg/kg dose (mean difference: +0.188s, $p = 0.02$). A power calculation with two-sided equality (Rosner, 2010) showed that, given the vehicle mean of 91.1% correct trials ($\pm 7.6\%$ S.D.), a sample size of 12 would be sufficient to detect an effect size of 0.37 with 80% power, which equates to a change of at least $\pm 6.1\%$ correct. Since none of the doses exceeded an apparent effect size of 0.03, it cannot be excluded that PCP causes significant changes at smaller effect sizes which would be detected with larger sample sizes.

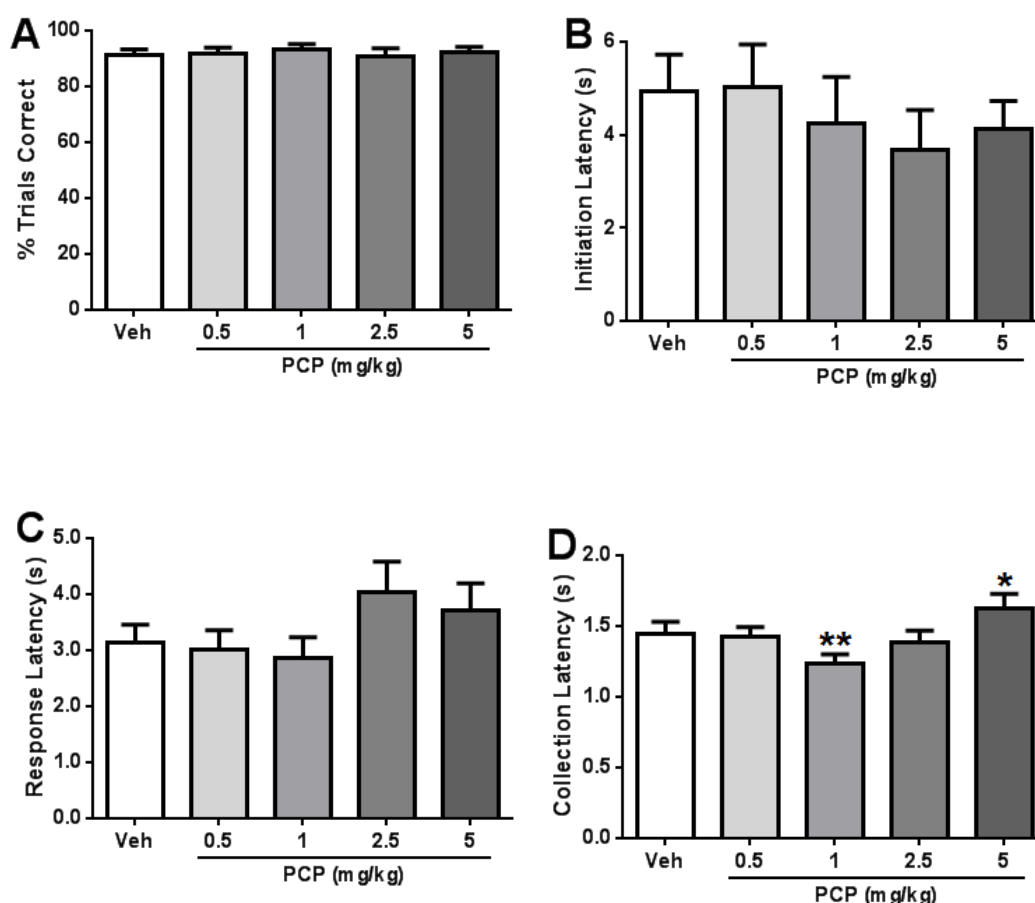


Figure 2.4 The effect of acute PCP administration (0 – 5mg/kg, n=12 for each condition) on various measures of visual discrimination performance.

Treatments were administered i.p. to male Lister-hooded rats 30 minutes prior to each session in a modified Latin-square design with at least 2 days washout between tested sessions. Each graph shows the mean ($\pm$ SEM) of: (A) percentage of trials correct in a session, (B) session initiation latency, (C) session response latency, and (D) session reward collection latency. * $p < 0.05$, ** $p < 0.005$; post-hoc Dunnett's test following one-way RM-ANOVA.

2.4 Discussion

The first experiment set out to determine a dose of PCP that would impair choice behaviour without simultaneous non-specific effects, e.g. motor or motivational impairments. The results of the testing phase showed that acute doses of PCP up to 5mg/kg (i.p.) had no effect on the ability to

express a previously learned association on a simple two-choice visual discrimination task, since there was no significant change in the percentage of correct responses or trials completed. PCP did not affect appetite or motivation to perform the task, as there was no difference in the time taken to initiate trials or on the number of trials completed. Pellet collection latency was significantly decreased by 1mg/kg PCP, but increased by 5mg/kg PCP. One caveat to this is that the 5mg/kg dose was always the last dose to be administered, and therefore drug tolerance or sensitisation may have developed (Flint and Ho, 1980).

In contrast, Talpos et al., 2012 found a significant increase in response latency with 2.5mg/kg PCP (s.c.), along with a significant decrease in percentage correct and trials completed. Disregarding minor methodological differences, the most likely reason we did not replicate this result was the dosage. Subcutaneous administration of PCP has been shown to produce a 3.5-fold higher brain concentration than i.p. administration (Kalinichev et al., 2008).

Choice behaviour in this task during the drug administration sessions was likely under habitual control (i.e. by the sensorimotor corticostriatal loop), given that all of the performance and latency measures had reached asymptote by this point (see Figure 2.3). Previous studies suggest that this is when the sensorimotor loop is engaged (Brigman et al., 2013; Keramati et al., 2011; Thorn et al., 2010). Therefore the inability of PCP to disrupt habit-driven choices is concordant with the hypothesis that its effects are more

specific to prefrontal circuitry at these doses, and render these doses more valid for assessing OFC-dependent reversal learning.

A major limitation of this experiment is that discrimination performance was only assessed in response to acute doses of PCP, whereas to assess reversal learning, consecutive doses need to be administered. Given that subchronic and chronic PCP administration results in different neural and behavioural changes to acute administration (see Janhunen et al., 2015; Jentsch and Roth, 1999), it is possible that discrimination performance would be affected with sustained dosing, although the findings in the literature suggest this is not the case with low doses (Dawson et al., 2012; Fellini et al., 2014; Goetghebeur and Dias, 2009; Janhunen et al., 2015).

In conclusion, PCP at doses below 5mg/kg (i.p.) should be suitable for assessing choice behaviour during reversal without affecting discrimination performance. These doses may produce specific learning deficits after rule reversal and avoid confounding motor or motivational impairments. In the following chapter, the effect of PCP on reversal learning will be investigated.

Chapter 3 – The effect of phencyclidine on learning after reversal of learned associations

3.1 Introduction

The NMDA-R antagonist phencyclidine (PCP) was previously shown not to affect discrimination of learned associations at **acute** doses up to 5mg/kg (see Chapter 2). However other studies have shown that it can affect choice behaviour in certain conditions, such as during the early goal-directed stage of learning (Fellini et al., 2014) or after a change in reward contingencies (see section 2.1 Introduction). This is proposed to be due to the sensitivity of different neural circuitry to NMDA-R antagonism. Briefly, **cortical** circuitry is more sensitive to lower doses of NMDA-R antagonists (Kocsis, 2012; Kocsis et al., 2013), possibly due to regional differences in NR2 subunit expression (Paoletti and Neyton, 2007) and reliance on NMDA receptor-driven gamma oscillations (Roopun et al., 2008), and so behaviour dependent on this circuitry – such as the flexible updating of action-outcome associations during early acquisition or post-reversal (see Figure 1.6) – is more likely to be impaired.

Therefore this experiment sought to determine the effects of PCP on two types of discrimination rule change: a simple reversal – where the conditioned rewarded stimulus (CS+) becomes unrewarded and the

conditioned unrewarded stimulus (CS-) becomes rewarded; and a novel acquisition – where the CS+ and CS- are replaced with a novel S+ and S-.

PCP was also administered prior to rule changes to validate that it does not affect performance of the learned rules (as per Chapter 2) and to prevent contextual effects of the drug potentially providing a reversal cue. A dose of 2.5mg/kg i.p. was used as this did not produce significant changes in any measure in Chapter 2. The drug was administered 20 minutes prior to ensure sufficient brain concentration for the entire 45 minute session (Gastambide et al., 2013). After completion of all touchscreen testing, a locomotor assay was conducted to test for PCP sensitisation/tolerance, which can occur with subchronic dosing (Phillips et al., 2001).

3.2 Methods

3.2.1 Animals

Sixteen male Lister-hooded rats (delivery weight: 150-200g, CRUK) were delivered, housed, and food restricted as described in section 2.2.1 Animals.

3.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity boxes

The locomotor activity boxes (39 × 23.5 × 24.5 cm) were made of clear Perspex surrounded by infrared photobeams (8 length-wise × 4 width-wise × 2 height-wise; Photobeam Activity System, San Diego Instruments, USA). An ambulatory count was signified by the consecutive breaking of two

adjacent **lower level** beams, and a rear was signified by the breaking of a **higher level** beam. The testing room was sound isolated and dimly lit (80 lux).

Compounds

See section 2.2.2 Apparatus.

3.2.3 Touchscreen task protocol

Phase 0 (Habituation)

Seven days after the start of food restriction, rats were placed into the operant chambers, with fans running and non-responding blue screens for 20 minutes. The pellet dispensers were baited with 3 sugar pellets (45mg 5TUL tablets, TestDiet) and the touchscreen apertures were baited with whipped cream. Halfway through the session, three pellets were dispensed simultaneously with an auditory cue. Rats were progressed to the next phase once they had learned to lick all of the cream from the screen, typically 2 sessions.

Phase 1 (White Square)

This was similar to Phase 0 but the rat had to initiate a trial by nosepoking the baited pellet dispenser. A white square then appeared in one of the apertures. Upon touching the white square, it would disappear, a tone would sound, the pellet receptacle would illuminate and a pellet would be dispensed. Upon receipt of the pellet, the receptacle light would extinguish, the white square would again display in one of the touchscreen apertures (pseudorandomised between trials) and this process continued for a maximum of 50 trials or 50 minutes. For the duration of the session, the

chamber house light remained on. Rats were progressed to the next phase once they had completed 50 trials within 50 minutes, typically 3 sessions.

Phase 2 (White Square plus Initiation and Interval)

This was similar to Phase 1 but after collecting the food pellet the rat had to wait 20 seconds (inter-trial interval) and then re-poke the pellet dispenser to initiate the next trial. If rats failed to respond to nosepoke the white square within 25 seconds of display then an omission was recorded and the trial incremented. Rats were progressed to the next phase once they had completed 50 trials within 50 minutes, typically 3 sessions.

Phase 3 (Acquisition without Punishment)

This was similar to Phase 2 but the apertures showed the rewarded stimulus (S+) and the non-rewarded stimulus (S-), with stimulus location pseudorandomised between trials. Touching the S+ led to sugar pellet reward, touching the S- did nothing. S+/S- image assignment was randomised and evenly balanced (n=8 Spider/Plane, n=8 Plane/Spider). Rats were progressed to the next phase once they had completed more than 50 trials within 45 minutes, typically 2 sessions.

Phase 4 (Acquisition with Punishment)

This was similar to Phase 3 but touching the S- led to a 5 second chamber blackout and counted as an incorrect trial (see Figure 2.2). The following trial was then a repeat of the S+/S- locations (correction trial, maximum of 6 consecutively). Rats were progressed to the next phase once they had completed >50 trials within 45 minutes and at least 90% of responses were correct (to the S+), typically 8 sessions.

Phase 4b (Pre-Reversal – Treated)

This was a continuation of Phase 4 except animals were administered their treatments 20 minutes prior to entering touchscreen chamber (PCP, n=8; 0.9% saline, n=8). Treatment assignment was randomised with counterbalancing across image assignment and rat performance (slow vs. fast learners). This phase lasted 3 sessions.

Phase 5 (Reversal – Treated)

This was a continuation of Phase 4b, except that prior to the phase change; the S+/S- assignment was reversed. This phase lasted 9 sessions, and was followed by a 2 day washout period (no behavioural testing).

Phase 5b (Reversal – No Treatment)

This was a continuation of Phase 5 except no treatment was administered prior to entering the touchscreen chamber. Rats were progressed to the next phase once they had completed >50 trials within 45 minutes and at least 80% of responses were correct.

Phase 5c (Pre-Novel Discrimination – Treated)

This was a continuation of Phase 5b except animals were administered their treatments 20 minutes prior to entering touchscreen chamber. This phase lasted 2 sessions.

Phase 6 (Novel Discrimination – Treated)

This phase was identical to Phase 5c, except that a novel pair of stimuli (Bomb/Flower) was introduced as the S+/S-. Animal assignment to S+/S- and drug/control was counterbalanced for each combination (S+_{Bomb}/S-_{Flower}

Vehicle, n=4; S+_{Flower}/S-_{Bomb} Vehicle, n=4; S+_{Bomb}/S-_{Flower} PCP, n=4; S+_{Flower}/S-_{Bomb} PCP, n=4). This phase lasted 8 sessions.

Phase 6b (Novel Discrimination – No Treatment)

This phase was identical to Phase 6 except no treatment was administered prior to entering the touchscreen chamber. This phase lasted one session, and after a six day washout period, animals underwent the locomotor activity assay.

3.2.4 Locomotor activity protocol

Rats were placed into locomotor boxes for 60 minutes to habituate. All rats then received 3.2mg/kg PCP i.p. at 1ml/kg (as this dose **has been shown to** detect PCP sensitisation; Phillips et al., 2001) and were placed back into the activity boxes for 90 minutes. Rearing and ambulatory activity was recorded in 5 minute time bins.

3.2.5 Measures and statistics

Multiple measures were statistically analysed: the number of correct and incorrect responses, number of sessions to reach criterion, and latencies during different trial stages (initiation, response and collection). For the locomotor activity assay, total ambulatory activity and rears were measured in five minute time bins.

For each phase of the touchscreen task, overall between-treatment effects were analysed using two-way repeated measures ANOVAs with treatment and session as factors. Sidak's post-hoc test was used to determine which sessions showed significant between-treatment effects,

since it is more powerful than the Bonferroni test when determining adjusted P-values but remains conservative when applied to non-independent tests (Abdi, 2007). Any animals that were not performing to criterion prior to the phase were excluded from analysis.

To analyse within-treatment effects, entire dosing phases were analysed by two-way ANOVA as above. Sidak's post-hoc test was used to determine which sessions showed significant within-treatment effects compared to the (treated) pre-reversal/pre-shift session.

To test for latency to reach criterion after the reversal, a Gehan-Breslow-Wilcoxon survival comparison test was performed. Where appropriate, measures were transformed for statistical analyses to produce normality. All analyses were performed using GraphPad Prism 6.

3.3 Results

3.3.1 Acquisition phases

For the acquisition phases (0 – 4), rats assigned to PCP treatment took a median of 15.5 sessions (range 14–18) to reach performance criteria, whilst rats assigned to vehicle treatment took a median of 14 sessions (range 12–17).

3.3.2 Pre-Reversal phase

For the three treated sessions prior to reversal, treatment had no effect on the number of correct ($F_{(1, 14)} = 0.10$, $p = 0.76$) or incorrect responses ($F_{(1, 14)} < 0.001$, $p > 0.99$). There was no overall effect of treatment on any latency (all $F < 0.1$, $p > 0.7$) but there was a significant treatment $\times$ session

interaction for all three measures (initiation latency: $F_{(2, 28)} = 5.2$, $p = 0.01$; response latency: $F_{(2, 28)} = 5.5$, $p = 0.01$; collection latency: $F_{(2, 28)} = 16.4$, $p < 0.0001$), with post-hoc tests showing a significant within-treatment difference between the first and last dosing session in the PCP-treated animals (all latencies $p < 0.05$; see Figure 3.3). This is likely due to the “first dose phenomenon” whereby some animals receiving PCP appeared to display severe ataxia and stereotypical movements during the first session, which disappeared on subsequent sessions (consistent with the repeated dosing literature; Gilmour et al., 2012).

3.3.3 Reversal phase

For the nine treated sessions after the image-reward contingencies had been reversed, there was a significant overall effect of treatment on the number of incorrect responses ($F_{(1, 14)} = 16.61$, $p = 0.0011$). Although there was no overall significant interaction between session and treatment ($F_{(8, 112)} = 0.23$, $p = 0.99$), post-hoc tests showing a significant difference during the second reversal session ($t_{126} = 3.3$, $p = 0.01$; see Figure 3.1). There was no effect of treatment on the number of correct responses ($F_{(1, 14)} = 1.50$, $p = 0.24$), suggesting that PCP treatment did not affect exploration/re-evaluation of the former S-, but did affect exploitation/extinction of the former S+. In addition, the PCP group took significantly longer to reach criterion after the reversal ($\chi^2 = 4.35$, d.f. = 1, $p = 0.04$; see Figure 3.2).

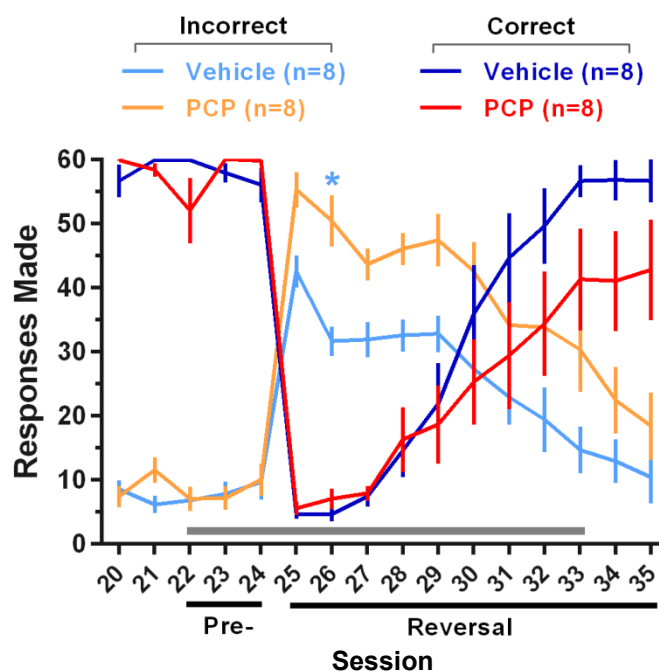


Figure 3.1 Responses made during a simple reversal test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number of correct and incorrect responses (mean $\pm$ SEM) for each session (x-axis) during the simple reversal test in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). * $p < 0.05$, post-hoc Sidak test following two-way RM-ANOVA.

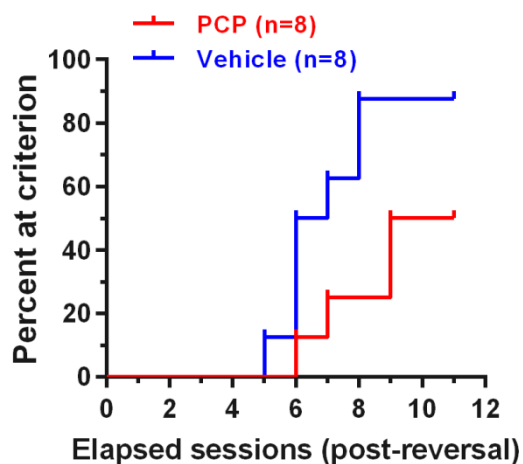


Figure 3.2 Sessions to criterion on a simple reversal test for PCP or vehicle treated animals.

A survival curve showing the effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the simple reversal test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

Latency measures during reversal

For trial initiation latency, there was a significant effect of session ($F_{(8, 112)} = 9.8$, $p < 0.0001$), treatment ($F_{(1, 14)} = 6.6$, $p = 0.02$), and a treatment $\times$ session interaction ($F_{(8, 112)} = 3.4$, $p = 0.002$). Post-hoc analysis showed a significant within-treatment increase during reversal sessions two ($t_{154} = 1.8$, $p = 0.007$) and three ($t_{154} = 2.15$, $p < 0.0001$) in vehicle treated rats compared to the session prior to reversal. PCP-treated animals showed no significant within-treatment increases during any reversal sessions (all $p > 0.3$). Similarly, there were significant between-treatment effects on the second ($t_{126} = 3.2$, $p = 0.015$) and third ($t_{126} = 4.4$, $p = 0.0002$) reversal sessions. This implies that vehicle treated animals were less motivated to

initiate trials after reversal (when lots of negative feedback was being received) whilst PCP animals were not.

For image response latency, there was a significant effect of session ($F_{(8, 112)} = 13.7, p < 0.0001$), with no effect of treatment or an interaction between the two ($F_s < 2.5, p > 0.1$). Post-hoc tests showed significant within-treatment increases for reversal sessions 1 – 3 in the vehicle treated animals, and significant increases in sessions 2 – 3 for PCP-treated animals (all $p < 0.05$, see Figure 3.3). This implies that both groups of animals increased image deliberation time in response to the negative feedback received during reversal.

Collection latency was not significantly affected by session or treatment ($F_s < 1.5, p > 0.2$), suggesting no disruption of consummatory motivation by either factor.

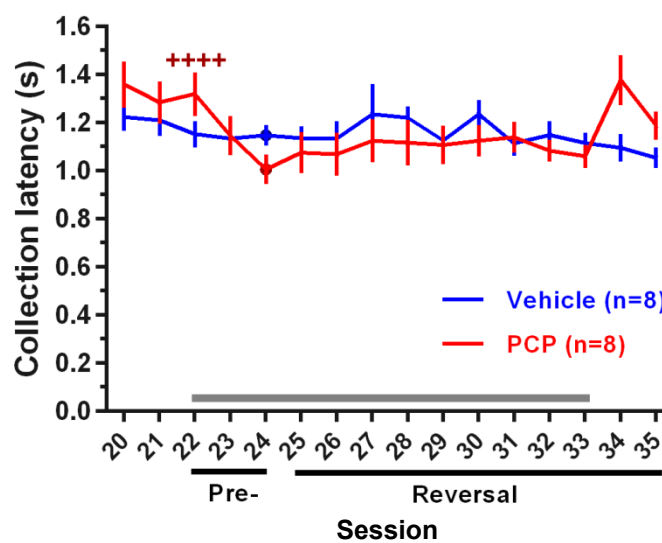
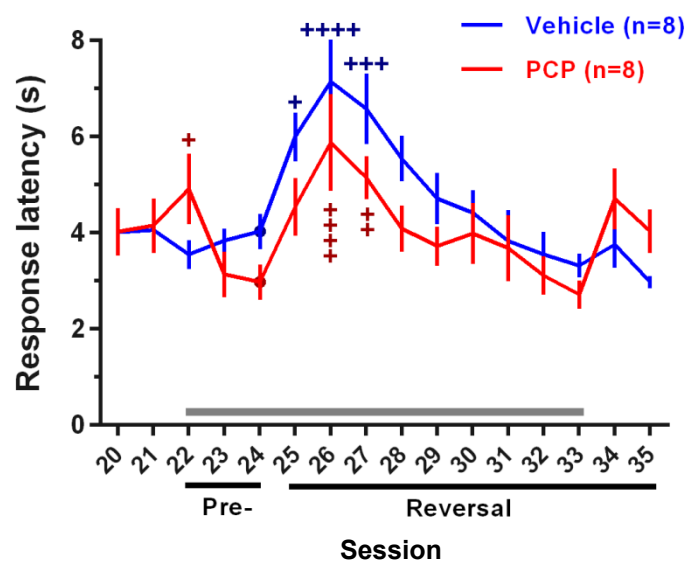
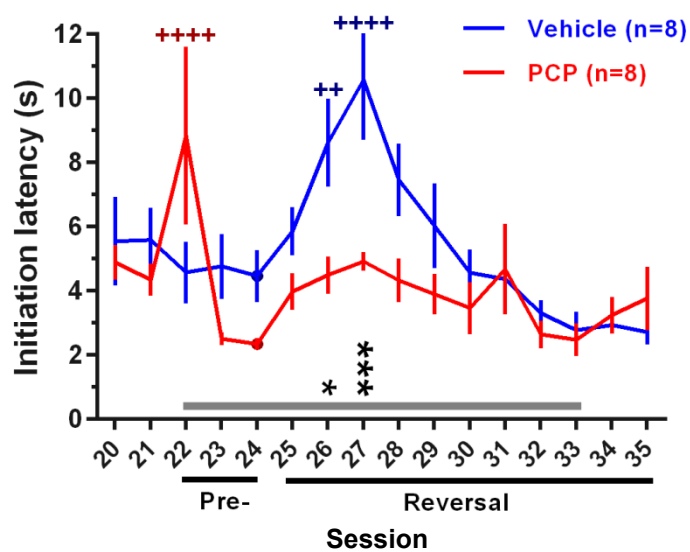


Figure 3.3 Latency measures during a simple reversal test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the trial latencies (mean $\pm$ SEM) for each session (x-axis) during the simple reversal test in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 min prior to session). * $p < 0.05$, *** $p < 0.001$ versus saline. + $p < 0.05$, ++ $p < 0.01$, ++++ $p < 0.0001$ versus pre-reversal session 24. Post-hoc Sidak test following two-way RM-ANOVA.

3.3.4 Novel Acquisition phase

There was no effect of treatment on the number of incorrect responses ($F_{(1, 14)} = 1.57$, $p = 0.23$) or correct responses ($F_{(1, 14)} = 2.46$, $p = 0.14$; see Figure 3.4). However, the PCP group took significantly longer to reach criterion after the shift ($\chi^2 = 5.88$, d.f. = 1, $p = 0.02$; see Figure 3.5), suggesting that PCP impaired acquisition of the discrimination, but that the former two measures were not sensitive enough to detect this.

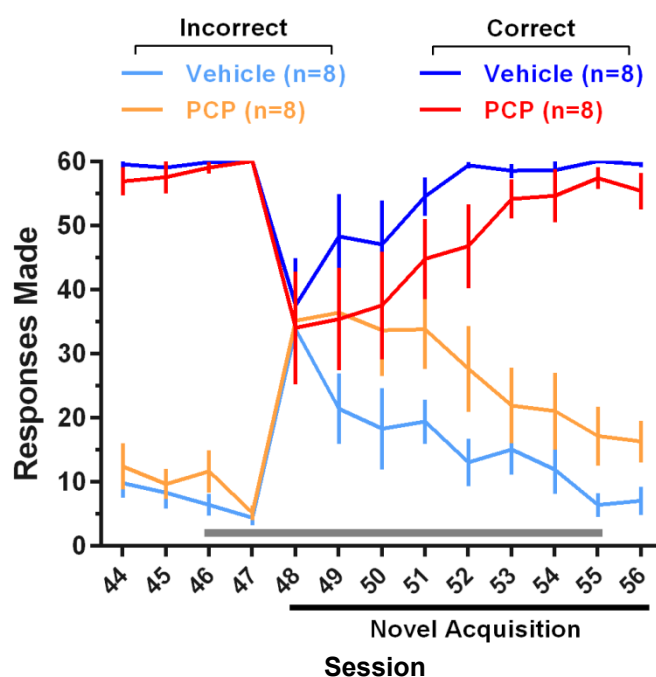


Figure 3.4 Responses made during a novel acquisition test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novel acquisition test **in male Lister-hooded rats**. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior).

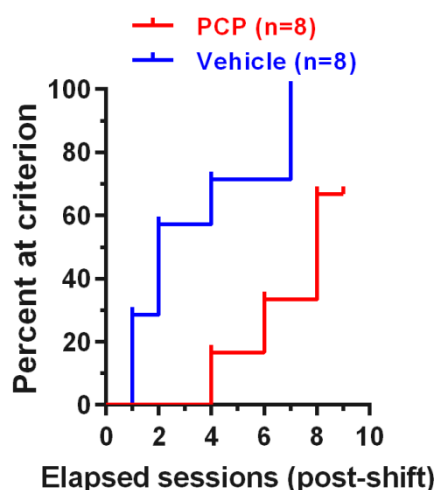


Figure 3.5 Sessions to criterion on a novel acquisition test for PCP or vehicle treated animals.

A survival curve showing the effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the novel acquisition test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

A separate analysis showed that there was a significant effect of S+/S- assignment on the percentage of correct trials ($F_{(3,12)} = 4.03$, $p = 0.03$), and interaction between image assignment and session ($F_{(30,120)} = 4.7$, $p < 0.0001$), with post-hoc tests showing a significant difference for the first four sessions between the S+Plane→S+Flower and S+Plane→S+Bomb groups (all $p < 0.05$, see Figure 3.6). There was insufficient statistical power to detect treatment × image assignment interactions ($n=2$ per condition). These findings suggest that the novel acquisition was harder for some animals due to a conditioned bias towards certain images.

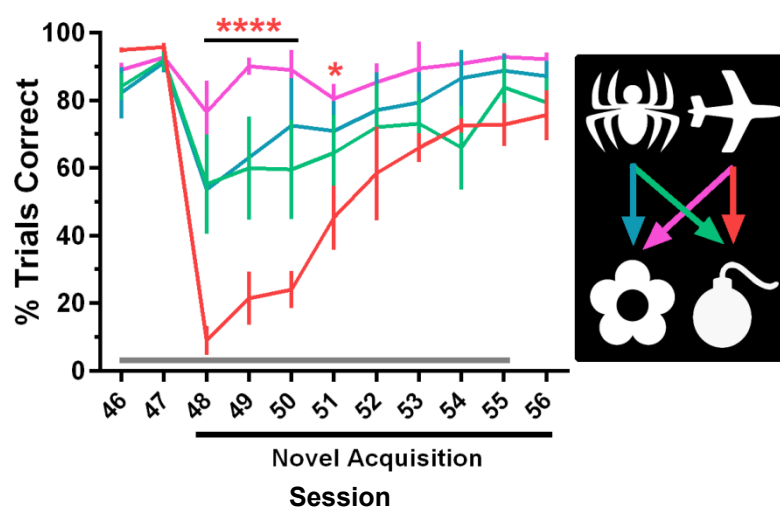


Figure 3.6 Responses made split by image assignment during a novel acquisition test for PCP or vehicle treated animals.

The effects of image assignment on the percentage (mean $\pm$ SEM) of correct trials for each session (x-axis) during the novel acquisition test. All **male Lister-hooded rats** went from discriminating Spider/Plane to Flower/Bomb. Blue indicates animals whose S+ changed from Spider to Flower (n=4). Green indicates S+Spider to S+Bomb (n=4). Pink indicates S+Plane to S+Flower (n=4). Red indicates S+Plane to S+Bomb (n=4). Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior). * $p < 0.05$, **** $p < 0.0001$, post-hoc Sidak test following two-way RM-ANOVA.

3.3.5 Locomotor activity

In the activity boxes during the 60 min prior to PCP administration, there was a significant effect of time on ambulation ($F_{(11, 154)} = 24.27$, $p < 0.0001$) and rears ($F_{(11, 154)} = 19.06$, $p < 0.0001$) in all animals. There was no significant difference in ambulation ($F_{(1, 14)} = 0.27$, $p = 0.61$) or number of rears ($F_{(1, 14)} = 0.19$, $p = 0.67$) between the cohort of rats that received PCP or vehicle during the study (see Figure 3.7). In the 90 minute period following PCP administration, the PCP cohort showed a significantly greater ambulatory response ($F_{(1, 14)} = 7.96$, $p = 0.01$) and more rears ($F_{(1, 14)} = 13.91$,

$p = 0.002$) than the vehicle cohort, suggesting that prior PCP treatment led to behavioural sensitisation to acute PCP challenge. There was a significant treatment $\times$ time interaction for ambulation ($F_{(17, 238)} = 2.436$, $p = 0.0016$) but not rears ($F_{(17, 238)} = 1.204$, $p = 0.26$).

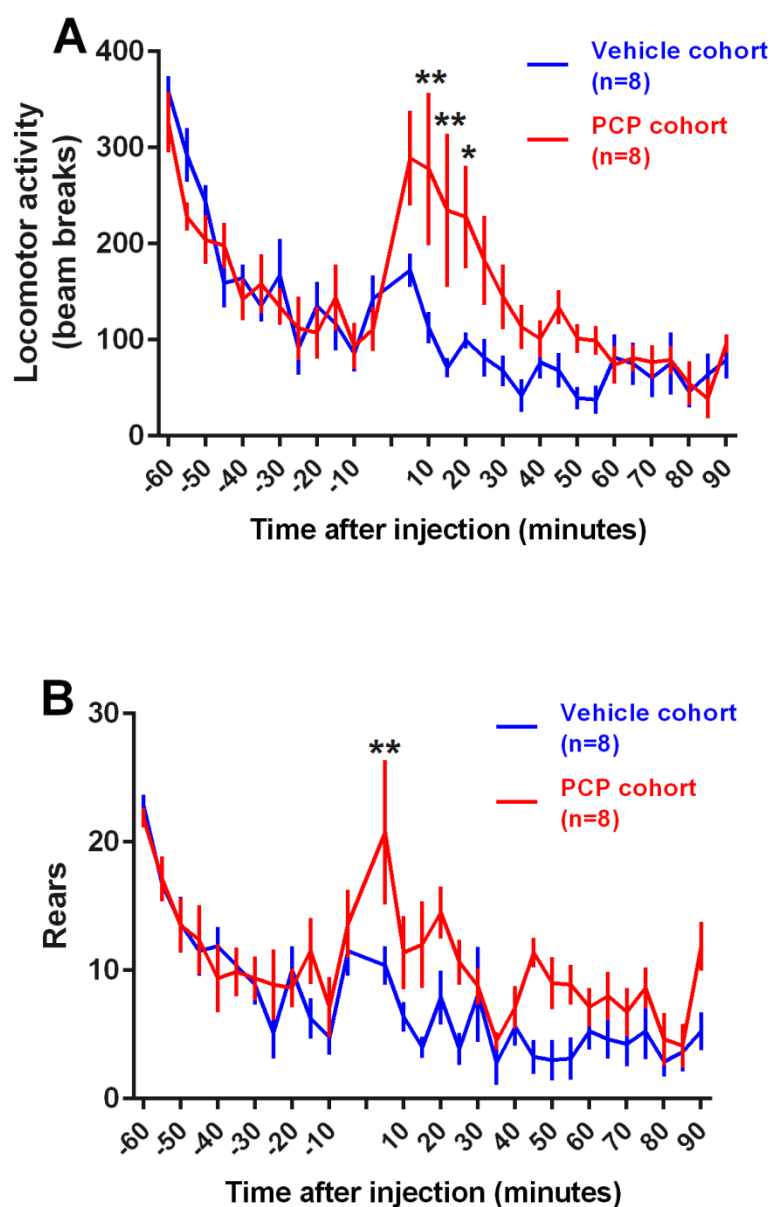


Figure 3.7 The effect of PCP on locomotor activity after receiving PCP during a simple reversal task.

The effect of a single PCP administration (3.2mg/kg i.p.) on (A) ambulatory activity and (B) rears in cohorts of male Lister-hooded rats that had previously received multiple daily doses of PCP (2.5mg/kg) or vehicle (0.9% saline) as part of a simple reversal learning task. Data shown as mean $\pm$ SEM. * $p < 0.05$, ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

3.4 Discussion

The primary goal of this experiment was to determine the effect of PCP on the re-evaluation of learned associations (reversal learning) and the learning of completely novel associations (novel acquisition). The results showed that PCP impaired both processes, but did not affect expression of a learned association (consistent with Chapter 2).

In computational terms, PCP did not affect exploration rate, since it had no effect on CS+ or CS- responding during the pre-reversal treatment period (although ceiling effects may have masked changes in CS+ sampling). On the first dosing session, PCP significantly increased latency measures in some animals. This was likely due to the dissociative anaesthetic effects of the drug causing ataxia, which developed tolerance on subsequent sessions (as reported in other studies; Gilmour et al., 2012). Additionally, on the remaining pre-reversal sessions, PCP decreased latency measures (i.e. accelerated movement). This was likely due to increases in tonic striatal dopamine (Adams and Moghaddam, 1998; Rafter et al., 2016), which promotes behavioural speeding (Niv et al., 2007).

Post-reversal, PCP animals decreased their responding to the former CS+ significantly more slowly than control animals, but were not significantly slower to increase their responding to the former CS-. This suggests that PCP impairs learning in response to negative, but not positive, reward prediction errors. Concordant with this, control animals significantly slowed the speed at which they initiated trials (i.e. appeared to be demotivated) in

the sessions following reversal, which could reflect the amount of negative feedback they were receiving, whereas PCP animals did not.

Most theories propose that a high average reward rate (such as that seen in the pre-reversal sessions) leads to high tonic dopamine release in the ventral striatum (Humphries et al., 2012; Niv et al., 2007). The ventral striatum is part of the limbic loop, and invigorates behaviour in response to Pavlovian cues (such as the magazine light that cues the opportunity to initiate a trial). In the early stages of reversal, the habitually rewarded actions lead to omission of expected reward (negative RPEs) and thus a decrease in ventral striatal dopamine release, which in turn leads to deceleration of all actions leading to reward omission (Hamid et al., 2015). Evidence suggests that PCP and other NMDA-R antagonists potentiate ventral but not dorsal striatal dopamine signalling at low doses (discussed in Marcus et al., 2001), and therefore may disrupt this behavioural slowing.

It should be noted that although PCP caused animals to deliberate the images for less time than controls, they still showed a significant increase in deliberation time induced by the reversal. This suggests that prior negative feedback is being processed and leading to increased image deliberation, but despite this, the animals still persevere with choosing the old CS+.

PCP treatment also impaired learning on the novel acquisition test. However interpreting the result of this test is problematic due to a conditioned bias towards certain images. Specifically, animals whose CS+ was previously Plane had a strong bias towards the novel stimulus Flower, regardless of whether it was the new S+ or S-. Meanwhile in animals whose

CS+ was previously Spider there was no bias. This suggests that the positive valuation of Plane had “contaminated” the valuation of Flower in such a way that the animals continued to sample it habitually. Overcoming this bias would require top-down inhibition by prefrontal circuitry that PCP would be expected to impair the functioning of (Walker et al., 2006). Disregarding the bias, it may be concluded that PCP impairs novel acquisition learning, which is dependent on the associative corticostriatal loop, and this finding has been reported previously (Fellini et al., 2014).

The results of the locomotor test showed that PCP animals were sensitised to the locomotor stimulating effects of PCP, six days after receiving their last dose. Reinforcing drugs are capable of eliciting behavioural sensitisation by enhancing the excitability of VTA dopamine neurons and nucleus accumbens dopamine release (reviewed in Steketee and Kalivas, 2011). NMDA receptor antagonists are known to alter synaptic plasticity via AMPA receptor stimulation (Maeng et al., 2008) and BDNF release (Gideons et al., 2014), both of which are associated with sensitisation in the accumbens (Boudreau and Wolf, 2005; Guillin et al., 2001). It is therefore possible that the effects of PCP on reversal learning were due to these plastic changes induced by repeated dosing rather than solely attributable to acute NMDA receptor antagonism. However, it should be noted that no control treatment was administered during the locomotor test, therefore it cannot be ruled out that the injection procedure itself acted as a contextual cue predictive of reward (since injection was always previously paired with PCP administration in those animals), and so the

enhanced locomotor activity would be due to anticipation of this reward rather than an augmented PCP response.

In conclusion, PCP impaired learning about contingency changes which appeared to be due to perseverance of previously acquired positive associations. The next chapter will examine reversal learning in a way that determines whether this perseveration arises due to a failure to devalue the former CS+ or re-value the former CS-.

Chapter 4 – The effect of phencyclidine on the reversal of positive and negative associations

4.1 Introduction

It has been previously shown that the NMDA-R antagonist PCP impairs simple reversal learning, and this appeared to be due to perseverant choice toward the previously rewarded stimulus (former CS+). However this perseverance may have arisen because of an affinity toward (overvaluation of) the former CS+, and/or an aversion toward (undervaluation of) the former CS-. Because both stimuli are present during the discrimination, simple reversal learning is inadequate for dissociating these factors.

Therefore this experiment sought to determine the effects of PCP on two types of reversal rule change (see Figure 4.1):

- The *perseverance* test – where the conditioned rewarded stimulus (CS+) becomes unrewarded and the conditioned unrewarded stimulus (CS-) is replaced with a novel S+. In order to perform well on this task, one must overcome the perseverance towards the old CS+.
- The *learned irrelevance* test – where the conditioned unrewarded stimulus (CS-) becomes rewarded and the conditioned rewarded stimulus (CS+) is replaced with a novel S-. In order to perform well on this task, one must overcome the learned irrelevance (aka avoidance) of the old CS- (Nilsson et al., 2015).

Both of these tests were followed by a novelty bias test where the current CS- is replaced by a novel S-, in order to confirm that PCP is not affecting the instinctive approach toward novel stimuli by rats.

As per the previous experiment, PCP was also administered prior to rule changes to validate that it does not affect performance of the learned discrimination and prevent contextual effects. The same dose and pre-injection time (2.5mg/kg i.p. 20 minutes prior) was used. At the end of testing, animals underwent a drug-free extinction test (where neither stimuli are rewarded) which is a sensitive test for chronic effects of PCP (Jentsch and Taylor, 2001), in addition to the locomotor assay.

4.2 Methods

4.2.1 Animals

Twenty-four male Lister-hooded rats (delivery weight: 150-200g, CRUK) were delivered, housed, and food restricted according to the conditions described in 2.2.1 Animals.

4.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity boxes

See section 3.2.2 Apparatus.

Compounds

See section 2.2.2 Apparatus.

4.2.3 Touchscreen task protocol

Phases 0 – 4

The first four phases of the reversal learning task were identical to that described in 3.2.3 Touchscreen task protocol, with the exception that S+/S- image assignment included a third stimulus and was randomised according to the following structure:

- n=4 Spider/Plane, n=4 Plane/Spider
- n=4 Spider/Bomb, n=4 Bomb/Spider
- n=4 Bomb/Plane, n=4 Plane/Bomb

Phase 4b (Pre-Reversal – Treated)

This was a continuation of Phase 4 except animals were administered their treatments 20 minutes prior to entering touchscreen chamber (PCP, n=12; 0.9% saline, n=12). Treatment assignment was randomised with counterbalancing across: image assignment, rat performance (slow vs. fast learners), and assigned reversal task. This phase lasted 2 sessions.

Phase 5 (Reversal – Treated)

This was a continuation of Phase 4b, except that prior to the phase change; the S+/S- assignment was changed depending on which arm of the task the rat was assigned (see Figure 4.1). Rats were assigned either to the Perseverance Test (n=6 per treatment) or to the Learned Irrelevance Test (n=6 per treatment). This phase lasted 9 sessions, and was followed by a 2 day washout period (no behavioural testing).

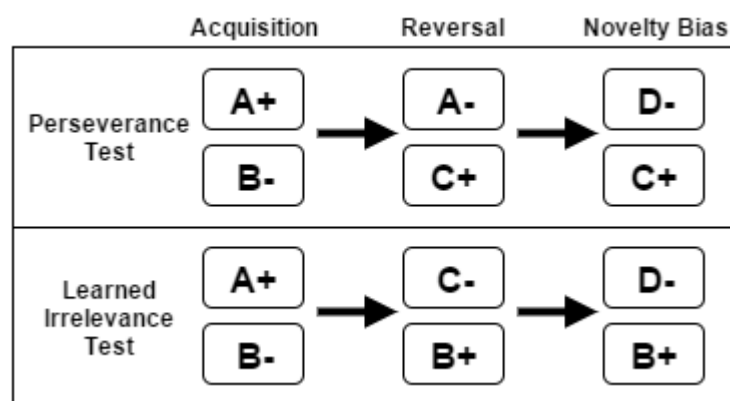


Figure 4.1 Stimulus-reward contingency changes for each phase of the split reversal task.

The reversal phase was split into two tests such that either the old CS+ remained but was now unrewarded (perseverance test), or the old CS- remained and was now rewarded (learned irrelevance test).

Phase 5b (Reversal – No Treatment)

This was a continuation of Phase 5 except no treatment was administered prior to entering the touchscreen chamber. Rats were progressed to the next phase once they had completed >50 trials within 45 minutes and at least 80% of responses were correct for 2 consecutive sessions.

Phase 5c (Pre-Novelty Bias – Treated)

This was a continuation of Phase except animals were administered their treatments 20 minutes prior to entering touchscreen chamber. This phase lasted 2 sessions.

Phase 6 (Novelty Bias – Treated)

This phase was identical to Phase 5c, except that the S- was replaced with a novel S- (Flower). This phase lasted 5 sessions.

Phase 6b (Novelty Bias – No Treatment)

This was a continuation of Phase 6 except no treatment was administered prior to entering the touchscreen chamber. This phase lasted 3 sessions.

Phase 7 (Extinction – No Treatment)

This phase was identical to Phase 6b except that now neither the S+ nor the S- was rewarded. This phase lasted 3 sessions. The following day animals were tested on the locomotor assay.

4.2.4 Locomotor activity protocol

See section 3.2.4 Locomotor activity protocol.

4.2.5 Measures and statistics

See section 3.2.5 Measures and statistics.

4.3 Results

4.3.1 Acquisition phases

For the acquisition phases (0 – 4), rats assigned to PCP treatment took a median of 15.5 sessions (range 14–23) to perform to criteria, whilst rats assigned to vehicle treatment took a median of 16.5 sessions (range 14–27).

4.3.2 Perseverance test

For the two treated sessions prior to reversal, treatment had no effect on the number of correct responses ($F_{(1,10)} = 1.4$, $p = 0.26$) but did significantly affect incorrect responses ($F_{(1,10)} = 6.0$, $p = 0.03$), with the vehicle group committing more errors on the first dosing session ($t_{20} = 3.17$, $p = 0.01$), although the overall interaction between treatment and session was just above the threshold of significance ($F_{(1,10)} = 4.8$, $p = 0.052$). This may be an

artefact of randomisation since the PCP group were committing less errors in the sessions prior to treatment ($p = 0.15$, data not shown).

For the treated reversal phase, there was no effect of treatment on the number of correct responses, i.e. choosing the novel S+ ($F_{(1, 10)} = 1.5$, $p = 0.25$), but a significant overall effect of treatment on the number of incorrect responses, i.e. PCP animals more frequently chose the old S+ ($F_{(1, 10)} = 16.1$, $p = 0.003$). Although there was no treatment $\times$ session interaction ($F_{(8, 80)} = 0.27$, $p = 0.98$), i.e. session did not alter the direction or magnitude of treatment effects overall, post-hoc tests showed significant treatment differences during reversal sessions 2–4 and 6–8 (all $p < 0.05$, see Figure 4.2). In addition, the PCP group took significantly longer to reach criterion after the reversal ($\chi^2 = 4.92$, d.f. = 1, $p = 0.03$; see Figure 4.3).

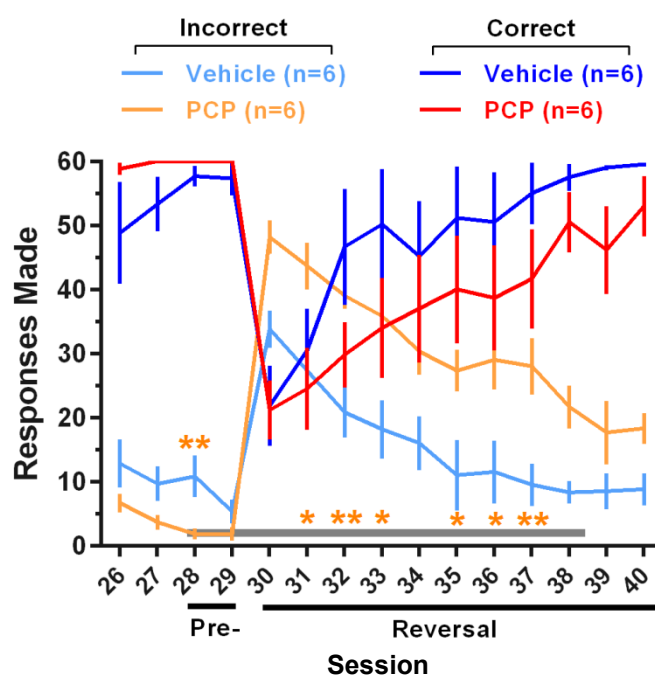


Figure 4.2 Responses made during the perseverance test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the perseverance test **in male Lister-hooded rats**. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). * $p < 0.05$, ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

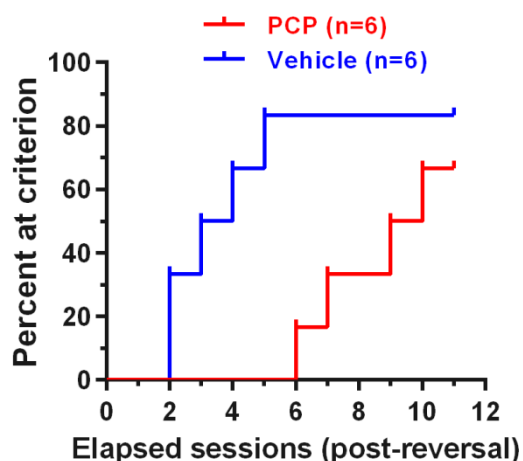


Figure 4.3 Sessions to criterion during the perseverance test for PCP or vehicle treated animals.

A survival curve showing the effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the perseverance test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

4.3.2 Learned Irrelevance test

One animal assigned to the vehicle group failed to reach criterion and so was excluded from analysis. For the two treated sessions prior to reversal, treatment had no effect on the number of correct ($F_{(1,9)} = 1.2$, $p = 0.30$) or incorrect responses ($F_{(1,9)} = 1.15$, $p = 0.32$).

For the treated reversal phase, there was no significant effect of treatment on the number of correct ($F_{(1,9)} = 0.27$, $p = 0.62$) or incorrect responses ($F_{(1,9)} = 1.74$, $p = 0.22$; see Figure 4.4) – i.e. PCP did not affect sampling of the novel S- or old S-. There was also no effect of treatment on latency to reach criterion after the reversal ($\chi^2 = 0.08$, d.f. = 1, $p = 0.78$; see Figure 4.5).

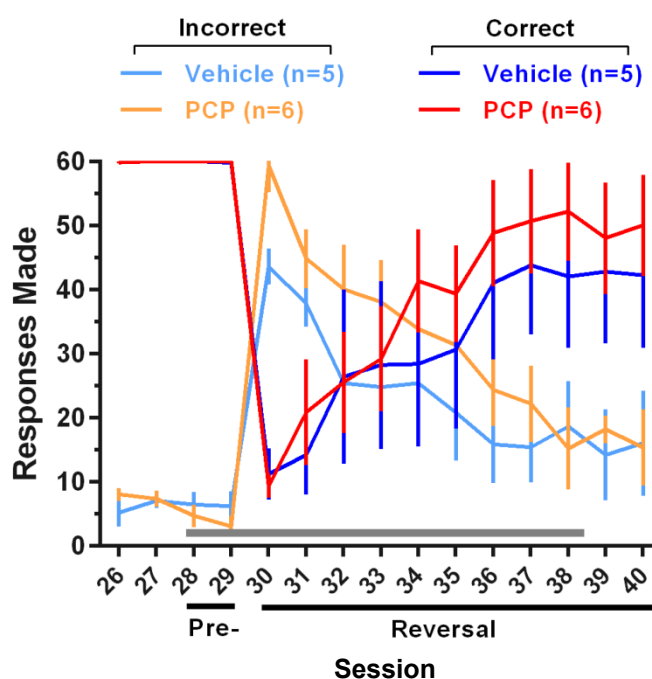


Figure 4.4 Responses made during the learned irrelevance test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the learned irrelevance test **in male Lister-hooded rats**. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session).

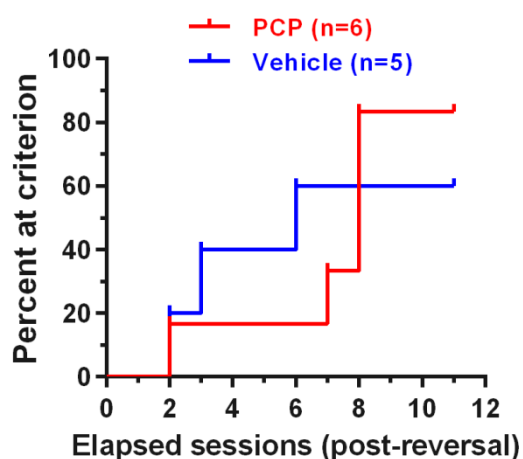


Figure 4.5 Sessions to criterion during the learned irrelevance test for PCP or vehicle treated animals.

A survival curve showing the effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the learned irrelevance test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

4.3.3 Novelty Bias

One animal from each treatment group failed to reach criterion prior to this phase and so were excluded from analysis. There was no effect of treatment on the number of incorrect responses during the novelty bias test ($F_{(1, 20)} = 0.4$, $p = 0.53$), i.e. PCP did not increase or decrease approach to the novel S-. There was a significant effect on the number of correct responses ($F_{(1, 20)} = 5.6$, $p = 0.03$), i.e. PCP animals sampled the S+ more, and although there was no overall significant treatment $\times$ session interaction ($F_{(4, 80)} = 2.2$, $p = 0.08$) post-hoc tests showed a significant difference for the first session of the novelty bias test ($t_{100} = 3.3$, $p = 0.006$, see Figure 4.6). Further subgroup analysis showed that there was a significant effect of treatment on

correct responses in animals who had previously experienced the learned irrelevance test ($F_{(1,9)} = 6.6$, $p = 0.03$) but not in those who had experienced the perseverance test ($F_{(1,9)} = 0.7$, $p = 0.4$; see Figure 4.7). The number of incorrect responses (sampling of the novel S-) was not affected by prior experience with the perseverance test ($F_{(1,9)} = 0.4$, $p = 0.56$) or learned irrelevance test ($F_{(1,9)} = 2.8$, $p = 0.13$).

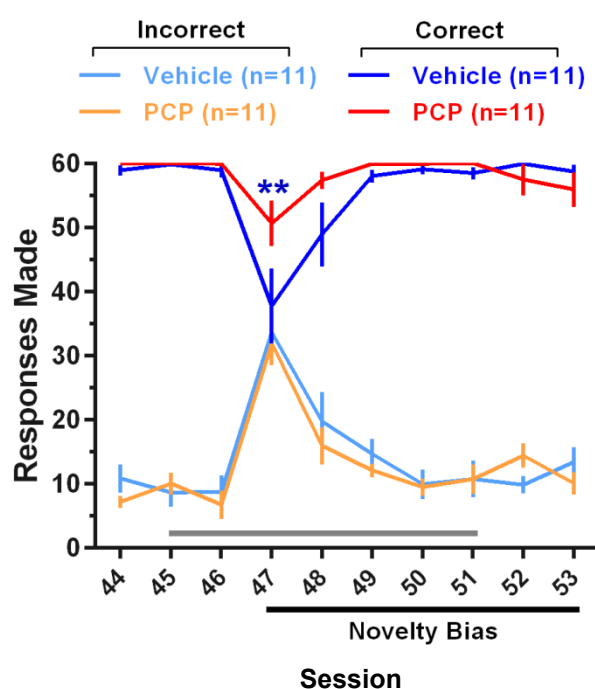


Figure 4.6 Responses made during the novelty bias test for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novelty bias test **in male Lister-hooded rats**. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

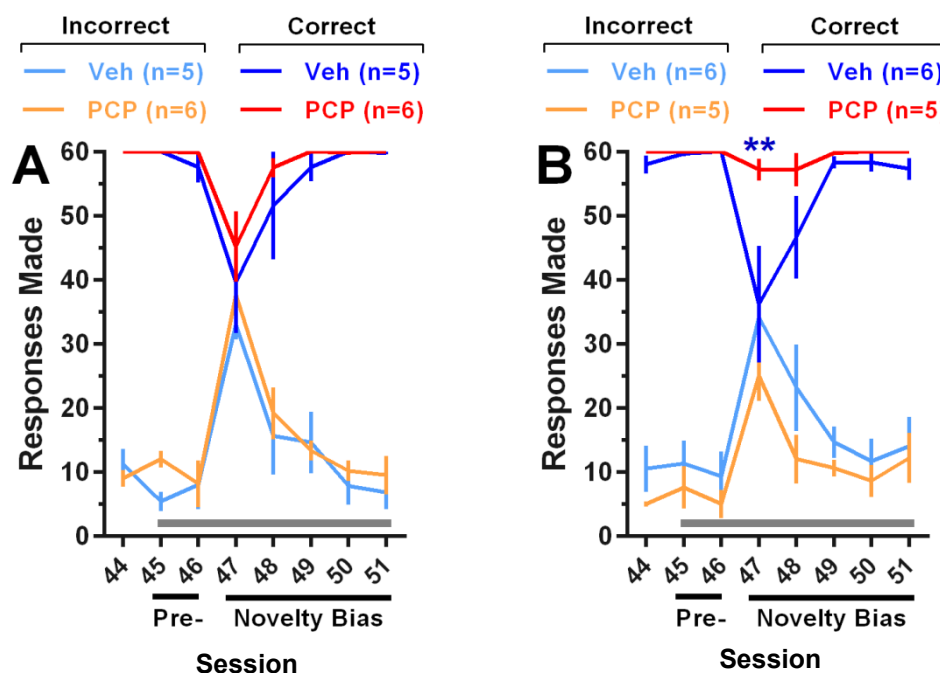


Figure 4.7 Responses made during the novelty bias test for PCP or vehicle treated animals, split by prior reversal contingency.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novelty bias test **in male Lister-hooded rats** who had previously undergone (A) the perseverance test, or (B) the learned irrelevance test. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). ** $p < 0.01$. Post-hoc Sidak test following two-way RM-ANOVA.

4.3.4 Extinction

During the non-treated extinction phase, there was no effect of treatment cohort on the number of responses to the formerly rewarded stimulus ($F_{(1, 21)} = 0.17$, $p = 0.68$) but there was an overall difference in the number of responses to the formerly unrewarded stimulus in the PCP cohort ($F_{(1, 21)} = 4.59$, $p = 0.04$, see Figure 4.8), i.e. PCP animals decreased responding to the S- during extinction.

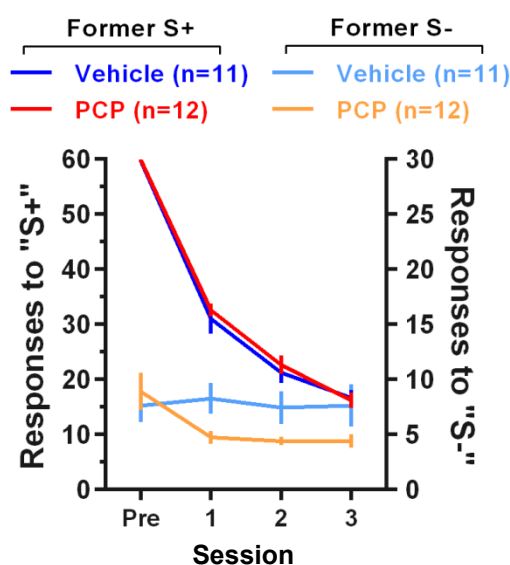


Figure 4.8 Responses made during the extinction for PCP or vehicle treated animals.

The number (mean $\pm$ SEM) of responses made to the formerly correct (S+) and incorrect (S-) stimulus for each session (x-axis) during the extinction test, by cohorts of **male Lister-hooded rats** that previously received PCP (n=12) or vehicle (n=11).

4.3.5 Latency Measures

During the two pre-reversal treated sessions, there was no effect of treatment or session on any latency measure (all $F_s < 3.2$, $p > 0.08$). There was a significant interaction between treatment and session on response latency ($F_{(1, 21)} = 6.0$, $p = 0.02$). These findings are similar to Chapter 3 except there was no evidence of ataxia on the first dosing day.

During the nine treated reversal sessions, there was a significant effect of session on initiation latency ($F_{(8, 168)} = 2.4$, $p = 0.02$). However unlike during the simple reversal test (Chapter 3), there was no significant within-treatment increase during any session for either treatment. There was a significant effect of treatment on initiation latency ($F_{(1, 21)} = 7.9$, $p = 0.01$), i.e.

PCP animals were quicker to initiate, with post-hoc tests showing a between-treatment difference on reversal session 7 ($t_{189} = 2.8$, $p = 0.046$; see Figure 4.9). There was no significant treatment $\times$ session interaction ($F_{(8, 168)} = 0.6$, $p = 0.8$).

Similarly, there was a significant effect of session on response latency ($F_{(8, 168)} = 2.3$, $p = 0.02$), and although no overall significant session $\times$ treatment interaction ($F_{(8, 168)} = 0.97$, $p = 0.46$) post-hoc tests showed significant within-treatment increases during reversal sessions 2 ($t_{210} = 2.99$, $p = 0.03$) and 6 ($t_{210} = 2.89$, $p = 0.04$) in the vehicle treated animals compared to the session prior to reversal. There were no significant within-treatment increases in the PCP treated animals. There was a significant effect of treatment on response latency ($F_{(1, 21)} = 13.3$, $p = 0.0015$), i.e. PCP animals deliberated the images for less time, and a significant treatment $\times$ session interaction ($F_{(10, 210)} = 2.7$, $p = 0.005$) with post-hoc tests showing between-treatment differences on reversal sessions 3–6 (all $t > 3.1$, $p < 0.02$) and session 8 ($t_{189} = 3.01$, $p = 0.027$). There was also a significant treatment effect on collection latency ($F_{(1, 21)} = 9.9$, $p = 0.005$), i.e. PCP animals collected pellets more rapidly, and a significant treatment $\times$ session interaction ($F_{(10, 210)} = 2.284$, $p = 0.015$) with a significant between-treatment effect on reversal session 1 ($t_{189} = 3.8$, $p = 0.002$).

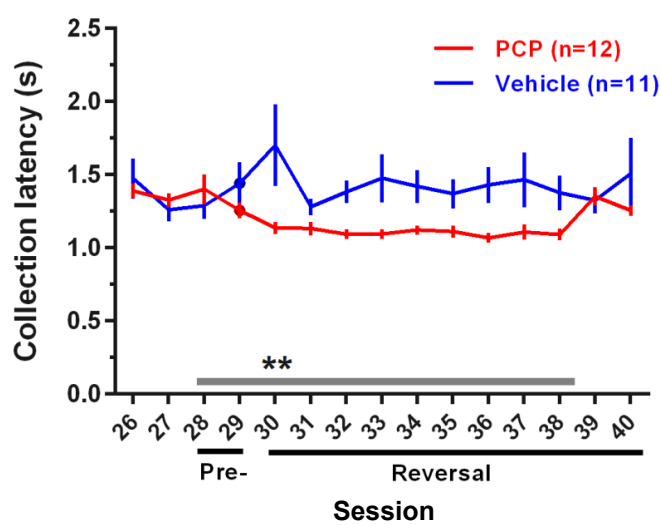
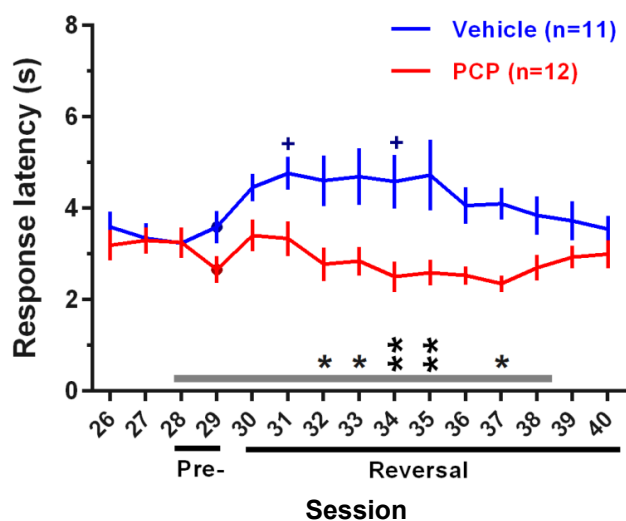
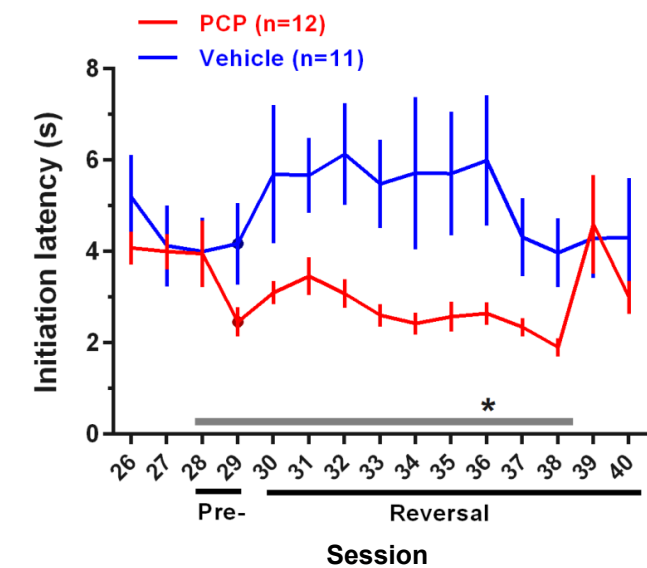


Figure 4.9 Latency measures during the perseverance and learned irrelevance test (collated) for PCP or vehicle treated animals.

The effects of PCP (2.5mg/kg) or vehicle (0.9% saline) on the trial latencies (mean $\pm$ SEM) for each session (x-axis) during the perseverance and learned irrelevance tests (collated) in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). * $p < 0.05$, ** $p < 0.01$ versus vehicle. + $p < 0.05$ versus pre-reversal session 29. Post-hoc Sidak test following two-way RM-ANOVA.

Across extinction sessions, animals took longer to initiate trials ($F_{(2, 44)} = 22.96$, $p < 0.0001$) and respond to the images ($F_{(2, 44)} = 11.63$, $p < 0.0001$). There was no effect of previous treatment on initiation latency ($F_{(1, 22)} = 0.02$, $p = 0.88$) or response latency ($F_{(1, 22)} = 0.51$, $P = 0.48$).

4.3.6 Locomotor activity

In the activity boxes during the 60 min prior to PCP administration, there was a significant effect of time on ambulation ($F_{(11, 242)} = 44.9$, $p < 0.0001$) and rears ($F_{(11, 242)} = 19.9$, $p < 0.0001$) in all animals. There was no significant difference in ambulation ($F_{(1, 22)} = 0.46$, $p = 0.5$) or rears ($F_{(1, 22)} = 0.41$, $p = 0.53$) between the cohort of rats that received PCP or vehicle during the study (see Figure 4.10). In the 90 minute period following PCP administration, the PCP cohort showed a significantly greater ambulatory response ($F_{(1, 22)} = 29.1$, $p < 0.0001$) and more rears ($F_{(1, 22)} = 4.9$, $p = 0.04$; see Figure 4.10). There was a significant treatment $\times$ time interaction for ambulation ($F_{(17, 374)} = 5.8$, $p < 0.0001$) but not rears ($F_{(17, 374)} = 0.84$, $p = 0.64$). This data replicates the findings of Chapter 3.

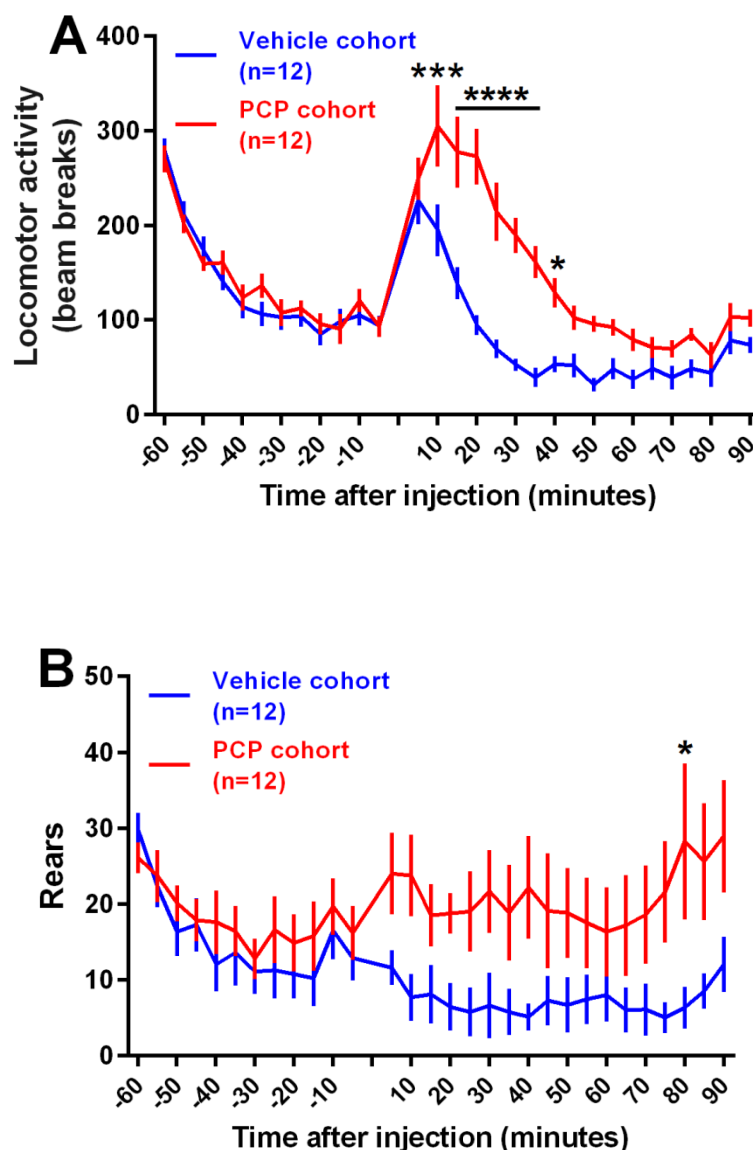


Figure 4.10 The effect of PCP on locomotor activity after receiving PCP during a perseverance or learned irrelevance task.

The effect of a single PCP administration (3.2mg/kg i.p.) on (A) ambulatory activity and (B) rears in cohorts of male Lister-hooded rats that had previously received multiple daily doses of PCP (2.5mg/kg) or vehicle (0.9% saline) as part of a perseverance/learned irrelevance task. Data shown as mean $\pm$ SEM. * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$, post-hoc Sidak test following two-way RM-ANOVA.

4.4 Discussion

This experiment sought to determine the mechanism by which PCP impairs simple reversal learning by examining the effects of PCP on tests where either the old CS+ remains (perseverance) or the old CS- remains (learned irrelevance).

The results showed that PCP significantly slows learning in the perseverance task but not the learned irrelevance task, implying that extinction of previously positive associations is impaired but re-evaluation of previous negative/neutral associations is not. This provides further evidence towards the idea proposed in Chapter 3 that PCP impairs negative, but not positive, prediction error signalling and/or processing, **an effect also seen with ketamine (Rychlik et al., 2017)**. However there are alternative interpretations because of nuanced differences between each arm of the task (see section 9.2.1).

By what neural mechanism might PCP selectively impair learning from negative feedback? One theory proposes that NMDA receptor antagonists selectively alter GABAergic interneuron firing in lateral OFC circuits important for processing non-reward (see Figure 1.8; Quirk et al., 2009). Alternately, NMDA receptor antagonists have been shown to increase burst firing of dopaminergic neurons in the VTA (Belujon and Grace, 2014), which would in turn decrease the signal-to-noise ratio of phasic dips in VTA firing that signal negative RPEs, and thus the magnitude of negative RPEs.

Interestingly, when the CS- was replaced with a novel S- (novelty bias task), all animals increased their exploration of the novel stimulus; however

only PCP-treated animals who had previously experienced the learned irrelevance test continued to exploit the old – and current – CS+, whereas the other groups slightly decreased their sampling of the CS+.

One interpretation of this effect is that the negative feedback from sampling the novel S- led to a generalised demotivational spillover effect. This also adequately explains why the control animals' latency measures slowed slightly, but fails to explain why only PCP-treated animals who performed the perseverance task also extinguished CS+ responding.

To explain this, a further speculative interpretation of learning transfer is required. For animals that previously experienced the perseverance test, the novelty bias test presents a familiar contextual change – a choice between the CS+ and a novel stimulus. Previously, these animals learned that when a novel stimulus is introduced, the CS+ becomes unrewarded. If this model of the task was implemented during novelty bias, these animals would automatically devalue the CS+ and extinguish their responding to it – which is exactly what occurred. Importantly, implementation of this model does not require inference/goal-directed tree-search since context acts as a cached – not computed – modulator; therefore PCP would not be expected to disrupt this process.

Meanwhile for animals that previously experienced the learned irrelevance test, the novelty bias test is less familiar, since in the last context where a novel stimulus was introduced, the CS+ was not an available choice. In this novel context, both groups of animals explored the S-, as would be expected. However PCP animals continued to exploit the CS+ (perseverance)

whereas control animals decreased exploitation of the CS+. Control animals may have erroneously inferred that the novel context meant that the value of the CS+ had changed, which would be a model-based, OFC-dependent (and thus PCP-disrupted) behaviour.

On the other hand, during the drug-free extinction test (where no contextual change occurs but the CS+ becomes unrewarded), both cohorts rapidly extinguished responding to the former CS+, but only the PCP cohort decreased their responding to the CS-. This suggests that PCP withdrawal induces the opposite effects of acute PCP, i.e. it enhances demotivation, although there were no significant differences in any latency markers indicative of motivation. Alternately, it is possible that prior PCP had neurotoxic effects on the OFC (Dawson et al., 2014; Elsworth et al., 2011) causing misattribution of RPE signalling so that negative RPEs in response to the old CS+ were also attributed to the CS- (see Walton et al., 2010).

Concordant with Chapter 3, PCP caused behavioural acceleration during the task. However in the current experiment, PCP treatment significantly attenuated the reversal-induced increase in image deliberation time. It is possible that latency differences arose because of task conditions – both reversal tests differed in the amount of negative feedback received compared to simple reversal learning; both tasks also introduced novel stimuli which may have provided a contextual cue or increased the salience of the task, and thus motivation (see Chapter 9 for other interpretations). Also consistent with Chapter 3 was the effect of prior PCP treatment on locomotor sensitisation to acute PCP challenge.

In conclusion, PCP impairs reversal learning by delaying the extinction of previously positive associations without affecting the re-evaluation of previously negative associations, which may be due to impaired negative RPE processing in the sensorimotor loop. The following chapter will examine the effects of a serotonin receptor antagonist on this form of reversal learning.

Chapter 5 – The effect of SB-242084 on the reversal of positive and negative associations

5.1 Introduction

In the previous chapter it was shown that PCP impairs reversal learning by delaying extinction of old positive associations without affecting negative associations. The orbitofrontal cortex is known to play a critical role in facilitating flexible learning by maintaining action-outcome expectancies and inferring new states of the environment/task (see section 1.5.2 Role of OFC). Intra-OFC infusion of NMDA-R antagonists is sufficient to impair reversal learning (Bohn et al., 2003; Brigman et al., 2013), and thus is a likely target of systemic PCP administration. PCP is also known to induce efflux of dopamine and serotonin (5-HT) in prefrontal regions (Quarta and Large, 2011). As previously mentioned (see Chapter 1), both of these neurotransmitters play important reinforcement signalling roles, and 5-HT has especially been implicated in OFC-dependent reversal learning.

Pharmacological studies in rodents have shown improvements in reversal learning with systemic 5-HT_{2c}R antagonism (Boulougouris et al., 2007b), and 5-HT_{2c}R blockade in the lateral OFC (but not medial PFC or nucleus accumbens) is sufficient for this effect (Boulougouris and Robbins,

2010). However other studies using different task conditions have found mixed results, including impairments (see Table 5.1). In addition, 5-HT_{2c}R agonists have been shown to promote compulsive behaviour in primates and rodents (reviewed in Fineberg et al., 2009).

Therefore an experiment was conducted to determine the effects of the selective 5-HT_{2c}R antagonist SB-242084 on the touchscreen perseverance and learned irrelevance task, as per Chapter 4. It was hypothesised that SB-242084 would enhance learning on the learned irrelevance test, as seen in Nilsson et al., 2012. A dose of 1mg/kg i.p. 20 minutes prior was chosen as this was sufficient to produce effects in previous studies.

Table 5.1 Effects of the selective 5-HT_{2c}R antagonist SB-242084 in rodent reversal learning tasks.

Dose	Route	Reversal task	ITI	Outcome
Boulougouris et al., 2007b:				
0.1-1 mg/kg	i.p.	2-choice, initial within-session reversal, distractor	10s	Decreased trials and errors to criterion on first reversal only.
Boulougouris and Robbins, 2010:				
0.1-1 µg/side	Intra-OFC	2-choice, initial within-session reversal, distractor	10s	Decreased trials and errors to criterion. Decreased perseveration and learning errors.
Nilsson et al., 2012:				
0.5 mg/kg	s.c.	2-choice, between-session reversal	15s	Decreased trials to criterion on first two reversals only in tasks where former S- was present.
Nilsson et al., 2013:				
0.5 mg/kg	s.c.	2-choice, between-session reversal	40s	No change in trials to criterion on simple reversal. Decreased trials to criterion when former S+ is present. Increased trials to criterion when former S- present.
Alsö et al., 2015:				
0.1-1 mg/kg	i.p.	3-choice, between-session reversal	5s	Increased errors towards previous S+ and constant S-. Increased errors during late reversal but decreased during early reversal. Increased overall errors and trials to criterion.
0.1-1 mg/kg	i.p.	2-choice, between-session reversal	5s	Increased late errors and decreased early errors. No overall effect on other measures.
1-3 µg/side	Intra-OFC	2-choice, between-session serial reversal	5s	Decreased errors during “perseveration” phases of reversal. No overall effect on other measures.

5.2 Methods

5.2.1 Animals

Twenty-four male Lister-hooded rats (delivery weight: 175-200g, CRUK) were delivered, housed, and food restricted according to the conditions described in section 2.2.1 Animals.

5.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity boxes

See section 3.2.2 Apparatus.

Compounds

SB-242084 dihydrochloride (Tocris, UK) was dissolved in 6% Tween80 / 0.9% saline solution. All solutions were made on the day of administration and administered at 1ml/kg via the intraperitoneal (i.p.) route, twenty minutes prior to touchscreen testing.

5.2.3 Touchscreen task protocol

All phases of the touchscreen reversal learning task were identical to that described in 4.2.3 Touchscreen task protocol, with the drug treatment group being administered SB-242084. Animals were not given an extinction phase or tested for locomotor activity as there is no evidence for chronic/sensitisation effects of SB-242084 treatment.

5.2.4 Measures and statistics

See section 3.2.5 Measures and statistics.

5.3 Results

5.3.1 Acquisition phases

For the acquisition phases (0 – 4), rats assigned to SB-242084 treatment took a median of 14 sessions (range 10–21) to perform to criteria, whilst rats assigned to vehicle treatment took a median of 12 sessions (range 10–26).

5.3.2 Perseverance test

One animal in the SB-242084 treatment group **was culled** during the reversal stage and so was excluded from RM-ANOVA analysis but had reached criterion and so was included in the latency to criterion analysis.

For the two treated sessions prior to reversal, treatment had no effect on the number of correct responses ($F_{(1, 9)} = 0.4$, $p = 0.6$) or incorrect responses ($F_{(1, 9)} = 2.6$, $p = 0.14$). For the treated reversal phase, there was no effect of treatment on number of correct responses ($F_{(1, 9)} = 0.35$, $p = 0.6$), or incorrect responses ($F_{(1, 9)} = 1.7$, $p = 0.23$; see Figure 5.1). In addition, there was no significant difference between treatments in the sessions taken to reach criterion after the reversal ($\chi^2 = 0.96$, d.f. = 1, $p = 0.3$; see Figure 5.2).

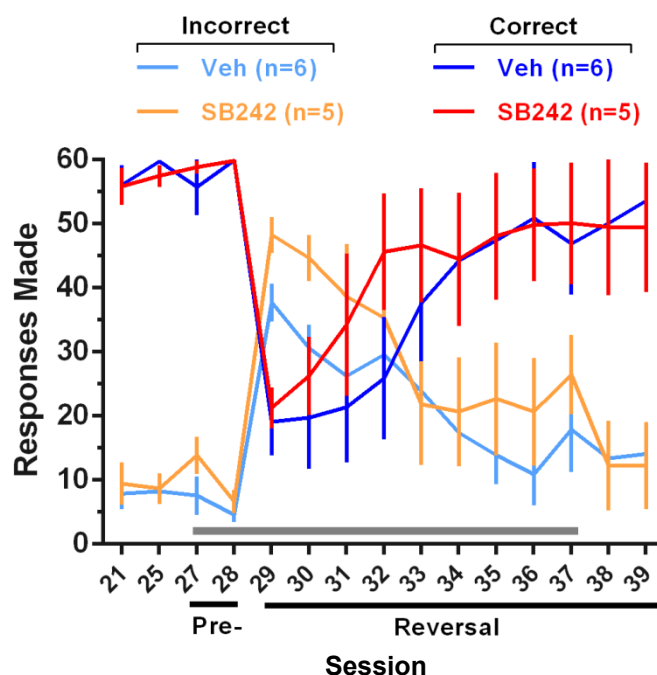


Figure 5.1 Responses made during the perseverance test for SB-242084 or vehicle treated animals.

The effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the perseverance test in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session).

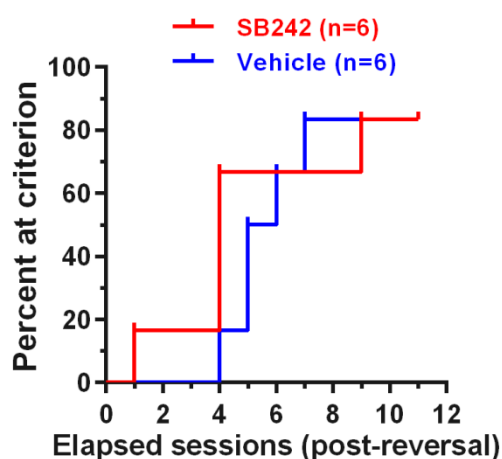


Figure 5.2 Sessions to criterion during the perseverance test for SB-242084 or vehicle treated animals.

A survival curve showing the effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the perseverance test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

5.3.3 Learned Irrelevance test

For the two treated sessions prior to reversal, treatment had no effect on the number of correct responses ($F_{(1, 10)} = 0.55$, $p = 0.5$) or incorrect responses ($F_{(1, 10)} = 0.54$, $p = 0.5$). For the treated reversal phase, there was no effect of treatment on number of correct responses ($F_{(1, 10)} = 0.96$, $p = 0.35$), or incorrect responses ($F_{(1, 10)} = 0.13$, $p = 0.7$; see Figure 5.3). In addition, there was no significant difference between treatments in the sessions taken to reach criterion after the reversal ($\chi^2 = 0.23$, d.f. = 1, $p = 0.6$; Figure 5.4).

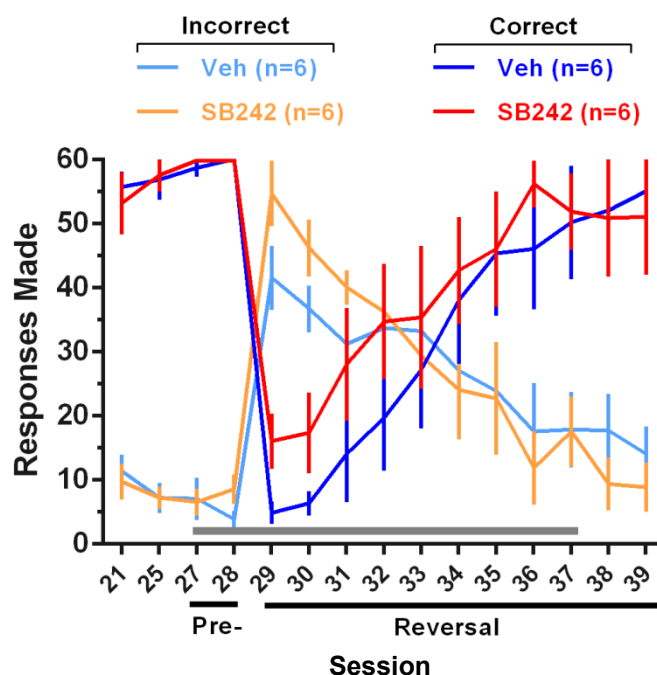


Figure 5.3 Responses made during the learned irrelevance test for SB-242084 or vehicle treated animals.

The effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the learned irrelevance test in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session).

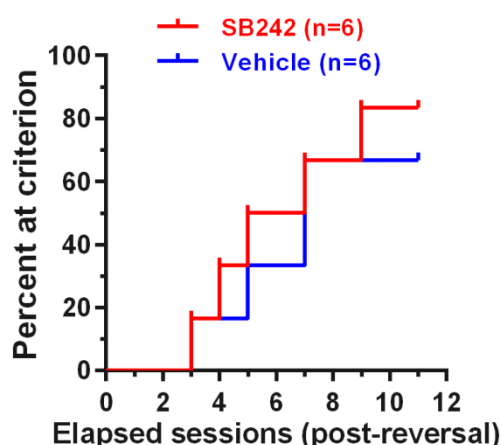


Figure 5.4 Sessions to criterion during the learned irrelevance test for SB-242084 or vehicle treated animals.

A survival curve showing the effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the learned irrelevance test. Treatments were administered i.p. at 1ml/kg, 20 minutes prior to sessions 0–9.

5.3.4 Novelty Bias

One animal in the SB-242084 group failed to reach criterion prior to this phase and so was excluded from analysis. There was no effect of treatment on the number of incorrect responses during the novelty bias test ($F_{(1, 20)} = 1.1$, $p = 0.31$), i.e. it did not affect approach to the novel stimulus. There was however a significant effect on the number of correct responses – i.e. the SB-242084 group sampled the S+ more ($F_{(1, 20)} = 4.5$, $p = 0.05$), and a significant treatment $\times$ session interaction ($F_{(4, 80)} = 3.8$, $p = 0.008$) with post-hoc tests showing a significant difference for the first ($t_{100} = 3.25$, $p = 0.008$) and second ($t_{100} = 2.8$, $p = 0.03$) session of the novelty bias test (see Figure 5.5). Further subgroup analysis did not show this effect on correct responses to be specific to animals who had previously received the perseverance test

($F_{(1,8)} = 3.0$, $p = 0.12$) or learned irrelevance test ($F_{(1,10)} = 1.6$, $p = 0.23$). However, caution should be made in comparing the effects of SB-242084 during the novelty bias test to PCP (see section 4.3.3 Novelty Bias) since the perseverance test group was underpowered ($n=4$).

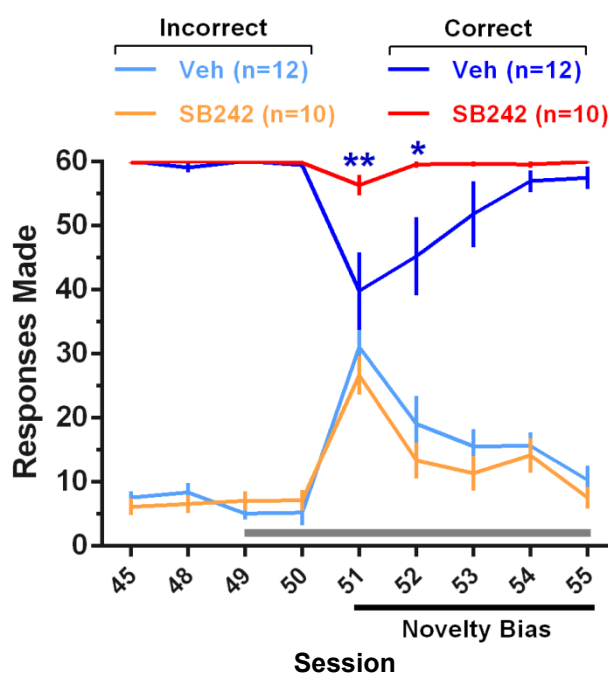


Figure 5.5 Responses made during the novelty bias test for SB-242084 or vehicle treated animals.

The effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novelty bias test **in male Lister-hooded rats**. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). * $p < 0.05$, ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

5.3.5 Latency Measures

During the two treated pre-reversal sessions, there was no effect of treatment or session on any latency measure (all $F_s < 3$, $p > 0.1$).

During the nine treated reversal sessions, there was a significant effect of session on initiation latency ($F_{(8, 168)} = 7.0, p < 0.0001$) and response latency ($F_{(8, 168)} = 11.0, p < 0.0001$), consistent with previous experiments showing reversal-induced task deceleration. The treatment $\times$ session interaction for initiation latency was insignificant ($F_{(8, 168)} = 1.7, p = 0.1$), however post-hoc tests showed a significant within-treatment increase during reversal sessions two ($t_{210} = 4.8, p < 0.0001$) and three ($t_{210} = 5.0, p < 0.0001$) in the vehicle treated animals compared to the session prior to reversal. SB-242084-treated animals showed no significant within-treatment increases during any reversal sessions (all $p > 0.7$), similar to that seen previously with PCP treatment.

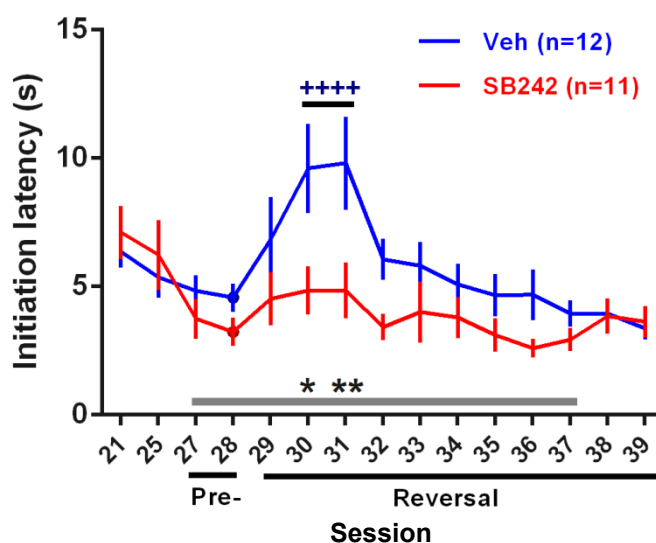
Consistent with this, there was a significant effect of treatment alone on initiation latency ($F_{(1, 21)} = 5.4, p = 0.03$), and although no significant treatment $\times$ session interaction ($F_{(8, 168)} = 1.7, p = 0.1$), post-hoc tests showed significant between-treatment effects on the second ($t_{189} = 3.2, p = 0.015$) and third ($t_{189} = 3.3, p = 0.009$) reversal sessions. This suggests that vehicle treated animals were less motivated to initiate trials after reversal (when lots of negative feedback was being received) whilst SB-242084 animals were not.

Similarly, there was a significant treatment $\times$ session interaction for response latency ($F_{(8, 168)} = 2.2, p = 0.03$), with post-hoc tests showing significant within-treatment increases during reversal sessions 1 – 4 (all $p < 0.05$; see Figure 5.6) for the vehicle-treated animals, compared to the pre-

reversal session. There were no significant within-treatment increases for the SB-242084-treated animals (all $p > 0.2$).

There was a significant effect of treatment on response latency ($F_{(1, 21)} = 5.6$, $p = 0.03$), with post-hoc tests showing significant between-treatment effects on reversal sessions 2 – 4 (all $p < 0.05$), suggesting that SB-242084-treated animals failed to increase image deliberation time in response to the negative feedback received during reversal.

There was no effect of session ($F_{(8, 168)} = 1.7$, $p = 0.1$) or any treatment $\times$ session interaction ($F_{(8, 168)} = 0.6$, $p = 0.8$) on collection latency, consistent with previous experiments. There was a significant effect of treatment ($F_{(1, 21)} = 13.05$, $p = 0.002$) – i.e. SB-242084 accelerated reward collection – with post-hoc significance on the reversal sessions 1 and 5 – 9 (all $p < 0.05$; see Figure 5.6).



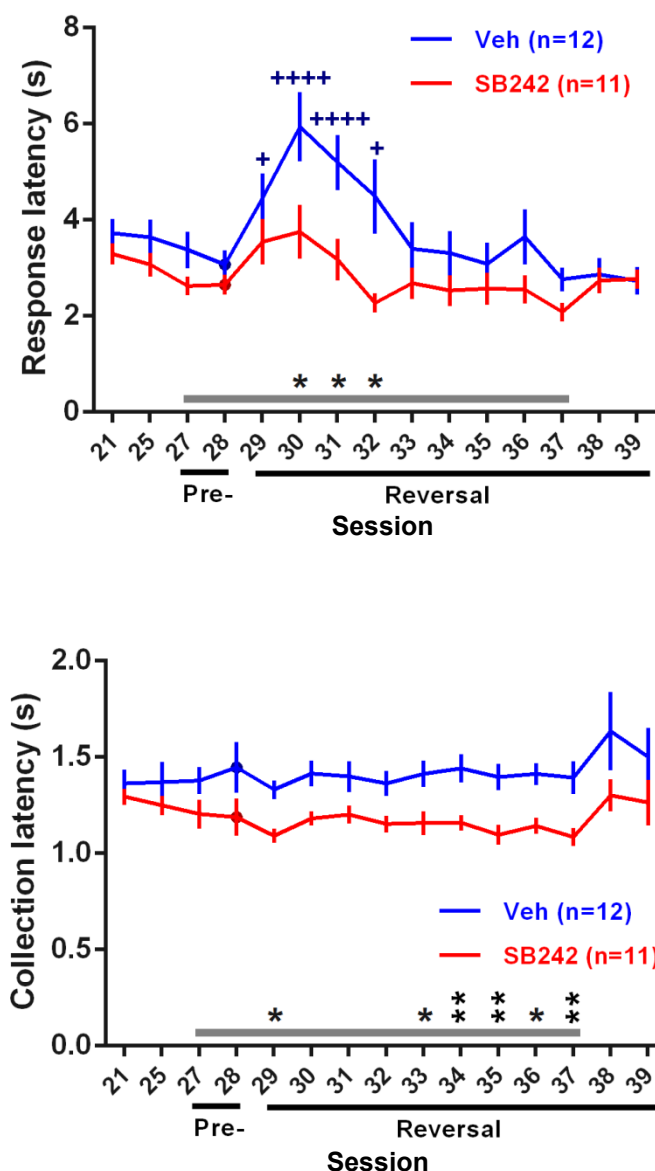


Figure 5.6 Latency measures during the perseverance and learned irrelevance test (collated) for SB-242084 or vehicle treated animals.

The effects of SB-242084 (1mg/kg) or vehicle (6% Tween80 in 0.9% saline) on the trial latencies (mean $\pm$ SEM) for each session (x-axis) during the perseverance and learned irrelevance tests (collated) in male Lister-hooded rats. Grey bar indicates treatment sessions (administered i.p., 1ml/kg; 20 minutes prior to the session). * $p < 0.05$, ** $p < 0.01$ versus vehicle. + $p < 0.05$, ++++ $p < 0.0001$ versus pre-reversal session 29. Post-hoc Sidak test following two-way RM-ANOVA.

5.4 Discussion

This experiment investigated the effect of the 5-HT_{2c}R antagonist SB-242084 on reversal learning tests that dissociate perseverance towards the former CS+ from learned irrelevance/avoidance of the former CS-.

The results showed that SB-242084 did not affect reversal learning rate in anyway, despite accelerating behaviour during the task. SB-242084 did not affect exploration rate since there was no change in choice behaviour during the treated pre-reversal period. Likewise, there were no significant differences in choice sampling between the treatments during reversal.

During the novelty bias task, control animals showed a similar choice profile to Chapter 4, namely increased exploration of the novel S- and decreased exploitation of the CS+. SB-242084-treated animals also increased exploration of the novel S- but did not decrease exploitation of the CS+, similar to PCP-treated animals in Chapter 4. SB-242084 treatment also induced behavioural acceleration similar to PCP. It is suggested that this occurs via 5-HT_{2c}R antagonism in the VTA leading to ventral striatal dopamine release (Alex and Pehek, 2007). Importantly, this speeding effect may be more specific to conditions where reward is available, i.e. is the result of an increase in goal-directed effort (independent of appetite, see Bailey et al., 2015), whereas PCP causes a generalised motor hyperactivity regardless of reward availability.

Overall, the finding that SB-242084 does not affect reversal learning and increases exploitation of the CS+ during the novelty bias test is in opposition to the hypothesis of Fineberg et al., 2009 stating that 5-HT_{2c}R antagonism

increases choice impulsivity. Likewise, our result is not consistent with numerous other studies (see Table 5.1) showing (ambivalent) effects on reversal learning.

Consolidating these discrepant findings is difficult since all of the studies used slightly different experimental parameters, reported different measures, and analysed the data in different ways. A consistent finding is that SB-242084 induces motoric impulsivity, i.e. speed of responding, which indicates that psychoactive doses of the drug were administered. It has also been further proposed that SB-242084 enhances choice impulsivity (Fineberg et al., 2009), i.e. randomness of responding, which facilitates early reversal learning by “breaking” compulsive habits. Inexplicably though, no effect on choice behaviour is found prior to reversal. It is therefore suggested that SB-242084 enhances motoric impulsivity via enhanced meso-accumbens dopamine signalling, and this occurs in response to Pavlovian cues (i.e. during all phases of the task). Meanwhile SB-242084 only enhances choice impulsivity during reversal learning, i.e. in response to negative prediction errors. It is during these contexts that OFC 5-HT signalling is important, and therefore most sensitive to modulation via 5-HT_{2c}R antagonism. Depending on experimental task conditions and type of data analysis (see Table 1.3), this impulsivity may facilitate, impair, or not affect choice behaviour.

In conclusion, the 5-HT_{2c}R antagonist SB-242084 did not affect choice behaviour after rule reversals where either the old CS+ remained or the old CS- remained, but did increase task motivation.

Chapter 6 – The effect of dual-hit neonatal PCP/post-weaning social isolation on the reversal of positive and negative associations

6.1 Introduction

So far it has been established that choice and motor behaviour during reversal learning are acutely affected by drugs targeting glutamate and serotonin systems. This chapter will focus on a developmental manipulation that produces chronic changes in glutamate and dopamine function, which is useful for providing a different baseline to test acute (or chronic) treatments against.

Social isolation of rodents after weaning has been shown to produce a host of behavioural changes compared to group housed animals. Most relevantly, some of these changes are similar to that seen with acute NMDA-R antagonism that are thought to model symptom dimensions in schizophrenia (see Rafter et al., 2016; Jentsch and Roth, 1999; Gilmour et al., 2012). For example, locomotor hyperactivity and sensorimotor gating deficits representing positive symptoms; social interaction abnormalities

representing negative symptoms; and memory deficits representing cognitive symptoms (reviewed in Fone and Porkess, 2008).

Isolates display accelerated Pavlovian conditioning (Harmer and Phillips, 1998b, 1998a; Phillips et al., 2002), which may be a result of augmented phasic dopamine signalling in the ventral and dorsomedial striatum (Yorgason et al., 2013, 2016). These animals are also quicker to acquire drug-related memories and take longer to extinguish these memories, which is dependent on NMDA-R mediated plasticity in the VTA (Whitaker et al., 2013). Despite this, isolates have been shown to be impaired at simple reversal learning (although findings are mixed, see Li et al., 2007 and Robbins, 2016).

Interestingly, one study found that amphetamine-induced dopamine release in isolates is augmented in the striatum, but attenuated in the prefrontal cortex (Fabricius et al., 2011). Since striatal dopamine hyperactivity coupled with prefrontal hypoactivity is seen in schizophrenia (Howes and Kapur, 2009), isolates have been proposed to be useful for modelling this dopaminergic state. However it should be noted that many isolation-induced neurochemical changes are not consistently found (Fone and Porkess, 2008), with some studies showing increases or no changes in prefrontal dopamine function (reviewed in Lukkes et al., 2009).

Similar to post-weaning social isolation, neonatal PCP administration on post-natal days 7, 9, and 11 has been shown to result in phenotype changes in adulthood, including working memory impairments (Boctor and Ferguson, 2010), sensitisation to the locomotor stimulation (Anastasio and

Johnson, 2008) and working memory impairments (Wiley et al., 2003) induced by PCP, and some of these effects are attenuated by drug treatment (e.g. antipsychotics). Neurochemically, neonatal PCP appears to induce long-term changes in glutamatergic signalling pathways, such as decreased parvalbumin expression (Kaalund et al., 2013), and altered NMDA receptor subunit expression (Anastasio and Johnson, 2008). These alterations in glutamatergic and GABAergic signalling may impair reversal learning by decreasing signal-to-noise ratios in the OFC, and rule maintenance in working memory in PFC (see Figure 1.8).

Meanwhile, dual-hit social isolation with neonatal PCP administration has been shown to produce a more disrupted phenotype than social isolation alone (Gaskin et al., 2014; Grayson et al., 2015). Furthermore, acute administration of a partial D₂/D₃ receptor agonist attenuates some of the behavioural changes (Watson et al., 2016). This suggests that PCP-isolates present a chronic model of dopamine dysfunction that is sensitive to acute pharmacological alleviation. Therefore an experiment was conducted to determine the effects of PCP-isolation on the touchscreen perseverance and learned irrelevance task, as per Chapter 4. Given previous evidence of enhanced phasic dopamine signalling, it was hypothesised that PCP-isolates would be enhanced on the perseverance task, which is more dependent on learning from positive RPEs. Meanwhile isolates would not learn quicker on the learned irrelevance test, which is more dependent on learning from negative RPEs, and may be impaired due to cortical glutamate dysfunction.

6.2 Methods

6.2.1 Animals

Twenty-eight male Lister-hooded rat pups (delivery age: PND4, CRUK) in four litters of seven were housed with their dams according to the conditions described in 2.2.1 Animals. On post-natal days (PND) 7, 9, and 11 half of the pups from each litter were administered a subcutaneous injection of either 0.9% saline vehicle (2 ml/kg) or PCP (10 mg/kg in vehicle). On PND 24 pups were weaned into non-ventilated cages (polycarbonate with metal grid lids) and the number reduced to 24. PCP-treated rats (n=12) were housed individually (41×24×20 cm³ cages) and vehicle treated rats (n=12) in groups of 4 (50×33×26 cm³ cages). PCP-isolates had no physical contact with conspecifics but retained visual, auditory and olfactory interaction in the same holding room. All animals received minimal handling (once weekly cage change and body weight measurement) until the start of behavioural testing five weeks later, described in Table 6.1. Touchscreen training was conducted for 6 days per week.

Table 6.1 A timeline of behavioural procedures in PCP-isolates and group-housed animals.

PND	Procedure
57	Subcutaneous insertion of temperature probes (not reported) *
59	Locomotor activity
60	Novel object discrimination
64 (am)	Social interaction
64 (pm)	Begin food restriction
67	Begin touchscreen training (reversal learning task)

* All animals carried an infra-red probe that was used for taking body temperatures before and after the social interaction test, which is not reported. No significant differences were found between groups.

6.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity / novel object discrimination boxes

See section 3.2.2 Apparatus.

Social interaction arena

Social interaction was assessed in a 75cm diameter metallic arena with a wooden floor bed and card walls (45cm high), in a dimly lit room (40 lux) to minimise anxiety. The arena was cleaned with 20% ethanol between uses to remove distractive odours.

6.2.3 Locomotor activity protocol

On PND 59, rats were placed into locomotor boxes for 60 minutes. Rearing and ambulatory activity was recorded in 5 minute time bins.

6.2.4 Novel object discrimination protocol

The day after LMA assessment (PND 60), recognition memory was tested in a novel object discrimination task. Animals were placed into the same LMA box as the day prior for 3 minutes to habituate. Animals were returned to their home cage for 1 minute before commencing the familiarisation trial. For the familiarisation trial, animals were placed back into the LMA box which now contained two identical objects (white cylindrical plastic bottles, diameter: 5cm, height: 8cm) fixated in the floor of the box. After 3 minutes, animals were placed back into their home cage for a 1 hour inter-trial interval, after which commenced the choice trial. For the choice trial,

animals were placed into the same LMA box as during familiarisation, but now one of the objects was replaced (at random) with a novel object (identically sized black and white striped cylindrical bottle). After 3 minutes of exploration, animals were returned to the home cage.

Active exploration of the objects was recorded using stopwatches and defined as: sniffing, licking, or moving vibrissae with the nose pointing towards and ≤ 1 cm from the object; climbing on top of objects or extensive chewing was not included.

6.2.5 Social interaction protocol

On PND 64, animals' social behaviour was measured. Animals were placed into the social interaction arena in pairs for 10 minutes whilst their behaviour was video recorded. Animals were paired with a novel conspecific of the same developmental treatment, with a similar weight, from a different litter and cage. Twenty minutes prior to the test, pairs were sprayed with blue/pink (randomly assigned) hair spray (Superdrug), to allow accurate tracking by the behavioural scoring software (Ethovision XT 8.5, Noldus). Seven individual behaviours were manually scored by observers blind to the animals' developmental condition: ano-genital sniffing, body sniffing, boxing and biting, crawling under-and-over, following, lying side-by-side and pinning (see Watson et al., 2016). Since the behaviour of each rat in the pair is not independent of the partner, the mean duration of each behaviour was calculated per pair and so each N value represents the number of pairs.

6.2.6 Touchscreen task protocol

On PND 67, rats began touchscreen training. All phases of the touchscreen reversal learning task were identical to that described in 4.2.3 Touchscreen task protocol, with the exception that no drug treatments were administered.

6.2.7 Measures and statistics

See section 3.2.5 for statistical analysis methods of reversal learning and locomotor activity data. For novel object discrimination, object exploration times were analysed by two-way RM ANOVA with treatment and object as factors. Between treatment effects for each social interaction measure was established using t-tests with Welch's correction for non-equivalent variance. The overall effect of treatment was assessed using a two-way RM-ANOVA with behaviour and treatment as factors. Social interaction video data was scored and analysed by two undergraduate students blind to treatment groups, James Dent and Urukarn Laohakanchanasiri.

6.3 Results

6.3.1 Locomotor activity

One group-housed animal was lost due to a tumour. In all animals, there was a significant effect of time on ambulation ($F_{(11, 231)} = 29.1$, $p < 0.0001$) and rears ($F_{(11, 231)} = 21.6$, $p < 0.0001$), i.e. animals became less active with time. There was no effect of developmental treatment on ambulation ($F_{(1, 21)} = 0.99$, $p = 0.33$) but a significant effect on rears ($F_{(1, 21)} = 5.7$, $p = 0.03$).

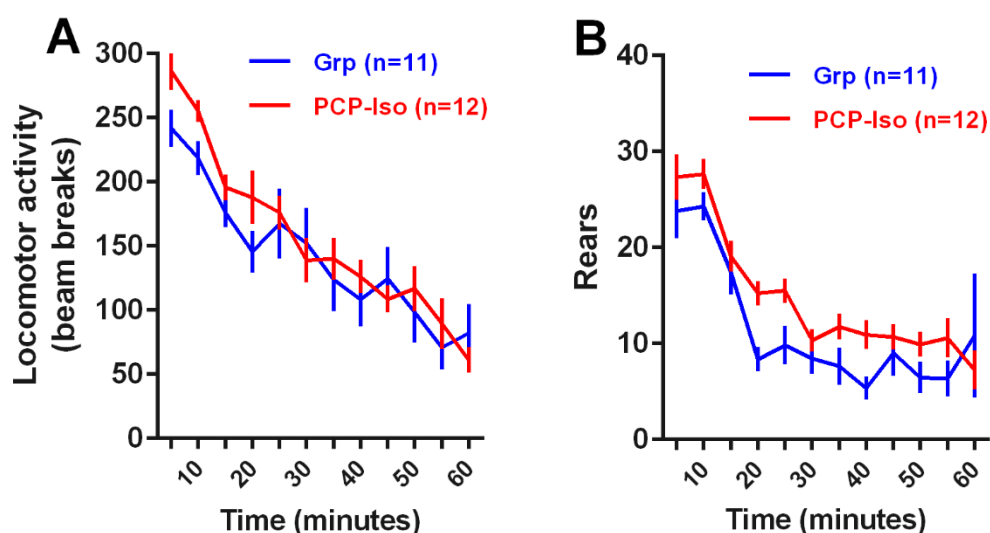


Figure 6.1 The effects of PCP-isolation or vehicle group-housing on (A) ambulatory activity and (B) rears during the locomotor activity test (PND 59).

Data points depict male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red) or group-housing (blue) after weaning (PND 24). Data shown as mean $\pm$ SEM.

6.3.2 Novel object discrimination

During the familiarisation trial, there was no effect of developmental treatment ($F_{(1, 21)} = 2.1$, $p = 0.16$) or object ($F_{(1, 21)} = 3.6$, $p = 0.07$) on exploration time. During the novel object (choice) trial, there was a significant effect of object ($F_{(1, 21)} = 14.6$, $p = 0.001$), i.e. the novel object was explored more, but no effect of developmental treatment ($F_{(1, 21)} = 0.86$, $p = 0.37$) nor any object $\times$ treatment interaction ($F_{(1, 21)} = 0.58$, $p = 0.46$), i.e. no differences in object exploration between treatments. Post-hoc tests showed that there was a significant difference in object exploration in the group-housed animals ($t_{21} = 3.17$, $p = 0.009$) but not in the PCP-isolates ($t_{21} = 2.21$, $p = 0.075$; see Figure 6.2), however this is insufficient evidence for a

difference between the groups (Nieuwenhuis et al., 2011), which is further evident by the lack of significant differences in the discrimination ratios (see Table 6.2).

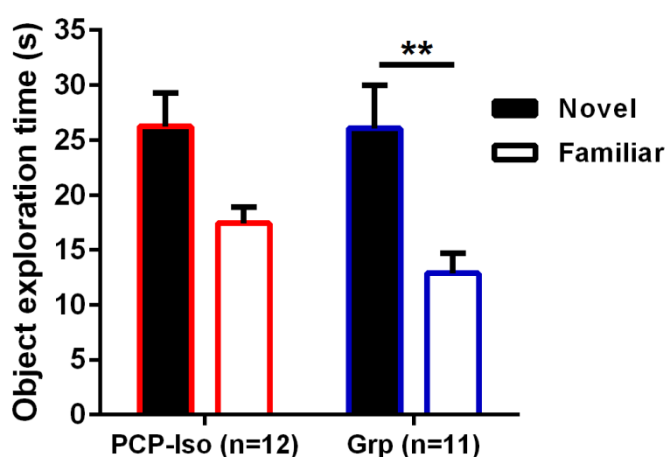


Figure 6.2 Novel object discrimination in PCP-isolates and group-housed animals.

Data points depict exploration time (mean $\pm$ SEM) during the choice trial of a novel object discrimination task (PND 60) in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red) or group-housing (blue) after weaning (PND 24). ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

Table 6.2. Discrimination measures for PCP-isolates and group-housed animals.

Measures are derived from object exploration times during the choice trial of a novel object discrimination task (PND 60) in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (**PCP-Iso**) or saline (**Grp**) on PND 7, 9, and 11 with concomitant social-isolation (**PCP-Iso**) or group-housing (**Grp**) after weaning (PND 24). As described by Akkerman et al., 2012, D1 is the absolute difference in object exploration (novel - familiar, in seconds), D2 is the discrimination index ($D1 \div \text{total exploration time}$), D3 is the discrimination ratio (novel object exploration $\div$ total exploration time). All data shown as mean $\pm$ SEM, p-values shown for two-tailed, unpaired t-tests.

	D1	D2	D3
PCP-Iso (n=12)	8.8 $\pm$ 3.1	0.18 $\pm$ 0.06	0.59 $\pm$ 0.03
Grp (n=11)	13.2 $\pm$ 4.3	0.31 $\pm$ 0.08	0.65 $\pm$ 0.04
p-value	0.46	0.23	0.23

6.3.3 Social interaction

The duration of each behaviour during the social interaction test is presented in **Table 6.3**. Of the seven behaviours, only crawling over-and-under differed significantly between developmental conditions ($t=2.86$, $d.f.=5.32$; $p=0.033$), and was significantly higher in PCP-isolates. A two-way repeated measures ANOVA with behaviour and developmental treatment as factors showed no significant effect of treatment overall ($F_{(1,9)} = 0.04$, $p = 0.85$).

Table 6.3. Social interaction measures for PCP-isolates and group-housed animals.

All data shown as mean $\pm$ SEM of time spent performing the behaviour during the social interaction test (PND 64), in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (**PCP-Iso**) or saline (**Grp**) on PND 7, 9, and 11 with concomitant social-isolation (**PCP-Iso**) or group-housing (**Grp**) after weaning (PND 24). p-values shown for two-tailed, unpaired t-tests with Welch's correction.

	PCP-Iso (n=6)	Grp (n=5)	t	d.f.	p
Anogenital sniffing	2.38 $\pm$ 1.26	15.6 $\pm$ 5.14	2.5	4.48	0.0604
Boxing and biting	2.60 $\pm$ 1.31	0.0 $\pm$ 0.0	1.98	5	0.1046
Crawling over and under	6.50 $\pm$ 2.07	0.48 $\pm$ 0.37	2.86	5.32	0.0330
Following	24.32 $\pm$ 5.36	41.5 $\pm$ 7.05	1.94	7.86	0.0883
Lying side-by-side	0.18 $\pm$ 0.18	0.89 $\pm$ 0.46	1.43	5.20	0.2107
Self-grooming	39.78 $\pm$ 5.83	48.7 $\pm$ 16.3	0.51	5.01	0.6296
Body sniffing	90.05 $\pm$ 8.96	72.6 $\pm$ 12.3	1.14	7.64	0.2877

6.3.4 Reversal learning task

Acquisition phases

One of the PCP-isolated animals was lost during this phase of the experiment due to the development of seizures. For the acquisition phases (0 – 4), group-housed animals took a median of 11 sessions (range 10–15) to perform to criteria, whilst PCP-isolates took a median of 12.5 sessions (range 10–16).

Perseverance test

For the 9 sessions post-reversal, there was a significant effect of developmental treatment on number of correct ($F_{(1,8)} = 7.9$, $p = 0.02$) and incorrect ($F_{(1,8)} = 9.2$, $p = 0.02$) responses, with post-hoc tests showing significantly more correct responses [treatment $\times$ session interaction: $F_{(8,64)} = 3.4$, $p = 0.003$] in the PCP-isolates during reversal sessions 3 – 5 and significantly fewer incorrect responses [treatment $\times$ session interaction: $F_{(8,64)} = 3.5$, $p = 0.002$] during reversal sessions 4 – 7 (all $p < 0.05$, see Figure 6.3). In addition, the PCP-isolates took significantly fewer sessions to reach criterion after the reversal ($\chi^2 = 7.39$, d.f. = 1, $p = 0.007$; see Figure 6.4).

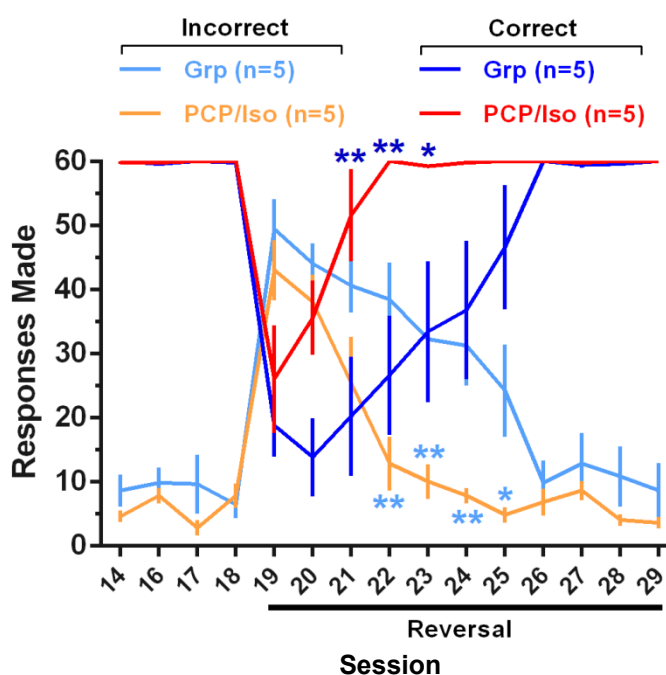


Figure 6.3 Responses made during the perseverance test for PCP-isolates or controls.

Data points depict the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the perseverance test in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red/orange) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red/orange) or group-housing (blue) after weaning (PND 24). * $p < 0.05$, ** $p < 0.01$, post-hoc Sidak test following two-way RM-

ANOVA.

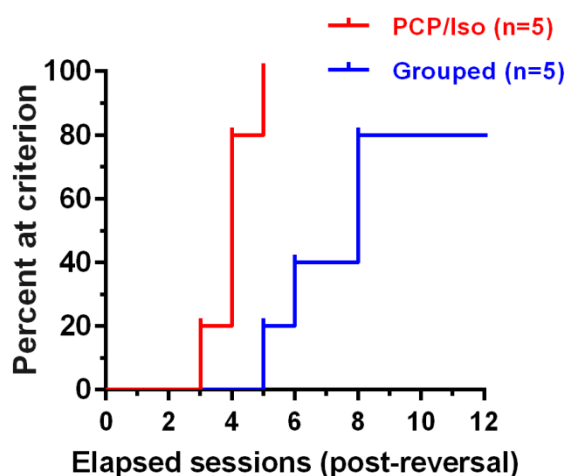


Figure 6.4 Sessions to criterion during the perseverance test for PCP-isolates or controls.

A survival curve showing the percentage of animals performing at criterion (>80% correct on subsequent sessions) during the perseverance test, in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red) or group-housing (blue) after weaning (PND 24).

Learned irrelevance test

For the 9 sessions post-reversal, there was no effect of developmental treatment on number of correct responses ($F_{(1, 10)} = 0.02$, $p = 0.9$) and incorrect responses ($F_{(1, 10)} = 0.001$, $p = 0.97$; see Figure 6.5). Similarly, there was no difference between the treatment groups in sessions taken to reach criterion after the reversal ($\chi^2 = 0.33$, d.f. = 1, $p = 0.57$; see Figure 6.6).

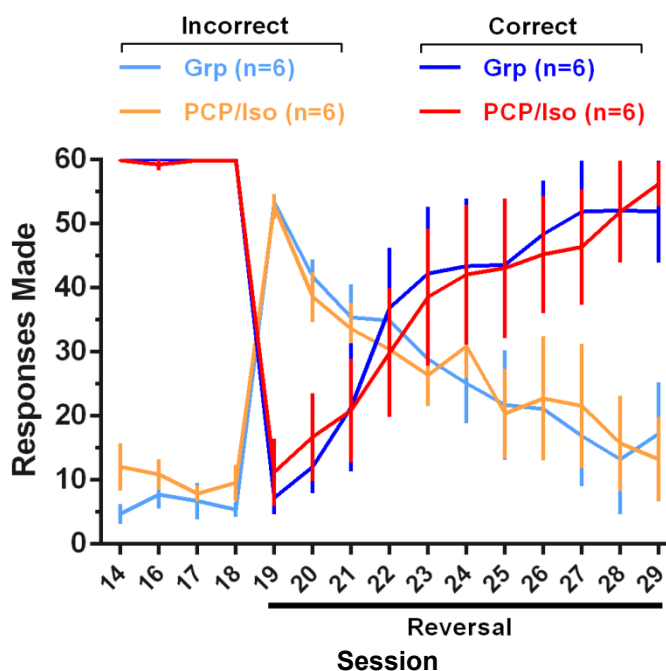


Figure 6.5 Responses made during the learned irrelevance test for PCP-isolates or controls.

Data points depict the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the learned irrelevance test in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red/orange) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red/orange) or group-housing (blue) after weaning (PND 24). * $p < 0.05$, ** $p < 0.01$, post-hoc Sidak test following two-way RM-ANOVA.

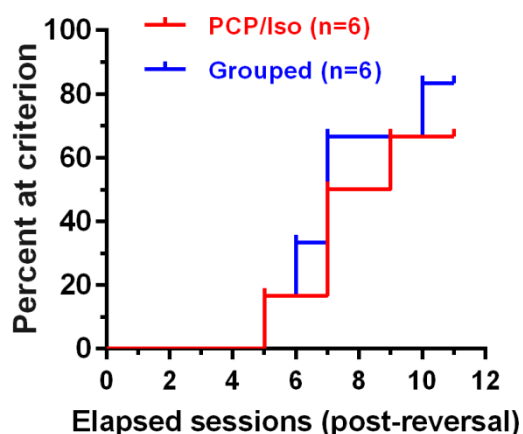


Figure 6.6 Sessions to criterion during the learned irrelevance test for PCP-isolates or controls.

A survival curve showing the percentage of animals performing at criterion (>80% correct on subsequent sessions) during the learned irrelevance test, in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red) or group-housing (blue) after weaning (PND 24).

Novelty bias

There was no effect of developmental treatment on the number of incorrect responses during the novelty bias test ($F_{(1, 20)} = 0.32$, $p = 0.58$), i.e. both groups approached the novel stimulus to a similar degree. There was also no significant effect on the number of correct responses ($F_{(1, 20)} = 0.4$, $p = 0.54$; see Figure 6.7).

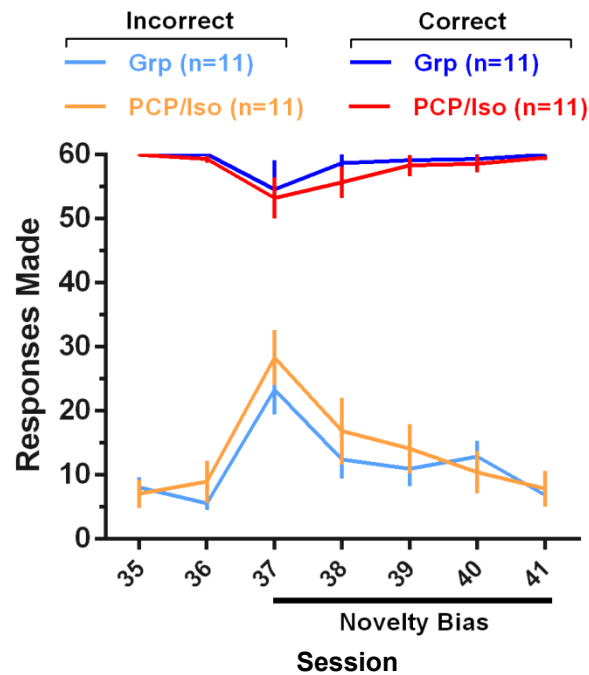


Figure 6.7 Responses made during the novelty bias test for PCP-isolates or controls.

Data points depict the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novelty bias test in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red/orange) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red/orange) or group-housing (blue) after weaning (PND 24).

Extinction

During the extinction phase, there was no effect of developmental treatment on the number of responses to the formerly rewarded stimulus ($F_{(1, 19)} = 0.26$, $p = 0.61$) or unrewarded stimulus ($F_{(1, 19)} = 0.16$, $p = 0.7$, see Figure 6.8).

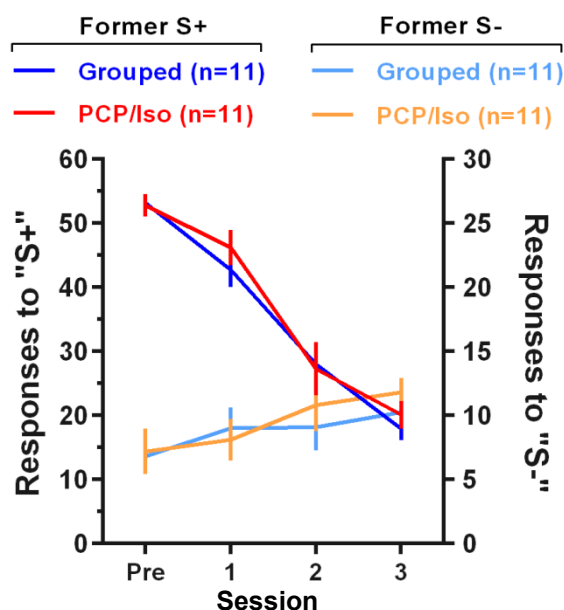


Figure 6.8 Responses made during the extinction test for PCP-isolates or controls.

Data points depict the number (mean $\pm$ SEM) of responses made to the formerly correct (S+) and incorrect (S-) stimulus for each session (x-axis) during the extinction test, in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red/orange) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red/orange) or group-housing (blue) after weaning (PND 24).

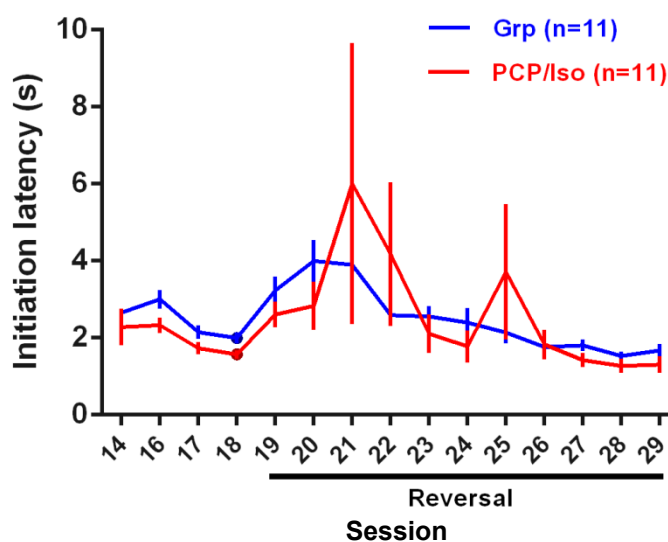
Latency measures

During the four sessions prior to reversal, there was no significant effect of developmental treatment or session on any latency measure, although the effect of treatment on initiation latency was at the threshold of significance ($F_{(1, 20)} = 4.15$, $p = 0.055$).

As with previous experiments, there was a significant effect of session on initiation latency ($F_{(8, 160)} = 2.66$, $p = 0.009$) and response latency ($F_{(8, 160)} = 7.22$, $p < 0.0001$) in the nine sessions post-reversal.

There was no effect of developmental treatment or any treatment $\times$ session interaction for these measures (all $F_s < 1$, $p > 0.5$), but post-hoc tests did show significant within-group increases in response latency for both treatment groups during reversal session one, and for sessions 2 – 3 in the PCP-isolates (see Figure 6.9).

During the reversal period, there was a significant effect of treatment on collection latency ($F_{(1, 20)} = 4.65$, $p = 0.04$), with no effect of session, session $\times$ treatment interaction, or any post-hoc significance. This implies that collection latency was consistently quicker in PCP-isolates (mean difference: 0.31 ± 0.03 s; 95% CI: 0.244 – 0.373), and may not have been detected pre-reversal due to lack of statistical power.



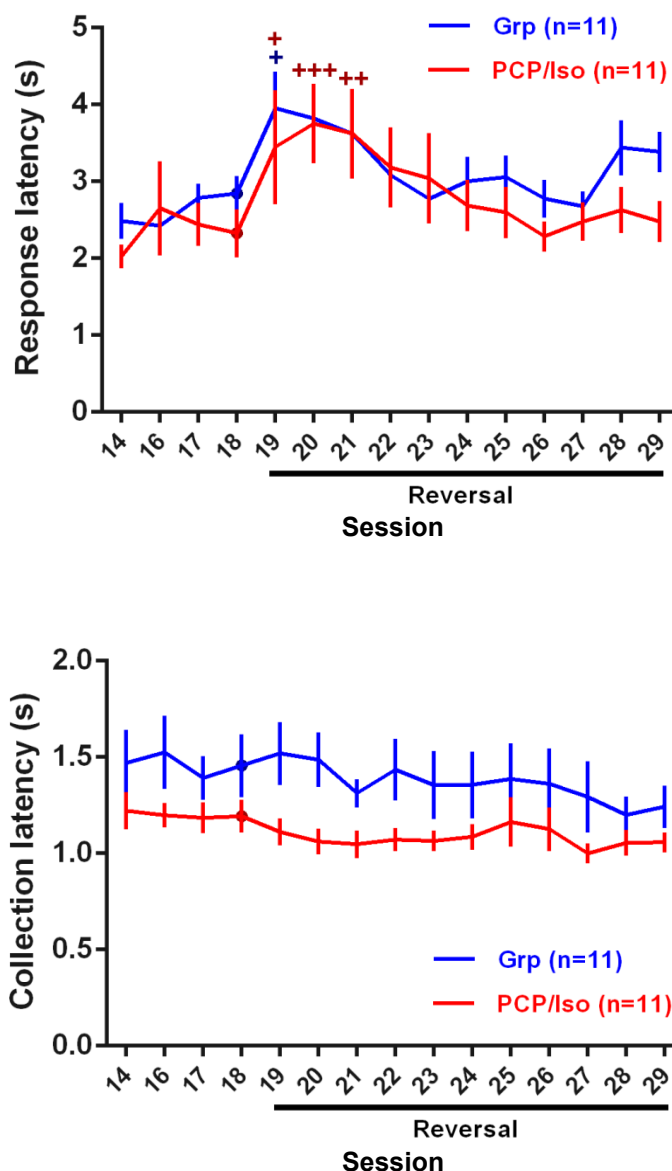


Figure 6.9 Latency measures during the perseverance and learned irrelevance test (collated) for PCP-isolates or controls.

Data points depict the trial latencies (mean $\pm$ SEM) for each session (x-axis) during the perseverance and learned irrelevance tests (collated), in male Lister-hooded rats who received subcutaneous injections of either 10mg/kg PCP (red) or saline (blue) on PND 7, 9, and 11 with concomitant social-isolation (red) or group-housing (blue) after weaning (PND 24). + $p < 0.05$, ++ $p < 0.01$, +++ $p < 0.001$ versus pre-reversal session 18. Post-hoc Sidak test following two-way RM-ANOVA.

Across extinction sessions, all animals took longer to initiate trials ($F_{(2, 40)} = 18.3$, $p < 0.0001$) and respond to the images ($F_{(2, 40)} = 7.7$, $p = 0.001$). There was no effect of developmental treatment on initiation latency ($F_{(1, 20)} = 0.4$, $p = 0.53$) or response latency ($F_{(1, 20)} = 0.005$, $p = 0.95$).

6.4 Discussion

This study investigated the effects of a putative animal model of schizophrenia – dual-hit post-natal PCP with post-weaning social isolation – on two forms of reversal learning. The most notable finding was that PCP-isolates display faster learning on the perseverance test.

Specifically, PCP-isolates were quicker to extinguish responding towards the old CS+ and quicker to exploit the new, novel S+. In the learned irrelevance test, PCP-isolates were no quicker at re-evaluating the old CS-, and sampled the novel S- to a similar degree as group-housed animals. PCP-isolates were no different on the novelty bias test or during extinction.

The lack of effect of PCP-isolation on learned irrelevance or extinction suggests that learning in response to negative RPEs is not altered, whilst the specific enhancement on the perseverance test can be explained by enhanced positive RPE processing. This is consistent with prior studies showing enhanced striatal phasic dopamine release in social isolates (Yorgason et al., 2013, 2016), which would amplify positive RPEs (see section 1.3.2 Neural evidence for a dual system). Thus, after the introduction of a novel stimulus, when animals are instinctively biased towards approaching the novel stimulus, PCP-isolates in the perseverance test experience larger phasic spikes in dopamine release after sampling the S+,

and so learn to exploit it more quickly. Meanwhile in the learned irrelevance and novelty bias tasks, there is no phasic dopamine spike to be amplified after sampling the novel stimulus (due to it being unrewarded). The data from this experiment does not support the proposal that isolates are hyperreactive to novelty (Robbins, 2016), since this would lead to increased sampling of the novel stimulus in all tests, which did not occur. Likewise, there were no differences on the novel object discrimination test. However, it cannot be ruled out that the post-natal PCP treatment somehow attenuated this phenotype, since isolation alone was not tested.

PCP-isolates were no different in their latencies to initiate trials or respond to images, suggesting no changes in task motivation or choice impulsivity, respectively. However PCP-isolates were consistently quicker to collect rewards indicating enhanced consummatory motivation. Previous studies have established that social isolates are more motivated to consume sucrose (Cuenya et al., 2015). Meanwhile the lack of effect on initiation latency is consistent with findings of no difference in tonic dopamine signalling in the ventral striatum (Yorgason et al., 2013, 2016).

This study found no change in locomotor activity with PCP-isolation, which agrees with two previous studies in this laboratory (Gaskin et al., 2014; McIntosh et al., 2012), but not another which found hyperactivity (Watson et al., 2016). This suggests that processes reliant on phasic dopamine signalling (e.g. associative learning) are more sensitive to disruption by PCP-isolation than processes reliant on tonic dopamine signalling (e.g. locomotor activity). Given that locomotor hyperactivity is a

consistent finding in social isolates (Fone and Porkess, 2008), it may be that the neonatal handling required for administering PCP protects against the formation of this phenotype in PCP-isolates.

The rationale for using the dual-hit model over social isolation alone is partly based on the assumption that dual-hit leads to a more severe phenotype that more accurately models schizophrenia (Gaskin et al., 2014). Although published data is limited, previous research suggests that dual-hit causes greater alterations in social interaction than isolation alone (McIntosh et al., 2012; Watson et al., 2016), and thus may better model the negative symptoms of schizophrenia. This study however only found a significant change in one measure, which was *increased* in PCP-isolates, and therefore cannot be interpreted as modelling social withdrawal in schizophrenia.

This study also found no impairment of novel object exploration in PCP-isolates. Although cross comparison with other experiments (see Appendix 2) shows that PCP-isolates sampled the novel S- less during the novelty bias test, this was also true of group-housed animals. Thus this effect could be attributable to the different rearing conditions between this experiment and others (e.g. delivery and injections as pups, pre-touchscreen assays, and non-ventilated cages). In addition, this decreased sampling was the result of faster learning toward the novel S- (i.e. not a memory deficit).

In conclusion, dual-hit neonatal PCP with post-weaning social isolation causes associative learning changes in adulthood reflective of enhanced positive RPE processing.

Chapter 7 – The effect of intra-OFC 6-hydroxydopamine on the reversal of positive and negative associations

7.1 Introduction

Previous research has strongly implicated the role of orbitostriatal circuits and dopamine signalling for flexible associative learning (see Chapter 1). Manipulating OFC glutamate (Brigman et al., 2013), GABA (Burke et al., 2009; Keiflin et al., 2013), or 5-HT (Boulougouris and Robbins, 2010) signalling all affect reversal learning, however less is known about OFC dopamine signalling. Primate studies have shown nuanced effects of OFC dopamine manipulations on associative learning. In a probabilistic discrimination task, OFC dopamine depletion with 6-hydroxydopamine (6-OHDA) leads to faster acquisition and decreased shifting after false feedback, possibly via enhancing tonic striatal dopamine levels (Clarke et al., 2014). However in deterministic discrimination tasks, OFC dopamine depletion has been shown to either impair acquisition and extinction (Walker et al., 2009), or have no effect on acquisition or reversal learning (Clarke et al., 2007; Roberts et al., 1994), and does not enhance tonic striatal dopamine signalling, but rather potassium-evoked (putatively phasic) signalling (Roberts et al., 1994).

In rodents, no study has looked at the effect of OFC dopamine depletion on reversal learning. However, aged rats had significantly lower tonic and phasic OFC dopamine levels compared to young rats. They were also impaired at reversal but not acquisition learning, and this deficit could be alleviated with a D₁R agonist (Mizoguchi et al., 2010). Meanwhile, intra-OFC infusion of a D₁R antagonist impairs reversal learning, whilst D₂R antagonism has no effect (Winter et al., 2009). Specifically, OFC D₁R antagonism increased post-correct switching errors suggesting a failure of rule stability, which may be due to impaired maintenance of recent A-O associations in working memory. Another study found no effect of OFC dopamine receptor antagonism on choice behaviour during reversal, but small effects on reaction times (Calaminus and Hauber, 2008). However animals in this study were not overtrained, thus the reversal may not have been particularly OFC-dependent (although NMDA-R antagonism did impair choice behaviour).

Dopamine is known to play an important role in stabilising working memory representations in other regions of PFC, with an inverted-U-shaped effect. High levels of PFC dopamine promote representation shifting via D₂R agonism, whilst intermediate dopamine levels promote representation stability via D₁R agonism (Durstewitz and Seamans, 2008). Interestingly, it has been proposed that mesocortical dopamine projection neurons fire in response to aversive salient events (Lammel et al., 2011), therefore frequent punishments (such as post-reversal or extra-dimensional shift) lead to disengagement of rule exploitation by destabilising rule representations in

working memory. Importantly, PFC contains low amounts of the dopamine transporter and therefore phasic dopamine release can accumulate (Sesack et al., 1998), possibly providing an “average punishment” signal.

Whether prefrontal dopamine signalling is more important for stabilising recent A-O associations in response to positive or negative RPEs is unknown. Therefore an experiment was conducted to determine the effects of OFC dopamine depletion (via the neurotoxin 6-OHDA) on the perseverance and learned irrelevance reversal learning tasks, as per Chapter 4. A bilateral infusion of 6µg 6-OHDA was used, as previous studies show this to be sufficient to produce >70% depletion in prefrontal regions (Nelson et al., 2010).

7.2 Methods

7.2.1 Animals

Twenty-eight male Lister-hooded rats (delivery weight: 150-200g, CRUK) were delivered, housed and food restricted according to the conditions described in 2.2.1 Animals.

7.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity boxes

See section 3.2.2 Apparatus.

7.2.3 Touchscreen task protocol

All phases of the reversal learning task were identical to that described in 4.2.3 Touchscreen task protocol, with the exception of no drug treatments being administered during the task. When all animals reached criterion ($\geq 90\%$ correct) during Phase 4, training and food restriction was halted and animals received surgery. Animals resumed food restriction seven days post-surgery and – three days later – resumed touchscreen testing at Phase 4. Once animals were performing at criterion for 3 consecutive days, they were moved onto Phase 5 (reversal). Upon completion of the remaining phases of the reversal learning task animals were sacrificed for tissue collection.

7.2.4 Surgery protocol

All animals were pre-treated with 20mg/kg desipramine hydrochloride (dissolved in saline, 2ml/kg i.p.) in order to protect noradrenergic nerve terminals from neurotoxic damage. Rats were then anaesthetised with isoflurane (induction: 4%, maintenance: 2%; constant 1l/min N₂O and 0.5l/min O₂) and treated with Rimadyl (4mg/kg s.c.) for analgesia. Rats had cannulae inserted bilaterally into the OFC at co-ordinates (from bregma): AP +3.8mm, ML ± 2.0 mm, DV (from dura) -4.2mm (see Boulougouris et al., 2007a). Rats then received a 0.6 μ l infusion of either vehicle (0.2% ascorbate in 0.9% saline; n=13) or 6-hydroxydopamine (6-OHDA, 6 μ g; n=14) hydrobromide (Sigma) in each side, at a rate of 0.3 μ l/min. The cannulae were removed from the brain after a further 5 minutes to ensure diffusion of the solution. Rats were sutured and treated with Lignol and saline. Seven

days later, animals resumed food restriction. One of the vehicle-assigned animals was lost during surgery.

7.2.5 Brain monoamine quantification

At the end of behavioural testing (approximately 2 months post-lesion), all rats were humanely decapitated and brains rapidly removed onto a cold plate. Brains were placed into an ice-cold acrylic coronal brain matrix (Harvard Instruments, USA) dorsal side up and held in position by two chilled razor blades. Additional razor blades were inserted to obtain 2mm slices at approximately +1mm and +4mm anterior to bregma (see Figure 7.1). Dorsal striatum was dissected bilaterally from the +1mm slice, and 0.84mm diameter micropunches of lateral OFC and medial PFC were taken bilaterally from the +3mm slice. Tissue was placed into 1.5ml eppendorfs, weighed, and frozen in liquid nitrogen for later analysis via HPLC.

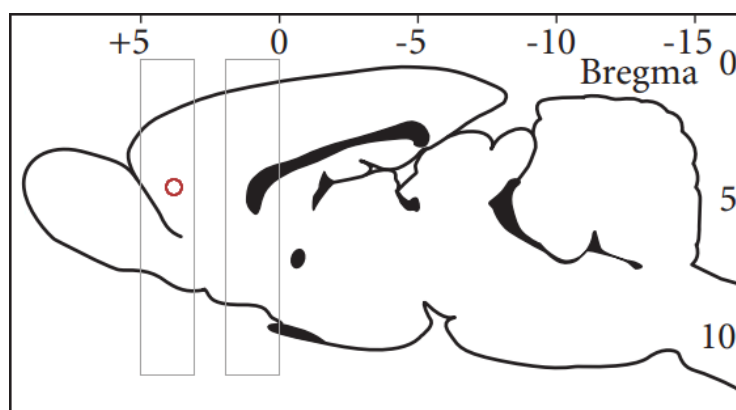


Figure 7.1 Approximate slice locations for monoamine analysis after 6-OHDA or vehicle infusion.

The slice spanning 3-5mm anterior of bregma was used to take OFC and mPFC samples. The slice spanning 0-2mm anterior of bregma was used to take dorsal striatum samples. The red circle represents the infusion target.

Brain samples were homogenized in 0.05M perchloric acid solution by sonication and centrifuged (17400g for 20 minutes; 4 °C). Supernatants were filtered through a 0.45 μm syringe tip and 5 μL samples injected onto the HPLC column. Monoamines (dopamine; 3,4-dihydroxyphenylacetic acid – DOPAC; homovanillic acid – HVA; 5-HT; and 5-hydroxyindoleacetic acid – 5-HIAA) were separated by a Targa C18 3 μm column (1mm $\times$ 10cm; Higgins) and were detected at a potential of +0.59 V using an electrochemical detector (Ag/AgCl cell). The mobile phase was delivered at 0.2 ml/min by an isocratic pump (Dionex P680) and consisted of 20 μM phosphoric acid, 20 μM acetic acid, 0.1 μM EDTA, 0.15 μM octanesulphonic acid and 10% v/v methanol adjusted to pH 4.07. Chromatograms were quantified using Clarity software (DataApex, Czech Republic) to produce

pmol/ml monoamine concentrations in each sample based on known standards. The lower limit of detection was 10 nM.

7.2.6 Measures and statistics

See section 3.2.5 for statistical analysis methods of reversal learning and locomotor activity data. Monoamine concentrations were quantified by dividing pmol/ml concentrations by region weight to produce pmol/mg values, and analysed by 2-way RM ANOVA with lesion and region as factors. Monoamine levels were correlated with sessions to criterion using a non-parametric Spearman analysis, since data was not normally distributed.

7.3 Results

7.3.1 Acquisition phases

For the acquisition phases (0 – 4), vehicle-assigned animals took a median of 11 sessions (range 9–15) to perform to criteria, whilst animals assigned to the 6-OHDA group took a median of 11 sessions (range 9–18). Post-surgery, there was no effect of lesion on any measure during the resumption of discrimination (all $F_s < 2$, $p > 0.3$; see Figure 7.2).

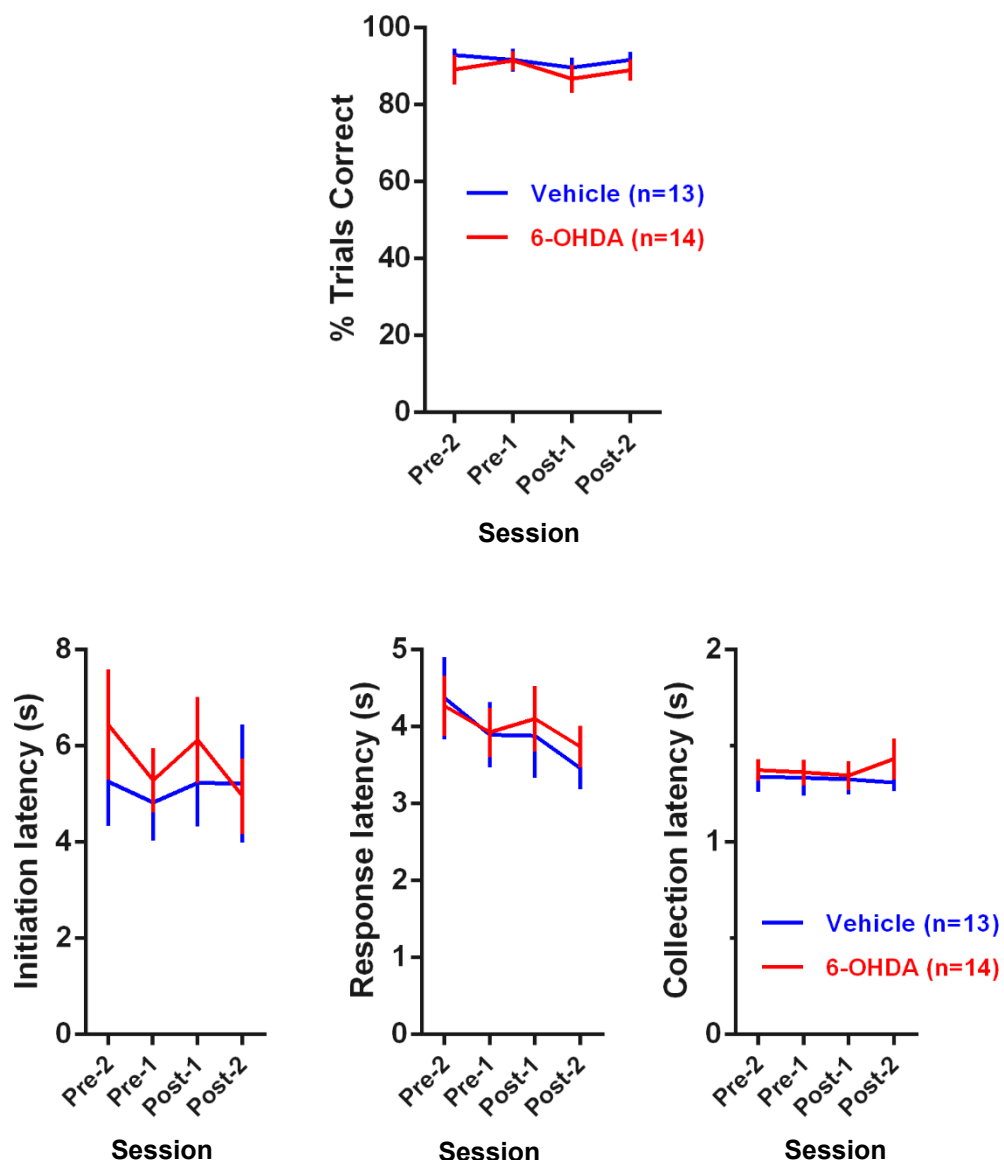


Figure 7.2 Response and latency measures before and after surgery in 6-OHDA or vehicle infused animals.

Graphs showing the percentage of correct trials and latency measures during the sessions prior and after surgery, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6µg 6-OHDA or vehicle (0.2% ascorbate saline). Data shown as mean ± SEM.

7.3.2 Perseverance test

For the 9 sessions post-reversal, there was no significant effect of lesion on number of correct responses ($F_{(1, 12)} = 0.6$, $p = 0.46$) or incorrect

responses ($F_{(1, 12)} = 0.98$, $p = 0.34$; see Figure 7.3). In conjunction, there was no effect of lesion on sessions to reach criterion after the reversal ($\chi^2 = 0.27$, d.f. = 1, $p = 0.6$; see Figure 7.4).

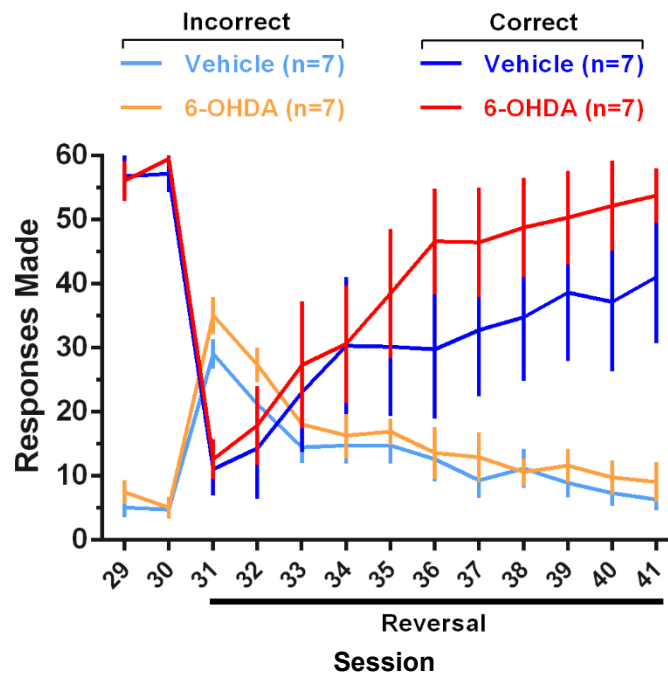


Figure 7.3 Responses made during the perseverance test for 6-OHDA or vehicle infused animals.

The number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the perseverance test, in male Listser-hooded rats that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline).

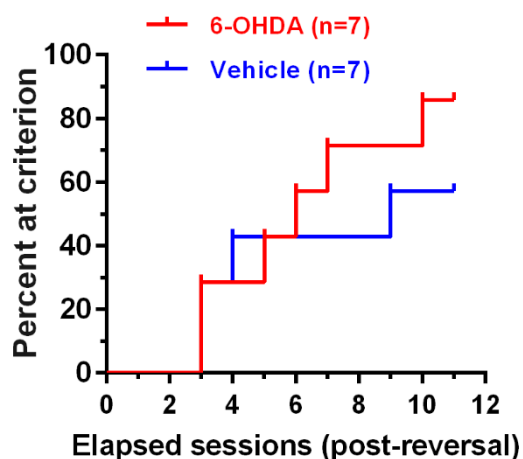


Figure 7.4 Sessions to criterion during the perseverance test for 6-OHDA or vehicle infused animals.

A survival curve showing the percentage of male Lister-hooded rats performing at criterion (>80% correct on subsequent sessions) during the perseverance test, in animals that had received intra-OFC infusion of either 6µg 6-OHDA or vehicle (0.2% ascorbate saline).

7.3.3 Learned irrelevance test

For the 9 sessions post-reversal, there was no significant effect of lesion on number of correct responses ($F_{(1, 11)} = 3.6$, $p = 0.08$) or incorrect responses ($F_{(1, 11)} = 1.7$, $p = 0.21$; see Figure 7.5). In conjunction, there was no significant effect of lesion on sessions to reach criterion after the reversal ($\chi^2 = 2.1$, d.f. = 1, $p = 0.15$; see Figure 7.6).

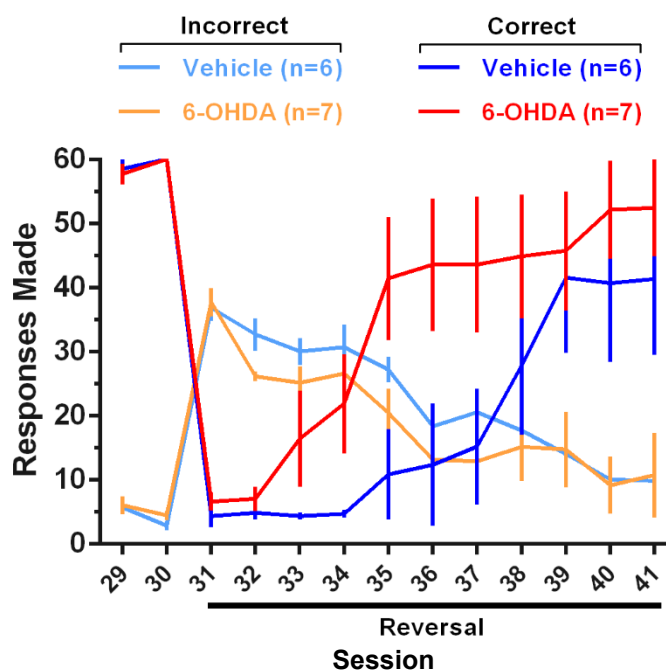


Figure 7.5 Responses made during the learned irrelevance test for 6-OHDA or vehicle infused animals.

The number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the learned irrelevance test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline).

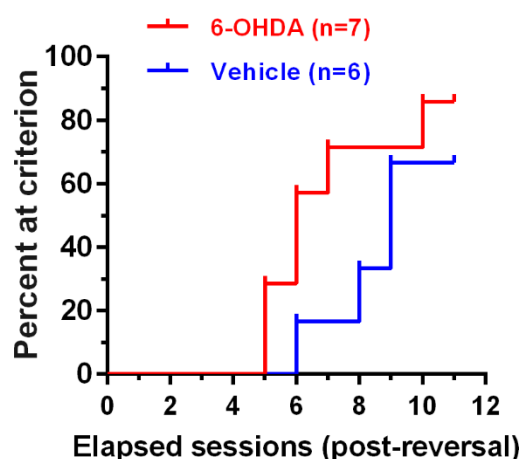


Figure 7.6 Sessions to criterion during the learned irrelevance test for 6-OHDA or vehicle infused animals.

A survival curve showing the percentage of **male Lister-hooded rats** performing at criterion (>80% correct on subsequent sessions) during the learned irrelevance test, in animals that had received intra-OFC infusion of either 6µg 6-OHDA or vehicle (0.2% ascorbate saline).

7.3.4 Novelty bias test

Three of the vehicle-assigned animals failed to reach criterion and so were excluded from analysis. There was no effect of lesion on the number of incorrect responses during the novelty bias test ($F_{(1, 22)} = 0.36$, $p = 0.55$), i.e. both groups approached the novel stimulus to a similar degree. There was also no significant effect on the number of correct responses ($F_{(1, 22)} = 0.18$, $p = 0.67$; see Figure 7.7).

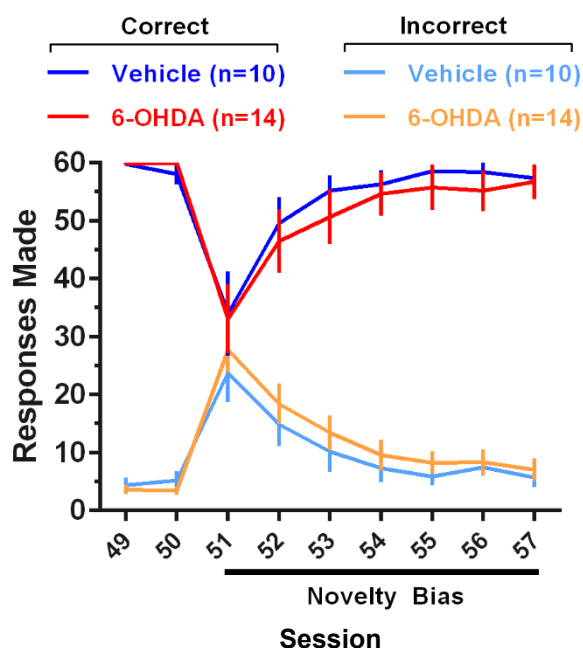


Figure 7.7 Responses made during the novelty bias test for 6-OHDA or vehicle infused animals.

The number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the novelty bias test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline).

7.3.5 Extinction

One 6-OHDA-treated and two vehicle-treated animals failed to reach criterion and so were excluded from analysis. During the extinction phase, there was no effect of lesion on the number of responses to the formerly rewarded stimulus ($F_{(1,22)} = 1.1$, $p = 0.3$) or unrewarded stimulus ($F_{(1,22)} = 0.03$, $p = 0.87$; see Figure 7.8).

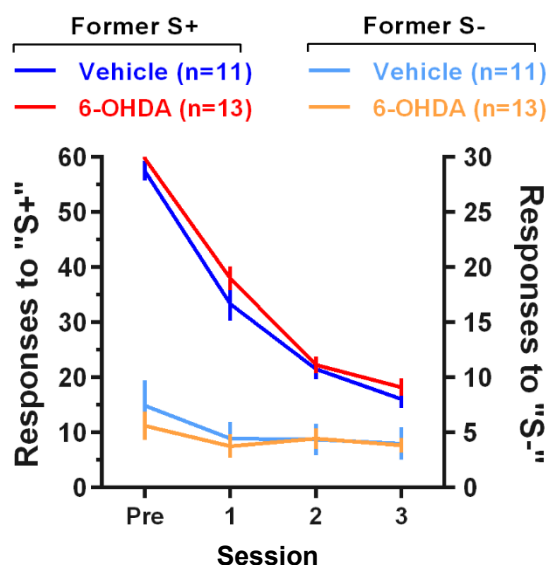


Figure 7.8 Responses made during the extinction test for 6-OHDA or vehicle infused animals.

The number (mean $\pm$ SEM) of responses made to the formerly correct (S+) and incorrect (S-) stimulus for each session (x-axis) during the extinction test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline).

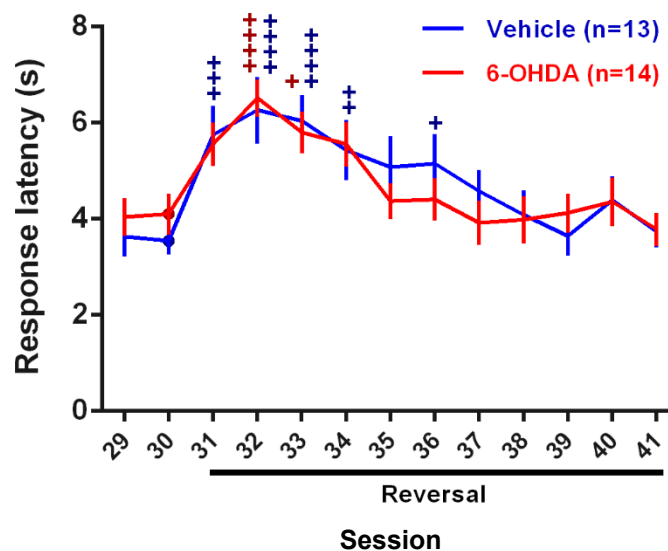
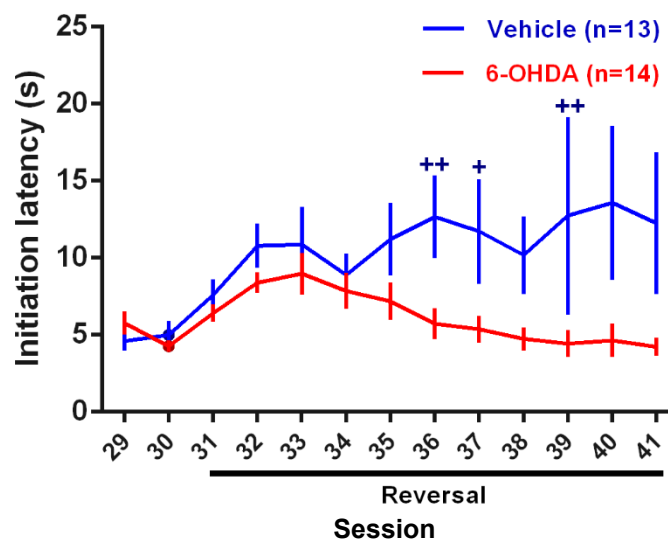
7.3.6 Latency measures

Unlike in previous experiments, there was no significant effect of session on initiation latency ($F_{(8, 200)} = 0.7$, $p = 0.67$) in the nine sessions post-reversal; there was however the expected effect on response latency ($F_{(8, 200)} = 10.1$, $p < 0.0001$).

There was no effect of lesion or any lesion $\times$ session interaction for these measures (all $F_s < 3.5$, $p > 0.07$), but post-hoc tests did show significant within-treatment increases in initiation latency in vehicle treated animals during reversal sessions 7, 8, and 9 (see Figure 7.9), which has not been seen in any prior experiments. There were also significant within-treatment increases in response latency in the sessions immediately after reversal in

both treatment groups (all $p < 0.05$), i.e. both cohorts increased image deliberation time after reversal.

There was no significant effect of session or lesion on collection latency, nor any interaction (all $F_s < 2$, $p > 0.05$).



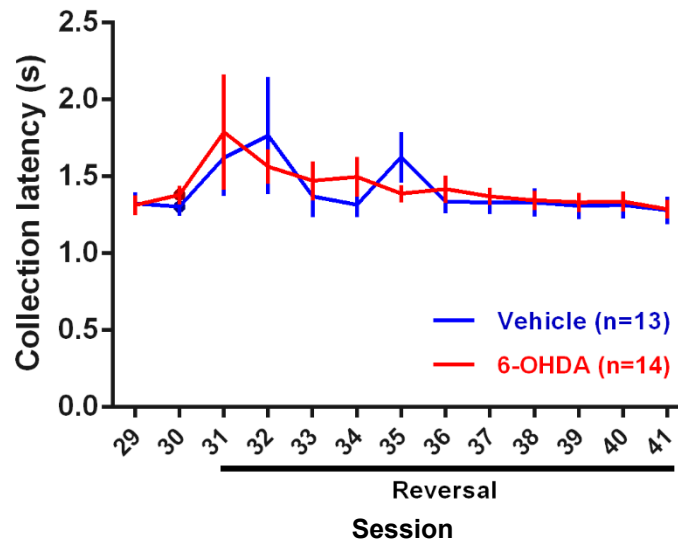


Figure 7.9 Latency measures during the perseverance and learned irrelevance test (collated) for 6-OHDA or vehicle infused animals.

The trial latencies (mean $\pm$ SEM) for each session (x-axis) during the perseverance and learned irrelevance tests (collated), in cohorts of **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline). + $p < 0.05$, ++ $p < 0.01$, +++ $p < 0.001$, ++++ $p < 0.0001$ versus pre-reversal session 30. Post-hoc Sidak test following two-way RM-ANOVA.

Across extinction sessions, all animals took longer to initiate trials ($F_{(2, 44)} = 16.4$, $p < 0.0001$) and respond to the images ($F_{(2, 44)} = 20.2$, $p < 0.0001$). There was no effect of lesion on initiation latency ($F_{(1, 22)} = 0.15$, $p = 0.7$) or response latency ($F_{(1, 22)} = 0.56$, $p = 0.46$).

7.3.7 Locomotor activity

There was a significant effect of time on ambulation ($F_{(11, 275)} = 92.6$, $p < 0.0001$) and rears ($F_{(11, 275)} = 28.1$, $p < 0.0001$) in all animals. There was no effect of cohort on ambulation ($F_{(1, 25)} = 0.3$, $p = 0.6$) but a significant effect on rears ($F_{(1, 25)} = 4.5$, $p = 0.04$; see Figure 7.10).

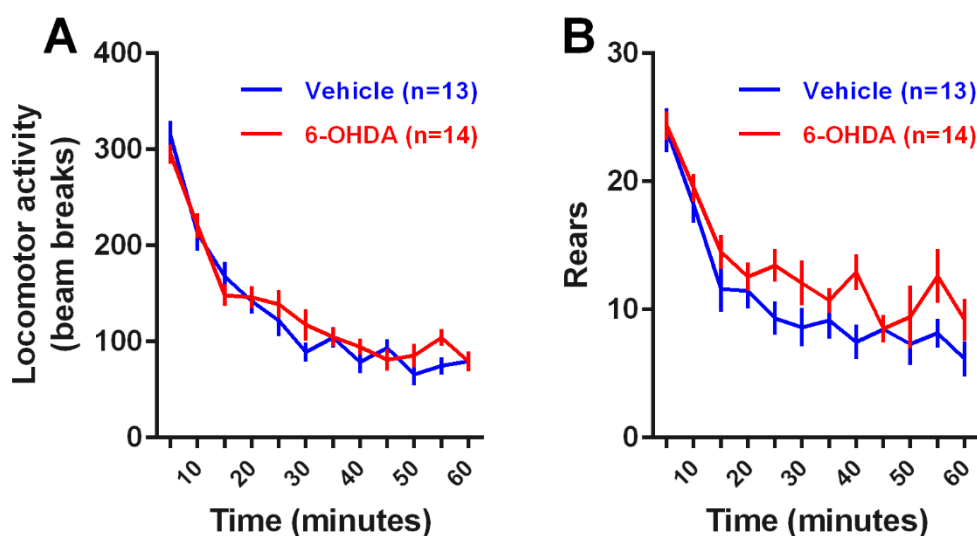


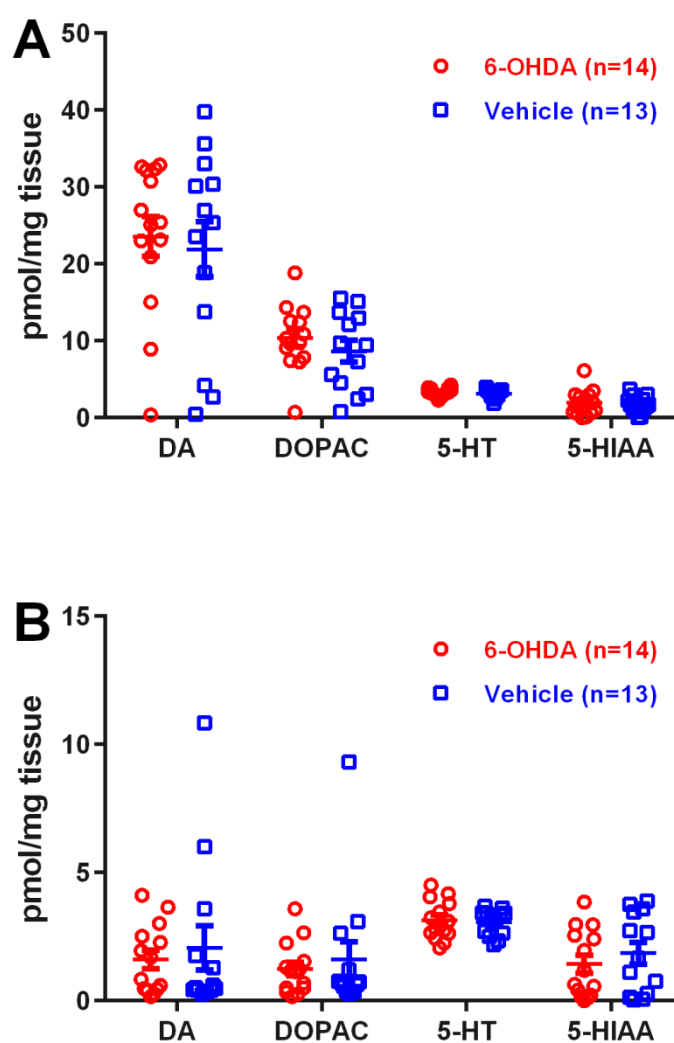
Figure 7.10 Locomotor activity for 6-OHDA or vehicle infused animals.

The ambulatory activity (A) and rears (B) in cohorts of male Lister-hooded rats that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline) approximately 8 weeks prior. Data shown as mean $\pm$ SEM.

7.3.8 Post-mortem neurochemistry

Post-mortem HPLC analysis of brain tissue showed no significant effect of lesion on any neurotransmitter in any of the sampled brain regions (lateral OFC, medial PFC, dorsal striatum; all $F_s < 0.5$, $p > 0.5$; see Figure 7.11). Compared to other published studies, the levels of all neurotransmitters and metabolites were below average, but still within 1 standard error of the mean, due to the large inter-study variance (see Table 7.1). This variance could be due to a host of factors, such as positional error when sampling tissue, tissue degradation during processing, or rat conditions such as age, strain, or the time of culling (i.e. due to circadian variation in monoamine levels). A Spearman's rank analysis showed no

correlation between OFC 5-HT levels (Figure 7.12) and the number of sessions taken to reach criterion in either the perseverance ($r = 0.04$, $p = 0.89$) or learned irrelevance test ($r = -0.08$, $p = 0.79$). Likewise, there was no correlation between OFC dopamine levels and sessions to criterion for either the perseverance ($r = 0.15$, $p = 0.61$) or learned irrelevance test ($r = 0.01$, $p = 0.97$).



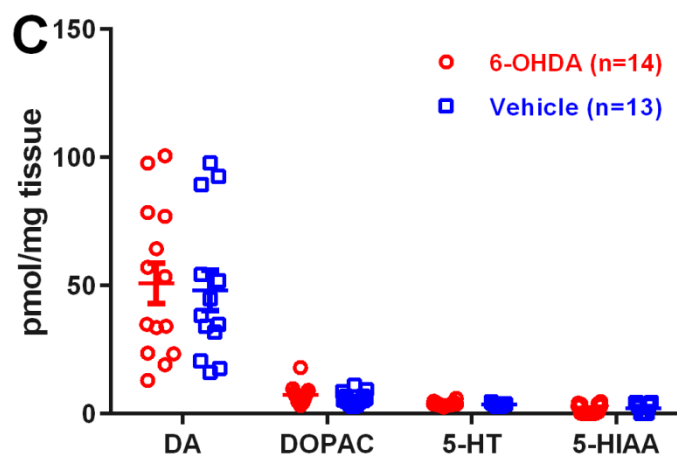


Figure 7.11 Neurochemistry for 6-OHDA or vehicle infused animals.

Graphs showing the concentrations (mean $\pm$ SEM with individual data points) of dopamine (DA), DOPAC, 5-HT, and 5-HIAA in the lateral OFC (A), medial PFC (B), and dorsal striatum (C), in cohorts of male Lister-hooded rats that had received intra-OFC infusion of either 6 μ g 6-OHDA or vehicle (0.2% ascorbate saline) approximately 8 weeks prior to tissue collection.

Table 7.1 Previously published levels of monoamines for OFC, mPFC and dorsal striatum in post-mortem rat brain tissue.

Region	[DA]	[DOPAC]	[5-HT]	[5-HIAA]
Barlow et al., 2015, Lister-hooded rats (pmol/mg):				
OFC	2.26	2.74	~0.7	~5
DS	28.73	15.79	-	-
Shortall et al., 2016a, Lister-hooded rats (pmol/mg):				
Frontal Cortex	0.72	-	3.5	-
DS	55.6	-	4.4	-
Shortall et al., 2016b, Lister-hooded rats (pmol/mg):				
Frontal Cortex	0.93		10.65	
DS	34.5		3.15	
Nelson et al., 2010, Wistar rats (pmol/μg):				
OFC	0.784	-	0.128	-
mPFC	0.07	-	0.126	-
DS	7.17	-	0.168	-
Cassaday et al., 2003, Dark Agouti rats (nmol/mg):				
Frontal Cortex	-	-	711*	-
DS	-	-	219*	-
González-Burgos et al., 2012, Sprague-Dawley rats (pg/mg):				
mPFC	-	-	~300	~500
Jaunaraes et al., 2012, Sprague-Dawley rats (pg/mg):				
mPFC	2361	-	6129	-
DS	23844	-	914	-
Piechal et al., 2012, Sprague-Dawley rats (ng/g):				
PFC	1251	189	2251	996
DS	11585	7458	2503	1208
Fitoussi et al., 2013, Wistar rats (ng/mg):				
mPFC	0.04	0.06	0.1	0.34
OFC	0.02	0	0.144	0.171
DS	7.1	1.53	0.17	0.25
Mean (± SEM) of published values (converted to pmol/mg):				
OFC	262 ± 261	n/a	43 ± 42	n/a
PFC	16 ± 11	n/a	27 ± 17	3.2 ± 1.0
DS	1081 ± 1015	23 ± 11	33 ± 27	n/a

*not included in mean calculation due to values being a factor of >1000 higher than other published levels, suggestive of errors in calculation or units reported.

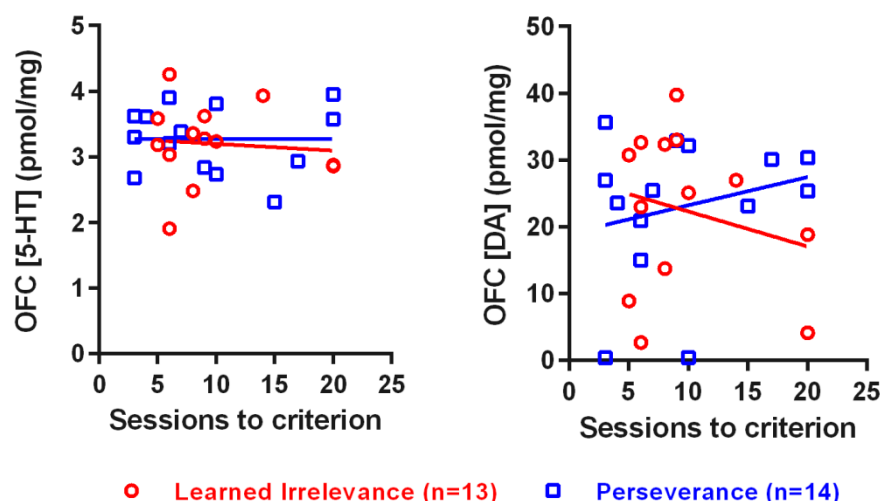


Figure 7.12 Neurochemical-behaviour correlates for 6-OHDA or vehicle infused animals.

Graphs showing the relationship between number of sessions taken to reach criterion after reversal, and post-mortem lateral OFC concentrations of 5-HT and dopamine, in cohorts of **male Lister-hooded rats** that had received intra-OFC infusion of either 6µg 6-OHDA or vehicle (0.2% ascorbate saline) **approximately 8 weeks prior to tissue collection**. Lines represent the least squares slope of best fit.

7.4 Discussion

This experiment investigated the effects of a neurotoxic lesion targeting orbitofrontal dopamine signalling on reversal learning. The primary finding was that animals treated with the toxin performed grossly similar to sham-treated animals.

Post-mortem neurochemistry was unable to confirm if the toxin had depleted dopamine levels, as has been previously reported (Nelson et al., 2010). Specifically, there was large inter-animal variance in dopamine levels whilst 5-HT levels were uniform which suggests the variance was caused by slicing error, with high prefrontal dopamine levels likely due to posterior slicing in the striatum. In addition, this variance was more prominent in the

OFC samples compared to the mPFC, which may be due to contamination in the OFC samples from the nearby nucleus accumbens (Paxinos and Watson, 2005). Another possible explanation is that the lesion had recovered, since brain tissue was collected approximately two months after surgery. Striatal 6-OHDA lesions are known to show dopamine recovery after 11 weeks (Stanic et al., 2003), however no recovery time course has been established for cortical lesions in rats.

Although there was no significant effect of 6-OHDA on choice behaviour as analysed by ANOVA on session-averaged data, when the session data was collated, 6-OHDA-treated animals were found to choose the former CS- more often in the learned irrelevance task (see Appendix 3). Assuming this is not a false positive, what could be the potential mechanisms behind this?

A previous study in marmosets showed that OFC dopamine depletion by 6-OHDA decreased “response stickiness” – the tendency to respond to the same stimulus after (probabilistic) reinforcement – and enhanced reinforcement sensitivity (i.e. learning rate) during acquisition of probabilistic discriminations. Since the experiment in this chapter is a deterministic discrimination, decreased response stickiness is the equivalent of increased exploration rate. Increased exploration (via decreasing valuation weighting) in 6-OHDA-treated rats may explain enhanced choice of the former CS- in the learned irrelevance task. This is because at the point of reversal, the values of stimuli are ranked as¹:

¹ The stimuli are valued as so because of evidence from the novelty bias test. Specifically, the introduction of a novel stimulus (in the presence of the conditioned

formerly rewarded stimulus (CS+) $\geq$ novel stimulus $\gg$ formerly unrewarded stimulus (CS-). Thus, increased exploration affects choice more drastically in the learned irrelevance task (where valuation disparity is large) than in the perseverance task (where valuations are similar). This hypothesis is incomplete however since it would predict 6-OHDA to slow devaluation of the novel stimulus during the perseverance and novelty bias test, which did not occur.

Increased exploration might also be expected to increase sampling of the former CS- in the extinction test, however this was not found. It is possible that the increased exploration only occurs in response to observable contextual changes, which are absent in the extinction test. However, another previous study (in marmosets) showed increased exploration and responding during extinction (Walker et al., 2009). Their findings may fit with the enhanced tonic striatal dopamine finding of Clarke et al. (2014), i.e. these animals learned about changes in stimulus values but persevered with choosing stimuli because of enhanced motivation regardless of reward omission.

In the locomotor test, 6-OHDA treated animals performed more rears, which was also noted in PCP-isolated animals in Chapter 6. However since

rewarded stimulus, CS+) consistently leads to changes in choice where the CS+ goes from being selected with a probability of $\sim .9$ (in the presence of CS-) to $\sim .4$, implying that the novel stimulus is (initially) valued just as highly as the CS+. See Appendix 2.

ambulatory activity was not increased, this may be a false positive (which rears are more prone to due to non-Gaussian data and heteroscedasticity).

In conclusion, an intra-OFC infusion of the dopamine neurotoxin 6-OHDA leads to subtle changes in choice behaviour suggesting enhanced re-evaluation of previously negative associations, without enhanced devaluation of previously positive associations. The following chapter will examine the effects of a serotonin neurotoxin in this task.

Chapter 8 – The effect of intra-OFC 5,7-dihydroxytryptamine on the reversal of positive and negative associations

8.1 Introduction

Previous research has established that serotonin (5-HT) signalling in the orbitofrontal cortex (OFC) is important for reversal learning (see sections 1.5.2 Role of OFC and 5.1 Introduction; Roberts 2011). Studies in primates have shown that OFC 5-HT depletion with 5,7-dihydroxytryptamine (5,7-DHT) impairs deterministic and probabilistic reversal learning (Clarke et al., 2007; Rygula et al., 2014). In rodents, numerous pharmacological and global 5-HT manipulation studies show bidirectional effects of brain 5-HT on reversal learning (reviewed in Izquierdo and Jentsch, 2012). The OFC may be particularly important, since markers of increased OFC 5-HT function correlate with reversal performance in primates and rodents (Barlow et al., 2015; Groman et al., 2013; Masaki et al., 2006). However no published studies have investigated the effect of OFC-specific 5-HT depletion on reversal learning in rodents.

Therefore an experiment was conducted to determine the effects of OFC 5-HT depletion (via the neurotoxin 5,7-DHT) on the perseverance and learned irrelevance reversal learning tasks, as per Chapters 4 – 7. A bilateral

infusion of 6µg 5,7-DHT was administered, as previous studies show similar doses to produce behavioural changes and >40% depletion of 5-HT in the amygdala (Ochoa et al., 2015; Tran et al., 2013).

8.2 Methods

8.2.1 Animals

Twenty-eight male Lister-hooded rats (delivery weight: 150-200g, CRUK) were delivered, housed and food restricted according to the conditions described in 2.2.1 Animals.

8.2.2 Apparatus

Touchscreens

See section 2.2.2 Apparatus.

Locomotor activity boxes

See section 3.2.2 Apparatus.

8.2.3 Touchscreen task protocol

All phases of the reversal learning task were identical to that described in 4.2.3 Touchscreen task protocol, with the exception of no drug treatments being administered during the task. In addition, due to a technical error, rats did not experience the novelty bias phase of the task. When all animals reached criterion ($\geq 90\%$ correct) during Phase 4, training and food restriction was halted and animals received surgery. Animals resumed food restriction seven days post-surgery and – three days later – resumed touchscreen testing at Phase 4. Once animals were performing at criterion for 3 consecutive days, they were moved onto Phase 5 (reversal). Upon

completion of the remaining phases of the reversal learning task animals were sacrificed for tissue collection.

8.2.4 Surgery protocol

Surgery was performed as described in 7.2.4 Surgery protocol, with the exception that 5,7-dihydroxytryptamine creatinine sulphate hydrobromide (5,7-DHT, Sigma; 6µg, n=14) was infused instead of 6-hydroxydopamine.

8.2.5 Brain monoamine quantification

At the end of behavioural testing (approximately 2 months post-lesion), samples of OFC and medial PFC were taken as described in 7.2.5 Brain monoamine quantification. In addition, samples of brainstem (for serotonergic raphe nuclei) were taken by removing the layer of cerebellum with a cold surgical blade and cutting out the midline of the brainstem. Brain samples were processed via HPLC as described in 7.2.5.

8.2.6 Measures and statistics

See section 7.2.6 Measures and statistics.

8.3 Results

8.3.1 Acquisition phases

For the acquisition phases (0 – 4), vehicle-assigned animals took a median of 13.5 sessions (range 9–23) to perform to criteria, whilst animals assigned to the 5,7-DHT group took a median of 13.5 sessions (range 10–20).

Post-surgery, there was no effect of lesion on any measure during the resumption of discrimination (all $F_s < 2.6$, $p > 0.1$; see Figure 8.1).

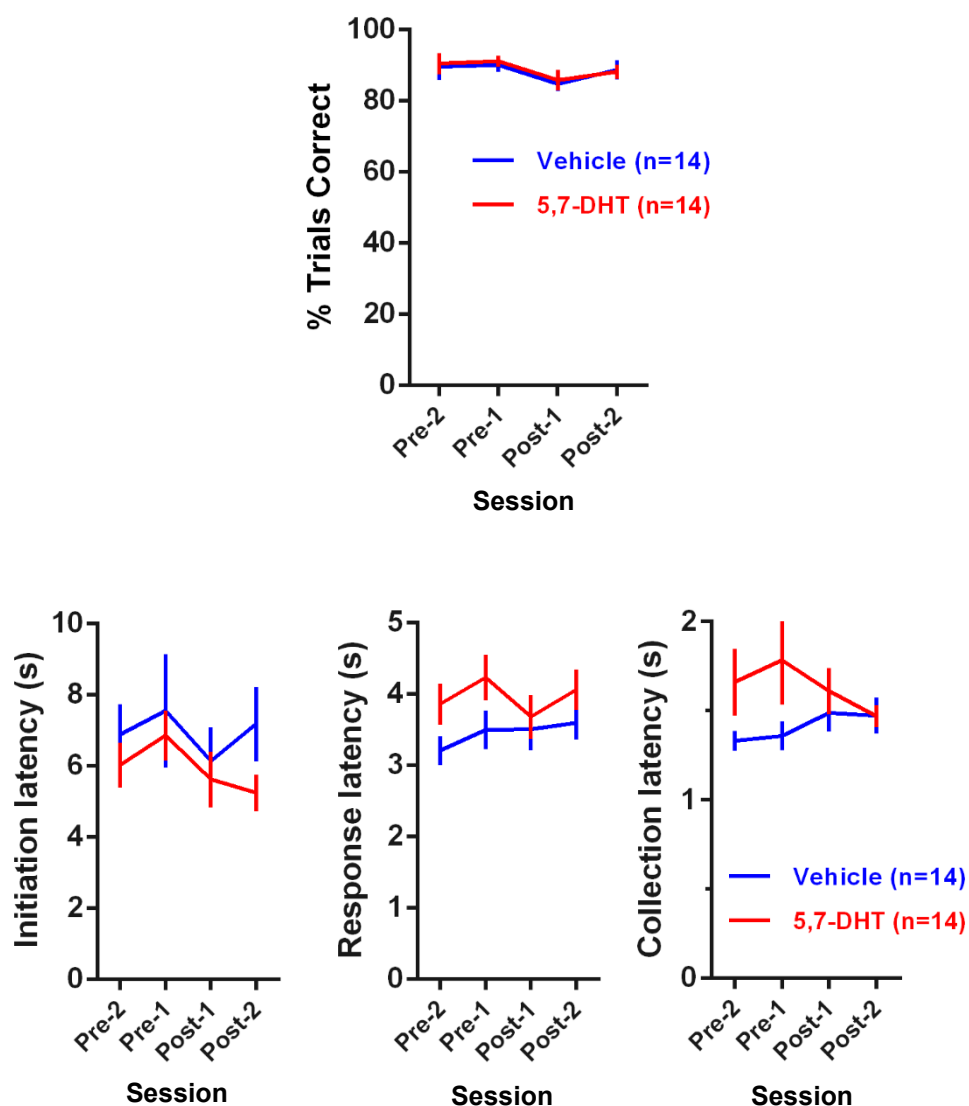


Figure 8.1 Response and latency measures before and after surgery in 5,7-DHT or vehicle infused animals.

Graphs showing the percentage of correct trials and latency measures during the sessions prior and after surgery, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline). Data shown as mean $\pm$ SEM.

8.3.2 Perseverance test

For the 9 sessions post-reversal, there was no significant effect of lesion on the number of correct ($F_{(1,12)} = 0.45$, $p = 0.52$) or incorrect ($F_{(1,12)} = 1.1$, $p = 0.31$; see Figure 8.2) responses. In conjunction, there was no effect of lesion on sessions to reach criterion after the reversal ($\chi^2 = 1.4$, d.f. = 1, $p = 0.24$; see Figure 8.3).

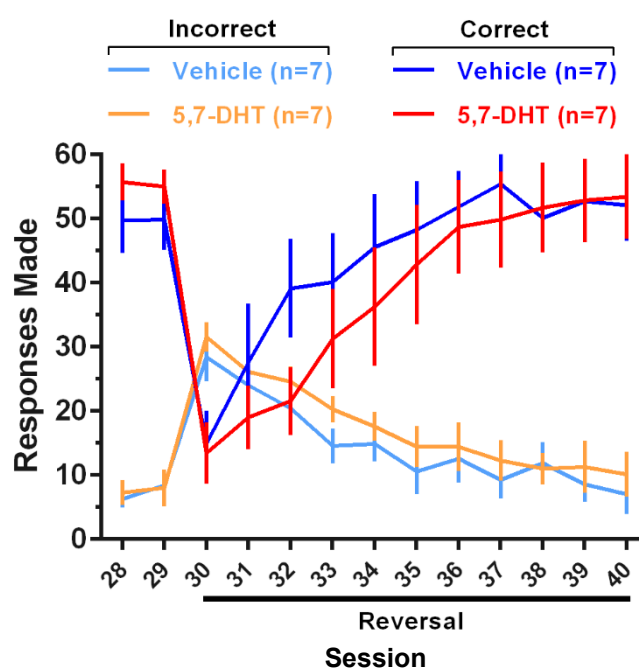


Figure 8.2 Responses made during the perseverance test for 5,7-DHT or vehicle infused animals.

A graph showing the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the perseverance test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline).

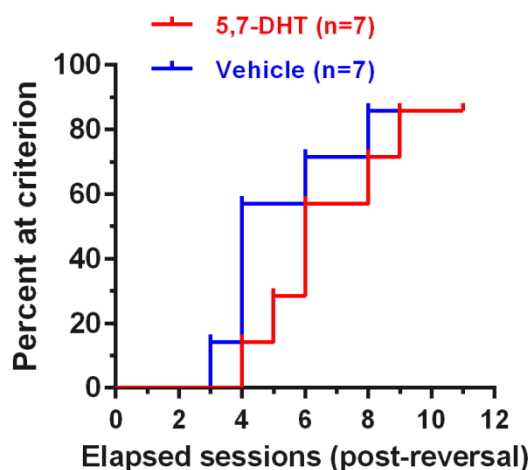


Figure 8.3 Sessions to criterion during the perseverance test for 5,7-DHT or vehicle infused animals.

A survival curve showing the percentage of animals performing at criterion (>80% correct on subsequent sessions) during the perseverance test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6µg 5,7-DHT or vehicle (0.2% ascorbate saline).

8.3.3 Learned irrelevance test

For the 9 sessions post-reversal, there was no significant effect of lesion on number of correct responses ($F_{(1, 12)} = 0.12$, $p = 0.74$) or incorrect responses ($F_{(1, 12)} = 0.009$, $p = 0.93$; see Figure 8.4). In conjunction, there was no effect of lesion on sessions to reach criterion after the reversal ($\chi^2 = 0.26$, d.f. = 1, $p = 0.61$; see Figure 8.5).

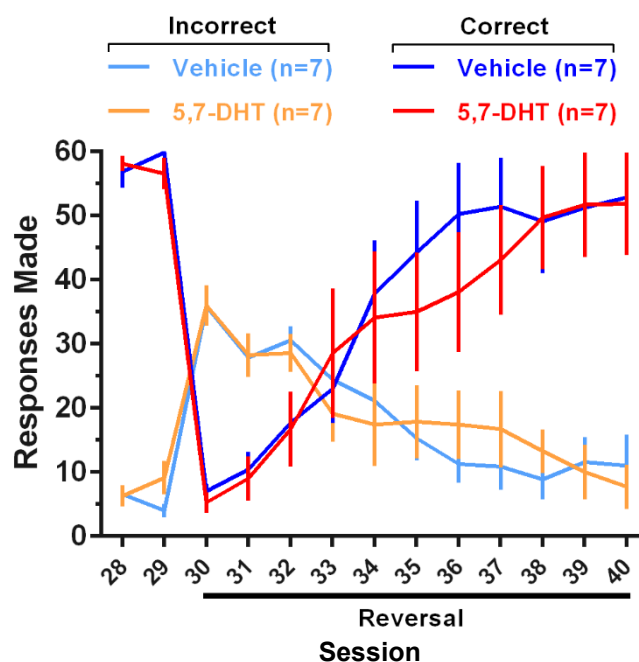


Figure 8.4 Responses made during the learned irrelevance test for 5,7-DHT or vehicle infused animals.

A graph showing the number (mean $\pm$ SEM) of correct and incorrect responses for each session (x-axis) during the learned irrelevance test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline).

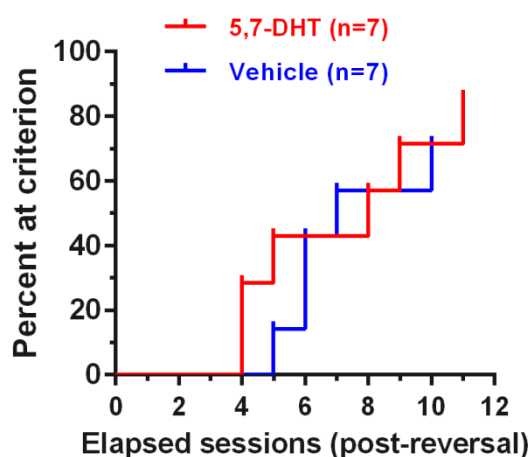


Figure 8.5 Sessions to criterion during the perseverance test for 5,7-DHT or vehicle infused animals.

A survival curve showing the percentage of animals performing at criterion (>80% correct on subsequent sessions) during the learned irrelevance test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6µg 5,7-DHT or vehicle (0.2% ascorbate saline).

8.3.4 Extinction

During the extinction phase, there was no effect of lesion on the number of responses to the formerly rewarded stimulus ($F_{(1,26)} = 0.59$, $p = 0.45$) or unrewarded stimulus ($F_{(1,26)} = 0.015$, $p = 0.91$; see Figure 8.6).

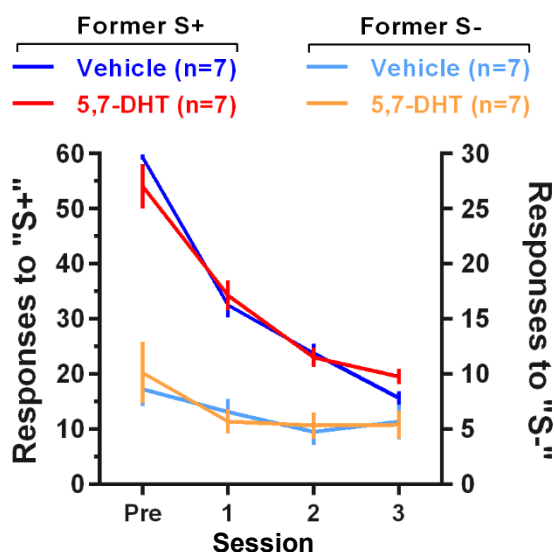


Figure 8.6 Responses made during the extinction test for 5,7-DHT or vehicle infused animals.

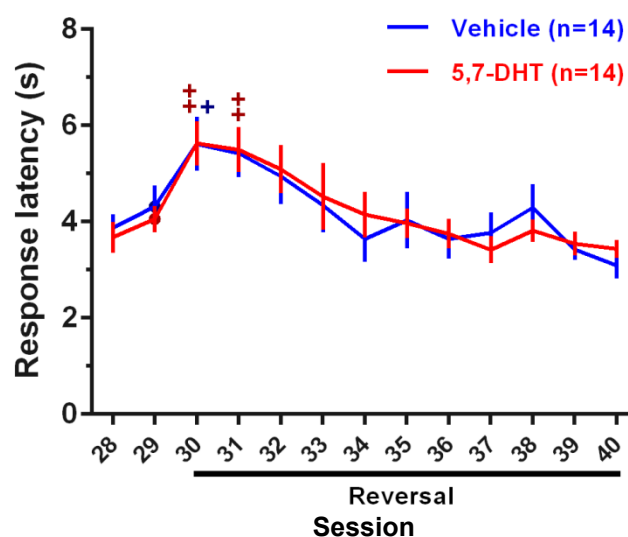
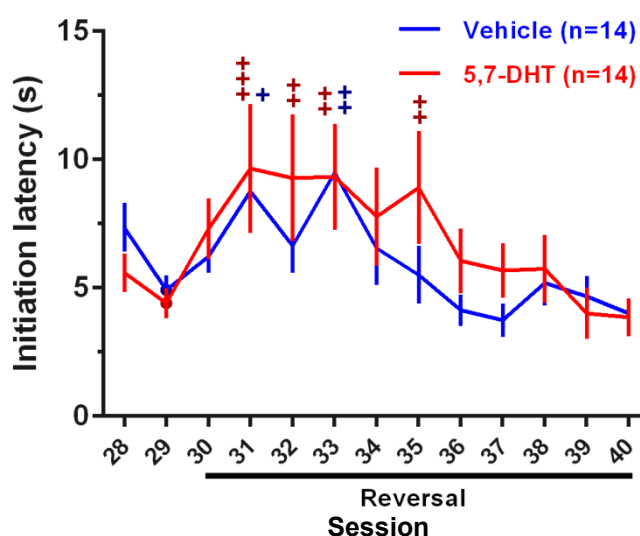
A graph showing the number (mean ± SEM) of responses made to the formerly correct (S+) and incorrect (S-) stimulus for each session (x-axis) during the extinction test, in **male Lister-hooded rats** that had received intra-OFC infusion of either 6µg 5,7-DHT or vehicle (0.2% ascorbate saline).

8.3.5 Latency measures

As per most of the previous experiments, there was a significant effect of session on initiation latency ($F_{(8, 208)} = 8.1, p < 0.0001$) and response latency ($F_{(8, 208)} = 13.1, p < 0.0001$) in the nine sessions post-reversal.

There was no effect of lesion or any lesion × session interaction for these measures (all $F_s < 1, p > 0.4$), but post-hoc tests did show significant within-treatment increases in initiation latency during reversal sessions 2, 3, 4, and 6 (see Figure 8.7). There were also significant within-treatment increases in response latency in the sessions immediately after reversal in both treatment groups (all $p < 0.05$).

There was no significant effect of lesion on collection latency ($F_{(1, 26)} < 0.0001$, $p > 0.99$), nor any lesion $\times$ session interaction ($F_{(8, 208)} = 0.4$, $p = 0.93$) however there was a significant effect of session with post-hoc tests showing a significant within-treatment increase in the vehicle group on reversal session 4 ($t_{234} = 3.6$, $p = 0.004$).



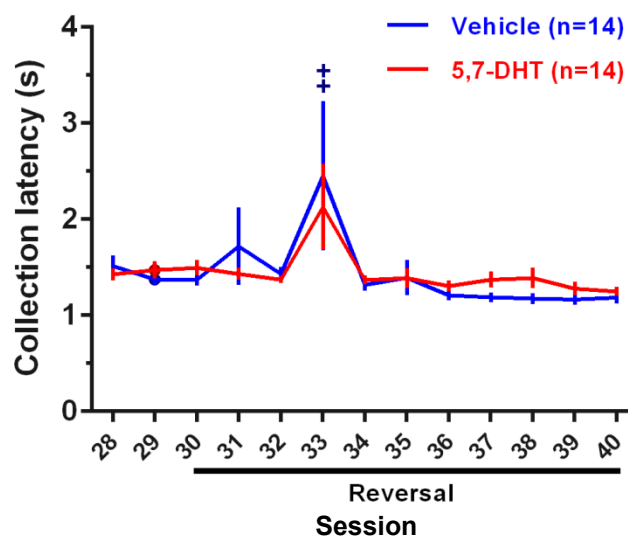


Figure 8.7 Latency measures during the perseverance and learned irrelevance test (collated) for 5,7-DHT or vehicle infused animals.

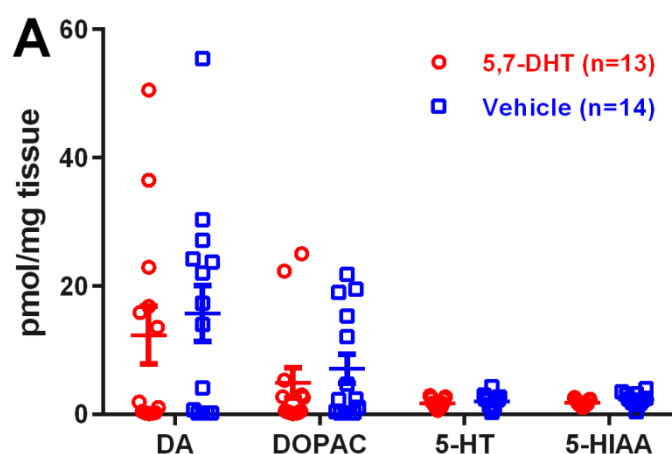
The trial latencies (mean $\pm$ SEM) for each session (x-axis) during the perseverance and learned irrelevance tests (collated), in cohorts of **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline). + $p < 0.05$, ++ $p < 0.01$, +++ $p < 0.001$ versus pre-reversal session 29. Post-hoc Sidak test following two-way RM-ANOVA.

Across extinction sessions, all animals took longer to initiate trials ($F_{(2, 52)} = 13.0$, $p < 0.0001$) and respond to the images ($F_{(2, 52)} = 29.0$, $p < 0.0001$). There was no effect of lesion on initiation latency ($F_{(1, 26)} = 0.1$, $p = 0.75$) or response latency ($F_{(1, 26)} = 0.1$, $p = 0.8$).

8.3.6 Post-mortem neurochemistry

For one of the 5,7-DHT-treated animals, the lateral OFC brain sample was lost. HPLC analysis of post-mortem brain tissue showed no significant effect of lesion on any neurotransmitter in any of the sampled brain regions (lateral OFC, medial PFC, brainstem; all $F_s < 2$, $p > 0.1$; see Figure 8.8).

Similar to OFC 6-hydroxydopamine lesions (see section 7.3.8), the neurotransmitter and metabolite levels were below the published average but still within 1 standard error of the mean (Table 7.1). A Spearman's rank analysis showed no correlation between OFC 5-HT levels (see Figure 8.9) and the number of sessions taken to reach criterion in either the perseverance ($r = 0.17$, $p = 0.59$) or learned irrelevance test ($r = -0.17$, $p = 0.56$). Nor did OFC dopamine levels correlate with sessions to criterion for the perseverance ($r = -0.21$, $p = 0.49$) or learned irrelevance test ($r = 0.47$, $p = 0.09$).



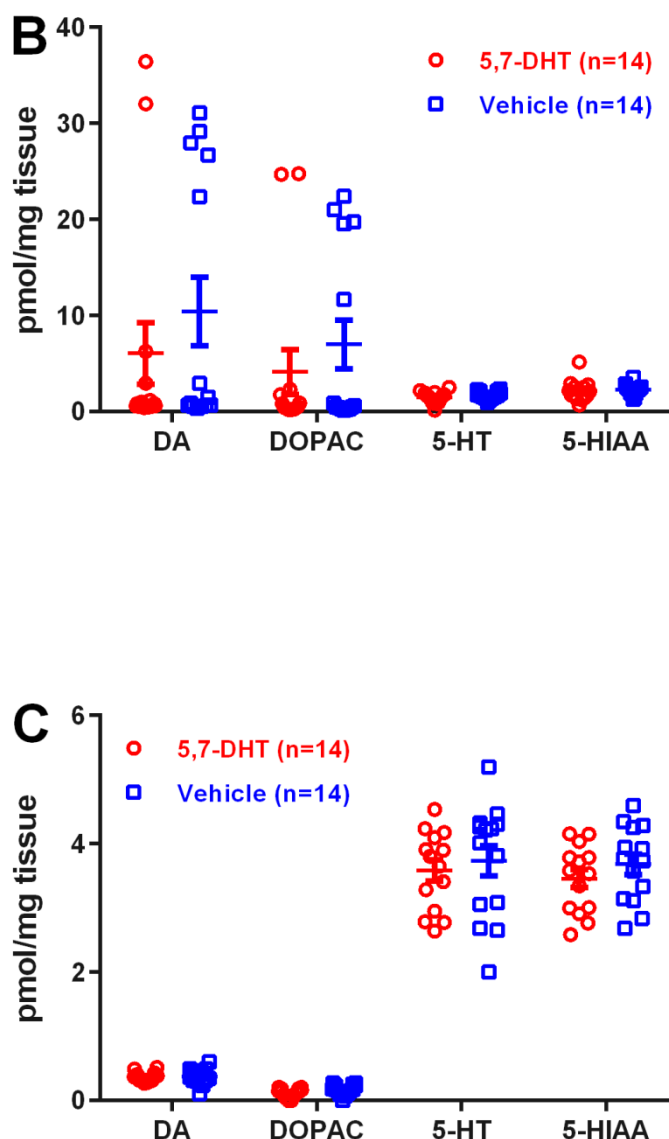


Figure 8.8 Neurochemistry for 5,7-DHT or vehicle infused animals.

Graphs showing the concentrations (mean $\pm$ SEM with individual data points) of dopamine (DA), DOPAC, 5-HT, and 5-HIAA in the lateral OFC (A), medial PFC (B), and brainstem (C), in cohorts of male Lister-hooded rats that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline) approximately 8 weeks prior to tissue collection.

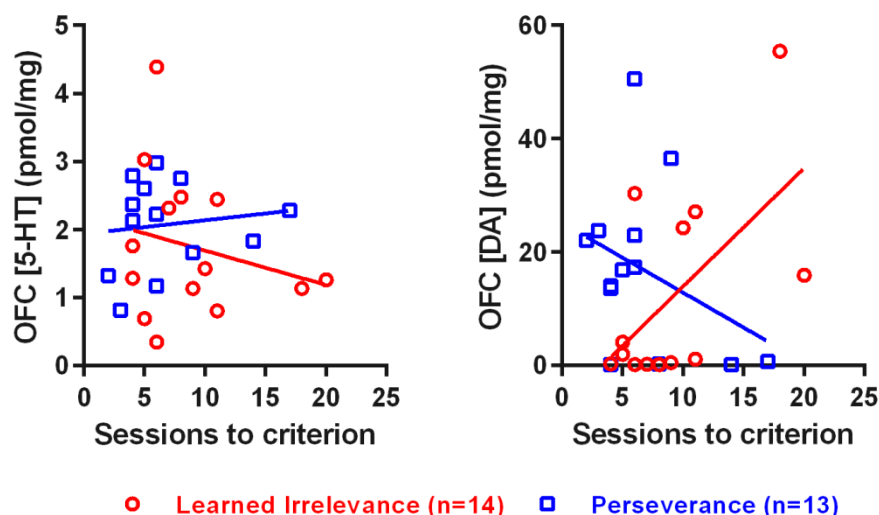


Figure 8.9 Neurochemical-behaviour correlates for 5,7-DHT or vehicle infused animals.

Graphs showing the relationship between number of sessions taken to reach criterion after reversal, and post-mortem lateral OFC concentrations of 5-HT and dopamine, in cohorts of **male Lister-hooded rats** that had received intra-OFC infusion of either 6 μ g 5,7-DHT or vehicle (0.2% ascorbate saline) **approximately 8 weeks prior to tissue collection**. Lines represent the least squares best-fit.

8.4 Discussion

This experiment investigated the effects of a neurotoxic lesion targeting orbitofrontal serotonin signalling on reversal learning. The main finding was that animals treated with the toxin performed similarly to sham-treated animals. Post-mortem neurochemistry was unable to confirm if the toxin had depleted serotonin levels, as has been previously reported (Tran et al., 2013).

One potential reason that the toxin failed to produce effects in this experiment may be due to the dose used. Although a previous study found that an 8 μ g 5,7-DHT OFC infusion produced impairments of reinforcer devaluation (West et al., 2013), the effect size was small and may be

attributable to the use of tenuous statistical reasoning (see Nieuwenhuis et al., 2011), since there was no significant difference between sham treated and lesioned animals, but shams had a devaluation index significantly different from zero (one-tailed) whereas lesions did not. Additionally, that study found no lesion effect on extinction, concordant with the results of this experiment. Another study found that a 6µg 5,7-DHT amygdala infusion failed to affect simple reversal learning, despite producing ~40% 5-HT depletion measured 68 days post-surgery (Ochoa et al., 2015).

Consistent with previous chapters, the rats in this experiment showed increases in the time to initiate trials and respond to images after the reversal, suggestive of demotivation due to reversal-induced negative feedback. There was no effect of OFC 5,7-DHT infusion on any latency measure, despite previous studies showing that global 5-HT depletion attenuates reversal-induced deceleration (Bari et al., 2010).

Implications for theories of OFC 5-HT function

In primate probabilistic reversal learning tasks, computational analyses show that reinforcement sensitivity (i.e. learning rate) is reduced by OFC 5-HT depletion (Rygula et al., 2014), but enhanced by OFC dopamine depletion (Clarke et al., 2014). In rats, global 5-HT depletion slows the rate of serial probabilistic reversal learning (Bari et al., 2010), which may in part be due to a slower learning rate, although they also found increased sensitivity to false negative feedback, which – in marmosets – has been shown to be due to 5-HT depletion in the amygdala, and not OFC (Rygula et al., 2014). Other studies show that disrupting amygdala function can improve or impair

reversal learning, depending on experimental conditions (Ochoa et al., 2015), leading to the proposition that the amygdala is important for comparing recent and long-term valuations (Jang et al., 2015).

In deterministic reversal learning tasks, OFC 5-HT depletion impairs serial reversal learning in primates (Clarke et al., 2007). In a split reversal with perseverance and learned irrelevance tests, there was no effect of lesion on the first reversal, but an increased number of errors in the perseverance test on the second reversal. This may explain the lack of effect in this experiment, as only one reversal was tested. Likewise, an earlier study found no effect of OFC 5-HT depletion on the first reversal of a simple reversal, but a large effect on subsequent reversals (Clarke et al., 2004).

In rodents, global 5-HT depletion via para-chlorophenylalanine (PCPA) also impairs deterministic reversal learning, which is ameliorated by the atypical SSRI vortioxetine (Wallace et al., 2014). Another study showed that reversal performance was correlated with the degree of para-chloroamphetamine induced 5-HT depletion in the OFC, mPFC and amygdala (Masaki et al., 2006). Meanwhile, chronic intermittent cold stress led to impaired reversal learning, and these rats failed to show a reversal-induced increase in OFC 5-HT release (Lapiz-Bluhm et al., 2009). Taken together, these studies strongly implicate the importance of 5-HT in reversal learning, but lack causal evidence for OFC 5-HT.

In conclusion, this experiment found that intra-OFC infusion of the serotonin neurotoxin 5,7-DHT failed to reduce 5-HT levels and to affect learning about formerly positive or negative associations in rodents. Given

the evidence that OFC 5-HT depletion perturbs reinforcement learning parameters in primates, and that global 5-HT depletion consistently disrupts reversal learning in rodents, it is suggested that incomplete lesion and/or task conditions were the reason for this null finding.

Chapter 9 – General Discussion

9.1 Summary of findings

The aim of this thesis was to further understanding of how associative corticostriatal circuitry is disrupted in compulsivity by manipulating brain neurotransmitter systems in ways that would disrupt reversal learning in specific ways (see Figure 1.8). Table 9.1 summarises the experimental findings. Briefly, Chapter 2 showed that acute administration of the NMDA receptor antagonist PCP did not affect rats' ability to perform a learned discrimination. Chapter 3 showed that PCP impaired rats' ability to reverse a learned association, and Chapter 4 showed that this occurred if the formerly rewarded stimulus (CS+) remained as a choice, but not when only the formerly unrewarded stimulus (CS-) remained. Chapter 5 showed that the 5-HT_{2c} receptor antagonist SB-242,084 did not affect discrimination ability or reversal learning, but did accelerate responding. Chapter 6 showed that the dual-hit early-life manipulation of PCP and social isolation led to accelerated reversal learning when the CS+ was present. Chapters 7 and 8 showed that infusion of the dopamine neurotoxin 6-hydroxydopamine or the serotonin neurotoxin 5,7-dihydroxytryptamine into the lateral orbitofrontal cortex **caused** no depletion of these neurotransmitters, and subsequently, had no effect on reversal learning.

Table 9.1 A summary of the experimental findings of this thesis.

Each chapter describes the effects of various treatments on forms of reversal learning. Arrows represent direction of performance changes (i.e. proportion of correct-to-incorrect responses) relative to control groups.

Chapter	Treatment	Acquisition		Reversal		
		Early	Late	Simple	Perseverance (CS+ → S-)	Learned Irrelevance (CS- → S+)
3, 4	PCP		=	↓	↓	=
5	SB-242084		=		=	=
6	PCP-Isolation	=			↑	=
7	OFC 6-OHDA		=		=	=
8	OFC 5,7-DHT		=		=	=

9.2 Interpretations for compulsivity

9.2.1 Split task interpretation

As described in Chapter 1, compulsive behaviour is thought to arise from a disrupted balance between goal-directed (model-based) and habitual (model-free) corticostriatal circuitry. In particular, disrupted signalling in the OFC, whether it is via 5-HT, glutamate, or dopamine systems, disrupts model-based behaviour, albeit in different ways (see Figure 1.8). 5-HT and glutamate have been implicated in OCD, whilst NMDA receptor and (prefrontal) dopamine hypofunction are associated with schizophrenia; and abnormal glutamate regulation in addiction. Dopamine medication-induced compulsivity in Parkinson's disease is more likely striatal-mediated (Voon et al., 2010), although some functional imaging evidence implicates impaired top-down modulation by OFC (van Eimeren et al., 2009).

The experiments in this thesis investigated treatments with relevant targets within this circuitry administered to rodents performing reversal learning tasks. These tasks encompass different stages/phases that rely differently on goal-directed and habitual learning (see Figure 1.6), therefore treatments were administered during the late acquisition stage to assess habitual learning, and during the reversal stage to assess goal-directed learning. Any significant change in choice behaviour during late acquisition would have been interpreted as a disruption to habitual circuitry; however this did not occur for any treatments in this thesis. Meanwhile slower learning post-reversal was interpreted as a disruption in the balance of corticostriatal circuitry, such that the associative (goal-directed) circuits were slower to infer a change in the state of the environment and/or infer outcomes of alternative choices, driving behaviour towards sampling those choices.

Measuring this uncertainty-induced exploration is problematic with a simple two-choice reversal learning task (where the CS+/CS- reward contingency flips), since it is not clear what factor is responsible for treatment-impaired learning. Continued sampling of the CS+ may be due to impaired devaluation of the CS+, or it may be due to aversion toward the CS-. The former process can be due to attenuated negative RPEs, whilst the latter process can occur with intact negative RPE signalling but impaired exploration. Some proxy measures shed light on this, e.g. omissions and latencies. For example, no increase in latencies or omissions implies that negative RPEs are not being registered and are not producing a

demotivating effect. Decreased selection of the CS+ with increased latencies/omissions is a sign that negative RPEs are being registered, but if not coupled with increased CS- sampling, implies that the reversal is due to impaired uncertainty-induced exploration, which is seen in schizophrenia (and correlates with negative symptoms, Strauss et al., 2011; see section 1.4.2 Schizophrenia). This interpretation may be better elucidated in future experiments with a three-choice task, since random responding toward the irrelevant third stimulus would also increase.

In order to circumvent this problem, treatments were tested in a split reversal design where either the CS+ remains (perseverance test) or the CS- remains (learned irrelevance test). A novelty bias test was also conducted to control for the interpretation that treatment effects were due to differences in novelty exploration. By measuring stimulus choice post-reversal (rather than percentage correct, which masks absolute sampling), less parsimonious interpretations could be made. For example, reversal deficits can be attributed to impairments in devaluation of the CS+ association in the perseverance test or to impairments in re-evaluation of the CS- association in the learned irrelevance test.

The mechanism behind treatment-disrupted choice revaluation is still subject to interpretation however. Ideally, certain permutations of choice patterns across the split test would indicate OFC dysfunction. However, isolating OFC dysfunction is difficult since OFC is implicated in a host of interlinked goal-directed behaviours (see section 1.5.2 Role of OFC), e.g. learning about non-outcomes, inferring outcomes that may have occurred

had an alternative choice been taken (see Figure 1.2), uncertainty-induced exploration, credit assignment (Walton et al., 2010; see Figure 1.7), and environmental state formation.

Thus, differences between the arms of the task could be due to altered exploration and/or altered learning as a result of the factors listed above. The following sections cover additional technical differences between the tasks that provide other interpretations.

9.2.2 Task difficulty

Due to rats' instinctive attraction to novel stimuli, the perseverance task is learned more quickly, since the novel stimulus is rewarded. Therefore, the perseverance task only requires the animal to extinguish a conditioned affinity toward the CS+, whereas the learned irrelevance test requires the animal to extinguish their instinctive affinity to the novel S- and their conditioned aversion toward the CS-. Task difficulty affects treatment sensitivity, since tasks that are too easy promote ceiling effects and tasks that are too difficult promote floor effects. Floor effects in the learned irrelevance test may explain why none of the experiments performed found any treatment-induced learning disturbances, despite variants of this task showing it to be sensitive to OFC and 5-HT disruption (Burke et al., 2009; Nilsson et al., 2012; Tait and Brown, 2007).

9.2.3 Stimulus associability

Until now, it has been assumed that the balance of behavioural control by model-based and model-free systems is single and continuous. The graph in Figure 1.6 implies that only one system can be dominant at any time, thus

the omega parameter (model-based/model-free balance) will be the same for all valuations at a given time. This clearly isn't the case since if a novel choice is introduced alongside a habitual choice (such as in the novelty bias test), the choice can't be under model-free control, since it hasn't been associated with any outcome yet (and thus doesn't have a valuation in the model-free system). Thus the balance of control between circuitry depends not just on choice familiarity, but also choice certainty, which in turn is determined by the amount of choice-outcome conditioning.

This introduces a potential confound for interpreting the two tests, since during the acquisition stage, the CS+ association is more strongly conditioned than the CS- association by virtue of the CS+ being chosen – and reinforced – more often. In these experiments, reversal occurred after all animals had reached criterion performance of $\geq 90\%$ correct on two consecutive sessions. Hypothetically, one animal could reach criterion quickly and continue to choose the CS+ 100% of the time, for many sessions, whilst another animal may hover at 50% for a long time before reaching criterion. The first animal would have a much stronger CS+ association and a relatively weak CS- association, since it has not been sampled for such a long period. Conversely the second animal has moderately strong CS+ and CS- associations.

How would these animals expect to differ on the two tests? Associability learning models (e.g. Pearce-Hall; Roesch et al., 2012) would expect animal one to be relatively impaired on the perseverance task – since the CS+ has lower associability (a valuation with higher certainty) – but relatively

enhanced on the learned irrelevance test – since the CS- has higher associability (a valuation with lower certainty). Whereas temporal difference reinforcement learning models (e.g. Rescorla-Wagner) would propose either: that animal one has a higher learning rate, and so would perform better on both tests; or that animal two has a higher exploration rate, and so would initially perform better on both tests. The two-system model described in Chapter 1 may reconcile these findings by considering the omega (model-based/model-free balance) parameter to not only modulate valuation balance between the systems, but also the beta (exploration rate) parameter. Table 9.2 below shows the choice probabilities for the hypothetical animals described above given their different values of the ω parameter toward conditioned stimuli.

Table 9.2 Hypothetical parameters for two animals in the split reversal task.

As described in the text, animal one achieved criterion quickly and exploited the CS+, so this association is strongly consolidated in the model-free system ($\omega_{CS} = 0.05$), whereas the CS- was rarely sampled and so is not strongly consolidated ($\omega_{CS} = 0.5$). Animal two sampled both stimuli roughly equally ($\omega_{CS} = 0.25$). Given the parameters shown plugged into the equation below, animal two would initially perform better on both reversal tests, choosing the now incorrect CS+ less on the perseverance task and choosing the now correct CS- more on the learned irrelevance test.

$$\text{Action selection equation: } p_t(CS | CS, NS) = \frac{e^{Q_t(CS)/\omega_{CS}\beta}}{e^{Q_t(CS)/\omega_{CS}\beta} + e^{Q_t(NS)/\omega_{NS}\beta}}$$

Where *CS* is the conditioned stimulus, ω_{CS} is the system balance for *CS* (viz. certainty), *NS* is the novel stimulus and ω_{NS} is the system balance for *NS*.

Animal	Test	p(CS)	ω_{CS}	ω_{NS}	β	Q _{CS}	Q _{NS}
1	Perseverance	0.999	0.05	0.8	1	0.9	0.7
2		0.938	0.25	0.8	1	0.9	0.7
1	Learned	0.337	0.5	0.8	1	0.1	0.7
2	Irrelevance	0.383	0.25	0.8	1	0.1	0.7

9.2.4 Solving problems of interpretation

Throughout the decision making literature, different interpretations are formulated into models and applied to the data. Models are then compared based on goodness-of-fit and other measures to produce an evidence weighting for how well they explain the data (Stephan and Mathys, 2014).

Unfortunately, the experiments in this thesis were not designed with this in mind, since the “Acquisition without Punishment” phase was added to accelerate learning, and some animals learn to select the S+ at a high rate immediately after this phase, which disallows computational models from accurately gauging the learning rate. In addition, performance can be

affected by other factors such as side bias and image bias, which models must adjust for, however this task used correction trials where the stimuli remained in the same positions after incorrect responses (for max. 6 trials). This was done to prevent animals using a “single side” strategy to receive reward 50% of the time whilst disregarding the visual content of the images. Additionally, the small sample sizes and image combinations meant that image bias couldn’t be assessed in the split reversal task since there was $n=4$ subjects per image pair.

9.3 Future questions and treatment implications

9.3.1 Serotonin

Serotonin is strongly implicated in compulsivity (see section 1.5.2 Role of OFC). However, the serotonergic manipulations in Chapters 5 and 8 failed to produce any effect on choice behaviour. Antagonists at the 5-HT_{2c} receptor have been touted as potential anti-compulsive agents based on previous work showing their ability to reduce compulsive behaviour in rodents (Flaisher-Grinberg et al., 2008; Schepisi et al., 2013). Likewise, SSRIs are efficacious in roughly half of OCD patients (Fineberg et al., 2012). How might SSRIs produce anti-compulsive effects in some people? Maia and Cano-Colino (2015) propose that OFC 5-HT hypofunction prevents neural ensemble shifting, which may lead to failures in new state inference and consequently, perseverative behaviour. It may be that this perseverative neural activity exists in most cases of OCD but is only significantly mediated by 5-HT dysfunction in those who go on to respond to SSRIs. Studies investigating predictors of treatment response in OCD generally show that

SSRIs are more favourable in the forbidden thoughts subtype (aggressive/sexual/religious), and are much less favourable in the contamination, symmetry, and hoarding subtypes (Hazari et al., 2016). It should be noted that serotonin is known to have strong modulatory roles on aggressive and sexual behaviour (Olivier, 2015), therefore the apparent efficacy of SSRIs for these subtypes of OCD may not be due to reduced compulsivity per se, but rather decreased aggressive/sexual thoughts or salience of these thoughts.

One rat study showed that repeated OFC stimulation led to the development of compulsivity that could be reversed by SSRI treatment (Ahmari et al., 2013). Given the model above, it would be intuitive to think that enhanced 5-HT transporter blockade in the OFC is preventing the deleterious effects of the optogenetic stimulation. However, another study showed that OFC lesion-induced compulsive behaviour was ameliorated by intra-striatal SSRI administration (Schilman et al., 2010). This could mean either: that the model is overly simplistic and must factor in striatal 5-HT, or that the compulsivity induced by OFC lesions and OFC disinhibition differ enough to merit different targets for amelioration. Currently, research investigating the role of 5-HT in the striatum is severely lacking. One study in marmosets found that serotonin depletions in the DMS have little effect on reversal learning, whereas dopamine depletions consistently impair it (Clarke et al., 2011).

9.3.2 Glutamate

Converging genetic and neurochemical evidence implicates abnormal glutamate signalling in OCD, schizophrenia, addiction, and other disorders of compulsivity (see Friston et al., 2016; Kalivas, 2009; Pauls et al., 2014). Although none of the mainstay treatments for these disorders target glutamate, glutamate modulators of varying mechanisms such as N-acetylcysteine, D-cycloserine, pomaglumetad, and riluzole have been trialled with mixed success (Olive et al., 2012; Pittenger et al., 2011; Singh and Singh, 2011).

The experiments in chapters 3 and 4 show that PCP consistently induces compulsive responding toward formerly reinforced stimuli. PCP is thought to induce a brain state similar to psychosis by decreasing overall NMDA receptor signalling, which in turn leads to increased AMPA receptor signalling via disinhibition of glutamatergic pyramidal neurons (see Rafter et al., 2016). Therefore, the results of these experiments predict that glutamate modulators that re-balance this signalling, either by AMPA-R antagonism or reducing synaptic glutamate, would facilitate the extinction of compulsive responding in psychosis. Although no studies have investigated this, many have shown that these modulators attenuate the behavioural and neurochemical effects of NMDA-R antagonists in humans (Deakin et al., 2008; Krystal et al., 2005) and rodents (Baker et al., 2008; Moghaddam and Adams, 1998; Moghaddam et al., 1997; Olszewski et al., 2012).

Paradoxically, recent studies in OCD have shown large symptom improvements with acute doses of ketamine (Rodriguez et al., 2013), and

chronic memantine as an augmentation to SSRI treatment (Ghaleiha et al., 2013). It should be noted that ketamine was most efficacious in very severe cases of OCD (intrusive obsessions for >8 hours/day). Memantine on the other hand has a much different pharmacodynamic and pharmacokinetic profile to ketamine (Gideons et al., 2014), and since it did not improve symptoms acutely, this would suggest that its efficacy is due to long-term plastic changes rather than acute NMDA-R antagonism.

Glutamate modulators are also being trialled in human addiction. N-acetylcysteine (NAC), which modulates synaptic glutamate by increasing activity of the glutamate-cysteine antiporter, has been shown to reduce cravings in cocaine-dependent individuals (reviewed in Deepmala et al., 2015). Rodent studies show that NAC is able to ameliorate behavioural and neurochemical alterations induced by previous cocaine exposure, such as outcome devaluation (Corbit et al., 2014), cocaine relapse (Moussawi et al., 2009; Reichel and See, 2012), and striatal glutamate hyperactivity (Corbit et al., 2014; Madayag et al., 2007). Interestingly, NAC has also been shown to attenuate the effects of acute NMDA-R antagonism (Baker et al., 2008), and post-weaning social isolation (Herrmann et al., 2014; Möller et al., 2013a, 2013b). Given that the results of Chapter 6 showed decreased compulsive responding in dual-hit PCP-isolated animals, an effect opposite of acute PCP in Chapters 3 – 4, what effect would NAC be expected to have? Since NAC is multimodal, it may normalise activity in both acute PCP and isolates. That is, NAC attenuates the hyperglutamatergic state induced by acute PCP by increasing extra-synaptic glutamate efflux, in turn stimulating pre-synaptic

metabotropic glutamate autoreceptors which reduce synaptic glutamate release. Meanwhile the glutamate normalising effects in the VTA (Whitaker et al., 2013) would reduce the sensitivity (learning rate) towards positively reinforced stimuli in PCP-isolates.

9.3.3 Reversal learning for modelling compulsivity

In the introductory chapter, it was proposed that reversal learning tasks are suitable for assessing compulsivity due to their reliance on associative and sensorimotor corticostriatal circuitry. However, it was also noted in Table 1.3 that the circuitry would be utilised differently depending on the parameters used. The experiments that followed used a two-choice, touchscreen visual discrimination with a 30 second ITI, based on the work by Bussey et al. (2008) and Mar et al. (2013) showing that Lister-hooded rats are adept at learning these discriminations. What was unknown prior to these experiments was how robust this task design is for assessing reversal learning, especially with the introduction of novel stimuli.

For example, it is assumed that this is a two-choice discrimination on the basis that there are two stimuli to choose from. However these stimuli are two-dimensional human-designed representations of single objects, such as a spider or a plane. Laboratory rats have no concept of white 2D spiders or planes; therefore, they likely see the images as shapes. These shapes may then be grouped together to form one whole, i.e. SpiderPlane on half of the trials and PlaneSpider on the other half. A rat whose S+ is Spider then learns that SpiderPlane = go left, and PlaneSpider = go right, rendering this a conditional discrimination rather than a two-choice task. However evidence

from an unreported pilot study weighs against this, as animals trained on two discriminations ($S_{+Spider}/S_{-Plane}$ and $S_{+Bomb}/S_{-Flower}$) showed transitive inference, e.g. when shown Spider/Flower would consistently choose Spider, regardless of image location. Therefore future experiments would preferably train on multiple discriminations simultaneously, and intermix stimuli on trials to ensure that learning is occurring separately towards each stimulus as a whole.

Alternatively, more choices can be added per trial. This further encourages the animal to engage the stimuli as random response strategies become less reliable at producing reward, i.e. random responding only results in reward $(100 \div n) \%$ of the time, where n is the number of choices. Further, having more than two choices means there is no need to split the reversal into perseverance and learned irrelevance arms, since choice of the CS_{+} and CS_{-} is no longer 100% anti-correlated due to the third stimulus. In a setup where the reversal switch is: $A_{+} B_{-} C_{-} \rightarrow A_{-} B_{+} D_{-}$, affinity toward the CS_{+} (A) and aversion toward the CS_{-} (B) can be gauged by measuring sampling of the novel S_{-} (D). Additionally, during the acquisition stage the irrelevant S_{-} (C) could be regularly switched with novel stimuli, so that when the reversal occurs, a novel context is not present, since animals have become habituated to novel stimuli being introduced.

The experiments in this thesis also only investigated a single reversal. This was on the basis that animals are capable of “learning the reversal” (e.g. acquiring a lose-shift heuristic; Jang et al., 2015), which may be less reliant on orbitofrontal circuitry. Future experiments should be conducted to see

what degree the OFC is necessary for serial reversal transfer. For example, if a rat learns a reversal in a maze, would it be quicker to learn a reversal in an operant task? This may be where OFC assists, since it applies association reversals learned earlier in life to new situations (Jang et al., 2015).

Another useful reinforcement learning measure to have would be decision confidence. Although in these experiments image response time was interpreted as a proxy for confidence, it may not be reliable since it can be affected by a host of other factors, e.g. decelerated when the animal is distracted or accelerated when the animal is in a hurry. In human tasks, participants can report their confidence directly, e.g. in one study, schizophrenics displayed more confidence by holding down a stimulus for a longer period of time (Corlett et al., 2007), fitting with the theories of delusion described earlier (see section 1.4.2 Schizophrenia). In a rodent design, it could be possible to have a gradient of blended images, and rats could report confidence by touching the further extremes of those gradients (with reinforcement being proportional to confidence).

9.4 Concluding remarks

Compulsive behaviours are a debilitating component of several psychiatric disorders. Although currently approved treatments are not particularly effective, better understanding of the neural circuitry of compulsivity has unlocked an abundance of other treatment targets and several of these show promise.

Several lines of evidence show that reversal learning is altered as a function of compulsivity in humans. This thesis showed that disruption to

glutamate systems in the rat brain reliably altered reversal learning, with bidirectional effects of acute disruption (Chapters 3 – 4) versus a developmental disruption (Chapter 6). Serotonergic and dopaminergic lesions did not produce such consistent effects, although there was no evidence that these lesions caused any neurochemical disruption of these systems. Systemic administration of the 5-HT_{2c} receptor antagonist SB-242084 increased task motivation, consistent with previous literature, suggesting a potential therapeutic role in motivational anhedonia.

References

- Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. In *Encyclopedia of Measurement and Statistics*, (Thousand Oaks (CA): Sage), pp. 103–107.
- Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L.S., Weiss, R., Cooper, T.B., Mann, J.J., Heertum, R.L.V., et al. (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8104–8109.
- Adams, B., and Moghaddam, B. (1998). Corticolimbic Dopamine Neurotransmission Is Temporally Dissociated from the Cognitive and Locomotor Effects of Phencyclidine. *J. Neurosci.* 18, 5545–5554.
- Adams, R.A., Stephan, K.E., Brown, H.R., Frith, C.D., and Friston, K.J. (2013). The Computational Anatomy of Psychosis. *Front Psychiatry* 4, 47.
- Ahmari, S.E., Spellman, T., Douglass, N.L., Kheirbek, M.A., Simpson, H.B., Deisseroth, K., Gordon, J.A., and Hen, R. (2013). Repeated Cortico-Striatal Stimulation Generates Persistent OCD-Like Behavior. *Science* 340, 1234–1239.
- Akkerman, S., Prickaerts, J., Steinbusch, H.W.M., and Blokland, A. (2012). Object recognition testing: Statistical considerations. *Behavioural Brain Research* 232, 317–322.
- Alex, K.D., and Pehek, E.A. (2007). Pharmacologic mechanisms of serotonergic regulation of dopamine neurotransmission. *Pharmacology & Therapeutics* 113, 296–320.
- Alsö, J., Nilsson, S.R.O., Gastambide, F., Wang, R. a. H., Dam, S.A., Mar, A.C., Tricklebank, M., and Robbins, T.W. (2015). The role of 5-HT_{2C} receptors in touchscreen visual reversal learning in the rat: a cross-site study. *Psychopharmacology* 232, 4017–4031.
- Anastasio, N.C., and Johnson, K.M. (2008). Atypical Antischizophrenic Drugs Prevent Changes in Cortical N-Methyl-D-Aspartate Receptors and Behavior Following Sub-chronic Phencyclidine Administration in Developing Rat Pups. *Pharmacol Biochem Behav* 90, 569–577.
- Atwood, B.K., Kupferschmidt, D.A., and Lovinger, D.M. (2014). Opioids induce dissociable forms of long-term depression of excitatory inputs to the dorsal striatum. *Nat Neurosci* 17, 540–548.
- Averbeck, B.B., Evans, S., Chouhan, V., Bristow, E., and Shergill, S.S. (2011). Probabilistic learning and inference in schizophrenia. *Schizophrenia Research* 127, 115–122.

- Bailey, M.R., Williamson, C., Mezas, C., Winiger, V., Silver, R., Balsam, P.D., and Simpson, E.H. (2016). The effects of pharmacological modulation of the serotonin 2C receptor on goal-directed behavior in mice. *Psychopharmacology (Berl.)* 233, 615–624.
- Baker, D.A., Madayag, A., Kristiansen, L.V., Meador-Woodruff, J.H., Haroutunian, V., and Raju, I. (2008). Contribution of cystine-glutamate antiporters to the psychotomimetic effects of phencyclidine. *Neuropsychopharmacology* 33, 1760–1772.
- Bari, A., Theobald, D.E., Caprioli, D., Mar, A.C., Aidoo-Micah, A., Dalley, J.W., and Robbins, T.W. (2010). Serotonin Modulates Sensitivity to Reward and Negative Feedback in a Probabilistic Reversal Learning Task in Rats. *Neuropsychopharmacology* 35, 1290–1301.
- Barlow, R.L., Alsiö, J., Jupp, B., Rabinovich, R., Shrestha, S., Roberts, A.C., Robbins, T.W., and Dalley, J.W. (2015). Markers of Serotonergic Function in the Orbitofrontal Cortex and Dorsal Raphé Nucleus Predict Individual Variation in Spatial-Discrimination Serial Reversal Learning. *Neuropsychopharmacology* 40, 1619–1630.
- Barnes, T.R. (2011). Evidence-based guidelines for the pharmacological treatment of schizophrenia: recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* 25, 567–620.
- Beck, K., Javitt, D.C., and Howes, O.D. (2016). Targeting glutamate to treat schizophrenia: lessons from recent clinical studies. *Psychopharmacology* 233, 2425–2428.
- Belujon, P., and Grace, A.A. (2014). Restoring Mood Balance in Depression: Ketamine Reverses Deficit in Dopamine-Dependent Synaptic Plasticity. *Biological Psychiatry* 76, 927–936.
- Biele, G., Rieskamp, J., Krugel, L.K., and Heekeren, H.R. (2011). The Neural Basis of Following Advice. *PLoS Biol* 9, e1001089.
- Bissonette, G.B., and Roesch, M.R. (2015). Rule encoding in dorsal striatum impacts action selection. *European Journal of Neuroscience* 42, 2555–2567.
- Björklund, A., and Dunnett, S.B. (2007). Dopamine neuron systems in the brain: an update. *Trends in Neurosciences* 30, 194–202.
- Boctor, S.Y., and Ferguson, S.A. (2010). Altered adult locomotor activity in rats from phencyclidine treatment on postnatal days 7, 9 and 11, but not repeated ketamine treatment on postnatal day 7. *NeuroToxicology* 31, 42–54.
- Bohn, I., Giertler, C., and Hauber, W. (2003). NMDA Receptors in the Rat Orbital Prefrontal Cortex are Involved in Guidance of Instrumental Behaviour under Reversal Conditions. *Cereb. Cortex* 13, 968–976.

- Bolam, J.P., and Ellender, T.J. (2016). Histamine and the striatum. *Neuropharmacology* 106, 74–84.
- Boudreau, A.C., and Wolf, M.E. (2005). Behavioral Sensitization to Cocaine Is Associated with Increased AMPA Receptor Surface Expression in the Nucleus Accumbens. *J. Neurosci.* 25, 9144–9151.
- Boulougouris, V., and Robbins, T.W. (2010). Enhancement of Spatial Reversal Learning by 5-HT_{2C} Receptor Antagonism Is Neuroanatomically Specific. *J. Neurosci.* 30, 930–938.
- Boulougouris, V., Dalley, J.W., and Robbins, T.W. (2007a). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research* 179, 219–228.
- Boulougouris, V., Glennon, J.C., and Robbins, T.W. (2007b). Dissociable Effects of Selective 5-HT_{2A} and 5-HT_{2C} Receptor Antagonists on Serial Spatial Reversal Learning in Rats. *Neuropsychopharmacology* 33, 2007–2019.
- Boureau, Y.-L., Sokol-Hessner, P., and Daw, N.D. (2015). Deciding How To Decide: Self-Control and Meta-Decision Making. *Trends in Cognitive Sciences* 19, 700–710.
- Bradfield, L.A., Bertran-Gonzalez, J., Chieng, B., and Balleine, B.W. (2013). The Thalamostriatal Pathway and Cholinergic Control of Goal-Directed Action: Interlacing New with Existing Learning in the Striatum. *Neuron* 79, 153–166.
- Bradfield, L.A., Dezfouli, A., van Holstein, M., Chieng, B., and Balleine, B.W. (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron* 88, 1268–1280.
- Brigman, J.L., Daut, R.A., Wright, T., Gunduz-Cinar, O., Graybeal, C., Davis, M.I., Jiang, Z., Saksida, L.M., Jinde, S., Pease, M., et al. (2013). GluN2B in corticostriatal circuits governs choice learning and choice shifting. *Nat Neurosci* 16, 1101–1110.
- Bromberg-Martin, E.S., Matsumoto, M., Hong, S., and Hikosaka, O. (2010). A Pallidus-Habenula-Dopamine Pathway Signals Inferred Stimulus Values. *Journal of Neurophysiology* 104, 1068–1076.
- de Bruin, N.M.W.J., van Drimmelen, M., Kops, M., van Elk, J., Wetering, M.M. de, and Schwienbacher, I. (2013). Effects of risperidone, clozapine and the 5-HT₆ antagonist GSK-742457 on PCP-induced deficits in reversal learning in the two-lever operant task in male Sprague Dawley rats. *Behavioural Brain Research* 244, 15–28.
- Burguière, E., Monteiro, P., Feng, G., and Graybiel, A.M. (2013). Optogenetic Stimulation of Lateral Orbitofronto-Striatal Pathway Suppresses Compulsive Behaviors. *Science* 340, 1243–1246.

- Burguière, E., Monteiro, P., Mallet, L., Feng, G., and Graybiel, A.M. (2015). Striatal circuits, habits, and implications for obsessive-compulsive disorder. *Current Opinion in Neurobiology* 30, 59–65.
- Burke, K.A., Takahashi, Y.K., Correll, J., Leon Brown, P., and Schoenbaum, G. (2009). Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with “disinhibition” of responding for previously unrewarded cues. *European Journal of Neuroscience* 30, 1941–1946.
- Bussey, T.J., Padain, T.L., Skillings, E.A., Winters, B.D., Morton, A.J., and Saksida, L.M. (2008). The touchscreen cognitive testing method for rodents: How to get the best out of your rat. *Learn. Mem.* 15, 516–523.
- Calaminus, C., and Hauber, W. (2008). Guidance of instrumental behavior under reversal conditions requires dopamine D1 and D2 receptor activation in the orbitofrontal cortex. *Neuroscience* 154, 1195–1204.
- Cardno, A.G., and Gottesman, I.I. (2000). Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am. J. Med. Genet.* 97, 12–17.
- Cassaday, H.J., Norman, C., Shilliam, C.S., Vincent, C., and Marsden, C.A. (2003). Intraventricular 5,7-dihydroxytryptamine lesions disrupt acquisition of working memory task rules but not performance once learned. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 27, 147–156.
- Chang, C.Y., Esber, G.R., Marrero-Garcia, Y., Yau, H.-J., Bonci, A., and Schoenbaum, G. (2016). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nat Neurosci* 19, 111–116.
- Clarke, H.F., Dalley, J.W., Crofts, H.S., Robbins, T.W., and Roberts, A.C. (2004). Cognitive Inflexibility After Prefrontal Serotonin Depletion. *Science* 304, 878–880.
- Clarke, H.F., Walker, S.C., Dalley, J.W., Robbins, T.W., and Roberts, A.C. (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cereb. Cortex* 17, 18–27.
- Clarke, H.F., Hill, G.J., Robbins, T.W., and Roberts, A.C. (2011). Dopamine, But Not Serotonin, Regulates Reversal Learning in the Marmoset Caudate Nucleus. *J. Neurosci.* 31, 4290–4297.
- 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., et al. (2014). Orbitofrontal Dopamine Depletion Upregulates Caudate Dopamine and Alters Behavior via Changes in Reinforcement Sensitivity. *J. Neurosci.* 34, 7663–7676.
- Collins, A.G.E., and Frank, M.J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral,

computational, and neurogenetic analysis. *European Journal of Neuroscience* 35, 1024–1035.

Collins, A.G.E., Brown, J.K., Gold, J.M., Waltz, J.A., and Frank, M.J. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *J. Neurosci.* 34, 13747–13756.

Cools, R., Altamirano, L., and D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* 44, 1663–1673.

Corbit, L.H., and Janak, P.H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *European Journal of Neuroscience* 31, 1312–1321.

Corbit, L.H., Chieng, B.C., and Balleine, B.W. (2014). Effects of repeated cocaine exposure on habit learning and reversal by N-acetylcysteine. *Neuropsychopharmacology* 39, 1893–1901.

Corlett, P.R., and Fletcher, P.C. (2015). Delusions and prediction error: clarifying the roles of behavioural and brain responses. *Cogn Neuropsychiatry* 20, 95–105.

Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R.F., Shanks, D.R., Robbins, T.W., Bullmore, E.T., Dickinson, A., and Fletcher, P.C. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* 130, 2387–2400.

Cuenya, L., Mustaca, A., and Kamenetzky, G. (2015). Postweaning isolation affects responses to incentive contrast in adulthood. *Dev Psychobiol* 57, 177–188.

Dalton, G.L., Wang, N.Y., Phillips, A.G., and Floresco, S.B. (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *J. Neurosci.* 36, 1996–2006.

Dawson, N., Thompson, R.J., McVie, A., Thomson, D.M., Morris, B.J., and Pratt, J.A. (2012). Modafinil Reverses Phencyclidine-Induced Deficits in Cognitive Flexibility, Cerebral Metabolism, and Functional Brain Connectivity. *Schizophr Bull* 38, 457–474.

Dawson, N., Xiao, X., McDonald, M., Higham, D.J., Morris, B.J., and Pratt, J.A. (2014). Sustained NMDA Receptor Hypofunction Induces Compromised Neural Systems Integration and Schizophrenia-Like Alterations in Functional Brain Networks. *Cereb. Cortex* 24, 452–464.

Dayan, P., and Huys, Q. (2015). Serotonin's many meanings elude simple theories. *eLife Sciences* 4, e07390.

- Deakin, J.F.W., Lees, J., McKie, S., Hallak, J.E.C., Williams, S.R., and Dursun, S.M. (2008). Glutamate and the neural basis of the subjective effects of ketamine: A pharmaco-magnetic resonance imaging study. *Arch Gen Psychiatry* 65, 154–164.
- Deepmala, Slattey, J., Kumar, N., Delhey, L., Berk, M., Dean, O., Spielholz, C., and Frye, R. (2015). Clinical trials of N-acetylcysteine in psychiatry and neurology: A systematic review. *Neuroscience & Biobehavioral Reviews* 55, 294–321.
- Dolan, R.J., and Dayan, P. (2013). Goals and Habits in the Brain. *Neuron* 80, 312–325.
- Durstewitz, D., and Seamans, J.K. (2008). The Dual-State Theory of Prefrontal Cortex Dopamine Function with Relevance to Catechol-O-Methyltransferase Genotypes and Schizophrenia. *Biological Psychiatry* 64, 739–749.
- van Eimeren, T., Ballanger, B., Pellecchia, G., Miyasaki, J.M., Lang, A.E., and Strafella, A.P. (2009). Dopamine Agonists Diminish Value Sensitivity of the Orbitofrontal Cortex: A Trigger for Pathological Gambling in Parkinson's Disease? *Neuropsychopharmacology* 34, 2758–2766.
- Elsworth, J.D., Morrow, B.A., Hajszan, T., Leranath, C., and Roth, R.H. (2011). Phencyclidine-induced Loss of Asymmetric Spine Synapses in Rodent Prefrontal Cortex is Reversed by Acute and Chronic Treatment with Olanzapine. *Neuropsychopharmacology* 36, 2054–2061.
- Eshel, N., Bukwich, M., Rao, V., Hemmelder, V., Tian, J., and Uchida, N. (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* 525, 243–246.
- Eshel, N., Tian, J., Bukwich, M., and Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nat Neurosci* 19, 479–486.
- Everitt, B.J., and Robbins, T.W. (2016). Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. *Annual Review of Psychology* 67, 23–50.
- Fabrizius, K., Steiniger-Brach, B., Helboe, L., Fink-Jensen, A., and Wörtwein, G. (2011). Socially isolated rats exhibit changes in dopamine homeostasis pertinent to schizophrenia. *International Journal of Developmental Neuroscience* 29, 347–350.
- Faure, A., Leblanc-Veyrac, P., and El Massioui, N. (2010). Dopamine agonists increase perseverative instrumental responses but do not restore habit formation in a rat model of Parkinsonism. *Neuroscience* 168, 477–486.
- Fellini, L., Kumar, G., Gibbs, S., Steckler, T., and Talpos, J. (2014). Re-evaluating the PCP challenge as a pre-clinical model of impaired cognitive

- flexibility in schizophrenia. *European Neuropsychopharmacology* 24, 1836–1849.
- Fellows, L.K. (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Annals of the New York Academy of Sciences* 1239, 51–58.
- Figee, M., Luijckes, J., Smolders, R., Valencia-Alfonso, C.-E., van Wingen, G., de Kwaasteniet, B., Mantione, M., Ooms, P., de Koning, P., Vulink, N., et al. (2013). Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. *Nat Neurosci* 16, 386–387.
- Fineberg, S.K., and Corlett, P.R. (2016). The doxastic shear pin: delusions as errors of learning and memory. *Cognitive Neuropsychiatry* 21, 73–89.
- Fineberg, N.A., Potenza, M.N., Chamberlain, S.R., Berlin, H.A., Menzies, L., Bechara, A., Sahakian, B.J., Robbins, T.W., Bullmore, E.T., and Hollander, E. (2009). Probing Compulsive and Impulsive Behaviors, from Animal Models to Endophenotypes: A Narrative Review. *Neuropsychopharmacology* 35, 591–604.
- Fineberg, N.A., Brown, A., Reghunandanan, S., and Pampaloni, I. (2012). Evidence-based pharmacotherapy of obsessive-compulsive disorder. *International Journal of Neuropsychopharmacology* 15, 1173–1191.
- Fitoussi, A., Dellu-Hagedorn, F., and De Deurwaerdère, P. (2013). Monoamines tissue content analysis reveals restricted and site-specific correlations in brain regions involved in cognition. *Neuroscience* 255, 233–245.
- Flaisher-Grinberg, S., Klavir, O., and Joel, D. (2008). The role of 5-HT_{2A} and 5-HT_{2C} receptors in the signal attenuation rat model of obsessive-compulsive disorder. *International Journal of Neuropsychopharmacology* 11, 811–825.
- Flint, B.A., and Ho, I.K. (1980). Tolerance development to phencyclidine by chronic administration. *Progress in Neuro-Psychopharmacology* 4, 233–239.
- Flint, J., and Munafò, M.R. (2013). Candidate and non-candidate genes in behavior genetics. *Current Opinion in Neurobiology* 23, 57–61.
- Fone, K.C.F., and Porkess, M.V. (2008). Behavioural and neurochemical effects of post-weaning social isolation in rodents—Relevance to developmental neuropsychiatric disorders. *Neuroscience & Biobehavioral Reviews* 32, 1087–1102.
- Frank, M.J. (2008). Schizophrenia: a computational reinforcement learning perspective. *Schizophr Bull* 34, 1008–1011.

- Frank, M.J., Seeberger, L.C., and O'Reilly, R.C. (2004). By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science* 306, 1940–1943.
- Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., and Hutchison, K.E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16311–16316.
- Franklin, N.T., and Frank, M.J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife* 4, e12029.
- Friston, K., Brown, H.R., Siemerikus, J., and Stephan, K.E. (2016). The dysconnection hypothesis (2016). *Schizophr. Res.* 176, 83–94.
- Gage, G.J., Stoetznner, C.R., Wiltschko, A.B., and Berke, J.D. (2010). Selective Activation of Striatal Fast-Spiking Interneurons during Choice Execution. *Neuron* 67, 466–479.
- Gallistel, C.R., Fairhurst, S., and Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13124–13131.
- Gaskin, P.L.R., Alexander, S.P.H., and Fone, K.C.F. (2014). Neonatal phencyclidine administration and post-weaning social isolation as a dual-hit model of “schizophrenia-like” behaviour in the rat. *Psychopharmacology (Berl.)* 231, 2533–2545.
- Gastambide, F., Mitchell, S.N., Robbins, T.W., Tricklebank, M.D., and Gilmour, G. (2013). Temporally distinct cognitive effects following acute administration of ketamine and phencyclidine in the rat. *Eur Neuropsychopharmacol* 23, 1414–1422.
- Gerdeman, G.L., Ronesi, J., and Lovinger, D.M. (2002). Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. *Nat Neurosci* 5, 446–451.
- Ghaleiha, A., Entezari, N., Modabbernia, A., Najand, B., Askari, N., Tabrizi, M., Ashrafi, M., Hajiaghaee, R., and Akhondzadeh, S. (2013). Memantine add-on in moderate to severe obsessive-compulsive disorder: Randomized double-blind placebo-controlled study. *Journal of Psychiatric Research* 47, 175–180.
- Gideons, E.S., Kavalali, E.T., and Monteggia, L.M. (2014). Mechanisms underlying differential effectiveness of memantine and ketamine in rapid antidepressant responses. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8649–8654.
- Gillan, C.M., Papmeyer, M., Morein-Zamir, S., Sahakian, B.J., Fineberg, N.A., Robbins, T.W., and de Wit, S. (2011). Disruption in the Balance Between Goal-Directed Behavior and Habit Learning in Obsessive-Compulsive Disorder. *Am J Psychiatry* 168, 718–726.

- Gillan, C.M., Morein-Zamir, S., Urcelay, G.P., Sule, A., Voon, V., Apergis-Schoute, A.M., Fineberg, N.A., Sahakian, B.J., and Robbins, T.W. (2014a). Enhanced Avoidance Habits in Obsessive-Compulsive Disorder. *Biological Psychiatry* 75, 631–638.
- Gillan, C.M., Morein-Zamir, S., Kaser, M., Fineberg, N.A., Sule, A., Sahakian, B.J., Cardinal, R.N., and Robbins, T.W. (2014b). Counterfactual Processing of Economic Action-Outcome Alternatives in Obsessive-Compulsive Disorder: Further Evidence of Impaired Goal-Directed Behavior. *Biological Psychiatry* 75, 639–646.
- Gillan, C.M., Apergis-Schoute, A.M., Morein-Zamir, S., Urcelay, G.P., Sule, A., Fineberg, N.A., Sahakian, B.J., and Robbins, T.W. (2014c). Functional Neuroimaging of Avoidance Habits in Obsessive-Compulsive Disorder. *Am J Psychiatry* 172, 284–293.
- Gillan, C.M., Kosinski, M., Whelan, R., Phelps, E.A., and Daw, N.D. (2016a). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* 5, e11305.
- Gillan, C.M., Robbins, T.W., Sahakian, B.J., van den Heuvel, O.A., and van Wingen, G. (2016b). The role of habit in compulsivity. *Eur Neuropsychopharmacol* 26, 828–840.
- Gilmour, G., Dix, S., Fellini, L., Gastambide, F., Plath, N., Steckler, T., Talpos, J., and Tricklebank, M. (2012). NMDA receptors, cognition and schizophrenia – Testing the validity of the NMDA receptor hypofunction hypothesis. *Neuropharmacology* 62, 1401–1412.
- Goetghebeur, P., and Dias, R. (2009). Comparison of haloperidol, risperidone, sertindole, and modafinil to reverse an attentional set-shifting impairment following subchronic PCP administration in the rat—a back translational study. *Psychopharmacology* 202, 287–293.
- Gold, J.M., Waltz, J.A., Matveeva, T.M., Kasanova, Z., Strauss, G.P., Herbener, E.S., Collins, A.G.E., and Frank, M.J. (2012). Negative symptoms and the failure to represent the expected reward value of actions: Behavioral and computational modeling evidence. *Arch Gen Psychiatry* 69, 129–138.
- González-Burgos, I., Fletes-Vargas, G., González-Tapia, D., González-Ramírez, M.M., Rivera-Cervantes, M.C., and Martínez-Degollado, M. (2012). Prefrontal serotonin depletion impairs egocentric, but not allocentric working memory in rats. *Neuroscience Research* 73, 321–327.
- Grant, J.E., Fineberg, N., van Ameringen, M., Cath, D., Visser, H., Carmi, L., Pallanti, S., Hollander, E., and van Balkom, A.J.L.M. (2016). New treatment models for compulsive disorders. *European Neuropsychopharmacology* 26, 877–884.
- Grayson, B., Barnes, S.A., Markou, A., Piercy, C., Podda, G., and Neill, J.C. (2015). Postnatal Phencyclidine (PCP) as a Neurodevelopmental Animal

- Model of Schizophrenia Pathophysiology and Symptomatology: A Review. (Springer Berlin Heidelberg), pp. 1–26.
- Green, A.E., Munafò, M.R., DeYoung, C.G., Fossella, J.A., Fan, J., and Gray, J.R. (2008). Using genetic data in cognitive neuroscience: from growing pains to genuine insights. *Nat Rev Neurosci* 9, 710–720.
- Groman, S.M., James, A.S., Seu, E., Crawford, M.A., Harpster, S.N., and Jentsch, J.D. (2013). Monoamine levels within the orbitofrontal cortex and putamen interact to predict reversal learning performance. *Biol. Psychiatry* 73, 756–762.
- Gruner, P., Anticevic, A., Lee, D., and Pittenger, C. (2015). Arbitration between Action Strategies in Obsessive-Compulsive Disorder. *Neuroscientist* 22, 188–198.
- Guillin, O., Diaz, J., Carroll, P., Griffon, N., Schwartz, J.-C., and Sokoloff, P. (2001). BDNF controls dopamine D3 receptor expression and triggers behavioural sensitization. *Nature* 411, 86–89.
- Gurney, K.N., Humphries, M.D., and Redgrave, P. (2015). A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. *PLoS Biol* 13, e1002034.
- Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy* 26, 317–330.
- Haber, S.N., and Knutson, B. (2009). The Reward Circuit: Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacology* 35, 4–26.
- Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J., and Berke, J.D. (2015). Mesolimbic dopamine signals the value of work. *Nat Neurosci* 19, 117–126.
- Hamilton, D.A., and Brigman, J.L. (2015). Behavioral flexibility in rats and mice: contributions of distinct frontocortical regions. *Genes, Brain and Behavior* 14, 4–21.
- Harmer, C.J., and Phillips, G.D. (1998b). Isolation Rearing Enhances Acquisition in a Conditioned Inhibition Paradigm. *Physiology & Behavior* 65, 525–533.
- Harmer, C.J., and Phillips, G.D. (1998a). Isolation Rearing Enhances the Rate of Acquisition of a Discriminative Approach Task but Does Not Affect the Efficacy of a Conditioned Reward. *Physiology & Behavior* 63, 177–184.
- Hartman, M., Steketee, M.C., Silva, S., Lanning, K., and Andersson, C. (2003). Wisconsin Card Sorting Test performance in schizophrenia: the role of working memory. *Schizophrenia Research* 63, 201–217.

- Hazari, N., Narayanaswamy, J.C., and Arumugham, S.S. (2016). Predictors of response to serotonin reuptake inhibitors in obsessive-compulsive disorder. *Expert Rev Neurother* 16, 1175–1191.
- Herrmann, A.P., Benvenutti, R., Pilz, L.K., and Elisabetsky, E. (2014). N-acetylcysteine prevents increased amphetamine sensitivity in social isolation-reared mice. *Schizophrenia Research* 155, 109–111.
- Histed, M.H., Pasupathy, A., and Miller, E.K. (2009). Learning Substrates in the Primate Prefrontal Cortex and Striatum: Sustained Activity Related to Successful Actions. *Neuron* 63, 244–253.
- Homayoun, H., and Moghaddam, B. (2008). Orbitofrontal cortex neurons as a common target for classic and glutamatergic antipsychotic drugs. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18041–18046.
- Howes, O.D., and Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III—The Final Common Pathway. *Schizophr Bull* 35, 549–562.
- Howes, O., McCutcheon, R., and Stone, J. (2015). Glutamate and dopamine in schizophrenia: An update for the 21st century. *J Psychopharmacol* 29, 97–115.
- Humphries, M.D., Khamassi, M., and Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Front. Neurosci.* 6, 9.
- Huys, Q.J.M., Eshel, N., O’Nions, E., Sheridan, L., Dayan, P., and Roiser, J.P. (2012). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Comput Biol* 8, e1002410.
- Idris, N.F., Repeto, P., Neill, J.C., and Large, C.H. (2005). Investigation of the effects of lamotrigine and clozapine in improving reversal-learning impairments induced by acute phencyclidine and d-amphetamine in the rat. *Psychopharmacology* 179, 336–348.
- Ito, M., and Doya, K. (2011). Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. *Current Opinion in Neurobiology* 21, 368–373.
- Izquierdo, A., and Jentsch, J.D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology* 219, 607–620.
- Jang, A.I., Costa, V.D., Rudebeck, P.H., Chudasama, Y., Murray, E.A., and Averbeck, B.B. (2015). The Role of Frontal Cortical and Medial-Temporal Lobe Brain Areas in Learning a Bayesian Prior Belief on Reversals. *J. Neurosci.* 35, 11751–11760.

- Janhunen, S.K., Svärd, H., Talpos, J., Kumar, G., Steckler, T., Plath, N., Lerdrup, L., Ruby, T., Haman, M., Wyler, R., et al. (2015). The subchronic phencyclidine rat model: relevance for the assessment of novel therapeutics for cognitive impairment associated with schizophrenia. *Psychopharmacology* 232, 4059–4083.
- Jaunarajs, K.L.E., George, J.A., and Bishop, C. (2012). L-DOPA-induced dysregulation of extrastriatal dopamine and serotonin and affective symptoms in a bilateral rat model of Parkinson's disease. *Neuroscience* 218, 243–256.
- Jentsch, J.D., and Roth, R.H. (1999). The Neuropsychopharmacology of Phencyclidine: From NMDA Receptor Hypofunction to the Dopamine Hypothesis of Schizophrenia. *Neuropsychopharmacology* 20, 201–225.
- Jentsch, J.D., and Taylor, J.R. (2001). Impaired Inhibition of Conditioned Responses Produced by Subchronic Administration of Phencyclidine to Rats. *Neuropsychopharmacology* 24, 66–74.
- Jones, J.L., Esber, G.R., McDannald, M.A., Gruber, A.J., Hernandez, A., Mirenski, A., and Schoenbaum, G. (2012). Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science* 338, 953–956.
- Kaalund, S.S., Riise, J., Broberg, B.V., Fabricius, K., Karlsen, A.S., Secher, T., Plath, N., and Pakkenberg, B. (2013). Differential expression of parvalbumin in neonatal phencyclidine-treated rats and socially isolated rats. *J. Neurochem.* 124, 548–557.
- Kalinichev, M., Robbins, M.J., Hartfield, E.M., Maycox, P.R., Moore, S.H., Savage, K.M., Austin, N.E., and Jones, D.N.C. (2008). Comparison between intraperitoneal and subcutaneous phencyclidine administration in Sprague–Dawley rats: A locomotor activity and gene induction study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 32, 414–422.
- Kalivas, P.W. (2009). The glutamate homeostasis hypothesis of addiction. *Nat Rev Neurosci* 10, 561–572.
- Kapur, S. (2003). Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *Am J Psychiatry* 160, 13–23.
- Keiflin, R., Reese, R.M., Woods, C.A., and Janak, P.H. (2013). The Orbitofrontal Cortex as Part of a Hierarchical Neural System Mediating Choice between Two Good Options. *J. Neurosci.* 33, 15989–15998.
- Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/Accuracy Trade-Off between the Habitual and the Goal-Directed Processes. *PLoS Comput Biol* 7, e1002055.

- Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M., and Altman, D.G. (2010). Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biol* 8, e1000412.
- Kim, H., Lee, D., and Jung, M.W. (2013). Signals for Previous Goal Choice Persist in the Dorsomedial, but Not Dorsolateral Striatum of Rats. *J. Neurosci.* 33, 52–63.
- Kimchi, E.Y., and Laubach, M. (2009). Dynamic Encoding of Action Selection by the Medial Striatum. *J. Neurosci.* 29, 3148–3159.
- Kish, S.J., Shannak, K., and Hornykiewicz, O. (1988). Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications. *N. Engl. J. Med.* 318, 876–880.
- Kocsis, B. (2012). Differential Role of NR2A and NR2B Subunits in N-Methyl-D-Aspartate Receptor Antagonist-Induced Aberrant Cortical Gamma Oscillations. *Biological Psychiatry* 71, 987–995.
- Kocsis, B., Brown, R.E., McCarley, R.W., and Hajos, M. (2013). Impact of Ketamine on Neuronal Network Dynamics: Translational Modeling of Schizophrenia-Relevant Deficits. *CNS Neurosci Ther* 19, 437–447.
- Koechlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences* 11, 229–235.
- Kohl, S., Schönherr, D.M., Luigjes, J., Denys, D., Mueller, U.J., Lenartz, D., Visser-Vandewalle, V., and Kuhn, J. (2014). Deep brain stimulation for treatment-refractory obsessive compulsive disorder: a systematic review. *BMC Psychiatry* 14, 214.
- Kriekhaus, E.E., and Wolf, G. (1968). Acquisition of sodium by rats: interaction of innate mechanisms and latent learning. *J Comp Physiol Psychol* 65, 197–201.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., and Baker, C.I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12, 535–540.
- Kringelbach, M.L., and Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology* 72, 341–372.
- Krugel, L.K., Biele, G., Mohr, P.N.C., Li, S.-C., and Heekeren, H.R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17951–17956.

- Krystal, J.H., Bennett, A., Abi-Saab, D., Belger, A., Karper, L.P., D'Souza, D.C., Lipschitz, D., Abi-Dargham, A., and Charney, D.S. (2000). Dissociation of ketamine effects on rule acquisition and rule implementation: possible relevance to NMDA receptor contributions to executive cognitive functions. *Biological Psychiatry* 47, 137–143.
- Krystal, J.H., Abi-Saab, W., Perry, E., D'Souza, D.C., Liu, N., Gueorguieva, R., McDougall, L., Hunsberger, T., Belger, A., Levine, L., et al. (2005). Preliminary evidence of attenuation of the disruptive effects of the NMDA glutamate receptor antagonist, ketamine, on working memory by pretreatment with the group II metabotropic glutamate receptor agonist, LY354740, in healthy human subjects. *Psychopharmacology* 179, 303–309.
- Lammel, S., Ion, D.I., Roeper, J., and Malenka, R.C. (2011). Projection-Specific Modulation of Dopamine Neuron Synapses by Aversive and Rewarding Stimuli. *Neuron* 70, 855–862.
- Lapiz-Bluhm, M.D.S., Soto-Piña, A.E., Hensler, J.G., and Morilak, D.A. (2009). Chronic intermittent cold stress and serotonin depletion induce deficits of reversal learning in an attentional set-shifting test in rats. *Psychopharmacology* 202, 329–341.
- Lee, S.W., Shimojo, S., and O'Doherty, J.P. (2014). Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron* 81, 687–699.
- Lex, B., and Hauber, W. (2010). The Role of Dopamine in the Prelimbic Cortex and the Dorsomedial Striatum in Instrumental Conditioning. *Cereb. Cortex* 20, 873–883.
- Li, C.-S.R. (2004). Do schizophrenia patients make more perseverative than non-perseverative errors on the Wisconsin Card Sorting Test? A meta-analytic study. *Psychiatry Research* 129, 179–190.
- Li, N., Wu, X., and Li, L. (2007). Chronic administration of clozapine alleviates reversal-learning impairment in isolation-reared rats. *Behav Pharmacol* 18, 135–145.
- Lucantonio, F., Stalnaker, T.A., Shaham, Y., Niv, Y., and Schoenbaum, G. (2012). The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci* 15, 358–366.
- Lucantonio, F., Caprioli, D., and Schoenbaum, G. (2014a). Transition from “model-based” to “model-free” behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology* 76, Part B, 407–415.
- Lucantonio, F., Takahashi, Y.K., Hoffman, A.F., Chang, C.Y., Bali-Chaudhary, S., Shaham, Y., Lupica, C.R., and Schoenbaum, G. (2014b). Orbitofrontal activation restores insight lost after cocaine use. *Nat Neurosci* 17, 1092–1099.

- Lukkes, J.L., Watt, M.J., Lowry, C.A., and Forster, G.L. (2009). Consequences of post-weaning social isolation on anxiety behavior and related neural circuits in rodents. *Front. Behav. Neurosci.* 3, 18.
- Madayag, A., Lobner, D., Kau, K.S., Mantsch, J.R., Abdulhameed, O., Hearing, M., Grier, M.D., and Baker, D.A. (2007). Repeated N-acetylcysteine administration alters plasticity-dependent effects of cocaine. *J. Neurosci.* 27, 13968–13976.
- Maeng, S., Zarate Jr, C.A., Du, J., Schloesser, R.J., McCammon, J., Chen, G., and Manji, H.K. (2008). Cellular Mechanisms Underlying the Antidepressant Effects of Ketamine: Role of α -Amino-3-Hydroxy-5-Methylisoxazole-4-Propionic Acid Receptors. *Biological Psychiatry* 63, 349–352.
- Maia, T.V., and Cano-Colino, M. (2015). The Role of Serotonin in Orbitofrontal Function and Obsessive-Compulsive Disorder. *Clinical Psychological Science* 3, 460–482.
- Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14, 154–162.
- Mar, A.C., Horner, A.E., Nilsson, S.R.O., Alsiö, J., Kent, B.A., Kim, C.H., Holmes, A., Saksida, L.M., and Bussey, T.J. (2013). The touchscreen operant platform for assessing executive function in rats and mice. *Nat. Protocols* 8, 1985–2005.
- Marcus, M.M., Mathé, J.M., Nomikos, G.G., and Svensson, T.H. (2001). Effects of competitive and non-competitive NMDA receptor antagonists on dopamine output in the shell and core subdivisions of the nucleus accumbens. *Neuropharmacology* 40, 482–490.
- Masaki, D., Yokoyama, C., Kinoshita, S., Tsuchida, H., Nakatomi, Y., Yoshimoto, K., and Fukui, K. (2006). Relationship between limbic and cortical 5-HT neurotransmission and acquisition and reversal learning in a go/no-go task in rats. *Psychopharmacology* 189, 249–258.
- Mathur, B.N., Capik, N.A., Alvarez, V.A., and Lovinger, D.M. (2011). Serotonin Induces Long-Term Depression at Corticostriatal Synapses. *J. Neurosci.* 31, 7402–7411.
- McIntosh, A.L., Ballard, T.M., Steward, L.J., Moran, P.M., and Fone, K.C. (2012). Clozapine reverses recognition memory but not social interaction deficits in a “dual-hit” rat model of schizophrenia. *Schizophrenia Research* 136, Supplement 1, S282–S283.
- McKee, B.L., Kelley, A.E., Moser, H.R., and Andrzejewski, M.E. (2010). Operant learning requires NMDA-receptor activation in the anterior cingulate cortex and dorsomedial striatum, but not in the orbitofrontal cortex. *Behav. Neurosci.* 124, 500–509.

- Menzies, L., Chamberlain, S.R., Laird, A.R., Thelen, S.M., Sahakian, B.J., and Bullmore, E.T. (2008). Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: The orbitofronto-striatal model revisited. *Neuroscience & Biobehavioral Reviews* 32, 525–549.
- van der Meulen, J.A.J., Bilbija, L., Joosten, R.N.J.M.A., de Bruin, J.P.C., and Feenstra, M.G.P. (2003). The NMDA-receptor antagonist MK-801 selectively disrupts reversal learning in rats. *Neuroreport* 14, 2225–2228.
- Mizoguchi, K., Shoji, H., Tanaka, Y., and Tabira, T. (2010). Orbitofrontal dopaminergic dysfunction causes age-related impairment of reversal learning in rats. *Neuroscience* 170, 1110–1119.
- Moghaddam, B., and Adams, B.W. (1998). Reversal of Phencyclidine Effects by a Group II Metabotropic Glutamate Receptor Agonist in Rats. *Science* 281, 1349–1352.
- Moghaddam, B., Adams, B., Verma, A., and Daly, D. (1997). Activation of glutamatergic neurotransmission by ketamine: a novel step in the pathway from NMDA receptor blockade to dopaminergic and cognitive disruptions associated with the prefrontal cortex. *J. Neurosci.* 17, 2921–2927.
- Möller, M., Du Preez, J.L., Viljoen, F.P., Berk, M., and Harvey, B.H. (2013a). N-Acetyl cysteine reverses social isolation rearing induced changes in corticostriatal monoamines in rats. *Metab Brain Dis* 28, 687–696.
- Möller, M., Du Preez, J.L., Viljoen, F.P., Berk, M., Emsley, R., and Harvey, B.H. (2013b). Social isolation rearing induces mitochondrial, immunological, neurochemical and behavioural deficits in rats, and is reversed by clozapine or N-acetyl cysteine. *Brain, Behavior, and Immunity* 30, 156–167.
- Morris, R.W., Quail, S., Griffiths, K.R., Green, M.J., and Balleine, B.W. (2015). Corticostriatal Control of Goal-Directed Action Is Impaired in Schizophrenia. *Biological Psychiatry* 77, 187–195.
- Morrison, P.D., and Murray, R.M. (2009). From Real-World Events to Psychosis: The Emerging Neuropharmacology of Delusions. *Schizophr Bull* 35, 668–674.
- Moussawi, K., Pacchioni, A., Moran, M., Olive, M.F., Gass, J.T., Lavin, A., and Kalivas, P.W. (2009). N-Acetylcysteine reverses cocaine-induced metaplasticity. *Nat Neurosci* 12, 182–189.
- Nelson, A.J.D., Thur, K.E., Marsden, C.A., and Cassaday, H.J. (2010). Catecholaminergic depletion within the prelimbic medial prefrontal cortex enhances latent inhibition. *Neuroscience* 170, 99–106.
- Nieuwenhuis, S., Forstmann, B.U., and Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci* 14, 1105–1107.

- Nieuwenstein, M.R., Aleman, A., and de Haan, E.H.F. (2001). Relationship between symptom dimensions and neurocognitive functioning in schizophrenia: a meta-analysis of WCST and CPT studies. *Journal of Psychiatric Research* 35, 119–125.
- Nilsson, S.R.O., Ripley, T.L., Somerville, E.M., and Clifton, P.G. (2012). Reduced activity at the 5-HT(2C) receptor enhances reversal learning by decreasing the influence of previously non-rewarded associations. *Psychopharmacology (Berl.)* 224, 241–254.
- Nilsson, S.R.O., Somerville, E.M., and Clifton, P.G. (2013). Dissociable Effects of 5-HT_{2C} Receptor Antagonism and Genetic Inactivation on Perseverance and Learned Non-Reward in an Egocentric Spatial Reversal Task. *PLoS ONE* 8, e77762.
- Nilsson, S.R.O., Alsö, J., Somerville, E.M., and Clifton, P.G. (2015). The rat's not for turning: Dissociating the psychological components of cognitive inflexibility. *Neuroscience & Biobehavioral Reviews* 56, 1–14.
- Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520.
- Ochoa, J.G., Stolyarova, A., Kaur, A., Hart, E.E., Bugarin, A., and Izquierdo, A. (2015). Post-training depletions of basolateral amygdala serotonin fail to disrupt discrimination, retention, or reversal learning. *Front. Neurosci* 9, 155.
- Okada, K., Nishizawa, K., Fukabori, R., Kai, N., Shiota, A., Ueda, M., Tsutsui, Y., Sakata, S., Matsushita, N., and Kobayashi, K. (2014). Enhanced flexibility of place discrimination learning by targeting striatal cholinergic interneurons. *Nat Commun* 5, 3778.
- Olive, M.F., Clewa, R.M., Kalivas, P.W., and Malcolm, R.J. (2012). Glutamatergic medications for the treatment of drug and behavioral addictions. *Pharmacology Biochemistry and Behavior* 100, 801–810.
- Olivier, B. (2015). Serotonin: A never-ending story. *European Journal of Pharmacology* 753, 2–18.
- Olszewski, R.T., Janczura, K.J., Ball, S.R., Madore, J.C., Lavin, K.M., Lee, J.C.-M., Lee, M.J., Der, E.K., Hark, T.J., Farago, P.R., et al. (2012). NAAG peptidase inhibitors block cognitive deficit induced by MK-801 and motor activation induced by d-amphetamine in animal models of schizophrenia. *Translational Psychiatry* 2, e145.
- van Os, J., and Kapur, S. (2009). Schizophrenia. *The Lancet* 374, 635–645.
- Otchy, T.M., Wolff, S.B.E., Rhee, J.Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S.M.H., and Ölveczky, B.P. (2015). Acute off-target effects of neural circuit manipulations. *Nature* 528, 358–363.

- Overton, D.D.A. (1966). State-dependent learning produced by depressant and atropine-like drugs. *Psychopharmacologia* 10, 6–31.
- Palencia, C.A., and Ragozzino, M.E. (2004). The influence of NMDA receptors in the dorsomedial striatum on response reversal learning. *Neurobiology of Learning and Memory* 82, 81–89.
- Palencia, C.A., and Ragozzino, M.E. (2005). The contribution of NMDA receptors in the dorsolateral striatum to egocentric response learning. *Behav. Neurosci.* 119, 953–960.
- Pallarés, M.A., Nadal, R.A., Silvestre, J.S., and Ferre, N.S. (1995). Effects Of ketamine, a noncompetitive NMDA antagonist, on the acquisition of the lever-press response in rats. *Physiology & Behavior* 57, 389–392.
- Panayi, M.C., and Killcross, S. (2014). Orbitofrontal cortex inactivation impairs between- but not within-session Pavlovian extinction: An associative analysis. *Neurobiology of Learning and Memory* 108, 78–87.
- Paoletti, P., and Neyton, J. (2007). NMDA receptor subunits: function and pharmacology. *Current Opinion in Pharmacology* 7, 39–47.
- Pauls, D.L., Abramovitch, A., Rauch, S.L., and Geller, D.A. (2014). Obsessive-compulsive disorder: an integrative genetic and neurobiological perspective. *Nat Rev Neurosci* 15, 410–424.
- Pavão, R., Tort, A.B.L., and Amaral, O.B. (2014). Multifactoriality in Psychiatric Disorders: A Computational Study of Schizophrenia. *Schizophr Bull* 41, 980–988.
- Paxinos, G., and Watson, C. (2005). *The Rat Brain in Stereotaxic Coordinates* (Elsevier Academic Press).
- Phillips, G.D., Harmer, C.J., and Hitchcott, P.K. (2002). Isolation rearing-induced facilitation of Pavlovian learning: Abolition by postsession intra-amygdala nafadotride. *Physiology & Behavior* 76, 677–684.
- Phillips, M., Wang, C., and Johnson, K.M. (2001). Pharmacological Characterization of Locomotor Sensitization Induced by Chronic Phencyclidine Administration. *J Pharmacol Exp Ther* 296, 905–913.
- Piechal, A., Blecharz-Klin, K., Wyszogrodzka, E., Kołomańska, P., Rok-Bujko, P., Krząścik, P., Kostowski, W., Widy-Tyszkiewicz, E., Filip, M., and Stefański, R. (2012). Neonatal serotonin (5-HT) depletion does not affect spatial learning and memory in rats. *Pharmacol Rep* 64, 266–274.
- Pittenger, C., Bloch, M.H., and Williams, K. (2011). Glutamate abnormalities in obsessive compulsive disorder: Neurobiology, pathophysiology, and treatment. *Pharmacology & Therapeutics* 132, 314–332.

- Prentice, K.J., Gold, J.M., and Buchanan, R.W. (2008). The Wisconsin Card Sorting impairment in schizophrenia is evident in the first four trials. *Schizophrenia Research* 106, 81–87.
- Quarta, D., and Large, C.H. (2011). Effects of lamotrigine on PCP-evoked elevations in monoamine levels in the medial prefrontal cortex of freely moving rats. *J Psychopharmacol* 25, 1703–1711.
- Quirk, M.C., Sosulski, D.L., Feierstein, C.E., Uchida, N., Mainen, Z.F., Quirk, M.C., Sosulski, D.L., Feierstein, C.E., Uchida, N., and Mainen, Z.F. (2009). A defined network of fast-spiking interneurons in orbitofrontal cortex: responses to behavioral contingencies and ketamine administration. *Front. Syst. Neurosci.* 3, 13.
- Rafter, M., Fone, K.C., and Moran, P.M. (2016). Chapter 10 - Glutamate Pharmacological Models Relevant to Schizophrenia and Psychosis: Can a Receptor Occupancy Normalization Approach Reduce the Gap between Animal and Human Experiments? In *Handbook of Behavioral Neuroscience*, M.V.P. and J.L. Waddington, ed. (Elsevier), pp. 139–174.
- Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology* 20, 262–270.
- Rayburn-Reeves, R.M., Stagner, J.P., Kirk, C.R., and Zentall, T.R. (2013). Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): qualitative differences in behavioral flexibility. *J Comp Psychol* 127, 202–211.
- Reddy, L.F., Waltz, J.A., Green, M.F., Wynn, J.K., and Horan, W.P. (2016). Probabilistic Reversal Learning in Schizophrenia: Stability of Deficits and Potential Causal Mechanisms. *Schizophr Bull* 42, 942–951.
- Redish, A.D. (2004). Addiction as a Computational Process Gone Awry. *Science* 306, 1944–1947.
- Reichel, C.M., and See, R.E. (2012). Chronic N-acetylcysteine after cocaine self-administration produces enduring reductions in drug-seeking. *Neuropsychopharmacology* 37, 298.
- Riceberg, J.S., and Shapiro, M.L. (2012). Reward Stability Determines the Contribution of Orbitofrontal Cortex to Adaptive Behavior. *J. Neurosci.* 32, 16402–16409.
- Robbins, T. (2016). Neurobehavioural sequelae of social deprivation in rodents revisited: Modelling social adversity for developmental neuropsychiatric disorders. *Journal of Psychopharmacology* 30, 1082–1089.
- Robbins, T.W., Gillan, C.M., Smith, D.G., de Wit, S., and Ersche, K.D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn. Sci. (Regul. Ed.)* 16, 81–91.

- Roberts, A.C. (2011). The Importance of Serotonin for Orbitofrontal Function. *Biological Psychiatry* 69, 1185–1191.
- Roberts, A.C., Salvia, M.D., Wilkinson, L.S., Collins, P., Muir, J.L., Everitt, B.J., and 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. *J. Neurosci.* 14, 2531–2544.
- Rodriguez, C.I., Kegeles, L.S., Levinson, A., Feng, T., Marcus, S.M., Vermes, D., Flood, P., and Simpson, H.B. (2013). Randomized Controlled Crossover Trial of Ketamine in Obsessive-Compulsive Disorder: Proof-of-Concept. *Neuropsychopharmacology* 38, 2475–2483.
- Roesch, M.R., Esber, G.R., Li, J., Daw, N.D., and Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience* 35, 1190–1200.
- Romanelli, R.J., Wu, F.M., Gamba, R., Mojtabai, R., and Segal, J.B. (2014). Behavioral therapy and serotonin reuptake inhibitor pharmacotherapy in the treatment of obsessive-compulsive disorder: a systematic review and meta-analysis of head-to-head randomized controlled trials. *Depress Anxiety* 31, 641–652.
- Roopun, A.K., Cunningham, M.O., Racca, C., Alter, K., Traub, R.D., and Whittington, M.A. (2008). Region-Specific Changes in Gamma and Beta2 Rhythms in NMDA Receptor Dysfunction Models of Schizophrenia. *Schizophr Bull* 34, 962–973.
- Rosner, B. (2010). *Fundamentals of Biostatistics* (Cengage Learning).
- Rudebeck, P.H., and Murray, E.A. (2014). The Orbitofrontal Oracle: Cortical Mechanisms for the Prediction and Evaluation of Specific Behavioral Outcomes. *Neuron* 84, 1143–1156.
- Rychlik, M., Bollen, E., and Rygula, R. (2017). Ketamine decreases sensitivity of male rats to misleading negative feedback in a probabilistic reversal-learning task. *Psychopharmacology (Berl)* 234, 613–620.
- Rygula, R., Clarke, H.F., Cardinal, R.N., Cockcroft, G.J., Xia, J., Dalley, J.W., Robbins, T.W., and Roberts, A.C. (2014). Role of Central Serotonin in Anticipation of Rewarding and Punishing Outcomes: Effects of Selective Amygdala or Orbitofrontal 5-HT Depletion. *Cereb. Cortex* 25, 3064–3076.
- Schepisi, C., De Carolis, L., and Nencini, P. (2013). Effects of the 5HT2C antagonist SB242084 on the pramipexole-induced potentiation of water contrafreeloading, a putative animal model of compulsive behavior. *Psychopharmacology (Berl)* 227, 55–66.
- Schilman, E.A., Klavir, O., Winter, C., Sohr, R., and Joel, D. (2010). The Role of the Striatum in Compulsive Behavior in Intact and Orbitofrontal-Cortex-

- Lesioned Rats: Possible Involvement of the Serotonergic System. *Neuropsychopharmacology* 35, 1026–1039.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427.
- Schmitz, Y., Benoit-Marand, M., Gonon, F., and Sulzer, D. (2003). Presynaptic regulation of dopaminergic neurotransmission. *Journal of Neurochemistry* 87, 273–289.
- Schoenbaum, G., Saddoris, M.P., and Stalnaker, T.A. (2007). Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann. N. Y. Acad. Sci.* 1121, 320–335.
- Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews* 95, 853–951.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Sesack, S.R., Hawrylak, V.A., Matus, C., Guido, M.A., and Levey, A.I. (1998). Dopamine Axon Varicosities in the Prelimbic Division of the Rat Prefrontal Cortex Exhibit Sparse Immunoreactivity for the Dopamine Transporter. *J. Neurosci.* 18, 2697–2708.
- Sharpe, M.J., Wikenheiser, A.M., Niv, Y., and Schoenbaum, G. (2015). The State of the Orbitofrontal Cortex. *Neuron* 88, 1075–1077.
- Shen, W., Flajolet, M., Greengard, P., and Surmeier, D.J. (2008). Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity. *Science* 321, 848–851.
- Shiflett, M.W., and Balleine, B.W. (2011). Molecular substrates of action control in cortico-striatal circuits. *Progress in Neurobiology* 95, 1–13.
- Shortall, S.E., Spicer, C.H., Ebling, F.J.P., Green, A.R., Fone, K.C.F., and King, M.V. (2016a). Contribution of serotonin and dopamine to changes in core body temperature and locomotor activity in rats following repeated administration of mephedrone. *Addiction Biology* 21, 1127–1139.
- Shortall, S.E., Green, A.R., Fone, K.C., and King, M.V. (2016b). Caffeine alters the behavioural and body temperature responses to mephedrone without causing long-term neurotoxicity in rats. *Journal of Psychopharmacology* 30, 698–706.
- Singh, S.P., and Singh, V. (2011). Meta-analysis of the efficacy of adjunctive NMDA receptor modulators in chronic schizophrenia. *CNS Drugs* 25, 859–885.
- Smith, G.D., and Ebrahim, S. (2002). Data dredging, bias, or confounding. *BMJ* 325, 1437–1438.

- Smith, K.S., and Graybiel, A.M. (2013). A Dual Operator View of Habitual Behavior Reflecting Cortical and Striatal Dynamics. *Neuron* 79, 361–374.
- Speechley, W.J., Whitman, J.C., and Woodward, T.S. (2010). The contribution of hypersalience to the “jumping to conclusions” bias associated with delusions in schizophrenia. *J Psychiatry Neurosci* 35, 7–17.
- Stalnaker, T.A., Roesch, M.R., Franz, T.M., Burke, K.A., and Schoenbaum, G. (2006). Abnormal associative encoding in orbitofrontal neurons in cocaine-experienced rats during decision-making. *European Journal of Neuroscience* 24, 2643–2653.
- Stalnaker, T.A., Cooch, N.K., and Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nat Neurosci* 18, 620–627.
- Stanic, D., Finkelstein, D.I., Bourke, D.W., Drago, J., and Horne, M.K. (2003). Timecourse of striatal re-innervation following lesions of dopaminergic SNpc neurons of the rat. *European Journal of Neuroscience* 18, 1175–1188.
- Steeds, H., Carhart-Harris, R.L., and Stone, J.M. (2015). Drug models of schizophrenia. *Therapeutic Advances in Psychopharmacology* 5, 43–58.
- Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., and Janak, P.H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* 16, 966–973.
- Steketee, J.D., and Kalivas, P.W. (2011). Drug Wanting: Behavioral Sensitization and Relapse to Drug-Seeking Behavior. *Pharmacol Rev* 63, 348–365.
- Stephan, K.E., and Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology* 25, 85–92.
- Stephan, K.E., Friston, K.J., and Frith, C.D. (2009). Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. *Schizophr Bull* 35, 509–527.
- Strauss, G.P., Frank, M.J., Waltz, J.A., Kasanova, Z., Herbener, E.S., and Gold, J.M. (2011). Deficits in Positive Reinforcement Learning and Uncertainty-Driven Exploration Are Associated with Distinct Aspects of Negative Symptoms in Schizophrenia. *Biological Psychiatry* 69, 424–431.
- Strauss, G.P., Waltz, J.A., and Gold, J.M. (2014). A Review of Reward Processing and Motivational Impairment in Schizophrenia. *Schizophr Bull* 40, S107–S116.
- Suzuki, Y., Jodo, E., Takeuchi, S., Niwa, S., and Kayama, Y. (2002). Acute administration of phencyclidine induces tonic activation of medial prefrontal cortex neurons in freely moving rats. *Neuroscience* 114, 769–779.

- Swets, M., Dekker, J., van Emmerik-van Oortmerssen, K., Smid, G.E., Smit, F., de Haan, L., and Schoevers, R.A. (2014). The obsessive compulsive spectrum in schizophrenia, a meta-analysis and meta-regression exploring prevalence rates. *Schizophrenia Research* 152, 458–468.
- Tai, L.-H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* 15, 1281–1289.
- Tait, D.S., and Brown, V.J. (2007). Difficulty Overcoming Learned Non-reward during Reversal Learning in Rats with Ibotenic Acid Lesions of Orbital Prefrontal Cortex. *Annals of the New York Academy of Sciences* 1121, 407–420.
- Tait, D.S., Phillips, J.M., Blackwell, A.D., and Brown, V.J. (2016). Effects of lesions of the subthalamic nucleus/zona incerta area and dorsomedial striatum on attentional set-shifting in the rat. *Neuroscience* 345, 287–296.
- Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A., and Schoenbaum, G. (2009). The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron* 62, 269–280.
- Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci* 14, 1590–1597.
- Talpos, J.C., Fletcher, A.C., Circelli, C., Tricklebank, M.D., and Dix, S.L. (2012). The pharmacological sensitivity of a touchscreen-based visual discrimination task in the rat using simple and perceptually challenging stimuli. *Psychopharmacology* 221, 437–449.
- Taylor, A.E., Burgess, S., Ware, J.J., Gage, S.H., Richards, J.B., Davey Smith, G., and Munafò, M.R. (2016). Investigating causality in the association between 25(OH)D and schizophrenia. *Scientific Reports* 6, 26496.
- Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron* 66, 781–795.
- Tran, L., Lasher, B.K., Young, K.A., and Keele, N.B. (2013). Depletion of serotonin in the basolateral amygdala elevates glutamate receptors and facilitates fear-potentiated startle. *Transl Psychiatry* 3, e298.
- van Wingerden, M., Vinck, M., Tijms, V., Ferreira, I.R.S., Jonker, A.J., and Pennartz, C.M.A. (2012). NMDA Receptors Control Cue-Outcome Selectivity and Plasticity of Orbitofrontal Firing Patterns during Associative Stimulus-Reward Learning. *Neuron* 76, 813–825.

- Vázquez-Borsetti, P., Celada, P., Cortés, R., and Artigas, F. (2011). Simultaneous projections from prefrontal cortex to dopaminergic and serotonergic nuclei. *International Journal of Neuropsychopharmacology* 14, 289–302.
- Vogel, S.J., Strauss, G.P., and Allen, D.N. (2013). Using negative feedback to guide behavior: Impairments on the first 4 cards of the Wisconsin Card Sorting Test predict negative symptoms of schizophrenia. *Schizophrenia Research* 151, 97–101.
- Voon, V., Pessiglione, M., Brezing, C., Gallea, C., Fernandez, H.H., Dolan, R.J., and Hallett, M. (2010). Mechanisms Underlying Dopamine-Mediated Reward Bias in Compulsive Behaviors. *Neuron* 65, 135–142.
- Voon, V., Derbyshire, K., Rück, C., Irvine, M.A., Worbe, Y., Enander, J., Schreiber, L.R.N., Gillan, C., Fineberg, N.A., Sahakian, B.J., et al. (2015). Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* 20, 345–352.
- Walker, S.C., Mikheenko, Y.P., Argyle, L.D., Robbins, T.W., and Roberts, A.C. (2006). Selective prefrontal serotonin depletion impairs acquisition of a detour-reaching task. *European Journal of Neuroscience* 23, 3119–3123.
- Walker, S.C., Robbins, T.W., and Roberts, A.C. (2009). Differential Contributions of Dopamine and Serotonin to Orbitofrontal Cortex Function in the Marmoset. *Cereb. Cortex* 19, 889–898.
- Wallace, A., Pehrson, A.L., Sánchez, C., and Morilak, D.A. (2014). Vortioxetine restores reversal learning impaired by 5-HT depletion or chronic intermittent cold stress in rats. *The International Journal of Neuropsychopharmacology* 17, 1695–1706.
- Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rudebeck, P.H., and Rushworth, M.F.S. (2010). Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. *Neuron* 65, 927–939.
- Watanabe, K., and Hikosaka, O. (2005). Immediate Changes in Anticipatory Activity of Caudate Neurons Associated With Reversal of Position-Reward Contingency. *Journal of Neurophysiology* 94, 1879–1887.
- Watson, D.J.G., King, M.V., Gyertyán, I., Kiss, B., Adham, N., and Fone, K.C.F. (2016). The dopamine D3-preferring D2/D3 dopamine receptor partial agonist, cariprazine, reverses behavioural changes in a rat neurodevelopmental model for schizophrenia. *European Neuropsychopharmacology* 26, 208–224.
- West, E.A., DesJardin, J.T., Gale, K., and Malkova, L. (2011). Transient Inactivation of Orbitofrontal Cortex Blocks Reinforcer Devaluation in Macaques. *J. Neurosci.* 31, 15128–15135.

- West, E.A., Forcelli, P.A., McCue, D.L., and Malkova, L. (2013). Differential effects of serotonin-specific and excitotoxic lesions of OFC on conditioned reinforcer devaluation and extinction in rats. *Behavioural Brain Research* 246, 10–14.
- Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record* 58, Article 9.
- Whitaker, L.R., Degoulet, M., and Morikawa, H. (2013). Social Deprivation Enhances VTA Synaptic Plasticity and Drug-Induced Contextual Learning. *Neuron* 77, 335–345.
- Wickens, J.R., Budd, C.S., Hyland, B.I., and Arbuthnott, G.W. (2007). Striatal Contributions to Reward and Decision Making. *Annals of the New York Academy of Sciences* 1104, 192–212.
- Wiley, J.L., Bühler, K.G., Lavecchia, K.L., and Johnson, K.M. (2003). Pharmacological challenge reveals long-term effects of perinatal phencyclidine on delayed spatial alternation in rats. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 27, 867–873.
- Winter, S., Dieckmann, M., and Schwabe, K. (2009). Dopamine in the prefrontal cortex regulates rats behavioral flexibility to changing reward value. *Behavioural Brain Research* 198, 206–213.
- Wit, S. de, Watson, P., Harsay, H.A., Cohen, M.X., Vijver, I. van de, and Ridderinkhof, K.R. (2012). Corticostriatal Connectivity Underlies Individual Differences in the Balance between Habitual and Goal-Directed Action Control. *J. Neurosci.* 32, 12066–12075.
- Yanike, M., and Ferrera, V.P. (2014). Interpretive monitoring in the caudate nucleus. *eLife Sciences* 3, e03727.
- Yin, H.H., and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7, 464–476.
- Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience* 22, 505–512.
- Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience* 22, 513–523.
- Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behavioural Brain Research* 166, 189–196.
- Yin, H.H., Ostlund, S.B., and Balleine, B.W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of

cortico-basal ganglia networks. *European Journal of Neuroscience* 28, 1437–1448.

Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* 12, 333–341.

Yorgason, J.T., España, R.A., Konstantopoulos, J.K., Weiner, J.L., and Jones, S.R. (2013). Enduring increases in anxiety-like behavior and rapid nucleus accumbens dopamine signaling in socially isolated rats. *European Journal of Neuroscience* 37, 1022–1031.

Yorgason, J.T., Calipari, E.S., Ferris, M.J., Karkhanis, A.N., Fordahl, S.C., Weiner, J.L., and Jones, S.R. (2016). Social isolation rearing increases dopamine uptake and psychostimulant potency in the striatum. *Neuropharmacology* 101, 471–479.

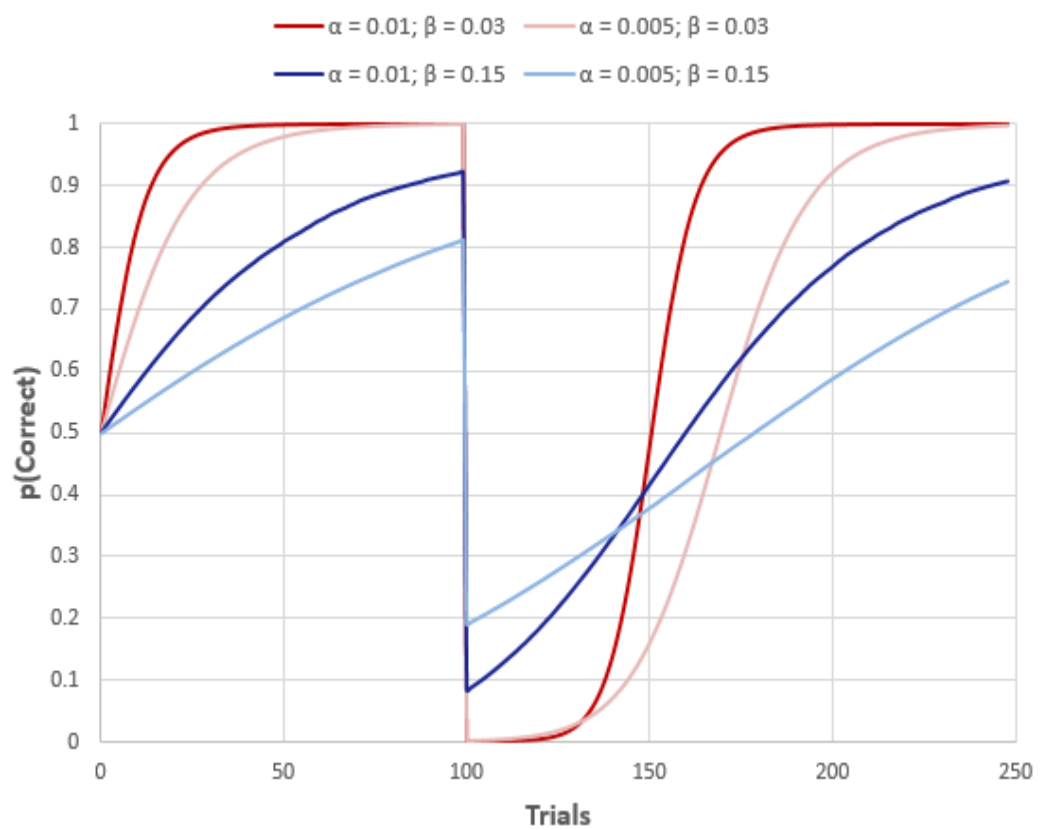
Zhang, L., Doyon, W.M., Clark, J.J., Phillips, P.E.M., and Dani, J.A. (2009). Controls of Tonic and Phasic Dopamine Transmission in the Dorsal and Ventral Striatum. *Mol Pharmacol* 76, 396–404.

Zohar, J., Greenberg, B., and Denys, D. (2012). Chapter 21 - Obsessive-compulsive disorder. In *Handbook of Clinical Neurology*, F.B. and D.F.S. Michael J. Aminoff, ed. (Elsevier), pp. 375–390.

Appendix

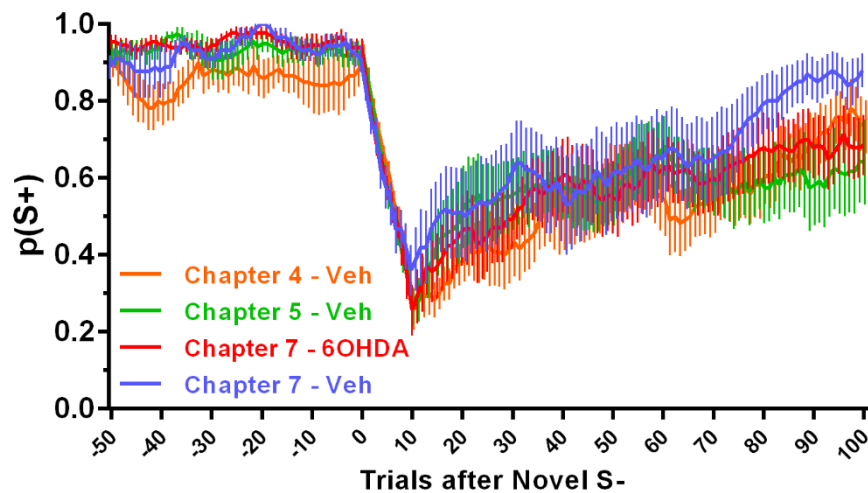
Appendix 1 – Simulated choice behaviour of four agents with different learning parameters in a two-choice reversal learning task.

See Figure 1.1 for more information.

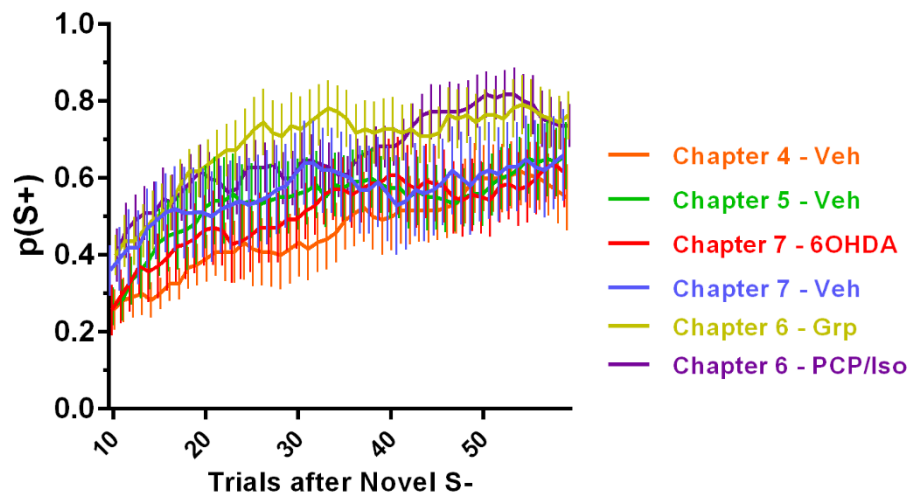


Appendix 2 – Novel stimulus choice across experiments.

The effect of replacing the CS- with a novel S- (novelty bias test) on probability of choosing the S+ (displayed as a moving average of the previous 10 trials). Initially, all animals explore the novel S- at a frequency of 60-70% of trials, but quickly switch back to exploiting the S+.



Below: The 60 trials after novel S- introduction, including data from Chapter 6. Animals in the experiment of Chapter 6 initially showed equivalent exploration of the novel S- but faster devaluation.



Appendix 3 – Chapter 7 results summated.

Responses made in the 9 sessions after reversal (mean $\pm$ SEM, collated) for the cohorts that received either intra-OFC 6-hydroxydopamine or vehicle (see Chapter 7). The overall effect of lesion on the *learned irrelevance* test was not quite significant ($F_{(1, 11)} = 3.5$, $p = 0.089$; 2-way RM-ANOVA), however post-hoc tests showed a significant effect of lesion on number of correct responses (i.e. choosing the former CS-; $p = 0.038$).

