Investigating reinforcement learning processes in depression and substance use disorder: translational, computational and neuroimaging approaches



Katharina Zühlsdorff

Department of Psychology University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

Trinity Hall

July 2022

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 60,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Katharina Zühlsdorff July 2022

Acknowledgements

Firstly, I would like to thank my supervisor, Prof Jeff Dalley, for his support throughout the last years, from helping to write my studentship application, to developing this project, always giving me invaluable advice and being easy to contact. I would like to thank the Institute for Neuroscience and Alan Turing Institute for funding my PhD, as well as Trinity Hall for providing a highly supportive environment. I am grateful for the advice and opportunities I received from my advisors, colleagues and collaborators: Prof Zoe Kourtzi, for the time that I got to spend collaborating with her research group as part of the Alan Turing Enrichment scheme, her helpful advice and always finding the time to meet with me; Prof Trevor Robbins, for discussing ideas and providing extensive feedback on my research; Dr Sharon Morein-Zamir, for collaborating with me on a project that she worked on during her postdoc and whom I have spent a lot of enjoyable time discussing cognitive flexibility with; Dr Jon Kanen for being an excellent friend and collaborator - without Jon, many of the projects presented in this thesis would have not been possible and he was a major driving force for pursuing this line of research; Dr Rudolf Cardinal, for introducing me to Bayesian methods and helping with computational modelling; Dr Tim Rittman, whom I had the chance to work on clinical projects with in the last months and who I have learnt a lot from; Dr Johan Alsiö and Dr Mona El-Sayed Hervig for giving me the opportunity to do some additional computational modelling on rodent data. I would like to also thank Prof Roshan Cools and her group, as well as Prof Antonio Verdejo-Garcia, Dr Juan Verdejo Roman, Dr Carles Soriano-Mas and colleagues for kindly sharing their data with me, which significantly contributed to this thesis. A major thank-you goes to my undergraduate project supervisor Prof Jennifer Bizley for making my time in her lab at UCL so enjoyable that I wanted to continue working in research and for encouraging me to undertake a PhD.

I would like to thank my family and friends: Saeed Kayhanian, who has helped me through the ups-and-downs of going through a PhD and helping me get to the finish line. To my parents Svetlana and Peter, without whom none of this would have been possible; they were always there for me and encouraged me to pursue the things in life that I enjoyed most. My siblings, Laura and Peter-Paul, who bring me a lot of happiness and always light

up my day. My friends Anna-Regina Entus, Tessa Hughes, Paulina Haas-Heger, Nicola Dümler and Emilia Sandoghdar for being my closest friends for many years and supporting me throughout my journey. Last but not least, all the wonderful friends I met in Cambridge - Emma Curran, Will Lloyd-Regan, Jack Smith, Maria Donde, and Emily Evans, who made my time in Cambridge one of the best and who were always there for me, especially in times of difficulty. A special shout-out to Clayton Rabideau, who has been one of my closest friends in Cambridge since day one and has been a tremendous support both as a friend as well as a collaborator. I am grateful for the people I got to meet in the Department of Psychology, including Jolyon Jones, Peter Zhukovsky, Chiara Toschi, Juila Sala-Bayo, Onno Kampman, Marjan Biria and many others who I had the chance to spend lots of hours discussing science with.

Abstract

Reinforcement learning (RL) is the process by which an animal utilises its previous experience to improve outcomes of future choices by maximising reward and minimising punishment. This thesis investigates how RL processes are altered in psychiatric disorders such as major depressive disorder (MDD) and substance use disorder (SUD). The neural basis underlying RL is investigated using brain neuroimaging techniques and translational approaches in both rats and humans. Given the importance of RL and implicated cognitive impairments in psychiatric disorders such as cognitive inflexibility, this PhD thesis sets out to integrate relevant computational and neurobiological substrates, an objective that hitherto has not been widely researched.

Chapter 3 presents the findings of a longitudinal study to investigate the behavioural and neural consequences of early-life maternal separation in rats as a way of simulating early life stress (ELS) in humans. The question addressed was whether early stress is necessary and sufficient for the development of stress-related behaviours relevant to depression. Animals underwent behavioural testing, including probabilistic reversal learning (PRL) to assess behavioural flexibility, and sequential fMRI to evaluate resting-state functional connectivity. Computational analyses revealed differences in reward and punishment learning rates in males arising from maternal separation (MS) and adulthood stress. In contrast, MS female rats showed differences in the 'stickiness' parameter, a latent variable aligned with a loss of flexibility and habit-like behaviour. Finally, MS females and MS males have opposite directional changes in connectivity, as females show lower functional connectivity from the amygdala to the anterior cingulate cortex, infralimbic cortex and insular cortex compared to males.

The subsequent chapter uses a computational approach to investigate latent vulnerability variables in cocaine addiction. A longitudinal dataset acquired in rats was analysed, which involved behavioural phenotyping for several addiction vulnerability traits, including behavioural inflexibility, together with high-resolution MRI brain scans. It was found that future drug-related compulsivity was predicted by higher values of the stickiness parameter, reflecting an increase in perseverative responding commonly found in stimulant-dependent

individuals. Structurally, a positive correlation between the volume of the anterior insular cortex and a parameter relating to how subjects explore versus exploit reward options was found.

The remaining results chapters involve the analysis of three datasets collected from human participants. Chapter 5 includes data from a study involving PRL run concurrently with fMRI scanning. The participants in this study included healthy controls (HCs), as well as individuals with cocaine use disorder (CUD) and gambling disorder (GD). Contrary to previously published findings, no significant differences in alpha, beta or kappa were observed between controls and the CUD group. However, in pathological gamblers, a significant increase in side stickiness was found, showing that gamblers tend to repeat responding in the same spatial location regardless of the outcome on previous trials. Neurally, there is an altered balance in the tracking of reward and punishment expected value (EV) in GD, as well as a shifted balance in processing positive and negative punishment prediction errors (PPE) in CUD. Reward EV tracking in GD involved greater activity in the middle temporal gyrus, cingulate gyrus, precuneus cortex and amygdala, whereas during punishment EV tracking there was lower activity in the postcentral gyrus, superior parietal lobule and precuneus cortex compared to HCs. In response to positive PPEs, the frontal pole, superior frontal gyrus and cingulate gyrus showed lower activity in patients with CUD than controls, but the same group showed greater activity following negative PPEs in the superior and middle frontal gyrus.

Chapter 6 includes behavioural and clinical data from samples of patients with SUD and/or MDD as well as healthy individuals. The main findings of this chapter were that patients with SUD have reduced reinforcement sensitivity and increased stimulus stickiness, as do patients diagnosed with both disorders. No evidence for an association between computationally derived variables and clinical measures (e.g., the Inventory of Depressive Symptomatology – IDS) was found.

The final results chapter presents a novel behavioural task that measures a different subtype of proactive cognitive flexibility, specifically, how healthy participants make decisions in the face of uncertainty and whether they shift their response when they are given the opportunity to repeat their choice following presentation of unreliable feedback. Participants changed their response more frequently following negative than positive feedback. Significant fMRI activations in the frontal pole, anterior cingulate cortex, frontal orbital cortex, and superior frontal gyrus were found when the response was changed rather than repeated. Furthermore, stronger connectivity between the anterior insula and parts of the occipital cortex was found during repeat trials. Finally, it was shown using a multivariate pattern fMRI analysis that behavioural responses on the next trial could be successfully predicted.

The results in this thesis demonstrate the importance of RL in preclinical and clinical psychiatric cohorts. The parameter κ is identified as a key behavioural marker across species. This parameter is altered as a result of ELS in rodents and can help predict rats that show high-compulsive behaviours on cocaine self-administration paradigms. In humans, κ is affected in individuals with GD as well as SUD. Brain regions underlying RL parameters, including κ , in both rodents and humans are identified, particularly highlighting the involvement of the cingulate gyrus in reinforcement learning across species. The results from the reversal learning task studies are then compared with findings from the behavioural and fMRI analyses of a new flexibility task, which extend our knowledge of cognitive flexibility beyond our current understanding of this construct.

Li	st of f	ìgures		xix	
Li	st of t	ables	2	xxiii	
Li	List of abbreviations x			xxv	
1	Intr	oductio	oduction		
	1.1	Introd	uction to Depression	1	
		1.1.1	Affective and cognitive changes in MDD	3	
		1.1.2	Neural changes in MDD	5	
		1.1.3	Neurochemical basis of MDD	7	
		1.1.4	Genetic & environmental contributors of MDD	8	
		1.1.5	Sex differences in MDD	10	
		1.1.6	Treatments for MDD	11	
	1.2	Introd	uction to Substance Use Disorders	13	
		1.2.1	Behavioural differences in SUD	13	
		1.2.2	Neural changes in SUD	16	
		1.2.3	Neurochemical basis of SUD	18	
		1.2.4	Genetic contributors to SUD	19	
		1.2.5	Treatments for SUD	22	
	1.3	What i	is cognitive flexibility?	23	
	1.4	Revers	sal learning: cognitive aspects, neural circuits and neurochemistry	25	
	1.5	Endop	henotypes in psychiatry	27	
		1.5.1	Preclinical models of depression	28	
		1.5.2	Preclinical models of SUD	30	
	1.6	Reinfo	preement Learning in Neuroscience and Psychiatry	31	
		1.6.1	Neural basis of RL	33	

		1.6.2	RL in MDD	35
		1.6.3	RL in SUD	35
	1.7	Thesis	hypotheses and objectives	36
2	Gen	eral me	ethods	41
	2.1	Rats		41
		2.1.1	Reinforcement learning modelling	41
		2.1.2	Bayesian Learner modelling	43
		2.1.3	Model comparison	45
		2.1.4	Drift Diffusion Modelling	46
		2.1.5	Conventional analysis of reversal learning task data	47
		2.1.6	Rodent image pre-processing	48
		2.1.7	Structural analysis	49
	2.2	Huma	ns	49
		2.2.1	Reinforcement learning modelling using a hierarchical Bayesian	
			approach	49
		2.2.2	Simulation of data using model parameters	51
		2.2.3	Reinforcement learning-drift diffusion model	51
		2.2.4	Image pre-processing	52
3	Reir	nforcem	ent learning in rodents exposed to maternal separation and adult-	
	hood	d stress		55
	3.1	Introd	uction	55
	3.2	Metho	ods	57
		3.2.1	Subjects	57
		3.2.2	Repeated Early Maternal Separation	58
		3.2.3	Adulthood stress	59
		3.2.4	Reversal learning task	59
		3.2.5	Statistical analysis of behavioural data	60
		3.2.6	Image acquisition	61
		3.2.7	Seed-based analysis	61
		3.2.8	Region-of-interest analysis	63
	3.3	Result	8	63
	3.4	Behav	ioural findings	63
		3.4.1	Probabilistic reversal learning and anhedonia	63
		3.4.2	Reinforcement Learning model comparison	64

		3.4.3	Reinforcement Learning results
		3.4.4	Bayesian Learner model comparison
		3.4.5	Bayesian Learning results
		3.4.6	Relationship between conventional PRL and RL parameters 70
		3.4.7	Drift Diffusion Model
	3.5	Imagin	g results
		3.5.1	Volumetric Analysis
		3.5.2	Seed-based analysis
		3.5.3	Region-of-interest analysis
	3.6	Discus	sion
4	Reir	nforcem	ent learning in rodents exposed to cocaine 87
	4.1	Introdu	uction
	4.2	Metho	ds
		4.2.1	Subjects
		4.2.2	Drugs
		4.2.3	Reversal learning task
		4.2.4	Intra-jugular surgery
		4.2.5	Drug self-administration
		4.2.6	Statistical analysis of behavioural parameters
		4.2.7	Image acquisition
	4.3	Results	93
	4.4	Behavi	oural findings
		4.4.1	Deterministic Reversal Learning
		4.4.2	RL model comparison
		4.4.3	Reinforcement learning
		4.4.4	Bayesian Learner model comparison
		4.4.5	Bayesian Learning results
		4.4.6	Simulations
		4.4.7	Behavioural correlations
	4.5	Imagin	g results
		4.5.1	Structural MRI
		4.5.2	Correlations between RL parameters and structural volumes 104
	4.6	Discus	sion

5	Reinforcement learning in patients with gambling disorder and substance use					
	diso	rder		111		
	5.1	Introd	uction	111		
	5.2	Metho	ods	114		
		5.2.1	Participants	114		
		5.2.2	Probabilistic Reversal Learning Task	115		
		5.2.3	Correlation analysis between conventional and modelling PRL mea-			
			sures	115		
		5.2.4	Imaging data acquisition	117		
		5.2.5	First-level models	117		
		5.2.6	Higher-level models	117		
	5.3	Result	ts	118		
	5.4	Behav	vioural findings	118		
		5.4.1	Demographic information	118		
		5.4.2	Conventional PRL measures	118		
		5.4.3	Choosing the winning model	119		
		5.4.4	RL analysis	119		
		5.4.5	Relating conventional parameters and depression score ratings to RL			
			parameters	121		
		5.4.6	Simulations	123		
		5.4.7	DDM analysis	123		
		5.4.8	RL-DDM	125		
	5.5	Imagin	ng results	131		
		5.5.1	Response to feedback/cue presentation	131		
		5.5.2	Brain activity during reward and punishment expected value tracking			
			in Gambling Disorder	131		
		5.5.3	Neural signal to positive and negative punishment prediction errors			
			is affected in Cocaine Use Disorder	134		
		5.5.4	Whole-brain connectivity analyses	137		
	5.6	Discus	ssion	138		
6	Reir	nforcem	ent learning in patients with Major Depressive Disorder	143		
	6.1	Introd	uction	143		
	6.2	Metho	ods	145		
		6.2.1	Participants	145		

		6.2.2	Probabilistic reversal learning task	146
		6.2.3	Correlation analysis between conventional and modelling PRL mea-	
			sures	147
	6.3	Results	3	148
		6.3.1	Demographic information	148
		6.3.2	Conventional PRL measures	148
		6.3.3	Choosing the winning RL model	149
		6.3.4	Group differences	150
		6.3.5	Relating conventional parameters and depression score ratings to RL	
			parameters	150
		6.3.6	Simulations	152
	6.4	Discus	sion	154
7	A no	vel pers	spective on cognitive flexibility	159
	7.1	Introdu	uction	159
	7.2	Metho	ds	161
		7.2.1	Participants	161
		7.2.2	Task procedure	162
		7.2.3	Neuroimaging acquisition	163
		7.2.4	Task performance	164
		7.2.5	First-level models	164
		7.2.6	Higher-level models	164
		7.2.7	Brain behaviour correlations	165
		7.2.8	Psychophysiological interaction analysis	165
		7.2.9	Multivariate pattern analysis	166
	7.3	Results	3	166
		7.3.1	Task performance	166
		7.3.2	Task-related brain activation - Changing versus repeating a response	168
		7.3.3	Activations associated with R1 and feedback	171
		7.3.4	Dorsal and ventral striatum are associated with learning to disregard	
			feedback	171
		7.3.5	Neural circuits implicated in the repeat and change conditions	173
		7.3.6	Multivariate Pattern Analysis: changing one's mind	176
	7.4	Discus	sion	178

8	Disc	ussion		183
	8.1	Reinfo	rcement Learning in Depression	183
		8.1.1	RL in rodents exposed to REMS and adulthood stress	183
		8.1.2	RL in patients with MDD and/or SUD	185
		8.1.3	Building bridges between human and rodent research studies	185
		8.1.4	Implications for the use of the maternal separation procedure to study	
			depression	187
		8.1.5	Implications for investigating sex-dependent differences in depression	n188
		8.1.6	Implications for the study of reinforcement learning and decision-	
			making in stress and depression	189
	8.2	Reinfo	preement Learning in Substance Use Disorder	190
		8.2.1	RL in rodents exposed to a cocaine self-administration paradigm	190
		8.2.2	RL in patients with Cocaine Use Disorder and Gambling Disorder .	191
		8.2.3	Building bridges between rodent and human studies	192
		8.2.4	Relating conventional parameters to RL parameters	194
		8.2.5	Implications for the study of reinforcement learning and decision-	
			making in disorders of compulsivity	196
	8.3	'Chang	ge Your Mind' - a novel task to investigate cognitive flexibility	197
		8.3.1	Implications for the study of cognitive flexibility	199
	8.4	Future	research directions	200
	8.5	Conclu	usions	201
Ар	pend	ix A A	ppendix for rodent chapters (chapters 3-4)	203
-	A.1	Reinfo	rcement Learning in rodents exposed to maternal separation and adult-	
		hood s	tress	203
		A.1.1	Additional conventional PRL measures	203
		A.1.2	Reaction time data	207
		A.1.3	Imaging quality checks	208
		A.1.4	Additional model comparison figures	209
		A.1.5	Separate t-tests - seed-based analysis	210
	A.2	RL in	rodents exposed to a cocaine self-administration paradigm	212
		A.2.1	Additional RL model comparison measures	212
		A.2.2	Structural analyses	213
Ар	pend	ix B A	dditional figures from human chapters (chapters 5-7)	215
-	B .1	fMRI o	quality checks	215

Bibliography			
B.3	RL in patients with MDD	220	
B.2	RL in patients with GD and SUD	217	

List of figures

1.1	Neurocircuitry underlying substance use disorder	18
1.2	Effects of different neurotransmitters on tasks commonly used to assess	
	impulsivity or compulsivity.	20
2.1	The Drift Diffusion Model.	47
3.1	REMS study timeline	58
3.2	Representation of the reversal learning task	60
3.3	Masks of the regions used for the seed-based analysis	62
3.4	Key findings of a conventional analysis of the PRL task	65
3.5	Q-learning model comparisons	66
3.6	Results of the Q-learning analysis.	68
3.7	Bayesian Learning model comparison	69
3.8	Results of the Bayesian Learner analysis.	71
3.9	Correlation matrix representing the relationship between conventional PRL	
	measures and parameters from the RL model across all sessions of the PRL	
	task for all animals	72
3.10	Results from the Drift Diffusion Model.	74
3.11	BLA seed-based analysis results	76
3.12	mOFC seed-based analysis results	77
3.13	Medial OFC and RL measures correlation results	78
3.14	Main results of the region-of-interest connectivity analysis	80
4.1	Study timeline.	92
4.2	Key findings of a conventional analysis of the deterministic reversal learning	
	task	95
4.3	RL model comparison.	96
4.4	Box-plots of the Q-learning parameters.	98

List of figures

4.5	BL model comparison.	99
4.6	Results of the Bayesian Learner analysis	101
4.7	Correlation matrix representing the relationship between conventional DRL	
	measures and parameters from the BL and RL models for all animals 1	103
4.8	Masks of the regions used for the volumetric analysis	104
4.9	Volumetric differences of the nucleus accumbens core volume between low-,	
	medium- and high-compulsive animals	105
4.10	Correlation matrix representing the relationship between RL parameters and	
	ROI volumes for all animals	106
5.1	Diagram of the probabilistic reversal learning task	116
5.2	Key findings of a conventional analysis of the PRL task	121
5.4	Results of the hierarchical Bayesian Q-learning model	121
5.3	Posterior distributions of parameters from the winning 5-parameter QL model.	122
5.5	Correlation matrix representing the relationship between conventional PRL	
	measures and RL parameters	124
5.6	Density plot of the latency distributions on the PRL task	126
5.7	Latencies across runs	127
5.8	Results from the drift diffusion model analysis	128
5.9	Testing for collinearity between the RL-DDM parameters	129
5.10	Results from the RL-DDM analysis	130
5.11	GD vs control contrast for reward EV	132
5.12	Control vs GD contrast for punishment EV	134
5.13	The control vs CUD contrast for positive PPE	136
5.14	The CUD vs control contrast for negative PPE	137
5.15	The GD vs control contrast - positive correlation between α_{rew} and response	
	to reward EV	138
6.1	Diagram of the probabilistic reversal learning task.	147
6.2	Key findings of a conventional analysis of the PRL task	149
6.3	Results of the hierarchical Bayesian Q-learning model	151
6.4	Correlation matrix representing the relationship between conventional PRL	
	measures, psychiatric questionnaire scores and RL parameters	153
7.1	Diagrammatic representation of the task structure	163
7.2	Effect of feedback and run on R2 change	167

List of figures

7.3	The R2 change versus repeat contrast.	169
7.4	Change in signal in response to positive and negative feedback across the	
	three runs.	172
7.5	Results of the AI psychophysiological interaction analysis	173
7.6	Results of the OFC psychophysiological interaction analysis	174
7.7	Results of the SFG psychophysiological interaction analysis	175
7.8	Results of the multivariate pattern analysis.	177
A.1	Additional findings of a conventional analysis of the PRL task	204
A.2	Additional findings of a conventional analysis of the PRL task	205
A.3	Figure reflecting the ability of the model to capture animal behaviour. The	
	yellow line represents the animal's true behaviour. When this value is 1, it	
	chooses the left stimulus, and when it is 2, it selects the right stimulus. The	
	blue line demonstrates the probability of choosing the left stimulus and the	
	orange line reflects the probability of choosing the right stimulus	206
A.4	Rection times on the PRL task.	207
A.5	Example of quality checks conducted on rodent imaging scans	208
A.6	Additional model comparison figures.	209
A.7	BLA seed-based analysis results - females	210
A.8	BLA seed-based analysis results - males	211
A.9	Additional model comparison measures for RL models	212
A.10	Volume differences across six brain regions	213
B .1	Brain mask and brain tissue segmentation of T1w image	215
B.2	Alignment of functional and anatomical MRI scans	216
B.3	BOLD summary.	216
B. 4	Correlations among nuisance regressors	217
B.5	The GD vs control contrast for feedback/cue presentation.	217
B.6	The SUD vs control contrast for feedback/cue presentation	218
B.7	The SUD vs control contrast - positive correlation between α_{rew} and response	
	to positive PPE	218
B.8	The GD vs control contrast - positive correlation between κ_{stim} and response	
	to cue/feedback presentation	219
B.9	The SUD vs control contrast - positive correlation between κ_{stim} and response	
	to cue/feedback presentation.	219
B.10	Simulations based on the RL parameters from the different groups	220

List of tables

2.1	Priors for model parameters	51
4.1	Summary table of statistics for the correlation analysis between conventional	
	PRL and modelling parameters.	102
5.1	Demographic information and results.	118
5.2	Model comparison summary.	119
5.3	Reward expected value tracking GD vs controls contrast summary	132
5.4	Punishment expected value tracking controls vs GD contrast summary	133
5.5	Positive PPE controls vs CUD contrast summary	135
5.6	Negative PPE CUD vs controls contrast summary	136
6.1	Demographic information	148
6.2	Model comparison summary.	150
6.3	Summary table of statistics from the correlation analysis between conven-	
	tional PRL and modelling parameters.	152
7.1	Change-repeat contrast summary	170
7.2	AI PPI repeat-change contrast summary	174
7.3	OFC PPI repeat-change contrast summary	175
7.4	SFG PPI change-repeat contrast summary	176
7.5	MVPA analysis summary of performance metrics	177
8.1	Summary of correlations between PRL and RL measures across all chapters.	195

List of abbreviations

ADHD - Attention-Deficit/Hyperactivity Disorder AI - Anterior Insula AIC - Akaike Information Criterion AMPA - α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid ANCOVA - Analysis of Covariance ANTs - Advanced Normalisation Software AQ - Autism Spectrum Quotient ASD - Autism Spectrum Disorder ASI - Addiction Severity Index ATD - Acute Tryptophan Depletion AUD - Alcohol Use Disorder AUROC - Area Under the Receiver Operating Characteristic Curve **BDNF** - Brain-Derived Neurotrophic Factor **BIC - Bayesian Information Criterion** BL - Bayesian Learner BLA - Basolateral Amygdala BOLD - Blood-oxygen-level-dependent CBGTC - Cortico-basal ganglia-thalamo-cortical CBT - Cognitive Behavioural Therapy **CRH** - Corticotrophin Releasing Hormone **CEN - Central Executive Network** CPu - Caudate Putamen CUD - Cocaine Use Disorder

CYM - Change Your Mind

DA - dopamine

DCCS - Dimensional Change Card Sort task

DDM - Drift Diffusion Model

- DIVA Diagnostic Interview for Adult ADHD
- dmPFC Dorsomedial PFC
- dlPFC Dorsolateral PFC
- DLS Dorsolateral Striatum
- DMS Dorsomedial Striatum
- DRL Deterministic Reversal Learning
- DS Dorsal Striatum
- DSM Diagnostic and Statistical Manual of Mental Disorders
- Dvars Spatial Standard Deviation of Successive Difference Images
- ECN Executive Control Network
- EEG Electroencephalogram
- ELS Early Life Stress
- EPI Echo Planar Imaging
- EV Expected Value
- EWA Experience-Weighted Attractor
- FI Fixed Interval
- FOV Field of View
- FR Fixed Ratio
- GWAS Genome-Wide Association Study
- HC Healthy Controls
- HDI Highest Density Interval
- IDS Inventory of Depressive Symptomatology
- IED Intra-/Extra-Dimensional Set Shift task
- IFG Inferior Frontal Gyrus
- IGT Iowa Gambling Task
- IL Infralimbic Cortex
- IPDE International Personality Disorder Examination
- IOFC Lateral Orbitofrontal Cortex
- IPFC Lateral Prefrontal Cortex
- MAOI Monoamine Oxidase Inhibitor
- MATE-crimi Measurements in the Addictions for Triage and Evaluation and criminality
- MDD Major Depressive Disorder
- MFG Middle Frontal Gyrus
- MNI Montreal Neurological Institute
- mOFC Medial Orbitofrontal Gyrus

- MS Maternal Separation
- MT Magnetisation Transfer
- MUD Methamphetamine Use Disorder
- MVP Multivariate Pattern Analysis
- NA Noradrenaline
- NAc Nucleus Accumbens
- NAcC Nucleus Accumbens Core
- NAcS Nucleus Accumbens Shell
- NART National Adult Reading Test
- NIDA Dutch Diagnostic Interview for Adult Autism Spectrum Disorders
- NMDA N-methyl-D-aspartic acid
- OCD Obsessive-Compulsive Disorder
- OFC Orbitofrontal Cortex
- PCC Pearson's Correlation Coefficient
- PCCx Posterior Cingulate Cortex
- PD Proton Density
- PG Pathological Gamblers
- PND Post-Natal Day
- PPC Posterior Parietal Cortex
- PPE Punishment Prediction Error
- PPI Psychophysiological Interaction Analysis
- PrL Prelimbic Cortex
- PRL Probabilistic Reversal Learning
- PTQ Perseverative Thinking Questionnaire
- pSMA Pre-Supplementary Motor Area
- **REMS Repeated Early Maternal Separation**
- RL Reinforcement Learning
- RL-DDM Reinforcement Learning-Drift Diffusion Model
- ROI Region-of-interest
- **RPE Reward Prediction Error**
- rs-FC Resting-state Functional Connectivity
- rs-fMRI Resting-state fMRI
- **RT** Reaction Time
- SA Self-Administration
- SCA Seed-based Correlation Analysis

SCID-I-CV - Structured Clinical Interview for DSM-IV Axis I Disorders

- SDI Stimulant-Dependent Individual
- SPT Sucrose Preference Test

SFG - Superior Frontal Gyrus

SN - Salience Network

SSRI - Selective Serotonin Reuptake Inhibitor

SSRT - Stop Signal Reaction Time

SSRTT - Stop Signal Reaction Time Task

SUD - Substance Use Disorder

- SVM Support Vector Machine
- TCA Tricyclic Antidepressants
- TE Time to Echo
- TR Time to Repetition
- T1w T1-weighted
- vmPFC Ventromedial PFC
- vlPFC Ventrolateral PFC
- VS Ventral Striatum
- VBRT Voucher-Based Reinforcement Learning
- WCST Wisconsin Card Sorting Test
- WFPT Wiener First-Passage Time
- 5-HT Serotonin
- 5-HTTLPR Serotonin Transporter-Linked Polymorphic Region

Chapter 1

Introduction

The high prevalence of neuropsychiatric disorders such as depression and drug addiction places a major burden on the provision of healthcare around the world (Trautmann et al., 2016). Worldwide, 264 million people suffer from depression and 35 million people are affected by drug addiction (UNODC, 2020; WHO, 2020). It is estimated that in middleand high-income countries approximately 50% of the population will be affected by at least one psychiatric disorder at some point in their life (Trautmann et al., 2016). The brain mechanisms underlying many of these disorders are not yet fully known, which is reflected in the fact that approximately one-half of patients do not respond to the first-line antidepressant medication prescribed to them (Garcia-Toro et al., 2012). Common treatments for substance use disorder (SUD) include behavioural therapies such as cognitive behavioural therapy (CBT) as well as medication-assisted therapies during detoxification, depending on the substance abused (Nelson et al., 2017). However, reports indicate that over 50% of individuals relapse within a year of treatment (Hasin et al., 2013). These statistics clearly highlight the unmet clinical need of treatments in psychiatric conditions such as for Major Depressive Disorder (MDD) and SUD and demonstrate the urgency of understanding the origin of these disorders to inform the development of new treatments. This thesis aims to contribute to the understanding of the behavioural and neural underpinnings of MDD and SUD in particular.

1.1 Introduction to Depression

Depressive disorders are debilitating conditions that can be classified into different subtypes, which according to the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) include: disruptive mood dysregulation disorder, major depressive disorder,

Introduction

persistent depressive disorder (also known as dysthymia), premenstrual dysphoric disorder, substance/medication-induced depressive disorder, and unspecified depressive disorder (APA, 2013). A key feature observed across all depressive disorders is low mood alongside cognitive difficulties. What differs between them is the onset, duration, and presumed aetiology. The most prevalent of depressive disorders is major depressive disorder, herein referred to as MDD or depression. In the United States alone, 7% of the population have experienced a twelve-month prevalence of MDD, and 8.4% have experienced at least one major depressive episode (APA, 2013; NIMH, 2022). Dysthymia, on the other hand, has a prevalence of 0.5% and premenstrual dysphoric disorder of 1.8%. MDD is identified by at least one episode that lasts a minimum of two weeks during which the individual's affect and cognition are significantly altered. This thesis will mostly focus on MDD; therefore, this introduction will serve to summarise existing knowledge on MDD specifically. Box 1 summarises the most up-to-date diagnostic criteria for MDD as specified by the DSM-5.

The prevalence of depressive disorder is higher in females than males across all subtypes, being nearly twice as high in women (Bogren et al., 2018). The incidence rate of MDD per 1000 people has been found to be as high as 2.96 in females, but only 1.56 in males. The rates for dysthymic disorder are 0.07 and 0.02, respectively. Bipolar disorder, however, has been found to be the same in both sexes, with females having an incidence rate of 0.04 and males of 0.05. Individuals that have experienced childhood trauma also have a more than twofold increased risk of MDD. Moreover, individuals affected by MDD that experienced childhood trauma demonstrate greater symptom severity and worse treatment outcomes than patients with MDD that have not experienced childhood trauma (Negele et al., 2015). MDD is the second highest contributor to chronic disease burden and is associated with an increased risk of developing other comorbidities such as diabetes mellitus, heart disease and stroke (Otte et al., 2016; Vos et al., 2015; Whooley and Wong, 2013). MDD also significantly increases the risk of suicide, which is reflected in the fact that 50% of worldwide suicides per year are committed during a depressive episode (WHO, 2016). Overall, MDD increases all-cause mortality risk by 60-70% (Walker et al., 2015). It is evident that a better understanding of MDD is imperative, particularly in order to guide the development of future treatments.

Box 1: DSM-5 criteria for Major Depressive Disorder

A) An individual will present with five (or more) of the following symptoms, which should be present during the same two-week period and should represent a change from previous functioning. Depressed mood or loss of interest or pleasures must be present as a symptom. Symptoms should be present most of the day, nearly every day.

- Depressed mood
- · Diminished interest or pleasure in all, or almost all, activities
- · Significant weight loss or weight gain, or decrease/increase in appetite
- Insomnia or hypersomnia
- Psychomotor agitation or retardation
- Fatigue or loss of energy
- Feelings of worthlessness of excessive or inappropriate guilt
- Diminished ability to think or concentrate, or indecisiveness
- Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

B) These symptoms cause clinical distress or impairment in social, occupational, or other important areas of functioning.

C) The episode is not attributable to the physiological effects of a substance or to another medical condition.

D) The occurrence of the major depressive episode if not better explained by schizoaffective disorder, schizophrenia, schizophreniform disorder, delusional disorder, or other specified and unspecified schizophrenia spectrum and other psychotic disorders.

E) There has never been a manic or hypomanic episode.

1.1.1 Affective and cognitive changes in MDD

The onset of a first major depressive episode often follows a negative life event (Monroe and Reid, 2008). Although most individuals initially respond to such an event with negative affect, only 20% of people will subsequently develop a depressive episode (Monroe and Harkness, 2005). Psychological theories propose that an individuals' inferences, thoughts,

Introduction

memory and recall of events are vulnerability factors for depression (Gotlib and Joormann, 2010; Mathews and MacLeod, 2005). Studies investigating emotion regulation have reported that depressed individuals use different strategies compared to non-depressed individuals when processing external emotional stimuli, as they tend to ruminate and suppress thoughts, whereas non-depressed individuals use strategies such as mood-incongruent recall (Joormann et al., 2007). Affective bias is defined as the tendency to focus on the processing of negative, rather than positive life events (Pulcu and Browning, 2017). Cognitive affective bias is a key component of MDD and has been observed in both rodent models of depression and humans with MDD (Murrough et al., 2011; Peckham et al., 2010; Robinson and Roiser, 2016; Stuart et al., 2013). Early effects of antidepressant treatments on affective bias have been reported (Harmer et al., 2009). Negative bias has been observed in depressed groups on perceptive tasks, memory, attention and working memory (Erickson et al., 2005; Joormann and Gotlib, 2008; Matt et al., 1992; Murphy et al., 1999). Overall, it is evident that emotional and cognitive processing are impaired in depressive disorders.

Cognitive dysfunction refers to deficits in attention, verbal and nonverbal learning, shortterm and working memory, visual and auditory processing, problem solving, processing speed, or motor functioning (McIntyre et al., 2013). In MDD, studies have reported moderate impairments in processing speed and learning in patients (Hasselbalch et al., 2012). Metaanalyses have found that executive function, memory, and attention are the most strongly affected cognitive domains and have large effect sizes (Rock et al., 2014). Cognitive theories have proposed that dysfunctional cognitive and affective processing are linked with each other (Joormann and Siemer, 2011). They state that processes such as attention, memory, and interpretation influence whether and which emotion is experienced. Employing effective emotion-regulation strategies may thus be impaired by cognitive deficits, subsequently leading to sustained negative affect resulting from negative life experiences. This may lead to depressed individuals being more likely to interpret ambiguous stimuli in a negative way. Moreover, people with MDD may struggle to disengage and inhibit attending and processing of negative information, as this requires cognitive functions such as working memory, which are impaired (Rose and Ebmeier, 2006). Inhibition, a key component of cognitive control, has been reported to be reduced on the negative affective priming task in depressed individuals (Joormann and Gotlib, 2010). Furthermore, participants with MDD exhibit difficulties removing irrelevant, negative information from their working memory (Joormann and Gotlib, 2008). These deficits in cognitive control may lead to an inability to disengage attention from negative and irrelevant stimuli, which can subsequently result in increased unwanted thoughts and difficulties forgetting that material (Joormann and Siemer,

2011). Depressed individuals also take longer to forget negative information, reflecting memory biases in depression (Power et al., 2000). The inability to inhibit the processing of negative information has been linked to increased rumination (Joormann and Gotlib, 2010). Thus, it appears that affective and cognitive processed are closely intertwined.

Decision-making, an important cognitive process, has also been found to be impaired in depression (Paulus and Yu, 2012; Trivedi and Greer, 2014). The Iowa Gambling Task (IGT), developed in the 1990s by Bechara and colleagues, reflects real-life decision-making processes by taking into account reward and punishment contingencies as well as uncertainty of outcomes (Must et al., 2013). Patients affected by MDD generally show altered sensitivity to reward and punishment on this task (Cella et al., 2010; Han et al., 2012). Specifically, MDD patients select the advantageous deck less frequently and tend not to shift their strategy following contingency changes. On a different decision-making task that involves betting on the location in which a token was hidden, depressed patients are also impaired and employ suboptimal betting strategies (Murphy et al., 2001). Additionally, individuals with MDD exhibit cognitive inflexibility, for example, on the Go/No-go task and the Wisconsin Card Sorting Test (WCST) (Murphy et al., 2012b; Purcell et al., 1997). These studies suggest that MDD has extensive effects on multiple aspects of brain function, including affective processing, cognitive control, flexibility, and inhibition.

1.1.2 Neural changes in MDD

Not only does MDD result in significant behavioural changes, but it also has widespread effects on brain structure, function and connectivity. Structurally, volumetric differences have been observed within the corticolimbic circuit, including the amygdala, anterior cingulate cortex (ACC), paracingulate cortex and dorsomedial PFC (dmPFC) (Sacher et al., 2012). This circuit is involved in regulating sensitivity to threat (Fisher and Hariri, 2013). Specifically, decreases in grey matter volume in the amygdala, paracingulate cortex and dmPFC have been observed during a depressive episode (Sacher et al., 2012). In the same analysis, increased glucose metabolism in the subgenual and pregenual ACC were highlighted. However, it must be noted that in this meta-analysis medication naïve, medication free and medicated patients were included. Decreased amygdala volume has been confirmed in post-mortem brains of patients with MDD (Altshuler et al., 2010; Bowley et al., 2002). However, conflicting evidence has suggested that in MDD, amygdala volumes are increased (van Eijndhoven et al., 2009; Weniger et al., 2006). Additional studies implied that this is dependent on whether patients are going through their first episode or have recurrent MDD, in the former there being

Introduction

increased volumes, and in the latter decreased (Frodl et al., 2002; Lorenzetti et al., 2009; Sheline et al., 1998). In a meta-analysis solely including first episode, medication-naïve MDD patients, grey matter volumes were increased in the thalamus, superior frontal gyrus (SFG), insula and middle frontal gyrus (MFG), but no amygdala changes were seen (Peng et al., 2016). The hippocampus has also been implicated in MDD, with significant reductions in hippocampal CA1-CA3 and total volumes having been reported (Huang et al., 2013; Malykhin et al., 2010).

Resting-state functional connectivity is also known to be altered in MDD. The default mode network (DMN), central executive network (CEN) and salience network (SN) are some of the most important neural circuits in the brain, and are responsible for brain function during rest, cognition, and emotion, respectively (Mulders et al., 2015). As discussed in the previous section, all of these functions have been found to be impaired in depression. Unsurprisingly, these networks are differentially recruited in patients. The DMN consists of the mPFC, posterior cingulate cortex (PCCx) and precuneus cortex (Andrews-Hanna et al., 2010). Its functions involve self-referential processing and emotion-regulation, as well as consciousness and memory processing and it has connectivity with the amygdala and hippocampus (Andrews-Hanna et al., 2014). Studies employing independent component analyses have found increased connectivity within the anterior DMN in medicated and unmedicated MDD patients compared to healthy controls and reported increased synchronisation between the mPFC and anterior DMN (Greicius et al., 2007; Li et al., 2013; Mulders et al., 2015). A more recent study, on the other hand, including 848 depressed and 794 healthy control participants, within-network connectivity of the DMN was reduced only in individuals with recurrent MDD, but not first-episode medication naïve patients (Yan et al., 2019). This reduction was further associated with medication usage, rather than illness duration. The CEN includes the lateral PFC (IPFC), PPCx, frontal eye fields and dmPFC (Rogers et al., 2004). It is active during cognitive tasks and is implicated in cognitive functions such as attention and working memory. Most studies report decreased connectivity of the dorsolateral PFC (dIFC) with areas of the CEN in depressed individuals, regardless of them being medicated or not (Alexopoulos et al., 2012; Ye et al., 2012). The SN includes the frontoinsular cortex, dorsal ACC, amygdala and temporal poles and is activated as a result of salient stimuli (Hermans et al., 2014; Seeley et al., 2007). Studies indicate increased connectivity of the SN with the ACC and increased connectivity between the insula and pregenual ACC as well as mOFC in MDD (Avery et al., 2014; Horn et al., 2010; Manoliu et al., 2014). The amygdala, which is part of the SN, is generally suggested to have decreased connectivity with multiple brain regions in patients (Ramasubbu et al., 2014; Tang et al., 2013). Overall, there are extensive
resting-state functional connectivity changes in MDD. However, the exact differences are dependent on a number of factors, including medication status and number of MDD episodes previously experienced. Future studies should aim to identify how these factors differentially influence connectivity.

Areas that are not part of the three mentioned circuits have also been implicated in in depression. Reduced functional connectivity between the ventral striatum (VS) and ventromedial PFC (vmPFC) and subgenual ACC has been observed in depressed females, both previously medicated and unmedicated (Furman et al., 2011). Connectivity between the dorsal caudate and dorsal PFC, on the other hand, was strengthened. Reduced striatal activity during reward selection, reward anticipation and reward feedback in MDD point to reduced activity of reward systems in MDD (Smoski et al., 2009). The mOFC, however, is hyperresponsive during reward selection. During visual emotive information processing, there is increased ACC activity to emotive stimuli and increased amygdala and hippocampal activity in response to negative stimuli (Jaworska et al., 2014). Antidepressant treatments are able to normalise many of these activation patterns. Moreover, MDD patients show impairments in memory for faces, which has been linked to greater amygdala activation on a face recognition task (Roberson-Nay et al., 2006). In summary, not only resting-state functional connectivity patterns are greatly altered in depression, but also brain responsiveness to various cognitive tasks is affected, providing potential explanations for the behavioural observations.

1.1.3 Neurochemical basis of MDD

The first neurotransmitters found to be implicated in MDD were the monoamine neurotransmitters serotonin (5-HT), noradrenaline (NA) and dopamine (DA) (Otte et al., 2016). This became evident after patients using certain antihypertensive drugs, which reduce monoamine levels, developed depression (Schildkraut, 1995). Iproniazid and imiprimine, which were developed for non-psychiatric conditions, were additionally found to have antidepressant effects in humans by increasing transmission of monoamines (Krishnan and Nestler, 2008). Many antidepressant treatments used nowadays target reuptake proteins or inhibit degradation (selective serotonin reuptake inhibitors (SSRIs) and monoamine oxidase inhibitors (MAOIs)), thus affecting monoamine transmission (Berton and Nestler, 2006). Nonetheless, decreased monoamine neurotransmission cannot be the sole cause of MDD, as MAOIs and SSRIs immediately increase monoamine transmission, but their effects on mood only become evident after few weeks (Krishnan and Nestler, 2008). Moreover, it

is still unknown which 5-HT receptors are involved in the action of SSRIs, but promising results have been found in rodents demonstrating that selective agonists of the 5-HT₄ receptor have strong antidepressant effects (Lucas et al., 2007). Acute tryptophan depletion (ATD) in healthy volunteers does not result in any significant mood changes (Ruhé et al., 2007). In patients recovered from MDD, ATD can re-introduce depressive symptoms in 30% of people. Therefore, serotonin reduction can induce depressive symptoms in individuals with a predisposition to depression, but not in healthy individuals.

GABA, the major inhibitory neurotransmitter in the brain, has also been implicated in MDD. In patients, lower levels of GABA in the ACC and occipital cortex have been found (Sanacora et al., 2004). However, thus far, no medication increasing GABA transmission has been demonstrated to significantly improve depressive symptoms (Cowen, 2016). Although studies on the involvement of the excitatory neurotransmitter glutamate have been inconclusive, recent research has found strong antidepressant effects of the glutamatergic N-methyl-D-aspartic acid (NMDA) receptor antagonist ketamine at very low doses (Corriger and Pickering, 2019). Preliminary results provide a positive outlook for future treatments, with studies demonstrating positive effects on reward-related symptoms (Kotoula et al., 2022).

Recent theories propose that acute increases of monoamine levels produce neuroplastic changes involving modifications in transcriptional and translational levels of in gene expression mediating molecular and cellular plasticity (Nestler et al., 2002; Wong and Licinio, 2001). One example of this is the interaction between the serotonin 5-HT_{1B} receptor, which interacts with the calcium-binding protein p11. This is upregulated in the cerebral cortex following chronic treatment with SSRIs and has been reported to be downregulated in the cingulate cortex of post-mortem samples from depressed individuals (Svenningsson et al., 2006). It must be be noted, however, that some of the used tissue used in this study was collected from individuals who died by suicide. Therefore, suicidality may be a potential confound. Moreover, chronic treatment with antidepressants has resulted in the upregulation of the transcription factor CREB in the hippocampus, which is downstream of 5-HT receptors (Nestler et al., 2002; Pittenger and Duman, 2007).

1.1.4 Genetic & environmental contributors of MDD

MDD is in part mediated by a genetic component. Heritability of this disorder is approximately 35%; first-degree relatives of individuals diagnosed with MDD have a three times greater likelihood of developing this disorder (Geschwind and Flint, 2015). Additionally, it has been determined that there is a genetic overlap between MDD and psychiatric disorders including schizophrenia and bipolar disorder (Lee et al., 2013). The exact genetic risk factors have not yet been identified, as MDD risk is strongly polygenic and consists of many genes with small effect sizes (Hyman, 2014). Nonetheless, some polymorphisms have been associated with MDD, including the glucocorticoid receptor gene NR3C1, the monoamine oxidase A gene, the glycogen synthase kinase- 3β gene and a group-2 metabotropic glutamate receptor gene GRM3 (Fan et al., 2010; Tsunoka et al., 2009; Van Rossum et al., 2006; Yoon and Kim, 2010). Serotonin transporter genes are also thought to increase susceptibility to depression, specifically, the serotonin transporter gene 5-HTTLPR (Caspi et al., 2010; Uher and McGuffin, 2010). Given that many antidepressant drugs target the serotonin receptor and that 5-HT levels are postulated to be altered in MDD, it is likely that 5-HT transporter genes are involved in MDD aetiology. Genome-wide association studies (GWAS) have also greatly contributed to our understanding of the genetics underlying MDD and have identified multiple risk variants associated with this disorder, e.g., DRD2 and CELF4, and highlighted the importance of PFC regions and the ACC in particular (Howard et al., 2019; Wray et al., 2018). Another GWAS study reported an interactive effect between 5HTTLPR and the DPF1 gene (Garvert et al., 2022). Many studies are aiming to predict pharmacological treatment response based on an individual's genetics (Lohoff and Ferraro, 2010). For example, glutamatergic genes such as GRIK4 have been linked to citalopram response and adverse effects (Paddock et al., 2007). Genes for corticotrophin-releasing hormone (CRH) receptor-1 CRHR1 and CRH binding protein can also help predict SSRI response (Binder et al., 2010; Liu et al., 2013).

Environmental factors also contribute strongly to MDD onset. For a long time, it has been known that negative life events, for example death of a relative, loss of employment, debt, and chronic disease, can lead to a depressive episode (Kessler, 1997). Equally important are stressful events occurring in adolescence, for example, neglect, domestic violence, or abuse (Li et al., 2016). The number of stressful events in adolescence can clearly predict severity and chronicity of MDD. Such traumatic experiences may lead to long-term neurobiological consequences, as demonstrated in animals and humans that have experienced early life stress (ELS) showing persistent increases of the hypothalamic-pituitary-adrenal axis, which involves CRH (Meaney, 2001; Stetler and Miller, 2011).

Gene-environment interaction studies have also helped to explain the aetiology of depression. In one study, participants with one or more copies of the short allele of the *5HTT* promoter variant had a greater chance of developing depression when exposed to a negative life event (Caspi et al., 2003). However, the interaction between stress and *5HTTLPR* was

not confirmed by GWAS (Culverhouse et al., 2018). In another study, individuals carrying one of two genotypes for the T102C polymorphism of the HTR2A gene expressed low levels of depressive symptoms in the presence of high maternal nurturance, which was not the case for carriers of the C/C genotype (Jokela et al., 2006). Epigenetic regulation also plays a key role, as it has been shown that DNA demethylation in glucocorticoid-response elements of a polymorphism in *FKBP5* is stress-dependent and can lead to increased expression of this gene in response to stress (Klengel et al., 2012). Therefore, MDD has genetic and environmental components to it, as well as an interaction of the two. However, the exact mechanisms remain unclear and are yet to be elucidated.

1.1.5 Sex differences in MDD

The prevalence, incidence and morbidity risk of depression is generally higher in females than males (Piccinelli and Wilkinson, 2000). These differences begin during puberty and continue throughout adulthood. Epidemiological studies suggest that women have a higher likelihood of experiencing greater symptom severity and that somatic-/cognitive-affective symptoms are more common in women (Marcus et al., 2008). Atypical depression is more frequently reported in women, and its symptoms include hypersomnia, weight gain and fatigue (Angst et al., 2002; Blanco et al., 2012). Comorbidities also differ between males and females, with males more commonly presenting with comorbid substance use disorder, whereas women commonly have a co-diagnosis of anxiety disorder (Marcus et al., 2008; Schuch et al., 2014).

Differences in responding to antidepressant medication between the sexes have also been observed. Although there is no consensus on this matter, cohort studies suggest that women respond better to SSRIs than men, whereas men tend to respond better to tricyclic antidepressants (TCAs) (Kornstein et al., 2000; Yang et al., 2011). Not only are these differences seen in humans, but animal models have also demonstrated differential efficacy of antidepressants (Eid et al., 2019). Healthy female rodents appear to be more sensitive to acute SSRI administration, including fluoxetine, paroxetine and sertraline; as well as non-SSRIs such as ketamine (Carrier and Kabbaj, 2013; David et al., 2001; Gómez et al., 2014; Kokras et al., 2015). Many medications interact with the ovarian hormones oestradiol and progesterone, influencing the animals' sensitivity to those drugs (Carrier and Kabbaj, 2013; Naito et al., 2007). On the other hand, animals that have undergone interventions to induce depressive-like phenotypes have shown no effects of SSRI treatment in females but

found behavioural improvements in males (Carrier and Kabbaj, 2012; Gobinath et al., 2018; Wainwright et al., 2016).

Biologically, a number of factors may explain the behavioural differences between men and women. Firstly, reductions in brain volume resulting from long-term depression are more common in men than in women (Yang et al., 2017). Evidence suggests that pre-menopausal women show fewer age-related volume reductions of regions such as the hippocampus, which are implicated in MDD, and which could be attributable to the neuroprotective effects of oestrogen (Bromberger and Epperson, 2018; Goto et al., 2011). Moreover, brain morphology in women with untreated depression has been reported to be abnormal in prefrontal-limbic areas, whereas in men the abnormalities are mainly found prefrontal-striatal regions (Kong et al., 2013). There are also differences in gene expression in depression that are sexdependent (Kang et al., 2020). In post-mortem studies, altered gene expression for glutamate and 5-HT receptors were mostly found in women (Goswami et al., 2010; Gray et al., 2015). In men, genes related to the brain-derived neurotrophic factor (BDNF) were more prevalent (Tripp et al., 2012). Depression-related transcriptional brain changes show little overlap between males and females and have been reported to show opposing trends in the two sexes (Seney et al., 2018). Sex differences in gene transcription have also been reported in rodents that were exposed to stress (Brivio et al., 2020). In summary, evidence suggests that depression-related biological profiles significantly differ between males and females across species, aligning with behavioural and pharmacological findings.

1.1.6 Treatments for MDD

The most common treatment options for MDD are psychotherapy and pharmacological interventions, either separately or combined (Otte et al., 2016). In most cases, mild MDD is initially treated with psychotherapeutic interventions, whereas treatment for more moderate or severe diagnoses involve medication or a combination of both treatment options. Extreme cases, such as treatment-resistant MDD, may require rarer treatments such as electroconvulsive therapy (Kellner et al., 2012). Psychotherapy can take a variety of forms, including CBT, behavioural activation therapy, psychodynamic therapy, problem-solving therapy, interpersonal therapy and mindfulness-based therapy. A summary of how these therapies work can be found in box 2. This section is intended to provide a brief overview of the available treatment options, and by no means provides a full list of these. For a more extensive review, please see (Karrouri et al., 2021; Otte et al., 2016).

Box 2: Types of psychotherapeutic interventions for MDD. Based on (Otte et al., 2016).

Cognitive behavioural therapy

CBT involves the identification of negative, distorted thinking patterns and helps the patient to develop skills to identify and challenge these thoughts.

Behavioural activation therapy

In this type of therapy, the therapist aims to increase the patient's positive activities that provide pleasure or mastery. This treatment may also involve the identification and confrontation of avoidance processes.

Psychodynamic therapy

Through psychodynamic therapy, the individual gains a deeper understanding of how emotions, thoughts and previous life experiences contributed to current problems. This enables the patient to deal with these patterns better.

Problem-solving therapy

This intervention is aimed at teaching the patient a set of skills that helps them generate creative methods to address problems, how to better identify and overcome hurdles and make more effective life decisions.

Interpersonal therapy

Interpersonal therapy can help to identify and resolve issues in relationships and deal with interpersonal conflict, role transitions and difficult relationships.

Mindfulness-based therapy

This involves frequent meditation and is aimed at the individual paying greater attention to thoughts, feelings, and experiences without judgement and accept things as they are without trying to change them.

A possible explanation for the efficacy of psychological therapies is the idea that the strategies employed can induce change via cognitive restructuring, behavioural activation, or better interpersonal functioning (DeRubeis et al., 2005).

As mentioned in previous sections, most drugs to treat MDD aim to increase transmission of neurotransmitters, particularly monoamines. These medications include SSRIs, TCAs, serotonin-norepinephrine reuptake inhibitors (SNRIs), serotonin antagonists and reuptake inhibitors (SARIs) and MAOIs. Beyond the increase of monoamine levels synaptically, these drugs are also known to induce neural plasticity (Hyman and Nestler, 1996). Specifically, they result in altered neural responses to biochemical perturbations in the synapse via downstream changes in intracellular signalling pathways, modifications in gene expression, neural as well as synaptic plasticity (Sharp, 2013). Newer medications for MDD that are currently being investigated include the NMDA receptor antagonist ketamine and 5-HT-2A receptor agonist psilocybin (Daws et al., 2022; Kotoula et al., 2021; Krystal et al., 2019). Moreover, researchers are trying to develop antidepressant drugs that do not focus on monoamine transmission, such as anti-inflammatory agents and opioid- κ antagonists (Ehrich et al., 2015; Noto et al., 2014).

1.2 Introduction to Substance Use Disorders

The DSM-5 defines SUDs by the presence of a minimum of two out of eleven possible criteria in a 12-month period (APA, 2013). Disorder severity is determined by the number of criteria that are present, with 2-3 criteria indicating mild SUD, 4-5 moderate and >6 criteria as severe. The criteria include craving, a persistent desire to cut down substance use, negative consequences arising from substance use, and withdrawal symptoms. These criteria generally apply to a variety of substances, such as alcohol, nicotine, stimulants, opioids, and cannabis. The cycle of addiction has been proposed to consist of three components: preoccupation-anticipation, binge-intoxication, and withdrawal-negative affect, which form part of the DSM-5 criteria (Koob and Le Moal, 2001). This thesis will mostly focus on stimulant use disorder, specifically cocaine use disorder (CUD). In box 3, the DSM-5 criteria for stimulant use disorder are outlined.

Epidemiologically, rates of SUDs are high across the general population (Deak and Johnson, 2021). Up to 30% of individuals will meet the criteria for alcohol or nicotine use disorder in their lives (Grant et al., 2016). 2.4% of the population will fulfil the criteria for CUD. SUDs are responsible for increased comorbidities and greater number of deaths, for example, nicotine use disorder results in 7 million deaths worldwide every year (UN, 2018). In 2018, approximately 50,000 people died as a result of opioid overdoses (Hedegaard et al., 2020). These facts underline the importance of gaining a deeper understanding of SUDs in order to guide development of preventative measures and treatments.

1.2.1 Behavioural differences in SUD

Substance use disorders affect a variety of brain processes and circuits, including those involving reward, inhibitory control, stress, emotional processing, learning and memory (Hayes et al., 2020). Deficits in decision-making provide a robust cognitive marker of substance abuse (Verdejo-Garcia et al., 2007). On the IGT, cocaine and cannabis users performed worse than non-substance users. Following 25 days of enforced abstinence,

cocaine users showed the smallest improvements in performance on the task compared to healthy controls and cannabis users, and heavier drug use across both cocaine and cannabis groups was inversely correlated with performance. Individuals with heightened risk for SUDs, as determined by the number of alcohol-dependent relatives, also showed worse performance on the IGT compared to individuals who are not at risk (O'Brien et al., 2014). Additionally, high-risk subjects learned less from experience than the low-risk group. Participants with the 5-HT transporter-linked polymorphic region (5-HTTLPR) short-allele homozygote have worse performance (Heils et al., 1996). 5-HTTLPR influences serotonin concentration in the synaptic cleft and serotonin transporter gene expression, whose dysregulation has been associated with SUD. In summary, not only individuals diagnosed with an SUD, but also people at high risk of developing an SUD, have difficulties on decision-making tasks such as the IGT.

Box 3: DSM-5 criteria for Stimulant-Related Disorders

A) Repeated amphetamine-type substance, cocaine, or other stimulant use leading to clinically significant impairment or distress, as highlighted by at least two of the listed symptoms occurring within a 12-month period:

- 1. Stimulant is frequently taken in larger amounts or over a longer period than intended.
- 2. Persistent desire or unsuccessful efforts to cut down or control stimulant use.
- 3. A lot of time is spent on activities to obtain the stimulant, use the stimulant, or recover from its effects.
- 4. Craving, or a strong desire or urge to use the stimulant.
- 5. Recurrent stimulant use resulting in inability to fulfil commitments at work, school, or home.
- 6. Continued stimulant use despite having persistent or recurrent social or interpersonal problems, which are caused or exacerbated by use of the stimulant.
- 7. Important social, occupational, or recreational activities are given up or not done as frequently due to stimulant use.
- 8. Repeated stimulant use in physically hazardous situations.
- 9. Stimulant use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that has probably been caused or exacerbated by the stimulant.
- 10. Tolerance as defined by either one of the following two:a) A need for substantially increased amounts of stimulant to achieve intoxication

or desired effect. b) A significantly reduced effect with continued use of the same amount of stimulant.

11. Withdrawal, which manifests itself as either of the following:

a) The characteristic withdrawal syndrome for the stimulant.

b) The stimulant is taken to relieve or avoid withdrawal symptoms.

Impulsivity is a multidimensional trait that reflects a predisposition for rapid, premature actions without forethought and aversion for delayed and uncertain rewards (Dalley and Robbins, 2017; Evenden, 1999; Jones et al., 2021). Impulsivity has been shown to be both a vulnerability marker and consequence of drug use; trait impulsivity has been found to predict compulsive drug-seeking in rats (Dalley et al., 2007; Ersche et al., 2010; Verdejo-Garcia et al., 2007). Genetically, the heritability of impulsive-compulsive behavioural phenotypes has been shown and that impulsive and compulsive traits in humans are strongly related (Chamberlain et al., 2018; Tiego et al., 2020). Impulsivity is multifaceted and can be sub-divided into impulsive choice and impulsive action (Dalley et al., 2011). Impulsive action reflects the failure of motor inhibition, whereas impulsive choice is the preference for immediate, rather than delayed reward. Impulsive action is oftentimes measured using the stop-signal reaction time task (SSRTT) or the Go/No-go task, impulsive choice is tested using paradigms such as the delay discounting task (DDT) (Jones et al., 2021). Overall task performance on the Go/No-go task is reduced in smokers, and especially performance on no-go trials is impaired (Luijten et al., 2017). In recreational cocaine users, the time required to inhibit responses to stop-signals on the SSRTT is significantly higher than in non-users (Colzato et al., 2007). Furthermore, alcohol, methamphetamine and cannabis-dependent subjects respond prematurely on the 5-choice serial reaction time task (5-CSRTT), another measure of impulsive action (Voon et al., 2014). Not only do people with SUDs show increased impulsive behaviour on tasks of motor impulsivity, but they also show increased impulsive choice. Cocaine users show greater delay discounting of monetary and other rewards, making it evident that outcome devaluation functions differ significantly between SUD patients and controls (Heil et al., 2006). Overall, these studies strongly suggest that impulsive action and choice are common behavioural markers of SUDs and are heightened across different SUD subtypes. Moreover, animal studies show that they may predispose to compulsive drug reinforcement.

1.2.2 Neural changes in SUD

The dopaminergic mesocorticolimbic pathway is responsible for reward processing and projects from the ventral tegmental area (VTA) to the striatum and PFC (Hayes et al., 2020). Other regions involved in reward processing include the OFC, ACC, amygdala, and hippocampus. The VS is important for signalling the reinforcing effects of a stimulus. The OFC and ACC attribute salience to reward; the vlPFC is responsible for inhibition of behaviour; the amygdala and hippocampus receive dopaminergic innervation and are

responsible for encoding memories related to rewarding stimuli. The MID task makes it possible to investigate brain activity associated with anticipated monetary loss, gain or neutral outcomes and the actual outcome. On this task, brain activity to reward anticipation was associated with activation clusters in the striatum and thalamus in healthy individuals (Oldham et al., 2018). In substance-dependent individuals (SDIs), however, there was blunted striatal activation during anticipation of reward (Luijten et al., 2017). In response to drug-related stimuli, on the other hand, participants with SUD show greater activation than healthy individuals. On so-called cue-reactivity tasks, SDIs show greater activity in the VS, ACC and amygdala (Chase et al., 2011; Kühn and Gallinat, 2011). A meta-analysis involving 25 task-based fMRI studies involving monetary reward anticipation (Luijten et al., 2017). Altered signals were also confirmed in the ACC, amygdala, OFC and dIPFC. During reward *outcome*, VS activation was greater than in healthy individuals, as well as in the insula and SFG. In Gambling Disorder (GD), the VS signal was decreased in the dorsal striatum (DS) following reward outcome.

Impulsivity is a manifestation of weakened inhibitory control, reflecting impaired topdown inhibitory control processes that are required to manage pre-potent motor or rewarddriven responses (Verdejo-García et al., 2008). Motor and choice impulsivity are thought to involve different neural pathways, with the former engaging the ACC, pre-supplementary motor area (pSMA) and inferior frontal gyrus (IFG), whereas the latter requires the MFG and frontal pole (Wang et al., 2016). On the stop signal task, alcohol-dependent individuals have shown greater response magnitude following a stop signal in the inferior parietal cortex and anterior pSMA, whereas healthy control participants showed a deactivation of the ACC (Hu et al., 2015). In participants with CUD, inhibitory control has been linked to decreased activation of the dmPFC and lateral occipital cortex (Ceceli et al., 2022). Moreover, there was a negative correlation in CUD between the stop signal reaction time (SSRT) and brain signal in the posterior cingulate gyrus, precuneus and cuneal cortices. On the Go/No-go task, participants with CUD had lower brain activations in the pSMA, insula and ACC during successful no-gos and commission errors (Kaufman et al., 2003). In heavy drinkers between the ages of 18 and 22, however, activity in the dlPFC, medial PFC, insula and ACC was greater than in light drinkers during no-go trials (Ames et al., 2014). These studies suggest that brain areas critical for cognitive control and inhibition are functionally impaired in SUD, reflecting behavioural observations. Figure 1.1 gives an overview of the neural circuitry involved.



Fig. 1.1 Neurocircuitry underlying substance use disorder.

Neurocircuitry underlying the three stages of the addiction cycle: preoccupation/anticipation, binge/intoxication and withdrawal/negative affect. From (Koob and Volkow, 2010).

1.2.3 Neurochemical basis of SUD

When an individual continuously uses a drug and slips into dependence, significant alterations in the mesocorticolimbic system occur, as DA levels are increased in the striatum (Koob and Volkow, 2009). Cocaine and amphetamines are blockers of the presynaptic dopamine transporter DAT, whereas other drugs such as opiates, alcohol and nicotine increase extracellular DA levels via the inhibition of GABA interneurons. DA is not the only substance whose transmission is altered in addiction and plays an important role. Endogenous endorphins are also involved in reward signalling, as shown via PET imaging studies in which increased endogenous endorphin levels following alcohol or stimulant intake have been reported (Colasanti et al., 2012; Turton et al., 2018). The transition to compulsion is thought to be a result of dysfunctional top-down inhibitory response control (Everitt and Robbins, 2016; Jentsch and Taylor, 1999). Behavioural paradigms have been developed in rodents to capture different aspects of addiction and study the transition to compulsion. Most studies investigating the neurochemistry underlying compulsive drug-seeking and taking has therefore been undertaken in rodents. Compulsive drug-taking can be assessed via response-contingent self-administration (SA) in the face of concurrent punishment, such as a foot-shock or unpleasant tastant such as quinine. Cue-controlled drug-seeking can be

investigated with second-order schedules of reinforcement paired with or without punishment, during which drug delivery occurs after a fixed amount of time alongside a classically conditioned stimulus. Later in this paradigm, the conditioned stimulus is presented without the drug, enabling the investigation of drug-seeking behaviours (Jones et al., 2021).

5-HT depletion in rats via intracerebroventricular infusions of 5-7-dihydroxytryptamine have led to increased drug-seeking under punishment (Pelloux et al., 2012). Citalopram, a selective 5-HT reuptake inhibitor, on the other hand, reduced compulsive drug-seeking. Genetic depletion of 5-HT_{2C} receptor, resulting in reduced 5-HT_{2C} receptor protein production in the mPFC, increased motor impulsivity and cocaine seeking during withdrawal (Anastasio et al., 2014). Moreover, a decrease in phasic DA release in the ventromedial striatum of non-human primates following weeks of cocaine exposure, alongside an increase in the dorsolateral striatum has been reported (Willuhn et al., 2012). DA release in the DS, but not VS, was observed following exposure to a drug-associated stimulus (Ito et al., 2002). Not only have DA and 5-HT been studied in relation to compulsive drug reinforcement, but the involvement of NA, GABA and glutamate has also been demonstrated. The α_2 receptor antagonist yohimbine, which causes NA release due to the inhibition of autoreceptors, increased cue-induced seeking for cocaine, an effect that was reduced via the α_2 receptor agonist clonidine (Lee et al., 2004). Infusions into the nucleus accumbens core (NAcC), which plays a key role in the reconsolidation of cue-drug memories, of the selective α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)/kainate receptor antagonist LY293558 dose-dependently decreased cocaine seeking in rats (Di Ciano and Everitt, 2001; Théberge et al., 2010). Furthermore, NMDA receptor antagonist infusions into the basolateral amygdala prevented the reconsolidation of drug-associated memories and impaired drug-seeking behaviours (Milton et al., 2008). The systemic administration of N-acetylcysteine reduced cocaine seeking during a second-order schedule of reinforcement, through increasing glutamate release (Murphy et al., 2012a). Finally, short hairpin RNA knockdown of GAT-3 in rats increased alcohol choice behaviour, and GABA receptor activation increased cocaine seeking under punishment (Augier et al., 2018; Sun and Yuill, 2020). In summary, drug seeking is a result of a complex interplay between various neurotransmitter systems, whose effects are yet to be fully understood.

1.2.4 Genetic contributors to SUD

SUDs have a strong genetic component (Yu and McClellan, 2016). However, we still have a poor understanding of the polygenic architecture responsible for traits predisposing



Fig. 1.2 Effects of different neurotransmitters on tasks commonly used to assess impulsivity or compulsivity.

Diagrams show the locations of cell bodies in the dorsal raphé nucleus (DRN), ventral tegmental area (VTA) and locus coeruleus (LC), as well as their ascending projections to the forebrain. The overall effects of systemically increasing or decreasing 5-HT (yellow) and dopamine (blue) or noradrenaline (orange) on different forms of impulsivity and compulsivity are summarised on the right. Upward and downward arrows denote increased and decreased impulsivity and compulsivity, respectively. Horizontal bidirectional arrows indicate no clear effect of the manipulation. The red asterisk (*) indicates that studies have reported both a decrease and null effects on SIP. A dash (-) indicates that the effects are unknown. Abbreviations: 5-CSRTT, 5-choice serial reaction time task; DDT, delay-discounting task; SIP, schedule-induced polydipsia; SSRTT, stop-signal reaction time task. Adapted from (Jones, Zuhlsdorff & Dalley, 2021).

an individual to SUD. Early studies employing genome-wide linkage and candidate gene association methods identified a number of genes that may predispose to alcohol use disorder (AUD). In families with high rates of alcohol dependence, regions of chromosome 4 coding for the alcohol dehydrogenase gene ADH were identified (Reich et al., 1998). This gene encodes ADH isozymes which metabolise alcohol into acetaldehyde, as well as several GABA receptor genes (Edenberg and Foroud, 2006). Association studies have related variants in or near alcohol metabolising genes, such as ADH1B and ALDH2, the chromosome 4 GABA gene cluster and the μ opioid receptor gene *OPRM1* (Edenberg et al., 2004; Ray and Hutchison, 2004). These studies have also highlighted the involvement of genes related to neural pathways important for stimulus-reward processing, for example dopaminergic (DRDS2, MAOA), serotonergic (HTR3A), glutamatergic (GRIN2C) and GABAergic (GABRA1, GABRA2) pathways (Prom-Wormley et al., 2017). Genetic association studies for nicotine dependence found genes important for metabolising drugs (ADH1B, CYP1JA1), synaptic transmission of various neurotransmitter systems (CHRNB4, GRM7, DRD1) and reuptake/vesicular packaging of neurotransmitters involved in learning and memory (SLC6A4, SLC18A2) (Harari et al., 2012; Liu et al., 2015; Prom-Wormley et al., 2017). For CUD, genetic association studies highlighted many genes that are involved in dopaminergic pathways. These include genes that are involved in the expression of the D2 dopamine receptor (DRD2), signalling of DA (ANK11, TTC12), DA degradation (COMT) and many other processes involved in DA transmission (Prom-Wormley et al., 2017).

In recent years, GWAS have provided a new insight into the genetics of SUDs. For opioid use disorder, GWAS have identified multiple important loci, such as the repulsive guidance molecule BMP co-receptor A (*RGMA*), cornichon family AMPA receptor auxiliary protein 3 (*CNIH3*), potassium voltage-gated channel subfamily C member 1 (*KCNC1*) as well as *FURIN* and *OPRM1* (Deak et al., 2022; Gelernter et al., 2014; Nelson et al., 2016). For AUD, the alcohol-metabolism gene alcohol dehydrogenase 1B and 1C (class 1) β -polypeptide (*ADH1B* and *ADH1C*) and aldehyde dehydrogenase 2 family member (ALDH2) are thought to predispose to alcohol abuse (Gelernter and Polimanti, 2021; Kranzler et al., 2019; Zhou et al., 2020). Only a few GWAS studies for cocaine dependence have been done thus far; nonetheless, some risk variants have been identified. One example is the family with sequence similarity 53 gene, member B (*FAM53B*) and transmembrane protein 51 (*TMEM51*) (Gelernter et al., 2014; Sun et al., 2020). GWAS have also identified an association between chromosome 7 and chromosome 8 loci (*FOXP2* and near *CHRNA2*, respectively) and Cannabis Use Disorder (Johnson et al., 2020).

More recent genetic methods have demonstrated the greatly polygenic nature of AUD and the heritability of traits relevant for AUD (Deak et al., 2019). Polygenic risk scores have been able to predict nicotine dependence, AUD, and cannabis use disorder (Belsky et al., 2013; Vink et al., 2014). Other studies have combined different types of 'omics' data, such as epigenomics or transcriptomics, to identify novel gene candidates. One study has been able to link a variant in an intronic region of the *CNTN4* gene and AUD (Clark et al., 2015). The advance of such novel techniques will hopefully help to develop a better understanding of genetic markers of SUDs and help guide future treatment discovery.

1.2.5 Treatments for SUD

The main treatments for SUD involve psychosocial or pharmacological interventions, or a combination of both (Hayes et al., 2020). For CUD, a commonly used treatment consists of group, individual and family therapy, known as intensive outpatient therapy (Kampman, 2019). Alternative psychosocial treatments include CBT and voucher-based reinforcement therapy (VBRT). The former strongly prevents relapse, and the latter is highly effective for promoting abstinence. In VBRT, the patient receives a voucher that can be redeemed for goods and services in the community, as long as the individual manages to fulfil a therapeutic goal, for example drug abstinence. It is most efficient when accompanied by another type of psychosocial treatment, such as CBT. A limitation of VBRT is that once the positive reinforcers are no longer present, abstinence is oftentimes short-lasting. However, when it is done in parallel with other therapies, long-lasting outcomes can be achieved (Rawson et al., 2002). CBT aims to reduce drug craving, by teaching patients to identify which situations they come across in daily life that induce craving. These situations can be avoided by using techniques such as distraction, recall of negative consequences and positive thought substitution (Carroll and Onken, 2005).

Pharmacological interventions for SUD depend on the drug, but generally target one of the following phases: acute intoxication, substitution, detoxification, or relapse prevention (Hayes et al., 2020). During withdrawal, an organism is unexpectedly deprived of a substance that it has been exposed to over prolonged periods of time (Ries et al., 2014). Therefore, one way of managing it is via substitution with an agent that has a similar effect on the brain. This can be done for a variety of SUDs, including, AUD, tobacco use and opioid use disorder. For AUD, benzodiazepines are used during the withdrawal/detoxification stage, naltrexone or acamprosate are the first line treatments for relapse prevention (Hayes et al., 2020). For opiate addiction, naloxone is used during overdose, methadone and buprenorphine are used for

substitution, lofexidine for detoxification and naltrexone for relapse prevention (Ling et al., 1996). During tobacco withdrawal, nicotine replacement is used (Wadgave and Nagesh, 2016). For stimulants, including cocaine, cannabinoids and other novel psychoactive substances, no pharmacological therapies have thus far been identified. However, treatment during all stages can be supported via medication for relief of symptoms. Withdrawal symptoms include sleep disturbances, restlessness, enervation, indigestion, and mood problems (Dakwar and Kleber, 2022). These can be addressed using medications such as zolpidem, benzodiazepines, stimulants, clonidine and guanfacine. The latter two are α_2 receptor agonists, which reduce the surge of NA observed in withdrawal states, leading to some of the named symptoms. For CUD, current drug classes that are being investigated include DA agonists, GABA receptor agonists and glutamate receptor antagonists (Kampman, 2019). Most agonist treatments target the same receptor as the drug of abuse, which is a difficult task for CUD, as cocaine targets more than one receptor. In summary, thus far only few treatments for drug dependence with high efficacy have been identified. Treatments do not exist for all types of SUDs. Further developments require a deeper understanding of the mechanisms underlying SUDs.

1.3 What is cognitive flexibility?

Cognitive flexibility is the ability to adapt to changes in the environment by switching task sets, responses, or strategies (Cools, 2015). It is a distinct yet overlapping construct with control processes such as inhibition and updating (Miyake et al., 2000). Flexibility makes it possible to adjust behaviour following changes in environmental or internal cues. Individuals with greater cognitive flexibility have improved life outcomes, better social functioning and reduced cognitive decline with age (Burke et al., 2019; Diamond and Lee, 2011; Koesten et al., 2009). Difficulties in flexible responding have been observed in various psychopathologies. For example, patients with obsessive-compulsive disorder (OCD) show impaired performance on tasks such as the Intra-/Extra- Dimensional Set Shift (IED) task, which is linked to reduced functional connectivity between the striatum and the prefrontal cortex (Remijnse et al., 2006; Vaghi et al., 2019). In patients with autism spectrum disorder (ASD), repetitive behaviours are correlated with reversal deficits on the IED task, further highlighting the importance of studying flexible behaviour (Yerys et al., 2009).

Cognitive flexibility is considered an executive function, or a process necessary for goal-directed behaviour (Miyake et al., 2000). Two commonly used paradigms to investigate flexibility include the WCST and the reversal learning task. The former requires the ability of an individual to infer rules in order to guide behaviour, create an attentional set and be able to

shift the set by adjusting behaviour when task rules are changed (Banich, 2009). This task is mostly used in humans, whereas the reversal learning task can be employed across species; in rodents, non-human primates and humans. This task assesses how easily the animal can adapt its behaviour in response to changes in stimulus-outcome or response-outcome contingencies (Izquierdo, 2017).

Different neural circuits are involved in mediating flexibility. The executive control network (ECN) and salience network (SN) are known to play important roles in supporting executive function and cognitive flexibility (Dajani and Uddin, 2015; Uddin, 2021). The ECN involves the dIPFC, vIPFC and inferior frontal junction (IFJ), midcingulate gyrus and parts of the temporal lobe and inferior parietal lobule. The SN consists of the anterior insula, the anterior midcingulate cortex, amygdala, and thalamus (Uddin et al., 2019). In monkeys, lesions of the lateral OFC have resulted in reversal learning impairments, whereas medial PFC lesions caused impairments in extradimensional shifts (Dias et al., 1996b). In rats, medial OFC inactivation has also resulted in reversal learning impairments, attributable to an increase in perseveration (Dalton et al., 2016). Inactivation of the dorsomedial striatum in rodents has also resulted in reversal learning impairments as well as difficulties in strategy switching (Ragozzino, 2007).

Deficits in cognitive flexibility are observed across a variety of psychiatric and neurological conditions; including, but not limited to, ASD, attention-deficit/hyperactivity disorder (ADHD), MDD, OCD, Alzheimer's disease and Parkinson's disease (Leeson et al., 2009; Remijnse et al., 2006; Swainson et al., 2000; Uddin, 2021). In ASD, executive function is affected across measures, with strong impairments observed in flexibility as assessed via the WCST (Demetriou et al., 2018; Lai et al., 2017). These behavioural changes have been associated with altered activity in the ECN and SN (Uddin, 2021). Excessive rumination in MDD is thought to be linked to inflexible thinking and difficulty engaging the ECN (Burrows et al., 2017; Uddin, 2021). In neurological disorders such as dementia, impaired task-switching and set-shifting is observed and has been linked to neurodegeneration of the PFC (McDonald et al., 2018). In summary, these findings underline the importance of studying cognitive flexibility in order to understand its neural basis and develop new treatments.

1.4 Reversal learning: cognitive aspects, neural circuits and neurochemistry

As mentioned previously, the reversal learning paradigm measures cognitive flexibility by evaluating behavioural adjustments following changes in stimulus-outcome or responseoutcome contingencies (Izquierdo, 2017). It is one of the main paradigms for assessing cognitive function and flexibility. Subjects are trained to discriminate between two stimuli, which can be visual, auditory, or spatial, and learn which one of the stimuli is rewarded and which one results in no reward or punishment. After having learnt the stimulus-outcome association, the contingencies are reversed, and subjects have to re-learn the contingencies (Fellows and Farah, 2003). These tasks can either be deterministic, which means that every correct response is rewarded, and every incorrect response is punished, or probabilistic, where only a certain percentage of correct trials are rewarded (e.g., 70% or 80 %) and some of the incorrect trials are spuriously rewarded (e.g., 30% or 20%) (DRL vs PRL). One of the reasons why reversal learning tasks have been so popular is that they can be employed across species, making it highly valuable for translational research (Izquierdo, 2017). In rats, the reversal learning task set-up includes levers, nose poke portals or a touchscreen, providing either solely spatial information or stimulus information if stimuli are presented on a screen (Jentsch and Taylor, 2001). In non-human primates, cards or a touchscreen can be used, similarly to the implementation of this task in rodents (Clarke et al., 2005; Walker et al., 2009). However, in monkeys, the number of reversals is usually different to rodents, as sometimes more than seven reversals need to be completed, whereas in rodents the number of reversals can be as low as one (Schoenbaum et al., 2002). In human studies, the reversal learning task can also take many forms, usually, a visual stimulus is presented on a screen and the participants needs to respond by touching the screen or pressing a button on the keyboard (Cools et al., 2002; Lawrence et al., 1999). In humans, reversal learning tasks are mostly probabilistic, as they are more difficult and can include multiple contingency changes (Izquierdo, 2017). PRL tasks significantly slow the rate of learning and enable researchers to study learning under uncertainty, which is relevant for various disorders such as MDD.

It is possible to extract a variety of measures from reversal learning tasks, not only including the number of correct responses and errors, but also the number of perseverative responses, the percentage of 'staying', i.e., selecting the same stimulus, or percentage of 'shifting', i.e., selecting a different stimulus. Moreover, it is possible to fit computational models on a trial-by-trial basis, which will be discussed in more detail in later sections. These tasks facilitate the study of at least three important cognitive processes: 1) learning from

reward or punishment following the selection of a stimulus; 2) estimation of the likelihood or prior probability of a reversal and 3) generating an understanding of the task or option space (Costa et al., 2015; Izquierdo, 2017; Stalnaker et al., 2015; Wilson et al., 2014).

Many regions are implicated in the mediation of reversal learning tasks. Using neuroimaging studies in humans, cortical areas such as the OFC and mPFC were found to be activated on an fMRI scan during a reversal (Cools et al., 2002; Remijnse et al., 2005). Excitotoxic lesions of the OFC in marmosets caused selective impairments in reversal learning without affecting any other forms of behavioural flexibility and single neuron activity of the OFC during reversal learning has revealed that these neurons track changes in reward contingencies (Dias et al., 1996b; Thorpe et al., 1983). In rodents, excitotoxic lesions or inactivations of the OFC impair reversal learning (Bissonette et al., 2008; Bohn et al., 2003; Izquierdo et al., 2013). The mPFC plays a slightly different role in reversal learning. Despite many rodent and monkey studies reporting that mPFC damage does not affect reversal learning (Bissonette et al., 2008; Floresco et al., 2008; Rudebeck et al., 2006), macaques with mPFC lesions show a small deficit in the ability to respond correctly following a reversal (Chudasama et al., 2013). Moreover, the ACC has been found to track reward and non-reward during reversal learning (Kawai et al., 2015). Therefore, it seems that the role of the mPFC is mostly linked to situations in which high attentional demands and performance monitoring are required (Izquierdo, 2017). Not only cortical regions are activated in reversal learning, but subcortical regions are, too. In humans, the dorsal and ventral striatum are active during a reversal (Rogers et al., 2000). Neurotoxic lesions of the dorsomedial striatum in marmosets and rats have resulted in reversal learning impairments (Castañé et al., 2010; Clarke et al., 2008). Ventral striatal neurotoxic lesions result in spatial reversal learning deficits in monkeys and PRL deficits in rats (Dalton et al., 2014; Stern et al., 1995). The amygdala receives cortical inputs and projects to both dorsal and ventral striatum, forming a circuit that supports reversal learning (Izquierdo, 2017). However, its exact involvement remains unclear. BLA and amygdala lesions in the monkey have improved reversal learning (Izquierdo, 2017). On the other hand, when the BLA is inactivated after training but prior to a reversal, performance is disrupted (Schoenbaum et al., 2003). Although no clear consensus can be reached on its role in reversal learning, it is evident that the amygdala is involved in mediating this cognitive process.

The involvement of neurotransmitters in reversal learning is unclear and complicated. What is known, is that 5-HT, glutamate and DA play an important role. Systemic 5-HT depletion impairs reversal learning and 5-HT levels in the vIPFC have been shown to be correlated with performance on the reversal learning task (Groman et al., 2013; Izquierdo

et al., 2012). Increased 5-HT transmission generally leads to improvements in reversal learning, whereas decreased 5-HT signalling results in impairments and increases perseverative responses (Clarke et al., 2007; Rogers et al., 1999; Rygula et al., 2015). Dopaminergic innervation in the brain also plays an important role in reversal learning. Optogenetic excitation of DA neurons in the ventral tegmental area and substantia nigra pars compacta enhance reversal learning in rats (Adamantidis et al., 2011; Rossi et al., 2013). DA depletions in the marmoset striatum have also resulted in impairments, which were not present after 5-HT depletions (Clarke et al., 2011). In humans, reversal learning improvements as a result of increased striatal DA via methylphenidate have been observed (Clatworthy et al., 2009). Finally, glutamate has also been implicated in reversal learning via its effects on the NMDA receptor, which is an important receptor for learning and memory (Rezvani, 2006). NMDA receptor antagonism impairs spatial reversal learning in rats; deletion of the GluN2A subunit of the receptor leads to the same effect in mice (Brigman et al., 2008; Marquardt et al., 2014). GluN2B antagonism and cortex-wide deletion have consistently resulted in increases in perseveration, leading to reversal learning deficits (Dalton et al., 2011; Radke et al., 2015). In summary, multiple neurotransmitter systems are involved in reversal learning, highlighting its underlying complexity.

1.5 Endophenotypes in psychiatry

Endophenotypes are measurable components that are present along the pathway between disease and distal genotype (Gottesman & Shields, 1972). They can be of neurophysiological, endocrinological, biochemical, neuroanatomical, cognitive, or neuropsychological nature, are heritable traits and can be derived from laboratory measures, e.g., neurocognitive performance deficits or impaired facial emotion recognition (Gottesman and Gould, 2003; Iacono, 2018). Endophenotypes enable researchers to 'decompose' or 'deconstruct' complex psychiatric diagnoses, making it possible to uncover the genetic underpinnings of these disorders more easily and develop animal models to investigate the respective endophenotype. They may also support prognosis and the development of treatments. Ideally, an endophenotype will have been or will be associated with a candidate gene or gene region. The criteria that need to be fulfilled for a component to be considered an endophenotype are the following (Gershon and Goldin, 1986; Gottesman and Gould, 2003; Leboyer et al., 1998):

- 1. Associated with illness in the population.
- 2. Heritable.

- 3. Primarily state independent.
- 4. Co-segregate with illness within families.
- 5. More prevalent in non-affected family members in the same family than a non-affected individual in the overall population.

Relevant endophenotypes have been found for different psychiatric disorders and are still being extensively researched. One example of successful endophenotype identification is for schizophrenia. Sensory motor gazing deficits, eye-tracking dysfunction and impaired working memory are a few of the endophenotypes associated with this disorder (Braff and Freedman, 2002; Diefendorf and Dodge, 1908; Goldberg and Green, 2002). Sensory motor gazing deficits have been linked to event-related electroencephalogram (EEG) markers, which have also been observed in unaffected relatives (Adler et al., 1982; Siegel et al., 1984). This endophenotype has been linked to a locus on chromosome 15, which may predispose to schizophrenia (Freedman et al., 1997). Similar successes have been reported for other schizophrenia endophenotypes as well as other psychiatric disorders; highlighting the importance of this approach in psychiatry. Experimental approaches in rodents have been widely used to investigate translationally relevant endophenotypes including impulsivity, anxiety and behavioural flexibility, with tasks developed to make objective assessments in different species (Benzina et al., 2014; Bergeron et al., 2017; Dalley and Robbins, 2017; Iacono, 2018). The greater neural accessibility of rodents enables deeper mechanistic explanations of relevant behavioural phenomena. One difficulty of using endophenotypes in psychiatry is that rather than being disorder-specific, they can be observed across different disorders. However, given that it is now known that psychiatric disorders are highly heterogeneous and since diagnostic criteria for psychiatric disorders are moving towards dimensional psychiatry (a dimensional approach in psychiatry aiming to identify core mechanisms of mental disorders across nosological boundaries), they may be very useful in the future (Feczko et al., 2019; Hägele et al., 2015; Robbins, 2012). The development and use of animal models for investigating endophenotypes relevant for drug addiction and depression will be further discussed in this section.

1.5.1 Preclinical models of depression

Although big steps in the understanding of MDD have been made in the last decades, these findings have not yet led to ground-breaking treatment and prevention strategies. Endophenotypes may provide a promising avenue for future research, especially those that can be translated into animals, allowing *in vivo* investigations of the neural circuits involved. Relevant endophenotypes for MDD include anhedonia, affective bias, and behavioural despair (Der-Avakian and Pizzagalli, 2018; Goldstein and Klein, 2014; Pizzagalli, 2014). Negative processing bias is a key feature of MDD, as discussed in previous sections (Gotlib and Joormann, 2010). In recent years, a variety of tasks have been developed to measure this bias for use in humans and animals, for example, probabilistic discrimination, approach-avoidance, and ambiguous cue tasks (Bari et al., 2010; Clarke et al., 2015; Murphy et al., 2003; Papciak et al., 2013). Tasks investigating behavioural despair, which is related to symptoms of helplessness and hopelessness in MDD, have not seen as much overlap between rodents and humans. This is primarily because in rodents, the forced swim and tail suspension tests are employed, which reflect acute stress rather than chronic depression (Porsolt et al., 1977; Steru et al., 1985). Anhedonia is most commonly assessed via the sucrose preference test (SPT), for which the analogous sweet taste test in humans exists, although results in rodents and humans seem to diverge, highlighting the need for novel tasks to measure anhedonia (Dichter et al., 2010; Katz, 1982).

In rodents, a persistent depressive-like phenotype can be elicited by exposing experimental animals to uncontrollable, chronic physical or psychological stress (O'Leary and Cryan, 2013). This can be done via the so-called Repeated Early Maternal Separation (REMS) procedure (Vetulani, 2013). The first days of an animal's life represent sensitive periods of development, as the experiences during this time shape cognitive, emotional, and physiological responses in adulthood (Kaffman and Meaney, 2007). Maternal care plays a critical role since it has a strong influence in the development of its offspring's molecular neurobiology. In REMS, pups are separated from their mothers over a prolonged period of time, exposing them to significant stress resulting in long-term neurobiological consequences, for example decreased glucocorticoid mRNA expression in the medial prefrontal cortex and alterations in the monoamine systems (Holmes et al., 2005; Sánchez et al., 2001). A widely used method of doing this is separating the pups for 3 hours daily from the second to twelfth postpartum day (Vetulani, 2013). Studies involving REMS have additionally shown that maternal sensory input increases oxytocin levels and stimulates the development of spatial learning and memory (Kojima et al., 2012; Liu et al., 1997). Chronic stress over a prolonged period has been shown to result in depression-like symptoms, such as behavioural despair, blunted reward processing and negative affective bias (Willner, 2017). 'Two-hit' models of depression include an additional stressor, shortly after maternal separation or in adulthood. They are based on the idea that animals that have experienced ELS are more vulnerable towards additional stressors and that behavioural differences are only observed after the

second stress (Murthy and Gould, 2018). In this thesis, results from a 'two-hit' model of depression will be presented to investigate depressive-like phenotypes in rats.

Despite the promising directions animal research may offer, difficulties in translation from rodent to human research must be acknowledged. The OFC, dlPFC, vlPFC and ACC are key regions implicated in MDD (Pizzagalli and Roberts, 2021). However, the dl- and vlPFC have no analogous regions in rodents, and for other regions that do have a homologous area in rodents, it is unclear whether functional analogy is present. Furthermore, many functionally analogous tasks for rodents and humans are in the process of being developed, as existing tasks do not always fully map behaviours from rodents to humans and vice versa. These factors should be kept in mind when translating findings between species.

1.5.2 Preclinical models of SUD

In recent years, several behavioural traits associated with predisposition to drug addiction have been identified (Meyer-Lindenberg and Weinberger, 2006). These endophenotypes include impulsivity and novelty/sensation-seeking, the latter being defined as the tendency to pursue novel and intense emotional experiences (Ersche et al., 2010; Nigg et al., 2006; Zuckerman and Neeb, 1979). In humans, investigating causal relationships between these predisposing factors and substance use disorder is difficult. However, animal models make it possible to investigate these behavioural traits and their neurobiological basis, which are present prior to compulsive drug-seeking (Jupp and Dalley, 2014). Both impulsivity and novelty/sensation-seeking have been linked to various SUDs, including cocaine, heroin and alcohol use disorder (Maremmani et al., 2009; Moeller et al., 2002; Petry, 2001). As described previously, impulsivity can be subdivided into impulsive action and impulsive choice, which can be measured by tasks such as the Go/No-Go task and delay discounting task, respectively. Increased impulsive action has been found in alcoholics, cocaine abusers and methamphetamine abusers (Dick et al., 2010; García-Marchena et al., 2018; Monterosso et al., 2001). Impulsive choice has also been observed in alcoholics and cocaine abusers, as well as nicotine addicts (Bickel et al., 1999; Claus et al., 2011; Simon et al., 2007). In rats, choice impulsivity can predict high rates of cocaine self-administration and has been linked to reduced D2/3 receptor availability in the nucleus accumbens (NAc) (Dalley et al., 2007). This also applies to nicotine and alcohol administration (Diergaarde et al., 2008; Radwanska and Kaczmarek, 2012). Similar observations have been made with regards to impulsive action, which can predict greater alcohol and nicotine administration (Diergaarde et al., 2008; Oberlin and Grahame, 2009). Novelty-seeking has been observed in individuals diagnosed with SUD, and can predict drug use risk (Noël et al., 2011; Stephenson and Helme, 2006). Furthermore, animals that are high responders on sensation-seeking tasks have been found to self-administer stimulant drugs more than low responders (Marinelli and White, 2000).

Compulsive cocaine-seeking can be investigated in rodents via a variety of tasks, which have helped to gain an insight into the neurobiological and neurochemical basis of drug addiction. Compulsive cocaine-seeking can be assessed by a response-contingent selfadministration of the drug, regardless of concurrent punishment (Jones et al., 2021). Variations of the self-administration paradigm exist, for example, drug delivery and foot-shocks can be concurrently delivered on a fixed-ratio 5 schedule of reinforcement, with an additional foot-shock given on the trial prior to drug delivery (Acosta et al., 2008). Alternatively, second-order schedules of reinforcement can be used, in which the reinforcer is paired with or without a foot-shock (Kelleher, 1958). Additionally, seeking-taking chain schedules can be used to assess compulsive drug-seeking, as they require animals to press a 'seeking' lever in order to gain access to a drug 'taking' lever, which results in the delivery of a drug. A mild foot-shock stress is introduced 50% of the time after the 'seeking' lever is selected (Pelloux et al., 2012). This is thought to reflect the uncertainty experienced during drug taking in humans. Various other paradigms have been developed for the purpose of investigating drug-seeking and drug-taking behaviours. In the following chapters, attention will mostly be focused on the fixed-interval (FI) 15 schedule of reinforcement.

1.6 Reinforcement Learning in Neuroscience and Psychiatry

In recent years, the use of computational modelling to support the study of neuroscience has become more popular, as computational models allow us to gain a new insight into latent mechanisms involved in processes such as reward learning and decision-making. A key issue in learning is how behaviour changes on a trial-by-trial basis in response to feedback, which is difficult to investigate using methods that do not involve computational modelling (Daw, 2009). These trial-by-trial computational analyses allow researchers to gain a more realistic and dynamic picture of the learning processes involved. In reward learning tasks such as bandit or reversal learning tasks, subjects select one of multiple available choices and get rewarded or punished for their response. When fitting a computational model, this information is collected on every trial, and is then fed into an algorithm, for example a reinforcement learning (RL) model. This model calculates the probability of selecting one

choice on the next trial based on previous choices and outcomes. RL is the process of learning how to act in order to achieve a goal, by taking into account previous rewards and losses with the aim of maximising overall rewards and minimising overall losses (Sutton and Barto, 1998). There are two types of RL: model-based and model-free. The former assumes an explicit model of the environment and the agent in that environment and describes the consequences of actions and outcomes in order to infer optimal policies (Huys et al., 2014). Model-free approaches do not include any explicit knowledge of the environmental dynamics or the outcomes of their actions, but instead evaluate their actions through trial-and-error learning. Through fitting model-free RL models, it is also possible to determine a number of free parameters for each subject independently. These parameters can include α , the learning rate, which reflects how quickly the subject learns from previous trials; β , a parameter representative of the animal exploiting previous information versus exploring other choices; and κ , the autocorrelation, also known as the 'stickiness', reflective of the subject's tendency to repeat a choice regardless of previous outcomes. These parameters frequently form the basis of such models, although there are other parameters that can be added or used to replace one of the aforementioned measures. The parameters included generally depend on the study design and the hypotheses set up by the experimenter, thus allowing flexibility to suit task demands.

Conventional analyses calculate performance metrics based on an average across all trials, without taking into account trial-by-trial learning dynamics. Therefore, neuroscientific and psychological tasks can greatly benefit from computational modelling. Moreover, these approaches make it possible to link the behavioural processes to neural responses such as the blood-oxygen-level-dependent (BOLD) signal. One example of the application of RL models to behavioural data and linking it to neural signals is from rats; multiple RL algorithms were compared to identify the model that could predict choice sequences generated on a conditional free-choice task the best, and the winning model was linked to neural activity in the NAc and ventral pallidum (VP) (Ito and Doya, 2009). The study found that the main role of the NAc and VP is to monitor information required for updating choice behaviours. In an additional study, neurons in the dIPFC were found to represent a monkey's past decisions on a free-choice task and reflected signals necessary to update expected reward estimates (Barraclough et al., 2004). These studies demonstrate the importance of computational modelling in neuroscience and how it can help gain a better understanding of decision-making and learning.

The phasic firing of dopaminergic midbrain neurons has been linked to the prediction error, a key component of the model-free RL update equation (Schultz and Romo, 1990).

Prediction errors represent the difference between expected and observed outcome. The magnitude of a positive prediction error is reflected in the magnitude of the phasic dopamineneuron bursts (Schultz et al., 1997). Excessive DA signalling is thought to result in a shift from model-based to model-free decision-making and involves cortico-basal ganglia-thalamocortical (CBGTC) loops (Huys et al., 2014; Maia and Frank, 2011; Robbins, 2012). This idea is relevant for the study of psychiatric disorders, as such a shift may underlie the transition from occasional recreational drug use to compulsive drug use, also reflecting a shift from goal-directed to habit-based behaviour (Huys et al., 2014). Disturbances of the dopaminergic system and CBGTC circuits also play an important role in other psychiatric and neurological conditions. For example, an imbalance between goal-directed and habit-based behaviour has been proposed for OCD and has been linked to dysfunction of orbitofrontal-striatal circuitry (Gillan et al., 2011; Gillan et al., 2014). ADHD is thought to involve a hypofunctioning dopaminergic system, reflecting excessive discounting of delayed rewards and executive dysfunction, which are prominent in this disorder (Frank et al., 2007; Sonuga-Barke, 2005). Discounting reflects the degree to which the value of future reinforcers declines relative to immediate ones. Tonic DA in the VS, which tracks changing reward values, is proposed to mediate the discount factor and reduced tonic DA in this area may explain the smaller discount factor in ADHD (Smith et al., 2005; Wang et al., 2021). Additionally, schizophrenia involves excessive DA signalling in the striatum and reduced DA in the PFC (Guillin et al., 2007). One theory of schizophrenia proposes that psychosis arises due to abnormal prediction errors, which are mediated by phasic DA signalling, resulting in inappropriate associations, causal attributions, and attentional salience (Corlett et al., 2007). In summary, computational RL accounts can help to build models of brain dysfunction in neurological and psychiatric disorders and gain additional insights into these diseases.

1.6.1 Neural basis of RL

Regions involved in RL are widespread across the brain. Neural signals reflecting reward expectancy can be separated into those reflecting action value and those that reflect state value (Lee et al., 2012). Neural activity related to action value has been found in the posterior parietal cortex (PPC), dlPFC, premotor cortex, mPFC and striatum (Kim et al., 2012; Lau and Glimcher, 2008; Pastor-Bernier and Cisek, 2011; Platt and Glimcher, 1999; Seo and Lee, 2007). During binary-choice tasks, the sum of action value functions for two choices represents the state value function (Lee et al., 2012). Signals for state and action values co-exist in the PPC and DS (Cai et al., 2011; Seo et al., 2012). Neurons solely representative

of state value are located in the VS, ACC, and amygdala (Belova et al., 2008; Cai et al., 2011; Seo and Lee, 2007). Post-decision state value functions have been associated with activity in the OFC, mPFC, dlPFC and striatum (Kim and Lee, 2011; Lau and Glimcher, 2008; Sul et al., 2011). Multiple areas are also implicated in the neural mechanisms for updating value functions, as would be required during a reversal learning task. Animals whose OFC is lesioned struggle to update their choices when reward contingencies are switched on such tasks (Fellows and Farah, 2003; Iversen and Mishkin, 1970). Neural signals to animal's previous choices have been recorded in the PFC and PPC in monkeys as well as in the frontal cortex and striatum of rodents (Barraclough et al., 2004; Fecteau and Munoz, 2003; Kim et al., 2007). Midbrain DA neurons were the first area reflecting prediction errors to be found (Schultz et al., 1997). However, other areas have been implicated since, such as the lateral habenula, globus pallidus, dlPFC, ACC, OFC, and striatum (Kim et al., 2009; Seo and Lee, 2007; Sul et al., 2010). Signals representative of chosen values are also observed in the mPFC, OFC and striatum (Kim et al., 2009; Lau and Glimcher, 2008; Sul et al., 2010). Therefore, is seems likely that the OFC and striatum, where both reward prediction errors and the chosen value signals co-exist, are important for updating value functions (Lee et al., 2012).

Not much is known about the neural basis of RL parameters such as α , β and κ , although a few studies in animals have identified potential mechanisms mediating these measures. The learning rate α is adjusted based on the level of uncertainty in the subject's immediate environment (Lee et al., 2012). Highly volatile environments require large learning rates in order to update value functions quickly, and they should be low when the environment is mostly stable. The level of volatility has been shown to be represented by activity in the ACC, suggesting that the learning rate is also modulated by this area (Behrens et al., 2007). The lateral PFC, OFC and amygdala show increased activity when making decisions under uncertainty, and therefore may also be involved in updating the learning rate (Hsu et al., 2005; Huettel et al., 2006). In rodents, reward learning rates are reduced following inactivation of the prelimbic cortex (PrL) and lateral OFC (IOFC) (Verharen et al., 2020). Inactivation of the PrL and infralimbic cortex (IL), as well as the lateral and medial OFC impacted punishment learning rates. Stimulus stickiness κ was reduced following inactivation of the IL and mOFC; β was unaffected by any inactivations of prefrontal cortical subregions. Therefore, it appears that the OFC, amygdala and some PFC regions are key to reward and punishment learning, and the mOFC and PFC areas are necessary for modulating κ . However, the exact circuits involved remain unknown and require further investigation.

1.6.2 RL in MDD

As mentioned previously, MDD is not uniquely a mood disorder, but also a disorder associated with dysfunctional decision-making and cognitive function (Gotlib and Joormann, 2010; Must et al., 2013). More recently, is has become evident that dysfunctional reinforcement learning is a key feature of MDD (Chen et al., 2015). This has been demonstrated by poor performance on the IGT, which relies on RL (Must et al., 2013). Performance on signal-detection tasks is also impaired, as MDD patients tend to know show a bias towards rewarded choices, which is not present in healthy individuals (Pizzagalli, 2014). The striatum and mPFC are both involved in RL, and these regions have shown altered responsivity in MDD during monetary anticipation and monetary outcomes (Steele et al., 2007; Zhang et al., 2013). Generally, MDD has been associated with hyposensitivity to reward and hypersensitivity to punishment, representing an imbalance between reward and punishment learning (Dombrovski et al., 2013; Steele et al., 2007). Neurally, MDD patients unresponsive to antidepressant treatments show a decreased reward prediction error (RPE) signal in the VS and dorsal ACC, as well as an increased RPE signal in the ventral tegmental area, rostral ACC, hippocampus and retrosplenial cortex (Gradin et al., 2011). Reduced signal following RPE in the striatum and midbrain was observed to be positively correlated with anhedonia. Moreover, MDD patients demonstrate a reduced signal when tracking reinforcement expected valued in the hippocampus and parahippocampal gyrus. The literature on how the reward and punishment learning rates are altered in MDD is still inconclusive, with some reports suggesting decreased reward and punishment learning rates, and some studies finding no effects (Chase et al., 2010; Gradin et al., 2011; Mukherjee et al., 2020). In summary, it is evident that RL processes are significantly impaired in depression, even though the exact mechanisms and effects are not entirely understood.

1.6.3 RL in SUD

It is thought that patients with SUDs have altered incentive salience to reward, which is related to higher reward properties of drugs due to alterations in mesolimbic DA signalling (Robinson and Berridge, 1993; Volkow et al., 2004). Due to the hyperactivity of the reward circuit, motivation and memory circuits are overactive and result in decision-making impairments. Reward circuits, which include the ventral tegmental area, amygdala, and OFC, are known to be permanently changed due to substance abuse (Koob and Le Moal, 2001; Volkow et al., 2004). Electrophysiological studies have highlighted deficits in salience detection in drug abusers. For example, event-related potentials related to DA production and

evaluation of salient information is increased in alcoholics, smokers, methadone-dependent individuals, and cocaine addicts (Iacono et al., 2002; Polich, 2007; Warren and McDonough, 1999). This signal is enhanced in response to drug-related cues but is reduced when nondrug-related cues are presented (Baskin-Sommers and Foti, 2015). Computational modelling of reversal learning task data has revealed increased stimulus stickiness κ_{stim} in SUD, as well as reduced reward learning rates (Kanen et al., 2019). The DA D2/3 receptor agonist pramipexole and DA D2/D3 receptor antagonist amisulpiride reduced stickiness and increased the reward learning rate, suggesting that these mechanisms are in part mediated by the DA system. Overall, it is observed that the reward systems and RL mechanisms are substantially altered in SUD, highlighting the importance of understanding RL in this patient sample better.

1.7 Thesis hypotheses and objectives

Broadly, this thesis will investigate how behavioural RL processes are altered in different psychiatric disorders, with a particular focus on depression and substance use disorder. Moreover, the neural basis underlying behavioural RL changes will be explored using neuroimaging techniques, and a novel insight into cognitive flexibility will be given via a newly developed behavioural task. Given the importance of RL, cognitive flexibility and decisionmaking in psychiatry and the lack of knowledge of the neural correlates underlying these processes, this thesis aims to contribute to our understanding of these topics. Computational modelling and neuroimaging approaches are used across two species: rats and humans. This project has benefitted from data availability from multiple sources in these two species, making it possible to translate these findings between them and build novel translational bridges. A common theme throughout this thesis will be the use of tasks measuring cognitive flexibility. In the first four chapters, the reversal learning task will be used. Data from this paradigm will be analysed using RL models in these four chapters. Additionally, the first three chapters contain structural and/or functional neuroimaging data, enabling the exploration of the neural basis of extracted RL parameters. In the last results chapter, a novel flexibility task will be presented. This will also include the analysis of both behavioural and task-based fMRI data, once again making it possible to link the two together.

In chapter 3, a number of important questions in relation to the REMS paradigm in rodents will be addressed. Data from a longitudinal study involving behavioural testing and neuroimaging acquisition throughout development is explored. Half of the rat cohort was exposed to REMS in adolescence, and all rats underwent repeated shock stress in adulthood.

The behavioural effects of REMS and repeated shock stress are investigated, with a particular focus on the probabilistic reversal learning task. Data from this task is analysed using a variety of models, including RL, Bayesian Learner and Drift Diffusion Models. Structural and functional MRI scans are compared between the different groups in order to identify differences in brain structure and connectivity arising from these interventions, as well as sex differences. To the best of my knowledge, this is a unique dataset as it consists of different data types, allowing the exploration of the link between RL parameters and functional connectivity in rats. This provides a unique insight into the neural processes underlying RL parameters and how they are affected as a result of early life and adulthood stress. Moreover, since only few previous studies included both male and female rats, this enables the exploration of sexual dimorphism in relation to stress and depressive-like phenotypes, which is highly relevant to the study of MDD. It was hypothesised that the reward learning rate would be decreased and the punishment learning rate would be increased in both male and female rats resulting from REMS and adulthood stress, since it is thought that MDD patients have an altered balance in reward processing, with hyposensitivity to reward and hypersensitivity to punishment. ELS and chronic stress both lead to alterations in these parameters, and it was expected that they would have an additive effect. Moreover, it was predicted that females would have stronger impairments in these two parameters resulting from the interventions, since they are more susceptible to the effects of stress (Kendler and Gardner, 2014; Kessler, 2003). Neurally, it was hypothesised that the amygdala would show altered connectivity to cortical areas, specifically the mPFC, since this has previously been observed in patients with MDD, as well as rats and humans that have been exposed to ELS. A further hypothesis is that the OFC would be associated with RL parameters, particularly reinforcement sensitivity, due to its involvement in valuation and outcome estimation.

In chapter 4, another dataset from rodents is presented. This study is also longitudinal, involving behavioural testing and brain scans throughout adolescence and adulthood prior to a cocaine self-administration paradigm. These data allow the investigation of the following questions: 1) what are the behavioural changes that define low-, medium- and high-compulsive rats, and could any behavioural markers be used to identify predisposition to high-compulsive drug-taking?, 2) are there any structural brain markers of high-compulsive animals?, 3) how is RL affected by compulsivity?, and 4) can structural changes be linked to RL parameter differences? Based on previous studies, it was hypothesised that highcompulsive animals would have greater values of the stickiness parameter κ , which would also be reflected in increased perseverative responding as extracted using conventional analyses. Structurally, volumetric differences in the ventral striatum were expected since this is a

key area for reward processing and has been shown to be altered in disorders of addiction. Similarly to the previous chapter, it was expected that RL parameters would be correlated with OFC volume.

The subsequent chapters present analyses from human data. Chapter 5 includes data from a study involving behavioural testing using the PRL task in an fMRI scanner. The participants in this study include control subjects, as well as individuals with cocaine use disorder and gambling disorder. This makes it possible to investigate the behavioural processes underlying these two compulsive disorders in humans and relate the observations to those reported in rodents in chapter 4. Further, through the task-based fMRI data, the behavioural processes can be linked to timeseries BOLD data, identifying any changes in functional connectivity between the different groups and discovering which brain mechanisms underlie RL parameters. Similarly to rats, it was expected that the GD and CUD groups would have higher values of stickiness, regardless of whether it was stimulus- or side-dependent, as well as increased perseverative responding as measured conventionally. It was hypothesised that increased value-free responding would be reflected via altered activity in the dorsal striatum and mPFC, which from part of the putative habit-based system.

Chapter 6 includes data from a naturalistic sample of patients with SUD and/or MDD as well as healthy individuals. It contains behavioural and clinical data from each of the participants. This allows the exploration of RL in a human sample of patients with MDD and makes it possible to link the findings made in rodents to human data and find out whether these are translatable. The hypothesis was that participants with MDD would have reduced reward and increased punishment learning rates, whereas participants with SUD would have greater κ values. On conventional measures, it was expected that the SUD group would exhibit greater perseveration and the MDD group would show lower win-stay and greater lose-shift behaviours. Patients with both diagnoses were expected to have a combination of altered learning rates and increased stickiness rates.

The final results chapter presents a novel behavioural task that measures a different subtype of proactive cognitive flexibility, specifically, how healthy participants make decisions in the face of uncertainty and whether they shift their response when they are given the opportunity to repeat their choice. This task reflects greater ecological validity than other commonly used tasks measuring cognitive flexibility. Task-based fMRI data was collected from these participants whilst they were completing the task. The main questions addressed in this chapter are: what are the behavioural mechanisms underlying decision-making when information in the environment is uncertain? Do they rely more on their own choices that were taken on previous trials or on the feedback they are presented with? What are the regions and circuits involved in shifting versus repeating a choice? These results can provide a novel perspective on cognitive flexibility, which is relevant for the study of psychiatric disorders. The following hypotheses were set up: 1) participants would shift more after their response was incorrect and feedback was negative and repeat their choice if their response was correct and feedback was positive; 2) response shifting was expected to be linked to activity in the ACC and SFG, due to their involvement in error monitoring, the anterior insula, as it is important for salience detection and the striatum, because it is required for behavioural inhibition and adjustments; 3) the task would measure a thus far unexplored aspect of cognitive flexibility, which would be accompanied by partially overlapping but also diverging neural substrates. It was expected that there would be more frontopolar activations due to the volitional nature of the task.

In summary, this thesis will aim to uncover the behavioural and neural changes associated with cognitive flexibility, with a particular focus on RL, across depressive disorders and disorders of compulsivity. Value-free behaviours are a key topic throughout the thesis, as it can be represented by the κ parameter in RL and can be separately studied through the novel task that will be presented. Specifically, it will be investigated how value-free responding is altered due to various interventions in rats and different diagnoses in humans. This will be achieved using a variety of neuroscientific methods, including computational modelling, neuroimaging and translating findings from rodents to humans.

Chapter 2

General methods

2.1 Rats

2.1.1 Reinforcement learning modelling

A Q-learning model is a model-free type of Reinforcement Learning (RL) algorithm (Daw, 2009). Different variations of the Q-learning model were fit to reversal learning task data from rats. The inputs to the model were the choice the animal had taken on each trial (either left or right stimulus), and whether the animal received a reward or not. In the two studies presented here, the rodent paradigm included either two white square stimuli presented on a touchscreen or two levers. Therefore, the animals had to learn to discriminate based on location, rather than discriminate between two different stimuli.

The Q-learning models fitted to the data were based on the following equation:

$$Q_{t+1}(c_t) = Q_t(c_t) + \alpha * (r - Q_t(c_t))$$
(2.1)

Where $Q_{t+1}(c_t)$ is the expected value of the stimulus on the next trial, $Q_t(c_t)$ is the previously expected value of the choice taken on the current trial (c_t) , α is the learning rate and r is the reward received on the current trial. $r - Q_t(c_t)$ is also known as the prediction error. Higher values of α represent faster adjustment of a stimulus' Q-value based on the prediction error. The Q-value is updated on a trial-by-trial basis.

The probability of choosing a stimulus given the Q-values for each option is calculated using the softmax decision rule:

$$P(c_t = L | Q_t(L), Q_t(R)) = \frac{exp(Q_t(L)/\beta + \kappa * L_{t-1})}{exp(Q_t(L)/\beta + \kappa * L_{t-1}) + exp(Q_t(R)/\beta + \kappa * R_{t-1})}$$
(2.2)

General methods

Where $Q_t(L)$ and $Q_t(R)$ are the Q-values of the left and right stimuli, respectively. In this example, the probability of choosing the stimulus on the left side is determined. β is the inverse temperature parameter, which can be interpreted as whether the animal exploits the information it has gained on the current trial (low β), or whether it explores the other options (high β) (Zhukovsky et al., 2019). β can also be interpreted as the value or reinforcement sensitivity (Huys et al., 2013). The α parameter was set to take values between 0.001 and 1, whereas the β parameter could take values between 0.005 and 5 (Daw, 2009). κ is the autocorrelation parameter, also known as 'stickiness' and represents the tendency to select the same choice regardless of reinforcement outcome of a given stimulus on previous trials. The step sizes for the α parameter were set to 0.05, whereas for β they were set to 0.15. L_{t-1} and R_{t-1} indicate whether the left or right stimulus was selected on the previous trial and take values of 1 or 0, depending on whether that side was chosen or not. κ could take values between -1 and 1 and had a step size 0.1. A grid-search algorithm tested all the possible parameter combinations for each animal and each session in order to determine the optimal parameter combination for the given set of trials in a session. The step size determines which values within the constrained parameter space will be used in the grid search.

Next, the probability of observing the sequence of actions in data D as observed in the reversal learning task given the model M and parameters θ is calculated using the product of probabilities:

$$P(Data \ D \mid Model \ M, \ parameters \ \theta) = \prod P(c_t = L | Q_t(L), Q_t(R))$$
(2.3)

After this, the free parameters are fit to achieve the maximum likelihood of the probability density function of data *D*:

$$\operatorname{argmax} P(D \mid M, \theta) \tag{2.4}$$

There are multiple variations of this model, five of which were tested.

- 1. 3-parameter model as described above, including α , β and κ parameters.
- 2-parameter model with no κ term
 This model is the same as described above but without a κ parameter.
- 3. 4-parameter model with two α terms This model introduces two separate learning terms for α , one for rewarded trials (α_{rew})
and one for non-rewarded trials ($\alpha_{non-rew}$). The Q-value is updated as seen in equation (2.1). The learning rate representing either reward or non-reward is chosen based on whether the action was rewarded or not.

4. 3-parameter model with a 'forgetting' mechanism and reward/aversion sensitivity In this variation of the model, $Q(c_t)$, the value of the chosen side, is updated similarly to equation (2.1):

$$Q_{t+1}(c_t) = Q_t(c_t) + \alpha * (\mu_1 * r - Q_t(c_t))$$
(2.5)

However, the Q value of the non-chosen stimulus in this model is also updated according to equation (2.6):

$$Q_{t+1}(c_t) = Q_t(c_t) - \alpha * (\mu_2 * r - Q_t(c_t))$$
(2.6)

This introduces a 'forgetting' mechanism for the non-chosen stimulus, in which α is used to update the Q-value of the chosen stimulus is equal to the α in the Q-value function of the non-chosen stimulus (Barraclough et al., 2004; Ito and Doya, 2009). This model does not include the β and κ parameters in the softmax function, but instead has two separate terms, reward sensitivity (μ_1) and aversion sensitivity (μ_2) in the update equation. This model was included because it has been reported to provide excellent model fit due to the additional 'forgetting' term.

5. 4-parameter model with a 'differential forgetting' mechanism

This model is the same as model number 4 containing a 'forgetting' term, except that in this variation, there is a separate α term for the Q-value function of the non-chosen stimulus, the α_{for} parameter.

$$Q_{t+1}(c_t) = Q_t(c_t) - \alpha_{for} * (Q_t(c_t))$$
(2.7)

These models were implemented in MATLAB R2019b (Mathworks) using the Q-learning code for a two-choice reversal learning task from (Zhukovsky, 2015b) as a template.

2.1.2 Bayesian Learner modelling

A further method of modelling how an animal's expectations are updated is with an update rule based on Bayes' Theorem (den Ouden and O'Reilly, 2015; Zhukovsky, 2015a).

General methods

At the beginning of each set of trials, it is assumed that the prior distribution function $P(q_t)$ is uniform (y=0.5). On every trial, the posterior distribution for each stimulus is updated according to the following equation:

$$P(Q_t(L) \mid S_t) = ((1-H) * P(S_t \mid Q_t(L))) * P(q_t) + (H * P(q))$$
(2.8)

Where S_t is the reward sequence up to the current trial, H is the Hazard rate, $P(S_t | Q_t(L))$ is the likelihood of the reward sequence given the Q-values for the stimulus, $P(q_t)$ is the prior distribution, which is also equivalent to the posterior from the previous trial due to a Markovian dependency, and P(q) is the uniform prior distribution from the beginning of the session. The prior distribution can be defined as the animal's reward expectation at the beginning of the trial, whereas the posterior distribution is the reward expectation at the end of the trial. The probability of choice c_t on the next trial is then determined using the softmax function as in equation (2.2), which includes the free parameters β and κ . The third free parameter in this model is the Hazard rate (H), and it determines how much the posterior changes based on the data-driven likelihood versus how much it depends on the uniform prior distribution (Zhukovsky, 2015a). A low hazard value indicates that the rat learns more from outcomes on previous trials than the prior distribution. This means that the rat adjusts more quickly to contingency changes, however, it also makes its choices less robust in response to random perturbations. The subsequent steps are identical to the ones described in the previous section for the Q-learning model.

Two additional variations of the Bayesian Learner (BL) model were fit to the data:

1. Varying memory size

It was possible to explore the effect of changing the number of trials the model has access to when calculating the likelihood, by increasing the memory size parameter M. In the instance where M=1, the likelihood distribution follows a Bernoulli distribution. Otherwise, it is reflected in a binomial distribution. Memory Sizes between 1 and 10 were tested to identify whether this improved model performance.

2. Flexible priors

In this model, flexible priors were employed to see whether this improved model fit. Compared to models one and two, where the prior distribution was uniform, in this model the prior distribution was biased towards either stimulus, which could be implemented using a beta function. This function has a higher probability density on either side of the function, reflecting a bias towards the right or the left. These models were implemented in MATLAB R2019b (Mathworks) using the Bayesian Learner code for a two-choice reversal learning task from (Zhukovsky, 2015a) as a template.

2.1.3 Model comparison

Model comparison of the models that were fit to the data was done using the following three measures:

1. Log-likelihood Ratio (d)

$$d = 2 * [log P (D | M_2, \hat{\theta}_{M_2}) - log P (D | M_1, \hat{\theta}_{M_1})]$$
(2.9)

Where M1 and M2 are the two models being compared, D is the sequence of data points, and $\hat{\theta}$ are the best fit parameters for each respective model.

 $P(D|M_1, \hat{\theta}_{M_1})$ represents the probability of data sequence *D* given the best fit parameters for the model. This measure compares two models given the respective likelihood of data sequence *D* and the best fit parameters for that model (Wu and Rao, 2006).

2. pseudo r^2

pseudo
$$r^2 = \frac{\log P(D \mid M, \hat{\theta}_M) - 0.5^n}{0.5^n}$$
 (2.10)

where *n* represents the number of trials. This measure contrasts the probability of observing the data with the chosen parameters to observing the data at random. It does not penalise overfitting (Cameron and Windmeijer, 1997).

3. Bayesian Information Criterion (BIC)

$$BIC = \log P(D|M)) \approx \log P(D|M, \hat{\theta}_M) - \frac{n}{2}\log m \qquad (2.11)$$

In this equation, n is the number of free parameters, and m represents the number of observations or in our case, trials. The BIC penalises large numbers of parameters in order to combat overfitting (Draper and Smith, 1981).

General methods

Based on these measures, the best models were selected for further analysis. All three measures were used to identify the best fitting model. The higher the value of the log-likelihood, the better the model fits a dataset. The significance threshold for comparing two models is 3.842 for p=0.05. Therefore, if the log-likelihood ratio crosses this threshold, the model with the higher log likelihood value is better fitting. Higher pseudo r^2 values also indicate better model fit, although no significance threshold needs to be reached. When using the BIC for model comparison, the lower BIC value indicates better model fit, unlike the other two measures. These values are calculated for every subject and for every session the subject completed, and the percentage of sessions/animals which favour one model over the animal is calculated. The model with the highest percentage of sessions/animals is selected as the winning model.

2.1.4 Drift Diffusion Modelling

The Drift Diffusion Model (DDM) can be used to model the cognitive processes involved in a two-choice task, such as a reversal learning task (Ratcliff and McKoon, 2008). It is a model of sequential sampling with diffusion signals, in which the subject accumulates evidence until a boundary is hit (Fudenberg et al., 2020). Once the boundary is hit, the choice corresponding to it is chosen. This process is shown in Fig 2.1. The starting point is labelled z, which we will refer to as β and the boundary separation is a, which we will label α , respectively. The information accumulation rate is known as the drift rate v or δ . There is also a parameter representing the non-decision time t_0 (τ). This model requires information on which stimulus the subject has taken on each trial and the reaction time for that trial. The data was modelled using the HDDM python toolbox, which uses hierarchical Bayesian parameter estimation of the DDM (Wiecki et al., 2013). A more detailed explanation of hierarchical Bayesian estimation of parameters can be found in section 2.2.1.



Fig. 2.1 The Drift Diffusion Model.

A graphical representation of the DDM, including the starting point z, boundary separation a and drift rate v. The non-decision time is also shown. From (Murata et al., 2014).

2.1.5 Conventional analysis of reversal learning task data

The study designs and detailed information on data collection are described in the respective chapters, as data from different studies were used. Conventional reversal learning measures, such as win-stay and lose-shift behaviour, were calculated for the reversal learning data from these studies. This was done for both the deterministic and probabilistic versions of this task. For the probabilistic reversal learning task, win-stay and lose-shift measures were calculated for previously correct and incorrect trials, separately (i.e., win-stay after a previously correct trial, lose-shift after a previously correct trial, win-stay after a previously incorrect trial and lose-shift after a previously incorrect trial). Additionally, the number of trials to reach a reversal (trials to criterion), the proportion of correct responses and number of perseverative responses were calculated. The number of perseverative responses were calculated as the percentage of trials during which the animal selected the same stimulus as on the previous trial if the previous trial was rewarded. Lose-shift, on the

General methods

other hand, was the percentage of trials during which the animal shifted its response to the other stimulus after not receiving a reward on the last trial.

2.1.6 Rodent image pre-processing

First, the structural and functional scans had to be manually aligned with a template image in standard orientation. Next, magnetisation transfer (MT), proton density (PD) and T1-weighted structural scans were bias-corrected and rigidly registered using SPMmouse (Sawiak et al., 2013), which runs via the Statistical Parametric Mapping software (SPM8) (Friston et al., 1994). Study-specific templates for each time point were generated using Diffeomorphic Anatomical Registration using Exponential Lie algebra (DARTEL) in SPMmouse (Ashburner, 2007). MT scans were selected for further analysis and fMRI registration, as they had the highest contrast and best quality.

The resting-state fMRI images were pre-processed using the FMRIB Software Library (FSL) and the Analysis of Functional NeuroImages software (AFNI) (Cox, 1996; Smith et al., 2004). As a multi-echo gradient echo sequence with three echoes was used, every three volumes were averaged and the scans were subsequently bias-corrected in FSL (*fast*) (Zhang et al., 2001). Next, the functional scans were slice-timing corrected (*slicetimer*, FSL) and despiked (*3ddespike*, AFNI) (Cox, 1996; Smith et al., 2004). The first 5 volumes were discarded to only include steady-state volumes. The scans then underwent motion correction (*3dvolreg*, AFNI) (Cox, 1996). Motion outliers were identified using dvars as a metric (*fsl_motion_outliers*) (Smith et al., 2004). Dvars is defined as the spatial standard deviation of successive difference images (Afyouni and Nichols, 2018). The volumes in which motion exceeded the box-plot cut-off were censored, however, this was only the case for 1% of total volumes (Smith et al., 2004). The voxel sizes of functional scans were then multiplied by 10 to be comparable to human voxel dimensions, which is required for FSL registration tools to work in the intended manner (Bajic et al., 2017).

Functional scans were co-registered to structural scans using 6 DOF affine registration (*FLIRT*, FSL) (Jenkinson and Smith, 2001). MT structural scans were non-linearly registered to the study-specific template (*FNIRT*, FSL) (Smith et al., 2004). The functional scans were warped into common template space using the warp field computed by FNIRT. This was done after applying the pre-transform using the affine matrix from FLIRT (*applywarp*) (Smith et al., 2004). Resulting images were high-pass filtered (100s) and smoothed using a full width at half maximum (FWHM) of 1.7 mm (*fslmaths*, FSL) (Smith et al., 2004). Quality checks

were conducted by assessing the estimated mean displacement and a visualising registration efficacy. Examples of how the quality checks were conducted can be found in the Appendix.

2.1.7 Structural analysis

Following normalisation of MT images to the study-specific brain template, segmentation of cortical and subcortical structures was achieved using the rat brain atlas, which contains the masks for all rat brain structures (Valdés-Hernández et al., 2011). The volumes of all structures from the atlas were calculated for each animal at each time point (*fslstats*, FSL) (Smith et al., 2004). The volumes for each ROI and each animal were adjusted for the total brain volume.

2.2 Humans

2.2.1 Reinforcement learning modelling using a hierarchical Bayesian approach

Seven different RL models using a hierarchical Bayesian approach were fitted to the human probabilistic reversal learning task data. All trials from the PRL tasks were used by the models and information on the subject ID, patient group, stimulus location, type of stimulus, response of the subject and whether the subject received a reward or not was also included. The hierarchical Bayesian inference was implemented using Stan (Stan Development Team, 2022). The top level of the hierarchy included a group-specific distribution for each RL parameter. In the next level, the RL parameters for each subject were drawn from a normal distribution that had a mean of the group from the highest level and whose variance represents inter-subject variability for that parameter. This allows the algorithm to select a combination of RL parameters for a set of trials, which in turn are used to govern an RL model trained by the sequence of stimuli and reinforcement. The highest density interval (HDI) was determined, which is the narrowest interval that contains a certain percentage of the sample values (e.g., 95%) (Kruschke, 2014).

The seven models explored were the following:

1. This model had two parameters: α and β . The Q-value was updated as shown in equation (2.1), which was then used to calculate the probability of each choice using the softmax decision rule, as in equation (2.2).

General methods

- 2. This model was similar to model 1, however, it also included a stimulus stickiness parameter κ_{stim} . This parameter represents the tendency to respond to the same stimulus as on the previous trial, irrespective of its location and of outcome (i.e. whether it was rewarded or not).
- 3. This model was similar to model 1, but it contained two learning rates, one for rewarded trials α_{rew} and one for non-rewarded, or punished, trials $\alpha_{non-rew}$.
- 4. Additionally to the parameters described in model 3, this model also contained a side stickiness parameter κ_{side} , which is the tendency to choose the same side as on the previous trial when responding.
- 5. As in model 4, there was a stickiness parameter, however, side stickiness was replaced with stimulus stickiness κ_{stim} .
- 6. In this model, both side and stimulus stickiness, κ_{side} and κ_{stim} , were added, resulting in five parameters.
- 7. The Experience-Weighted Attractor Model (EWA) (Camerer and Ho, 1999) This model compares the value of incoming information against one's beliefs. The model contains an experience weight for each stimulus, which modulates learning from reinforcement and is updated whenever that stimulus is chosen. Over time, it changes according to a decay factor. Similarly to previous models, there is still a β parameter in the softmax function. This model makes it possible for the learning rate to change over time.

The hierarchical Bayesian method was applied to the human data as there has been a move towards using this approach in the literature, rather than using a grid-search approach. However, this modelling approach does not allow an in-depth analysis of how parameters change **across** sessions. As this is not relevant for human experiments, but very important in rodent experiments, this method was only used for human data.

Table 2.1 shows a summary of the priors for the respective parameters. Model fitting was done using Hamiltonian Markov Chain Monte Carlo sampling via Stan 2.17.2 (Carpenter et al., 2017). The potential scale reduction factor measure \hat{R} was using to ensure convergence (Brooks and Gelman, 1998; Gelman, 2013). When the \hat{R} is close to 1 it indicates perfect convergence. A value of below 1.2 is used as a cut-off for convergence, and a cut-off of 1.1 is used as a stringent criterion for convergence (Brooks and Gelman, 1998). A cut-off of 1.1 was selected.

Model comparison was done with a bridge sampling estimate of the marginal likelihood via the "bridgesampling" R package (Gronau et al., 2017; Gronau et al., 2020). This method estimates the marginal likelihood and posterior probability of each model based on the data and prior model probabilities.

The computational modelling was done with the help of Dr Rudolf Cardinal and his scripts. The procedure used is similar to what is described in (Kanen et al., 2019).

	Prior	Reference
Reward learning rate, α_{rew}	Beta(1.2, 1.2)	den Ouden et al., 2013
Punishment learning rate, α_{pun}	Beta(1.2, 1.2)	den Ouden et al., 2013
Combined learning rate α	Beta(1.2, 1.2)	den Ouden et al., 2013
Reinforcement sensitivity, β	Gamma(α =4.82, β =0.88)	Gershman, 2016
Side stickiness, κ_{side}	Normal(0,1)	Christakou et al., 2013
Stimulus stickiness, K _{stim}	Normal(0,1)	Christakou et al., 2013
Experience decay factor, ρ	Beta(1.2, 1.2)	den Ouden et al., 2013
Decay factor for previous payoffs, ϕ	Beta(1.2, 1.2)	den Ouden et al., 2013
Softmax inverse temperature, β	Gamma(α =4.82, β =0.88)	Gershman, 2016

Table 2.1 Priors for model parameters

2.2.2 Simulation of data using model parameters

In order to find out whether the winning model was capable of reproducing the behavioural observations, the data were simulated using the posterior group mean parameters from the winning model, followed by a conventional PRL analysis as described in section 2.1.5. For every group, data for 100 participants was simulated. Inter-subject variability in parameters was not included in the simulation, as the aim of this analysis was to figure out whether the winning model is sufficient to reproduce the behavioural measures on a group level.

2.2.3 Reinforcement learning-drift diffusion model

To analyse choice dynamics underlying behaviour on the reversal learning task in more depth, the novel reinforcement learning-drift diffusion model (RL-DDM) was fitted to the data. In this model, the softmax decision rule is replaced with the DDM (Wiehler and Peters, 2020). The benefit of this model compared with a standard Q-learning model is that it integrates the response time distributions of the decisions into the model.

The DDM uses the Wiener first-passage time (WFPT) instead of a softmax function to calculate the probability of the RT of a choice x (Pedersen et al., 2017):

$$RT(x) \sim WFPT[\alpha, \beta, \delta, \tau]$$
 (2.12)

The non-decision time τ , bias β and boundary separation α , as described previously, are trial-independent free parameters. The drift rate δ , on the other hand, varies on a trial-by-trial basis as a function of the difference in expected rewards. It is multiplied by scaling factor *m*, which represents the ability of the subject to use knowledge of the reward probabilities:

$$\delta(t) = [V_{upper}(t) - V_{lower}(t)] * m$$
(2.13)

 $V_{upper}(t)$ and V_{lower} (t) are the reward expectations for the different choices. The values V are updated according to equation 2.1, using the prediction error and learning rate α . The parameter β in the model remains constant over time. The data was modelled using the HDDM python toolbox, which uses hierarchical Bayesian parameter estimation of the RL-DDM (Fontanesi et al., 2019).

2.2.4 Image pre-processing

Data acquisition for the different studies is described in detail in the respective chapters, as the datasets were collected at different centres. However, the pre-processing steps for both studies involving human data were the same. The FMRIB Software Library (FSL) and FMRIPREP, a Nipype based tool were used to pre-process the fMRI scans (Esteban et al., 2018; Smith et al., 2004). FMRIPREP employs a combination of softwares, including FSL and the Advanced Normalisation Tools (ANTs) software (Tustison et al., 2010). Each T1-weighted image was bias-field corrected using *N4BiasFieldCorrection* and skull-stripped using *antsBrainExtraction* with the OASIS template from the ANTs software. The fMRI scans were spatially normalised to the ICBM 152 Nonlinear Asymmetrical template version 2009c through non-linear registration with the antsRegistration tool using brain-extracted versions of both the T1-weighted volume and template (Avants et al., 2008).

The brain-extracted T1w images were segmented into cerebrospinal fluid, white matter and grey matter using *fast* (FSL) (Zhang et al., 2001). Functional MRI scans were slicetiming corrected using *slicertimer* (FSL) and subsequently motion corrected with *mcflirt* (FSL) (Jenkinson et al., 2002). For scans that had associated field maps, distortion correction was performed using *fugue* (FSL) (Jenkinson, 2003). Then, fMRI images were co-registered to their corresponding T1w scan using boundary-based registrations with six degrees of freedom with *flirt* (FSL) (Greve and Fischl, 2009). The field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using *antsApplyTransforms* using Lanczos interpolation.

Frame-wise displacement was calculated for each fMRI scan with Nipype (Power et al., 2014). The first five volumes were discarded to avoid T1 saturation effects. The functional scans were high-pass filtered (128 s) and spatially smoothed with a 6 mm full-width, half-maximum 3D Gaussian kernel. A canonical haemodynamic response function was modelled to the onsets of the explanatory event types. Quality checks were conducted by checking successful registration, ensuring that none of the participants showed excessive motion using DVARS and framewise-displacement measures (excessive motion threshold being 10% of the total number of volumes) and by inspecting their respective carpet plots. More details on quality checks can be found in the Appendix.

Chapter 3

Reinforcement learning in rodents exposed to maternal separation and adulthood stress

3.1 Introduction

Increasingly, evidence suggests that patients with MDD are impaired on tasks during which behaviour needs to be flexibly adjusted in response to unexpected changes in reward contingencies. An example of such a task is the reversal learning paradigm, during which the subject's performance depends on previous reward and punishment to update behaviour (Chen et al., 2015; Murphy et al., 2003). By fitting RL models to probabilistic reversal learning task data from patients with depression, lower reward learning rates and reward sensitivity compared to control participants have been reported (Chase et al., 2010; Mukherjee et al., 2020). This suggests that patients learn more slowly from previous feedback and have blunted responses to positive and negative outcomes. Furthermore, it is known that there are sex differences in MDD, with some studies showing that disease prevalence is twice as high in women than in men (Bromet et al., 2011; Salk et al., 2017). Females are more sensitive to reward and punishment feedback on decision-making tasks than males, which may contribute to increased MDD prevalence in females (Fattore, 2015; van den Bos et al., 2013).

Due to the heterogeneous nature of depressive disorders and their complexity, research on these disorders has proven to be challenging (Wang et al., 2017). Rodent models have enabled researchers to investigate endophenotypes that are translationally relevant to the study of psychiatric disorders and thus depressive disorders, such as anxiety, behavioural

flexibility, and anhedonia (Cope et al., 2012; Gottesman and Gould, 2003; Iacono, 2018; Laughlin et al., 2011; Moreno-Fernández et al., 2017). One intervention that has been used to induce depression-like symptoms in rodents is repeated early-life maternal separation. During this procedure, pups are intermittently separated from their mothers from an early postnatal age until the start of adolescence (Vetulani, 2013). Researchers have found that REMS results in increased depression-like behaviours and anxiety in rats, lowered preference for sucrose on the SPT, as well as reduced appetitively-conditioned locomotor activity in a novel environment (Cui et al., 2020; LeDoux and Pine, 2016; Matthews et al., 1996). The consequences of REMS are also known to be sex-dependent. For example, females demonstrate greater depressive-like symptoms on the SPT and there are sex-specific differences on object recognition tasks after REMS (Bath et al., 2017; Goodwill et al., 2019).

Early life stress in both humans and rodents has a profound negative impact on cognitive flexibility and other behavioural measures. In mice and rats, maternal separation and chronic stress lead to reduced cognitive flexibility on reversal learning tasks (Goodwill et al., 2019; Hyer et al., 2021; Thomas et al., 2016). Interestingly, acute stress interventions in rats have led to improvements in reversal learning performance (Bryce and Howland, 2015; Thai et al., 2013). In human adolescents that have experienced ELS, both cognitive flexibility as well as instrumental learning are impaired (Harms et al., 2018). Furthermore, adults that have been exposed to ELS but do not have any psychiatric diagnoses show deficits in positive feedback sensitivity and reward learning (Wilkinson et al., 2021). Drift diffusion models have also helped to gain an insight into how decision-making processes are disrupted as a result of stress and anxiety. A rodent study found altered boundary separation and drift rate resulting from acute restraint stress and administration of an anxiogenic drug, accompanied by a starting point closer to the low reward boundary (Hales et al., 2016). In humans with MDD, a less biased starting point than in control participants has been reported, alongside slower response times, a slower drift rate and wider boundary separation, aligning with the findings in rodents (Lawlor et al., 2020; Pedersen et al., 2021). In summary, the literature demonstrates that there are extensive changes in decision-making that arise from acute or chronic stress and depression. These changes can be quantified using RL and DDM models, which further enable us to dissect the mechanisms underlying behavioural changes.

Through resting-state fMRI studies, researchers have been able to identify which brain regions have altered functional connectivity in MDD. In medication-naïve patients, decreased resting-state functional connectivity (rs-FC) from the amygdala to the PFC has been reported (Satterthwaite et al., 2016; Wu et al., 2020). This was also found in individuals exposed to ELS (Fan et al., 2014; Herringa et al., 2013; Thomason et al., 2014). Additionally,

studies in humans have reported that following 8 weeks of SSRI treatment, PFC to amygdala connectivity is normalised and indistinguishable from controls (Fales et al., 2009; Zhang et al., 2020). In rodents, on the other hand, amygdala-PFC activity had been shown to be hyperactive following ELS (Cohen et al., 2013; Honeycutt et al., 2020; Johnson et al., 2018). Even though the changes in connectivity in rodents and humans do not show the same trend, these studies support the idea that amygdala-PFC circuitry plays a major role in MDD aetiology and may be relevant for treatment.

In this chapter, data from a study investigating behavioural and neurobiological consequences of early-life maternal separation and repeated adulthood stress in rodents is discussed. A seed-based fMRI approach is used to investigate functional connectivity changes in animals exposed to REMS between post-natal days (PND) 5-19 compared to control animals. Additionally, RL models were fit to behavioural data from a PRL task, which includes spurious feedback, as only 80% of correct trials and 20% of incorrect trials are rewarded (Murphy et al., 2003). Therefore, this task includes uncertainty and stimulus ambiguity, which depressed individuals are less tolerant towards (Saulnier et al., 2019). The Q-learning parameters were related to conventional measures of RL performance, such as win-stay/lose-shift behaviour and perseverative errors. It was predicted that animals exposed to REMS would exhibit a lower reward learning rate, a higher non-reward learning rate and lower reinforcement sensitivity. The aim of this study was to gain a more in-depth understanding of the functional connectivity and RL differences in rodents exposed to REMS and relate the behavioural changes to neurobiological consequences.

3.2 Methods

3.2.1 Subjects

Fourteen pregnant female Lister-Hooded rats were delivered by Envigo (Blackthorn, UK). The rats received nest material upon arrival and gave spontaneous birth to litters on gestational days 22-24. Each litter size had four to six offspring, containing both males and females. Pups were mixed when two litters were born within 24 hours of each other to minimise epigenetic and genetic effects. The pups were randomly assigned to the control group or the group that would undergo REMS (n=30 and n=28, respectively). The day of birth was defined as postnatal day zero (PND0). Food and water were provided ad libitum up to PND73-78 when behavioural testing commenced. Throughout behavioural training, animals were kept to 85% of their free-feeding weight. The temperature in the home cages

was kept constant at 23±0.2 °C and relative humidity was constant at 60±5%. The animals were kept on a 12-hour reverse light/dark cycle. All experiments adhered to the guidelines of the United Kingdom 1986 Animals Scientific Procedures Act (Project Licence PA9FBFA9F, held by Prof Amy Milton) and were approved by the local ethics committee (AWERB). The data were collected by other members of the Dalley group; the analyses of PRL behavioural data and rs-fMRI presented in this chapter were conducted by myself.

3.2.2 Repeated Early Maternal Separation

The pups in the REMS group were separated from the mothers for six hours every day between PND5 and PND19. During this time, pups were taken to a separate room and were kept together in a ventilated cage. Bedding was provided and the surface temperature of the bedding was kept between 30 and 35 °C with a heat pad and warmed air. Control animals were kept in their home cages. After PND20, offspring were maintained in same-sex pairs. A timeline of the study can be seen in Fig. 3.1.



Fig. 3.1 REMS study timeline.

Timeline of the REMS study, involving a two-hit model of depression, behavioural testing (including the PRL task) and multiple neuroimaging scans, before and after exposure to adulthood stress.

3.2.3 Adulthood stress

Between PND260-300, animals in both groups were exposed to a regime of mild, inescapable foot-shock stress over a continuous 20-day period. On 14 to 16 of those days, animals were placed in an operant chamber for 30 minutes during which time they received 1 or 2 temporally unpredictable foot-shocks. Every animal received 20 mild shocks throughout the 20-day time period. During the first 8 days, the number of shocks delivered was 2, 1, 2, 0, 1, 2, 0, 1. On the remaining days, the distribution of shock delivery of the remaining 11 shocks varied between animals. The number of shocks in one session would never exceed 2. The timing onset of each shock was random but did not occur during the first five or last ten minutes of the sessions. Shock intensity was 0.5 mA with a duration of 0.5 seconds.

3.2.4 Reversal learning task

Animals were trained in one of eight identical operant chambers (Med Associates, St. Albans, VT, USA). There was a 15-inch LCD touchscreen on one side of the chamber. On the opposite side of the chamber, a pellet receptacle with a head-entry detector and accompanying light was located. Small 50% sucrose pellets were delivered by an electronic pellet dispenser (TestDiet, St. Louis, MO, USA) when trials were completed. K-Limbic software (Conclusive Marketing LTD., High Wych, UK) was used to control the touchscreens.

On each day, every animal conducted either a training or a test session. The session ended after 40 minutes or when 200 trials were completed by the animal. After habituation to the test environment, animals were trained to respond to a solid white square stimulus on the right or left side of the touchscreen. When the stimulus was correctly touched, a 0.5 s tone was played, and the pellet receptacle light turned on and a food pellet was delivered. Once the animal took the pellet from the receptacle, the light was turned off, followed by a 5 s inter-trial interval.

After receiving at least 100 rewards in each of two consecutive sessions, the rats progressed to the next stage, during which animals had to respond to the solid white square stimulus, but any touch outside of the stimulus was punished by a 5 s time-out. After receiving 100 rewards on two additional consecutive sessions, subjects progressed to the deterministic reversal learning task. During this stage, two white square stimuli on opposite sides of the screen were presented simultaneously. One was the 'correct' stimulus and the other one was 'incorrect'. When the former was touched, the trial was rewarded. If the 'incorrect' option was touched, there was a 5 s time-out punishment. When the animal selected the 'correct' stimulus 8 times consecutively, the contingencies were reversed, and the other side was the

'correct' choice. Animals progressed to the probabilistic reversal learning task stage after achieving four reversals on two consecutive days. This task was similar to the deterministic reversal learning task, except that in PRL only 80% of trials were rewarded when a 'correct' response was made, and 20% of 'incorrect' responses were rewarded randomly.

PRL testing started at PND230 and was repeated daily over the course of 9 days. After adulthood stress, PRL testing was repeated over 3 days (Fig. 3.1.). Since there were only 3 sessions after adulthood stress, the results prior to the stressor were constrained to the first 3 sessions. For each session, trials up until the fourth reversal were included. Forty-six animals were included in the final analysis, as some animals had to be excluded due to data loss or the animal not completing the task correctly.



Fig. 3.2 Representation of the reversal learning task.

Diagram of a touchscreen-based system for testing reversal learning in rodents. A) Animals learn to associate the rewards with the visual stimulus presented on the touchscreen. Both the reward and stimulus are presented at the same time. B) The animal learns to touch the screen to receive the reward. C) The animal learns to associate one of the stimuli with reward and learns to not select the other stimulus. D) The initial stimulus-reward contingency is reversed. Adapted from (Brigman et al., 2010).

3.2.5 Statistical analysis of behavioural data

The conventional and modelling parameters from the DDM, RL and Bayesian Learner models were centred. Linear mixed-effects (LME) models were fit to the measures individually, including REMS, sex, adulthood stress and their interaction as fixed factors. The LME also contained a random intercept for each subject. Subsequent post-hoc pairwise comparisons of estimated marginal means were run (Lenth et al., 2018; Pinheiro et al., 2021). The residuals were checked for normality using the Shapiro-Wilk test, which was verified in the data and confirmed an assumption required for fitting an LME model (Schielzeth et al., 2020). The relationship between the RL and conventional PRL measures was investigated through a correlation analysis using Pearson's correlation coefficient (PCC), accounting for

multiple comparisons with FDR correction of 5% (Benjamini and Hochberg, 1995). Prior to running this analysis, it was ensured that all measures had a normal distribution using the Shapiro-Wilk test, which is a necessary condition when calculating PCC. All statistical analyses were run in R: A Language and Environment for Statistical Computing (R Core Team, 2020). For all analyses, significance was considered to be below p=0.05.

3.2.6 Image acquisition

Animals were scanned using a 9.4 T horizontal bore MRI scanner (BioSpec 94/20, Bruker, Coventry, UK). Structural imaging was achieved using multi-parametric mapping, from which T1-weighted and PD-weighted sequences were obtained. Contrast between grey and white matter was enhanced by applying MT pulses (10μ T, 2 kHz off-resonance) within each repetition. The field of view of $30.72 \times 25.6 \times 20.48 \text{ mm}^3$ was constrained within a matrix of $192 \times 160 \times 128$ voxels, yielding an in-plane isotropic resolution of 0.16 mm². To acquire the functional MRI data, a three-dimensional multi-echo gradient echo sequence was used with a relaxation time (TR) of 1832 ms, an echo time (TE) of 5.5/16.5/27 ms for the three echoes, a flip angle of 6 degrees, and RF spoiling of 117 degrees. The field of view of $28.8 \times 21.6 \times 21.6 \text{ mm}^3$ was constrained within a matrix of $64 \times 48 \times 48$ voxels, resulting in an in-plane resolution of 0.45 mm with a slice thickness of 0.5 mm. 655 frames were acquired for the resting-state fMRI scans acquired contiguously and interleaved acquisition. An acceleration factor of 1.55 was used.

Rats were kept under 1-2% isoflurane anaesthesia which was delivered in 100% oxygen at 1 L/min throughout the duration of the structural sequences, which was lowered to 20% oxygen/80% air during resting-state fMRI acquisition. The animals were continuously monitored using a pulse oximeter, respiratory tracer pad, and rectal temperature probe (SA Instruments, Stony Brook, NY, USA) to ensure that vital signs were within normal limits. Whenever the signs exceeded the limits, anaesthesia depth and temperature (using a heat pad) were adjusted, however, isoflurane was aimed to be kept at low levels (1-2 %). The data were collected at the Small Animal Neuromaging Facility, West Cambridge, as part of a collaborative study between the Dalley and Bullmore groups, funded by GlaxoSmithKline. Detailed information on image pre-processing can be found in chapter 2.

3.2.7 Seed-based analysis

Region-of-interest (ROI) masks were based on the atlas from (Valdés-Hernández et al., 2011) and were warped into the study-specific template space using warp fields from atlas

space to the study-specific template space generated by DARTEL. The following ROIs were selected: prelimbic cortex, infralimbic cortex, BLA, medial orbitofrontal cortex and lateral orbitofrontal cortex (Fig. 3.3). The seed-based analyses were run in FSL (*dual_regression*) (Beckmann et al., 2009). The resulting voxel-wise connectivity maps were compared using randomise with 5000 permutations and threshold-free cluster enhancement (Winkler et al., 2014). Multiple comparisons were corrected for using the family-wise error rate (Benjamini and Hochberg, 1995). The contrast tested was a sex×MS interaction, at each time point separately. Furthermore, the Q-learning parameters $\alpha_{non-rew}$ and κ from the RL model were centred and were separately used as regressors in the basolateral amygdala (BLA) and mOFC analyses (which was based on the results of the first seed-based analyses), and contrasts tested whether there was a positive or negative correlation with these parameters for all animals. In the correlation analyses, REMS and sex were accounted for. FSLeyes was used to generate the results figures (Smith et al., 2004). In these fMRI figures, the right and left hemispheres of the brain are inverted from the observer's perspective.



Fig. 3.3 Masks of the regions used for the seed-based analysis.

A) infralimbic cortex (X:78; Y:148,150,152; Z:48), B) prelimbic cortex (X:76; Y:146,152,158; Z:59), C) medial orbitofrontal cortex (X:78; Y:164,168; Z:52), D) lateral orbitofrontal cortex (X:62; Y:156,162; Z:54), E) basolateral amygdala (X:104; Y:116; Z:26,30).

3.2.8 Region-of-interest analysis

Based on the results from the seed-based analysis, timeseries from pre-selected regions were extracted and the BLA timeseries were correlated with these ROIs. The ROIs included: anterior insula (AI), caudate putamen (CPu), mOFC, lOFC, PrL and IL. The correlation values were converted using Fisher's r to z transformation. An LME was fitted to the data, containing a sex×REMS×adulthood stress interaction and a random intercept for each subject, followed by post-hoc pairwise comparisons between groups using estimated marginal means (Lenth et al., 2018; Pinheiro et al., 2021). Additionally, the connectivity values for each animal were correlated with their respective Q-learning parameters. Multiple comparison testing was corrected for with FDR correction (Benjamini and Hochberg, 1995). The statistical analyses were run in R (R Core Team, 2020).

3.3 Results

3.4 Behavioural findings

3.4.1 Probabilistic reversal learning and anhedonia

A conventional analysis of the PRL task, as described in sections 2.1.5 and 3.2.5, showed that there are sex-specific effects of REMS and adulthood stress on some of the measures. Firstly, there were no statistically significant differences in the percentage of correct responses and the number of trials required to reach criterion. Win-stay behaviour, which is the proportion of trials during which an animal selects the same response after receiving a reward regardless of the previous trial being correct or not, was found to be lower in female than in male animals (previous response correct: F(1,46)=4.50, p=0.030; previous response incorrect F(1,46)=5.85, p=0.019).

A three-way sex×MS×adulthood stress interaction reached significance for lose-shift behaviour after a correct trial (F(1,236)=4.16, p=0.043). Lose-shift behaviour increased in control males after exposure to adulthood stress (t(236)=-2.56, p=0.011). Adulthood stress did not affect female lose-shift measures. A statistically significant sex×MS×adulthood stress interaction was also seen for lose-shift behaviour after an incorrect trial, which is the proportion of trials in which the animal changes its response after a loss and an incorrect response (F(1,236)=7.49, p=0.0067). Before exposure to the second stressor, MS females shifted less after a loss than control females (t(44)=3.06, p=0.0040). After the stressor, an

increase in lose-shift behaviour was observed only in MS females (t(236)=-2.02, p=0.044) (Fig. 3.4). No such changes were observed in males.

The LME fitted to the number of perseverative responses had a significant sex×adulthood stress interaction (F(1,236)=7.01, p=0.0086) and a near-significant MS×adulthood stress interaction (F(1,236)=3.75, p=0.054). Through post-hoc comparisons it was found that this was attributable to control males perseverating more than MS males prior to the second stressor (t(46)=2.71, p=0.0097) and that there was a reduction across the two time points in control males only (t(236)=3.41, p=0.0078). Results for the non-affected conventional PRL measures can be found in the Appendix.

Anhedonia, as measured via the SPT, was not affected by MS or adulthood stress, and no interactive effects reached significance. Only sex affected behaviour on the SPT (F(1,45)=6.85, p=0.012), as females showed greater preference for the sucrose solution (t(45)=-3.70, p<0.001).

3.4.2 Reinforcement Learning model comparison

The BIC, log-likelihood ratio and pseudo r^2 (section 2.1.4) were the measures used to identify which of the computational models tested fitted the data the best. The QL 4-parameter model with no 'forgetting' was used as a baseline for comparison. All the Q-learning models (2-parameter, 3-parameter, 'forgetting' and 'differential forgetting' models) were compared to this baseline. The 4-parameter model showed significantly better fit to the data than the 2-parameter model, however, evidence of better fit than the 3-parameter model and 'forgetting' model was small. 80% of the animals and sessions had a lower BIC for the 4-parameter model than the 2-parameter model, whereas only 55% had a BIC lower for the 4-parameter model than the 3-parameter 'forgetting' model. For the 3-parameter model, all sessions had a BIC lower than the 4-parameter model, but only 1% had a difference greater than 4. Furthermore, none of the sessions crossed the log likelihood ratio threshold of 3.842, indicating that neither model is significantly better. Model fit for the model with the 'differential forgetting' mechanism was substantially worse than the baseline, as only 18% of sessions had a lower BIC than for the 4-parameter model. An example of model comparison can be found in Fig. 3.5, with extensive information on all model comparisons in the Appendix.





Males (green) and females (orange) and control and maternally separated (MS) animals prior to (pre) and after (post) the adulthood stressors. The behavioural measures that changed as a result of MS and adulthood stress were lose-shift behaviour after an incorrect trial in females, lose-shift behaviour after a correct trial and perseverative responding in males. A) Lose-shift behaviour after a correct response; B) Lose-shift behaviour after an incorrect response; C) Perseverative responding. * - p < 0.05.



Fig. 3.5 Q-learning model comparisons.

A) BIC comparison of the 4-parameter and 2-parameter model; B) BIC comparison of the 4-parameter and 3-parameter model; C) the log likelihood ratio test for the 4- and 2-parameter models; D) the log likelihood ratio test for 4- and 3-parameter models.

The lower the BIC, the more evidence is provided for the model. The dotted lines represent the threshold for a small, but insignificant improvement in model fit of one model over the other, and the continuous line represents the threshold for substantial evidence for one model over the other. In the log likelihood ratio test, the significance threshold for comparing models is at 3.842 for p=0.05. Above this threshold, there is a significant difference in data likelihood. Based on these figures, it can be seen that the 4-parameter model fits the data significantly better than the 2-parameter model, as for the majority of sessions, the BIC for this model is lower than that of the 4-parameter model. The likelihood ratio test confirms this, as the difference between the two models crosses the significance threshold for most sessions. The comparison between the 4-parameter and 3-parameter model suggests that the 3-parameter model provides slightly improved, but not significantly better model fit. Due to our interest in identifying differences in learning from reward and non-rewarded trials in this study, the 4-parameter model was chosen.

3.4.3 Reinforcement Learning results

A statistical analysis of the RL parameters, as described in section 3.2.5, found that the learning rate parameter for non-rewarded trials, $\alpha_{non-rew}$, had a MS×adulthood stress effect (F(1,236)=5.09, p=0.025). After adulthood stress, the $\alpha_{non-rew}$ parameter decreased in control males (t(236)=2.33, p=0.021), which was not found in MS males (p=0.41) (Fig. 3.6). $\alpha_{non-rew}$ was the same in control and MS females and was not affected by adulthood stress. The stress×adulthood stress interaction for α_{rew} , which is the learning rate for rewarded trials, did not reach significance (F(1,236)=3.34, p=0.069). The β parameter (how much the animal exploits previous information vs exploring other options) was not significantly affected by either intervention (p=0.71 and p=0.64). Therefore, sex, MS and adulthood stress did not impact the Q-learning measures α_{rew} and β .

However, the κ ('stickiness') LME contained a significant sex×MS×adulthood stress interaction (F(1,236)=8.20, p=0.0046). Post-hoc analyses found that MS females had higher values than control females (t(44)=-2.23, p=0.031), with no differences seen in males. In control females, stickiness increased after exposure to the foot-shock stressors (t(236)=-1.93, p=0.055), whereas in MS females there was no difference (p=0.12) (Fig. 3.6). In males, no differences due to REMS or adulthood stress were found (p=0.18 and p=0.43, respectively). These analyses have thus revealed a sexual dimorphism, which could be observed both prior to and post adulthood stress, as well as an effect of REMS followed by a second stressor later in life.



Fig. 3.6 Results of the Q-learning analysis.

RL parameters in males (left) and females (right) and control and maternally separated animals prior to (pre) and after (post) repeated shock-stressors over multiple days. A) α_{rew} , the learning rate from rewarded trials on the PRL task; B) $\alpha_{non-rew}$, the learning rate from non-rewarded trials; C) β , the inverse temperature parameter and D) κ , the autocorrelation, otherwise known as the stickiness parameter. * – p<0.05; ** – p<0.10.

3.4.4 Bayesian Learner model comparison

The three different BL models were also compared with each other. The model with a memory size of one and a flat prior was used as the standard for comparison. It was found that neither memory size nor different priors improved model fit, indicating the baseline model provided the best explanation of the data. This model and the 4-parameter Q-learning model were compared with each other, and the BL model showed better model fit. This suggests that the BL model provides a better mechanistic explanation of the data. However, due to the fact that the BL proposes a different learning mechanism than the Q-Learning model and since it has not yet been explored in the literature in relation to addiction and depression, it was decided to focus further analyses on the Q-Learning models. Specifically, the 4-parameter model was selected and related to the imaging data. Nonetheless, the BL parameters were statistically analysed to identify whether there were any differences in the parameters between the groups.



Fig. 3.7 Bayesian Learning model comparison.

A) Bayesian Information Criterion (BIC) comparison of the Bayesian Learning models with a memory size of 1 and 4; B) BIC comparison of the BL models with a memory size of 1 and 7; C) BIC comparison of the BL models with a memory size of 1 and 10. Dashed lines show a BIC difference of two, continuous lines represent a BIC difference of four.

3.4.5 Bayesian Learning results

The Hazard rate, which determines how much the posterior changes based on the datadriven likelihood versus how much it depends on the uniform prior distribution, was solely affected by the adulthood stressor, regardless of sex or maternal separation (F(1,44)=48.09, p<0.001). The Hazard rate decreased as a result of stress, indicating greater reliance on the data-driven likelihood, rather than the prior distributions (t(44)=6.57, p<0.001). There was a significant sex×TP interaction for the β parameter and a significant effect of adulthood stress on κ (F(1,44)=4.97, p=0.031 and F(1,44)=16.87, p<0.001, respectively). The β parameter also decreased following the stressor, although this was only the case in control females (t(44)=3.03, p=0.073), and κ was found to be reduced after the stressor across all animals (t(44)=4.22, p<0.001). The results can be found in Fig. 3.8.

3.4.6 Relationship between conventional PRL and RL parameters

Pearson's correlation coefficient was used to correlate the PRL and RL measures. There was a positive correlation between κ and win-stay behaviour (r(46)=0.51, p=0.002) (Fig. 3.9). κ is representative of the tendency to select the same response regardless of outcome. However, it seems plausible that higher 'stickiness' would lead to higher 'staying'. Furthermore, a positive correlation between α_{rew} and win-stay correct as well as proportion of correct responses was observed (r(46)=0.52, p=0.002 and r(46)=0.56, p=0.001, respectively). This implies that the higher the learning rate from rewarded trials, the greater the animals' tendency to stay after receiving a reward. No other correlations between the RL and conventional behavioural measures survived correction for multiple comparisons.



Fig. 3.8 Results of the Bayesian Learner analysis.

BL parameters prior to and post adulthood stress. The Hazard rate, which determines what the posterior is most affected by; β is the inverse temperature parameter and κ is the autocorrelation. All three parameters decreased following the stressor, although the β parameter only decreased in control females. * – p<0.05



Fig. 3.9 Correlation matrix representing the relationship between conventional PRL measures and parameters from the RL model across all sessions of the PRL task for all animals.

The colour bar on the right-hand side represents the correlation strength, the numbers within each square are the p-values. Prop correct – proportion of correct responses; Perseverative resp – number of perseverative responses; Alpha rewarded – α_{rew} parameter representing the learning rate from rewarded trials; Alpha non-rewarded – $\alpha_{non-rew}$ parameter representing the learning rate from non-rewarded trials; Beta – the exploitation/exploration parameter β ; Kappa – κ , also known as the stickiness parameter; correct and incorrect indicate the outcome of the previous response.

3.4.7 Drift Diffusion Model

To explore further how the decision-making processes were affected by sex, REMS and adulthood stress, a DDM was fit to the data. Prior to the DDM analysis, an LME was fit to the reaction time (RT) data. The effect of REMS and a sex×adulthood stress interaction on the RT became evident (F(1,44)=10.26, p=0.0025 and F(1,44)=30.62, p<0.001, respectively). The MS group had overall greater RTs (t(44)=-3.27, p=0.0021). RTs in males decreased as a result of adulthood stress (t(28)=3.81, p<0.001), but increased in females (t(19)=-4.12, p<0.001). Figures for these results and can be found in the Appendix.

The statistical analysis of DDM parameters highlighted that the parameter α , which is the boundary separation, was solely affected by the adulthood stress (F(1,44)=861.7, p<0.001). Specifically, it caused an increase in the values of α (t(44)=-28.49, p<0.001) (Fig. 3.10). There was also an effect of adulthood stress on the β parameter (F(1,44)=3067.1, p<0.001), which decreased (t(44))=17.99, p<0.001). Furthermore, there was a near-significant sex×adulthood stress interaction for β (F(1,44)=3.23,p=0.079). The only factor affecting δ and τ was the second stressor (F(1,44)=2668.62, p<0.001; F(1,44)=16.06, p<0.001), which resulted in decreased and increased values, respectively (t(44)=50.76, p<0.001 and t(44)=-3.90, p<0.001). Even though the DDM model did not highlight any changes in decision-making due to REMS, it showed that all DDM parameters were altered due to repeated shock stress exposure in adulthood, regardless of sex or REMS.

3.5 Imaging results

3.5.1 Volumetric Analysis

A structural imaging analysis was conducted to identify any volumetric changes arising from either early-life maternal separation or the repeated shock stressor. The analysis focused on six pre-defined regions, which were subsequently also used in the functional seed-based analysis: BLA, mOFC, IOFC, IL and PrL.

Through fitting an LME to the volumes of each ROI independently, it was found that the BLA size was affected by a sex×MS×adulthood stress interaction (F(1,29)=7.87, p=0.0089). This effect was largely driven by an increase in BLA volume between the two time points (t(29)=-4.28, p<0.001) and larger BLA structures in males (t(44)=3.47, p<0.001). The IL, PrL, mOFC and IOFC were solely affected by sex and time, which was expected as male rodents have greater overall brain volumes and since rodent brains develop between PND60 and PND300 (Semple et al., 2013).



Fig. 3.10 Results from the Drift Diffusion Model.

DDM parameters in all animals prior to and after repeated shock-stressors over multiple days. α is the boundary separation, β is the bias, δ is the drift rate and τ is the non-decision time. Due to repeated stress in adulthood, α and δ increased in all animals, whereas β and τ decreased (p<0.001 for all parameters). * – p<0.05.

3.5.2 Seed-based analysis

Six ROIs were used as seeds in a whole-brain seed-based analysis (SCA): BLA, mOFC, IOFC, IL, PrL (Fig. 3.3). The SCA revealed changes in the rs-FC arising from REMS and the repeated adulthood stress. At the first time point, PND62, which was after REMS but before adulthood stress, no rs-FC differences between the ROIs and other regions were highlighted. In order to investigate whether adulthood stress resulted in connectivity changes, the same SCAs were repeated for the second time point. None of the ROIs except the BLA and mOFC had significantly altered connectivity. When the BLA was used as the ROI, there was a statistically significant sex×MS interaction at this time point. Connectivity from the BLA to regions such as the AI, cingulate cortex (Cg), IL and dorsolateral striatum (DLS) were affected bilaterally (Fig. 3.11). Specifically, connectivity from the BLA to these regions was weaker in MS females and control males than control females and MS males, respectively. T-tests confirmed that in females, connectivity to all of the mentioned areas was affected, but in males only the connectivity to the DS and AI was altered. Results from the separate t-tests can be found in the Appendix.

When the mOFC was the seed region and the voxel-wise connectivity maps were compared, weaker connectivity between the mOFC and right VS in MS females and control males compared to their counterparts was found (Fig. 3.12). Therefore, it seems plausible that these changes in rs-FC arose as a result of the regimen of unpredictable foot-shocks.



В	Seed-region	Number of voxels in cluster(s)	Volume of cluster(s) (mm ³)	Degrees of Freedom	Range of t- statistics
	BLA	14041	47389	43	2.70-5.44

Fig. 3.11 BLA seed-based analysis results

A) Areas to which the rs-FC from the BLA is altered due to a sex×maternal separation interaction. These areas are: the cingulate cortex, infralimbic cortex, dorsal striatum and anterior insular cortex. The connectivity strength is higher in control females and MS males than MS females and control males. The colour bar represents the t-statistic of the respective voxels. B) A table summarising the characteristics of the highlighted clusters.



В	Seed-region	Number of voxels in cluster(s)	Volume of cluster(s) (mm ³)	Degrees of Freedom	Range of t- statistics
	Medial OFC	113	382	30	3.94-5.92

Fig. 3.12 mOFC seed-based	analysis	results
---------------------------	----------	---------

A) Areas to which the rs-FC from the mOFC is altered due to a sex×maternal separation interaction. The cluster is located in the right ventral striatum, and the connectivity strength between the two regions is higher in control females and MS males than MS females and control males. The colour bar represents the t-statistic of the respective voxels. B) A table summarising the characteristics of the highlighted clusters.

To identify the relationship between altered connectivity and the observed changes in the RL parameters, the parameters were included as additional regressors in the mOFC and BLA analyses at both time points. At the first time point, PND62, values of κ positively correlated with connectivity between the mOFC and a cluster spanning the Cg, DS, IL, PrL and the granular/dysgranular insular cortex in all animals, regardless of sex or REMS (Fig. 3.13). This suggests that there was stronger connectivity between the mOFC and these areas in animals that showed higher levels of κ , or stickiness. No other significant correlations were found.



В	Seed-region	Number of voxels in cluster(s)	Volume of cluster(s) (mm ³)	Degrees of Freedom	Range of t- statistics
	Medial OFC	9136	30834	30	2.55-4.62
	Medial OFC	9130	30834	30	2.33-4.02

Fig. 3.13 Medial OFC and RL measures correlation results

A) Areas with which the connectivity to the mOFC positively correlated with the reinforcement learning parameter κ , also known as behavioural stickiness. The cluster spans the cingulate cortex, infralimbic/prelimbic cortices, dorsal striatum and parts of the granular/dysgranular insular cortex. This effect is seen bilaterally but is more pronounced in the left hemisphere. The colour bar on the right-hand side represents the t-statistic of the respective voxels. B) A table summarising the characteristics of the highlighted clusters.

These results show that REMS on its own did not result in any functional connectivity differences. At the first time point, prior to adulthood stress, we observed a positive correlation between the rsFC from the mOFC to cortico-striatal regions, such as the PFC, Cg and DS and the RL parameter κ . After adulthood stress, the connectivity between the BLA and the same regions was weakened in MS females and control males, and there was reduced connectivity between the mOFC and VS.
3.5.3 Region-of-interest analysis

The more targeted correlation analysis, which was based on the results from seed-based analysis, revealed a more in-depth insight into the regions affected by maternal separation, and how they may relate to Q-learning. The LMEs found a significant effect only of sex on the connectivity from the BLA to AI (F(1,45)=14.7, p<0.001), IL (F(1,45)=7.44, p=0.0091) and mOFC (F(1,45)=15.0, p<0.001). There was a significant effect of sex (F(1,45)=9.52, p=0.0035) and adulthood stress (F(1,22)=5.52, p=0.028) on the BLA to CPu connectivity. This was also the case for the BLA to PrL connectivity (F(1,45)=6.05, p=0.018 and F(1,22)=11.5, p=0.0026, respectively) and BLA to IOFC functional connectivity (F(1,45)=8.03, p=0.0069 and F(1,22)=7.26, p=0.013, respectively). Pairwise comparisons revealed there was a reduction in the connectivity between the BLA and the CPu from TP1 to TP2 (t(22)=2.12, p=0.045), an increase in BLA to IOFC connectivity (t(22)=-2.65, p=0.015) and in the connectivity from the BLA to the PrL (t(22)=-3.16, p=0.0045) across all animals (Fig. 3.14).

Moreover, the rs-FC between the mOFC and nucleus accumbens shell (AcbS) was affected by a sex×adulthood stress interaction (F(1,22)=6.27, p=0.020), as was the mOFC to CPu connectivity (F(1,22)=5.64, p=0.027), the former arising due to stronger connectivity in females than males a the first time point (t(22)=-3.55, p=0.0089) and a reduction in connectivity between the two time points in females (t(22)=2.57, p=0.076). The latter was due to the connectivity strength being greater in females than males at both time points (TP1: t(22)=-3.21, p=0.019; TP2: t(22)=3.18, p=0.021). Furthermore, rs-FC decreased between the two time points in females (t(22)=2.86, p=0.042).



Reinforcement learning in rodents exposed to maternal separation and adulthood stress

Fig. 3.14 Main results of the region-of-interest connectivity analysis.

Connectivity changes arising from exposure to repeated adulthood shock stress. The stressor resulted in decreased functional connectivity between the BLA to CPu, as well as the mOFC to AcbC. The opposite trend was observed for the BLA-PrL and BLA-IOFC connectivity. PCC - Pearson's Correlation Coefficient; BLA - basolateral amygdala; CPu - caudate putamen; PrL - prelimbic cortex; IOFC - lateral orbitofrontal cortex; mOFC - medial orbitofrontal cortex. * $\frac{80}{80}$

3.6 Discussion

In this chapter, it was explored how sex, early-life maternal separation and repeated foot-shock stress in adulthood alter RL in rats. Furthermore, it was investigated what the neural basis of potential changes in RL might be. The behavioural analysis revealed a sexual dimorphism, as the stickiness parameter κ was differentially altered in males and females due to REMS and adulthood stress. Females had higher values of κ as a result of REMS than control females prior to the second stressor. κ increased in control females after the adulthood stress but was unaffected in MS females. Thus, both REMS and foot-shock stress increased stickiness in females, but these effects were not additive. Stickiness was unaffected in males. Changes in lose-shift behaviour prior to the second stress aligned with these results, since MS females shifted less after a loss on the previous trial than controls. Given that κ provides a measure of stickiness, this suggests that higher κ values result in less contingency shifting after a loss. The conventional perseveration measure was greater in control males than MS males, but there was a decrease in perseverative responding after adulthood stress in the control group. In parallel, there was a decrease in $\alpha_{non-rew}$ and an increase in lose-shift, solely in this group. The sexual dimorphism was not seen in the DDM results. The four DDM parameters, α , β , δ and τ , were unaffected by sex and maternal separation. However, all decision-making parameters were altered by repeated adulthood stress. α , which is the boundary separation, increased post-stressor, as did the drift rate δ . This indicates a higher threshold for decision-making but reaching that threshold more rapidly. Additionally, the starting bias β decreased, indicating a starting bias closer to the non-rewarded outcome, as did the non-decision time τ .

Thus far, the stickiness parameter κ , which is the tendency to make the same choice regardless of previous reward or punishment, has not been investigated in the two-hit model of depression in rodents, and only scarcely in relation to early-life stress or MDD in humans. Habitual behaviour results from the direct elicitation of the instrumental response to the environmental stimuli repeatedly associated with it, autonomous of the goal (Robbins and Costa, 2017). κ provides a measure of habitual behaviour, which is manifested as value-free perseverative responding (Ito and Doya, 2009; Miller et al., 2019). Habitual responding and κ are both defined by performing the same action irrespective of the outcome of that action. ELS in humans has been linked to greater habitual responding in an avoidance learning task (Gordon et al., 2020; Patterson et al., 2019). Furthermore, it is known that acute stress in rodents and humans results in actions being more habit-based than goal-directed (Dias-Ferreira et al., 2009; Schwabe and Wolf, 2009, 2010). These studies support the

Reinforcement learning in rodents exposed to maternal separation and adulthood stress

observations reported in this chapter, specifically those in females, as there was an increase in κ due to REMS and adulthood stress. However, this does not explain why κ was not affected in male rats.

Sex-dependent effects of stress have been reported in rodents (Bath et al., 2017; Goodwill et al., 2019; Vetulani, 2013). For example, differences in behavioural profiles such as anticipatory responding to reward, anxious behaviours and emotionality have been found (Couto et al., 2012; Renard et al., 2007). It is unclear which sex is more sensitive to the effects of ELS in rodents. However, in humans, depression and anxiety are more frequently diagnosed in females. Moreover, women exposed to childhood trauma are more susceptible to mood disorders, aligning with the observations reported in this study (Albert, 2015; Gallo et al., 2018).

The changes reported in lose-shift behaviour and $\alpha_{non-rew}$ due to adulthood stress in control males and not MS males was contrary to the initial hypotheses and studies from humans, as patients with MDD generally have lower learning rates than control participants (Mukherjee et al., 2020; Safra et al., 2019). Although REMS has caused persistent changes in anxious and depressive phenotypes in rodents in some studies, other groups have not observed any effects (Cui et al., 2006; Raineki et al., 2012; Rice et al., 2008; van der Kooij et al., 2015). Therefore, it has not yet been established what changes to an animal's behavioural profile arise due to maternal separation. The observations in males could be interpreted as MS males being more resilient, agreeing with the concept of stress immunization proposed by (Levine, 1957), i.e., mild and predictable early-life stress can result in a protective effect on subsequent stress responses.

Fitting DDMs to behavioural paradigms such as approach-avoid conflict and reversal learning tasks have shown that patients with MDD have lower starting point biases, lower drift rates and wider decision thresholds (Lawlor et al., 2020; Pedersen et al., 2021). Acute treatment of an anxiogenic drug and social isolation in rats has previously led to increased boundary separation and an altered decision starting point (Hales et al., 2016). Although in this study no changes in any of the DDM parameters solely due to REMS were seen, the repeated foot-shock stress intervention resulted in increased boundary separation and a lower bias starting point, aligning with previous studies. However, our findings show an increased drift rate and reduced non-decision time, which is contrary to other reports. This may be attributable to differences in interventions, as well as cross-species differences. Increased boundary separation implies more conservative decision-making, which may be attributable to higher levels of anxiety or stress in animals after repeated stress exposure (Hales et al., 2016). The altered starting bias reported here, which was closer to the non-rewarded outcome

as a result of adulthood stress in all animals, is also likely to be attributable to higher levels of anxiety and stress after this intervention. These findings together may be interpreted as more conservative decision-making which is biased to the negative outcome due to repeated foot-shock stress, accompanied with faster accumulation of evidence and less time taken by stimulus encoding and motor processes (Pedersen et al., 2021).

The functional MRI analysis also revealed a sexual dimorphism. At TP1, which was before the repeated stress, there were no differences in the rs-FC as highlighted by the seed-based analysis. At TP2, however, MS females and control males had lower connectivity between the BLA seed and the AI and the DS than their counterparts. Additionally, MS females had lower connectivity to the IL and Cg. There was a similar trend for the connectivity between the mOFC and the VS. Furthermore, a positive correlation between κ and the connectivity from the mOFC to a cluster spanning the IL, DS and Cg was found. It has previously been reported that inactivation of the mOFC in rats reduces κ , supporting the findings presented here and providing additional evidence for the neural substrates underlying this RL parameter (Verharen et al., 2020). There are some limitations that need to be kept in mind, such as the fact that the number of animal scans available at PND62 was substantially lower than at PND300 (N=31 vs N=44). Moreover, although PND62 is considered adulthood in rats, the animals were scanned approximately 170 days before PRL testing commenced at PND230, which means that there may have been developmental changes in that time.

The volumetric analysis of the structural MRI data did not highlight any significant changes resulting from REMS. Although it was found that male rats had larger brain volumes than female rats, and that volumes were larger at PND300 than at PND62, these differences were unsurprising. The larger brain volumes at TP2 are likely not attributable to the repeated foot-shock stressor, but instead to the developmental changes that will have taken place between these two time points. The ROI correlation analysis of the functional MRI data revealed connectivity differences due to sex and adulthood stress, but not due to REMS or interactive factors. The mOFC to VS rs-FC connectivity decreased between the two time points. Altered connectivity between these regions was also highlighted in the seed-based analysis. rs-FC from the BLA to cortical areas such as the IOFC and PrL were strengthened following the stress intervention, but the connectivity between the BLA and CPu decreased. This more targeted ROI analysis gives an indication of the changes that may have occurred as a result of the second stress, whereas the seed-based analysis looked at the time points separately.

When the IL is lesioned or inhibited in rats, habitual responding is impaired, which suggests that the IL is a key component of the network giving rise to habitual behaviour

Reinforcement learning in rodents exposed to maternal separation and adulthood stress

(Killcross and Coutureau, 2003; Smith et al., 2012). Lesions of the DLS, also a part of the putative habit system, reduce perseverative responding and facilitate flexible behaviour (Graybiel, 2008; Miller et al., 2019; Yin and Knowlton, 2006). Chronic stress in rodents can result in frontostriatal reorganisation, resulting in a shift from goal-directed to habitual actions (Dias-Ferreira et al., 2009; Sharp, 2017). Even though the number of human studies on this matter is small, it is known that habitual action selection activates the insular cortex (Eryilmaz et al., 2017). Additionally, activations in the putamen, which is the human analogue of the rodent DLS, have been linked to tracking habit development (Tricomi et al., 2009). It therefore seems possible that the increased κ observed in control females after the second stressor are due to strengthened circuits involving the DLS, insular and prefrontal cortices.

The reduction in the learning rate parameter $\alpha_{non-rew}$ in control males between the two time points could be explained by the weakened connectivity from the BLA to the DS and IL. Lesioning the IL in rats has resulted in impairments of $\alpha_{non-rew}$, and both rodent and human studies have proven that the DS and BLA are essential for reward learning (Balleine et al., 2007; Jog et al., 1999; Verharen et al., 2020; Wassum and Izquierdo, 2015). In human studies, the rs-FC between the amygdala and PFC was lower in participants that have experienced ELS (Colich et al., 2017; Fan et al., 2014). Furthermore, patients with MDD exhibit reduced amygdala-PFC and AI-PFC connectivity, with sex differences also having been reported (Kandilarova et al., 2018; Kong et al., 2013; Colich et al., 2017).

Given that no functional connectivity differences were observed at TP1, it can be inferred that the connectivity changes at the last time point were due to repeated foot-shock stress and that REMS and sex caused differential predispositions to behavioural and neurobiological consequences. The changes in connectivity were confirmed by both the seed-based and ROIs analyses. Even though no effects of sex or REMS were found in the ROI analysis, this may be due to methodological differences, given that the ROIs were selected based on existing rodent atlases, whereas the seed-based analysis was done on a whole-brain level. Furthermore, the ROI analysis included adulthood stress as an effect, which could not be included in the SCA. Another reason for this discrepancy may be the statistical bias that is introduced through selecting the ROIs based on previous analyses. Nonetheless, the ROI analysis gave us an indication of what might have been a result of the repeated stressor; connectivity from the BLA to cortical areas, including the PrL and IOFC, increased. On the other hand, connectivity to striatal areas from both the BLA and mOFC decreased, suggesting that overall input to the striatum was reduced after chronic stress. Altered connectivity from the BLA to PFC regions due to chronic stress has frequently been reported in both rodents and humans (Hultman et al., 2016; Robinson et al., 2014; Shin et al., 2005). This is due to

the importance of the BLA in the control of emotions, particularly fear and anxiety (LeDoux, 1992). The reduced connectivity to the striatum could be due to hypoactivity in prefrontal areas resulting from stress, which is also a consequence observed across species (Hayes et al., 2012; Lim et al., 2017; Sailer et al., 2008).

In summary, this chapter explored how RL substrates are changed in a two-hit model of depression in rats. Further, it was assessed how the rs-FC was altered due to REMS and repeated foot-shock stress. The data also made it possible to pinpoint the neurobiological substrates of RL parameters, particularly the stickiness parameter κ . For the first time, it was identified that this behaviour is modulated by circuits involving the mOFC, dorsal striatum, PFC and insular cortex. In the next chapters, RL and its underlying biology will be explored in an additional rodent cohort as well as human data.

Chapter 4

Reinforcement learning in rodents exposed to cocaine

4.1 Introduction

Substance use disorder is a brain disorder manifesting itself through compulsive drug seeking in spite of adverse outcomes (APA, 2013). In the most up-to-date version of the Diagnostic and Statistical Manual of Mental Disorders, compulsive drug-seeking behaviour is a key feature of a SUD diagnosis. Specifically, it refers to the continued use of drugs even though it endangers the individual's life. In order to align with the human definition and diagnosis of SUD, animal studies have in recent years moved to procedures that replicate the persistent drug-seeking behaviour in the face of punishment (Lüscher et al., 2020). For example, compulsive drug-taking in rodents can be studied via drug self-administration in the face of punishment such as a foot-shock. This can be implemented using a fixed ratio schedule of reinforcement, in which the drug and punishment are delivered at the same time, following a specified number of responses (e.g., after 5 responses, FR-5). Another method of assessing compulsive drug-seeking is to use a second-order schedule of reinforcement. Initially, the drug is delivered after a fixed-interval and is paired with a classically conditioned stimulus (Everitt and Robbins, 2000). Subsequently, only the conditioned stimulus is presented after a FI, and following an additional interval of the same length, the drug is delivered. The conditioned stimulus elicits drug-seeking behaviours in the drug-free period, making it possible to study this important feature of addiction.

The transition from controlled to compulsive drug-seeking is thought to arise from impaired inhibitory response control, which is accompanied by a shift of behavioural control from the VS to DS and altered balance in frontostriatal connectivity (Everitt and Rob-

bins, 2016; Kalivas and Volkow, 2005; Lüscher et al., 2020). These alterations in limbiccorticostriatal circuits are thought to be a result of excessive DA modulation which is associated with drug abuse. Support for this has been provided by studies showing that non-addicted individuals which occasionally use stimulants have an increased release of DA in the VS when presented with drug-associated cues, whereas individuals that are addicted have an increased DA release in the putamen (Boileau et al., 2006; Wong et al., 2006). This may explain the accumulating evidence proposing that addiction is accompanied by a shift from goal-directed to habitual behaviour (Lüscher et al., 2020; Zapata et al., 2010). The former has been linked to activity in the DLS, whereas the latter is thought to be mediated by the ventral and dorsomedial striatum (DMS) (Balleine and O'Doherty, 2009).

Reversal learning impairments following prolonged drug use and drug withdrawal have been reported in both rats and humans. For example, withdrawal from cocaine SA results in long-lasting deficits on reversal learning in rats as measured by the trials required to reach a criterion (Calu et al., 2007). This deficit parallels reversal learning impairments following lesions of the OFC, suggestive of its involvement (Schoenbaum et al., 2003). Furthermore, in rats exhibiting high rates of cocaine SA, decreased lose-shift behaviour on a deterministic reversal learning task following SA has been reported (Zhukovsky et al., 2019). However, there were no differences on the DRL task prior to SA. In the same study, RL models were fitted to the DRL data. A comparison of the winning model's parameters found that there were no differences prior to cocaine SA but found increased β and κ values in the high-compulsive animals after SA. This suggests increased exploration, rather than exploitation, of lower Q-values, as well as an increased tendency to repeat the same choice regardless of outcome. Reductions in lose-shift behaviour following repeated drug use have also been reported in rats exposed to methamphetamine and in patients diagnosed with SUD (Groman et al., 2017; Parvaz et al., 2015).

The OFC, BLA and striatum are areas thought to mediate reversal learning (Izquierdo et al., 2017). The OFC is responsible for inhibiting responses that are no longer relevant, as is required in a reversal learning paradigm, and keeps track of associative information necessary for behaviour (Jones and Mishkin, 1972; Rolls et al., 1996). The BLA is involved in learning of stimulus-response contingencies, and the DLS encodes stimulus-response associations (Jog et al., 1999; Schoenbaum et al., 1999). Interestingly, all areas that are key for reversal learning have been shown to be impaired in stimulant-dependent individuals (SDIs). Altered metabolism in the OFC is impaired in both cocaine and methamphetamine users and neural responses to drug-associated cues are abnormal in cocaine users (Volkow and Fowler, 2000). Altered neural signals in response to drug-associated cues are also found in the BLA (Carelli

and Wondolowski, 2003). In cocaine-treated rats, cue-selective neurons fail to reverse their cue-selectivity after a reversal (Stalnaker et al., 2009). As mentioned previously, the DS and VS are also key regions involved in drug addiction; altered metabolism has been found in the striatum of subjects reporting craving (Volkow et al., 1991, 2001, 1999). In summary, it is expected that in both rats and humans that have repeatedly consumed a drug, the OFC, BLA and striatum would show altered activity, which is linked to reversal learning impairments.

In this chapter, a study is presented in which rats underwent a second-order schedule of reinforcement with cocaine accompanied by a mild electric foot-shock in later stages of training. Prior to cocaine SA, animals were extensively tested on a variety of behavioural tasks, including DRL. Additionally, the animals' brains were scanned at multiple time points throughout their development, both in adolescence and adulthood, prior to the drug-seeking procedure. This allowed the exploration of behavioural and neural features that may underlie compulsive drug-seeking. The DRL data was analysed using a computational approach, making it possible to relate RL measures to compulsivity and identify which areas of the brain were associated with these parameters. These results extend the findings from chapter 3 to another rodent cohort and help identify which behavioural and structural markers predict compulsive drug-seeking.

4.2 Methods

4.2.1 Subjects

Fifty-two male lister-hooded rats (Charles River, Kent) were used in this study. After arrival, pregnant dams were housed in temperature and humidity controlled, ventilated cages. After giving birth spontaneously, male offspring of the same dam were housed in groups of four and the dam sacrificed. All animals were kept on a 12-hour reverse light/dark cycle. Animals had ad libitum access to food and water. Throughout behavioural training, animals were kept to 85% of their free feeding weight. Upon completion of behavioural testing and intravenous catheter surgery, animals were singly housed for the duration of the experiment. All experiments were carried out in accordance with the UK Animals Scientific Procedures act (1986) under the UK Home Office project licenses (PPL 70/7587 & PPL 70/8072) and were approved by the University of Cambridge Ethics Committee. The data were collected by other members of the Dalley group, including Dr Jolyon Jones and Dr Peter Zhukovsky; the analysis of reversal learning behavioural data was done together with Dr Peter Zhukovsky and the analysis linking structural MRI to RL parameters was done by myself.

4.2.2 Drugs

Cocaine hydrochloride (NIDA Drug Supply Programme) was dissolved in sterile 0.9% saline. Drug doses are reported in the salt form.

4.2.3 Reversal learning task

Rats completed training and testing of a spatial discrimination reversal learning task at PND64. The task was completed in one of twelve operant chambers places in ventilated, sound-attenuated cubicles. Subjects were habituated to the apparatus over two days, twenty minutes per day. Afterwards, they were trained to enter the magazine and trigger the onset of a single stimulus light, either on the left or on the right. They learnt how to respond to the stimulus via a lever press and were rewarded with a food pellet. Once a subject achieved fifty correct responses, they were only rewarded on every second correct response (FR2), and subsequently every third response (FR3). Failure to respond within a 30 s period resulted in a 5 s time-out. As soon as the subject achieved criterion, they were tested on a spatial discrimination task. On the next day, there was a reversal of the stimulus-reward contingency. Animals had 1 h to complete the discrimination task.

4.2.4 Intra-jugular surgery

A home-made indwelling catheter was implanted into the right jugular vein of rats under isoflurane anaesthesia (O2 carrier gas; 2 L/min; 5% for induction and 2-3% for maintenance and analgesia) (Metacam, 1mg/kg, sc., Boehringer Ingelheim). After surgery, rats received daily oral treatment with an analgesic for three days and an antibiotic (Baytril, 10mg/kg, Bayer) for one week. Catheters were flushed with 0.1 ml of heparinized saline (50 U/ml, Wockhardt®) in sterile 0.9% NaCl every other day after surgery and then before and after each daily self-administration session.

4.2.5 Drug self-administration

The experiments were conducted in thirty-six standard operant conditioning chambers (Med Associates, St. Albans, VT, USA), which were placed within a sound-attenuating box to eliminate background noise. Each chamber was equipped with two retractable levers (4 cm wide, 12 cm apart, and 8 cm from the grid floor), a cue light (2.5 W, 24 V) above each lever, and a white house light (2.5 W, 24 V) at the back of the chamber, in front of the levers. Silastic tubing shielded with a metal spring extended from each animal's IV catheter

to a liquid swivel (Stoelting, Wood Dale, IL, USA) mounted on an arm fixed outside of the operant chamber. Tygon tubing extended from the swivel to a Razel infusion pump (Semat Technical, Herts, UK) located adjacent to the external chamber. The MED-PC IV software was used to control lever presses, light stimuli presentation, reward delivery and was used for data collection.

Rats were trained to acquire cocaine self-administration (0.25 mg/100µl/5.7s/infusion) under continuous reinforcement over four 2-hour sessions on different days. Each active lever press led to drug infusion alongside a 20 s timeout, a 20 s illumination of the cue light, which was placed above the active lever (conditioned stimulus), followed by the offset of house light and retraction of both levers. When the inactive lever was pressed by the animal, this was recorded, but did not result in any further outcomes. Active and inactive lever assignment was counterbalanced, and a maximum of 30 infusions were available at this stage.

After four days of sessions with continuous reinforcement, the schedule of reinforcement was changed to fixed intervals. These increased daily from 1 min (FI1), to 2 min (FI2), FI4, FI8, FI10 and FI15. Following three FI15 schedule or reinforcement sessions, rats were trained to seek cocaine under the control of the drug-paired conditioned stimulus for thirty sessions under a FI15 second-order schedule of reinforcement.

After five daily sessions under the second-order schedule of reinforcement, drug-seeking behaviour was punished in the last seven minutes of the interval if the animal was actively engaging in responding for the drug. During this time period, mild electric foot-shocks (1 s duration, 0.25-0.45 mA) were dispensed by a scrambler (Med Associate St. Albans, USA) connected to the grid floor of the operant boxes were delivered on every sixteenth lever press. The aversive stimuli were paired with a cue light, which was located on the top middle region of the wall. Following these punishment sessions, rats were re-introduced to five baseline second-order schedules of reinforcement sessions to explore whether they would recover their initial drug-seeking behaviour, or whether they would show long-term behavioural adaptations to successive punishment sessions. A K-means cluster analysis was carried out on the number of shocks rats were willing to receive during the first, drug-free, interval of the last two punishment sessions to identify non-overlapping populations stratified on their propensity to persist in seeking cocaine in the face of adverse consequences.



Fig. 4.1 Study timeline.

Fifty-two rats were scanned on PND63. They then underwent behavioural phenotyping, which included the deterministic reversal learning task. Following the behavioural phenotyping in adulthood, rats were scanned. During self-administration, the rats were trained on a fixed ratio of one (one lever press equates to one reward) during the acquisition stage. In the fixed interval stage, the rats would have to wait a certain period of time before being rewarded with cocaine (e.g., 15 minutes). In the second-order stage, the subjects had to wait 15 minutes to receive a reward, but they had to press the lever 10 times to receive cocaine. During reward presentation after the 10th lever press, a flash of light was presented. During the shock stage, the FI15(FR10:S) schedule still applied but if the subjects pressed the lever 16 times, an electric shock was given. However, this was only done in the last 7 minutes of each 15-minute interval. The rats underwent 5 days of shock training, receiving a 0.25 mA shock on the first day, 0.35 mA on the second and third days and 0.45 mA on the last two days. After the shock stage, the rats were re-trained using a FI15(FR10:S) schedule.

4.2.6 Statistical analysis of behavioural parameters

The conventional and modelling parameters were centred. Linear mixed-effects models were fit to the measures, containing which compulsivity group (low, medium, high) they were assigned to as a fixed factor. The LME also contained a random intercept for each subject. Subsequent post-hoc pairwise comparisons of estimated marginal means were run (Lenth et al., 2018; Pinheiro et al., 2021). The residuals were checked for normality using the Shapiro-Wilk test, which was verified in our data and confirmed an assumption required

for fitting an LME model (Schielzeth et al., 2020). The relationship between the RL and conventional measures was investigated through a correlation analysis using Spearman's rank correlation coefficient, as the data were non-normally distributed. Normality was tested for using the Shapiro-Wilk test. Multiple comparisons were accounted for using FDR correction (Benjamini and Hochberg, 1995). All statistical analyses were run in R: A Language and Environment for Statistical Computing (R Core Team, 2020). For all analyses, significance was considered to be at p=0.05.

4.2.7 Image acquisition

Animals were scanned using a 9.4 T horizontal bore MRI system (BioSpec 94/20, Bruker, Coventry, UK). Structural images were obtained using on a three-dimensional multi-gradient echo sequence (TR/TE 25/2.4 ms with 6 echo images spaced by 2.1 ms, flip angle 6° with RF spoiling of 117°). Contrast between grey and white matter was enhanced by applying MT pulses (10 µT, 2 kHz off-resonance) within each repetition. The field of view of 30.72 $\times 25.6 \times 20.48$ mm³ was constrained within a matrix of $192 \times 160 \times 128$ voxels, yielding an in-plane isotropic resolution of 0.16 mm² with a total scan time of 6 min 36 s with zerofilling acceleration (25% in the readout direction; 20% in each phase encoding direction). Throughout all scans, rats were anaesthetised with isoflurane (1-2% in 1L/min, O2:air 1:4). Respiratory rate and pulse oximetry (SA instruments; Stony Brook, NY) were measured and anaesthetic dose rates were adjusted keeping these within physiological range. Body temperature was measured with a rectal probe and a heated water system was adjusted to maintain 36-37°C. The data were collected at the Department of Psychology and the Small Animal Neuromaging Facility, West Cambridge, as part of an MRC-funded study awarded to Barry Everitt, Jeffrey Dalley, Trevor Robbins, Amy Milton and David Belin (MR/N02530X/1).

4.3 **Results**

4.4 Behavioural findings

4.4.1 Deterministic Reversal Learning

Behaviour on the DRL task was initially assessed using measures such as win-stay, lose-shift, perseverative errors, and total trials to criterion. LMEs were fit to these data. There

were no statistically significant differences between the low-, medium- and high-compulsive rats on win-stay (p=0.94), lose-shift (p=0.15), trials to criterion (p=0.55) or perseverative errors (p=0.65). A summary of these results can be found in Fig. 4.1.

4.4.2 RL model comparison

The BIC, log-likelihood ratio and pseudo r^2 (section 2.1.4) were used as measures to compare the computational models tested and identify which model fitted the data the best. In total, five models were run and compared with each other; they included a 2-parameter model (including α and β parameters); a 3-parameter model (α , β and κ parameters); a 4-parameter model (α_{rew} , $\alpha_{non-rew}$, β and κ parameters); a 3-parameter 'forgetting' model, which introduces a 'forgetting' mechanism for the non-chosen stimulus; and a 4-parameter 'differential forgetting' model, which has a separate learning rate term for the non-chosen option.

The QL 3-parameter model with no 'forgetting' term was used as the baseline for comparison. The 2-parameter, 4-parameter, 'forgetting' and 'differential forgetting' models were compared with the 3-parameter model. The 3-parameter model was found to fit the data the best, as the BIC values were lower for the 3-parameter model than for the other models (Fig. 4.2). Furthermore, on the log-likelihood test, most animals crossed the significance threshold (which is 3.842 for p=0.05) when comparing the 3-parameter model to another model, contributing additional evidence for the 3-parameter model providing the best fit. Additionally, the pseudo r^2 was the highest for the 3-parameter model for most animals. Figures for all model comparisons can be found in the Appendix (section A.2).



Fig. 4.2 Key findings of a conventional analysis of the deterministic reversal learning task. Conventional measures in low-, medium- and high-compulsive rats. No statistically significant differences were found. LC - low-compulsive; MC - medium-compulsive; HC - high-compulsive.



Fig. 4.3 RL model comparison.

A) BIC comparison of the 3-parameter and 2-parameter model, B) BIC comparison of the 3-parameter and 4-parameter model, C) BIC comparison of the 3-parameter and 3-parameter 'forgetting model', D) BIC comparison of the 3-parameter and 4-parameter 'differential forgetting' model. Dashed lines indicate a BIC difference of 2; continuous lines represent a BIC difference of 4.

The dotted lines in Fig. 4.2 represent the threshold for a small, but insignificant evidence for one model over the other, and the continuous lines represent the threshold for substantial evidence for one model compared to the other. Based on these figures, it can be seen that the 3-parameter model fits the data the best. Model fit for the 3-parameter model is significantly better than the 4-parameter model, since for most animals, the BIC is greater for the 4-parameter (92%) (Fig. 4.2.B). The likelihood ratio test confirms this, as the difference between the two models crosses the significance threshold for the majority of animals. When comparing the 2- and the 3-parameter models, fewer animals have a greater BIC on the 2-parameter model (51%), however, the animals that do, show significantly greater evidence for the 3-parameter model (23% have a BIC difference greater than 4, whereas none of the animals that have a lower BIC for the 2-parameter model cross that threshold) (Fig. 4.2.A).

The log-likelihood test crossed the threshold of 3.842 in favour of the 3-parameter model for some animals, indicating that this model provides a better fit (for more details, see Appendix). This trend can also be observed when comparing pseudo r^2 values. The evidence for the 3-parameter vs the 3-parameter 'forgetting' model is insignificant across all measures (50% of animals show better fit of the 3-parameter model), indicating that both models provide good fit. The BIC values for most animals are greater for the 'differential' forgetting model than for the 3-parameter model, suggesting that the 3-parameter model fits the data better (81% show better model fit of the 3-parameter model). It was decided to proceed with the analysis of the 3-parameter model without a 'forgetting' term, as it is has been published in the context of cocaine-seeking in rats previously (Zhukovsky et al., 2019).

4.4.3 Reinforcement learning

Following the model comparisons, the RL parameters from the winning model with three parameters were statistically analysed in order to explore whether they differed between low-, medium- and high-compulsive rats. The RL parameter α , which is the combined learning rate for reward and punishment, did not significantly differ based on compulsivity (p=0.81). β , the exploration vs exploitation parameter, was also the same across the three groups (p=0.54). However, κ , the autocorrelation parameter representing the tendency to select the same stimulus regardless of previous outcomes, was found to be significantly affected by compulsivity (F(2,35)=3.34, p=0.047). Specifically, the high-compulsive group had the highest levels of κ (M=0.82, SD=0.26), followed by the medium-compulsive animals (M=0.68, SD=0.33) and low-compulsive animals having the lowest κ values (M=0.50, SD=0.23). Pairwise comparisons found that this effect was driven by low- and high-compulsive groups having significantly different κ values (t(35)=2.46, p=0.049). A summary of these results can be found in Fig. 4.4.





RL parameters in low-, medium- and high-compulsive rats. Alpha (α) is the learning rate from rewarded and non-rewarded trials on the DRL task; beta (β) is the inverse temperature parameter and kappa (κ) is the autocorrelation. LC - low-compulsive; MC - medium-compulsive; HC - high-compulsive. * – p<0.05

4.4.4 Bayesian Learner model comparison

For the comparison of the BL models, different models with varying memory sizes were tested. The memory size affects how many previous trials the model has access to when calculating the likelihood of the reward sequence. The model with a memory size of 1 was used as the baseline for comparison. The results can be found in Fig. 4.5.



Fig. 4.5 BL model comparison.

A) Bayesian Information Criterion comparison of the Bayesian Learning models with a memory size of 1 and 4, B) BIC comparison of the BL models with a memory size of 1 and 7, C) BIC comparison of the BL models with a memory size of 1 and 10. The dashed lines represents a BIC difference of 2; the continuous lines represent a BIC difference of 4.

As can be seen in the figures, increasing the memory size improved model fit, as reflected by lower BIC values. This was further supported by pseudo r^2 and log-likelihood measures. Although no animals crossed the threshold of 3.842 for the model with a memory size of 4, as the memory sizes were increased to 7 and 10, more animals crossed that threshold. When comparing the model with memory size 1 versus model with memory size 4, 79% of animals had a lower BIC for the model with memory size 4, however, the BIC difference was never greater than 2. When increasing the memory size to 7 and 10, the same percentage of animals had a lower BIC for the model with a larger memory size, but more animals had a BIC difference close to or greater than 4. The model with a memory size of 10 was chosen, as this was found to be the best-fitting model.

4.4.5 **Bayesian Learning results**

The BL model did not reveal any additional insights into the behavioural mechanisms underlying compulsivity. The Hazard rate, which determines how much the posterior changes based on the data-driven likelihood versus how much it depends on the uniform prior distribution, was unaffected by higher levels of compulsivity (p=0.40). The β and κ parameters were also unaffected (p=0.39 and p=0.14, respectively). The results can be found in Fig. 4.6.

4.4.6 Simulations

The behavioural DRL data were simulated with the extracted parameters from the winning 3-parameter RL model. The simulated data were then analysed using a conventional approach. No statistically significant differences between any of the simulated conventional DRL measures were found, including win-stay, lose-shift, perseverative responding and total trials to criterion. This suggests that the winning model could successfully replicate the original DRL data.

4.4.7 Behavioural correlations

In order to identify how the RL and DRL measures relate to each other, these measures were correlated using Spearman's correlation coefficient, as the measures were non-normally distributed. After correcting for multiple comparisons using FDR correction, a significant positive correlation between α and win-stay was found (r(48)=0.40, p=0.013). Furthermore, α was negatively correlated with total trials to criterion and number of perseverative errors (r(48)=-0.36, p=0.024 and r(48)=-0.67, p<0.001, respectively). It was also positively correlated with lose-shift (r(48)=0.39, p=0.015). β was also positively correlated with lose-shift behaviour and negatively correlated with the number of perseverative errors (r(48)=0.54, p<0.001; r(48)=-0.36, p=0.024). κ had a significant positive correlation with win-stay (r(48)=0.57, p<0.001) and a negative correlation with lose-shift behaviour (r(48)=-0.43, p=0.007). Moreover, it was found to be negatively correlated with the trials needed to reach criterion (r(48)=-0.36, p=0.024).





Bayesian Learner parameters in low-, medium- and high-compulsive rats. The Hazard rate, which determines what the posterior is most affected by; β is the inverse temperature parameter and κ is the autocorrelation. LC - low-compulsive; MC - medium-compulsive; HC - high-compulsive. * – p<0.05

As discussed in the previous chapter, κ is representative of the tendency to select the same response regardless of previous reinforcers, therefore, greater κ values may lead to higher 'stickiness' and thus higher 'staying'. The positive correlation between α and win-stay indicates that animals that learn more quickly from the feedback on previous trials tend to stay after a win rather than shifting their response and exploit the information they already know about that stimulus. Additionally, these animals reach a criterion more quickly and make fewer perseverative errors.

Interestingly, correlations between BL parameters and conventional PRL measures were also observed. The BL β parameter was correlated with lose-shift, number of perseverative errors and total trials to criterion (r(48)=-0.40, p=0.013; r(48)=0.52, p=0.001; r(48)=0.55, p<0.001). Both the BL β parameter and Hazard rate were correlated with the RL α parameter, although the former had a negative correlation and the latter a positive one (r(48)=-0.60, p<0.001; r(48)=0.43, p=0.007). The BL κ parameter, on the other hand, had a positive and negative relationship with win-stay and lose-shift (r(48)=0.45, p=0.004; r(48)=-0.76, p<0.001), as well as a negative and positive correlation with the RL β and κ parameters (r(48)=-0.40, p=0.013; r(48)=0.61, p<0.001). A summary of all the statistics of these correlations can be found in Table 4.1 and Fig. 4.7.

Table 4.1 Summary table of statistics for the correlation analysis between conventional PRL and modelling parameters.

	Win-stay	Lose-shift	Perservative errors	Total trials to criterion	Alpha	Beta	Kappa	Hazard Rate	Beta BL	Kappa BL
Win-stay	-	0.173	0.317	0.001	0.013	0.317	0.000	0.319	0.103	0.004
Lose-shift	-0.224	-	0.000	0.776	0.015	0.000	0.007	0.505	0.013	0.000
Perservative errors	-0.164	-0.595	-	0.098	0.000	0.024	0.107	0.124	0.001	0.059
Total trials to criterion	-0.523	0.042	0.273	-	0.024	0.204	0.024	0.059	0.000	0.136
Alpha	0.398	0.386	-0.672	-0.36	-	0.001	0.688	0.007	0.000	0.234
Beta	-0.163	0.541	-0.361	0.207	0.513	-	0.688	0.015	0.074	0.013
Карра	0.567	-0.427	0.263	-0.362	0.069	-0.063	-	0.053	0.6	0.000
Hazard Rate	0.16	0.112	-0.251	-0.306	0.431	0.389	0.316	-	0.000	0.688
Beta BL	-0.267	-0.403	0.523	0.547	-0.603	-0.291	-0.089	-0.719	-	0.204
Kappa BL	0.452	-0.757	0.308	-0.242	-0.194	-0.398	0.608	0.063	0.208	-

Summary of r-values and p-values for the correlation analysis. The top right half of the table represents the p-values, and the bottom left shows the r-values. Cells highlighted in red show p-values < 0.05. BL - Bayesian Learner.



Fig. 4.7 Correlation matrix representing the relationship between conventional DRL measures and parameters from the BL and RL models for all animals.

The colour bar on the right-hand side represents the correlation strength. Alpha – α parameter representing the learning rate from rewarded and non-rewarded trials; Beta – the exploitation/exploration parameter β ; Kappa – κ the stickiness parameter; BL - Bayesian Learner.

4.5 Imaging results

4.5.1 Structural MRI

Seven ROIs were selected to conduct a volumetric analysis on. The brain volumes for all animals from those ROIs were extracted for two time points, at PND 63 and between PND220-240, the latter being shortly before cocaine SA. This would highlight which structures may have been different before cocaine exposure between high-compulsive animals and low-compulsive animals. The seven ROIs included the NAcC, NAcS, CPu, AI, IL, PrL as well as the BLA. Figures for the latter three ROIs can be found in chapter 3 (Fig. 3.2), the NAcc core and shell, CPu and AI can be found in Fig. 4.8.



Fig. 4.8 Masks of the regions used for the volumetric analysis.

A) nucleus accumbens core (X:91; Y:140,142,144; Z:32), B) nucleus accumbens shell (X:91; Y:138,140,142; Z:32), C) caudate putamen (X:101; Y:130,135,140; Z:44), D) anterior insular cortex (X:104; Y:141; Z:30).

LMEs were fitted to the volumes from each ROI separately. For all selected ROIs, only an effect of TP was observed, which is attributable to brain development still taking place between PND63 and PND220. No compulsivity×TP interactions were found, except for the NAcC (F(2,33)=3.17, p=0.055). This was mostly driven by an effect of brain growth between the two time points. The NAcC volumes were lower in the high-compulsive group (M=6.30, SD=0.45) at the second TP than the medium- and low-compulsive groups (M=6.63, SD=0.30 and M=6.46, SD=0.35, respectively), but this effect was non-significant (Fig. 4.9). Detailed information on the LME results for all ROIs can be found in the Appendix.

4.5.2 Correlations between RL parameters and structural volumes

In order to find out whether any of the selected ROIs at PND63 correlated with the Q-learning parameters, the ROI volumes for each animal were correlated with the parameters α , β and κ using Spearman's correlation coefficient. The brain volumes only from TP1 were used, as that is when behavioural testing took place and therefore any behavioural differences would relate to structural volumes at that time point. Following FDR correction for multiple comparisons, only the AI volume significantly correlated with one of the parameters, namely, the exploitation vs exploration parameter β . There was a positive correlation between these two measures (r(45)=0.37, p=0.031). This suggests that the AI is involved in mediating behaviours underlying the β parameter. A follow-up of the relationship between the AI and



Fig. 4.9 Volumetric differences of the nucleus accumbens core volume between low-, mediumand high-compulsive animals.

NAcbC volumes for low-, medium- and high-compulsive animals, at PND63 and PND220-240. There is a significant compulsivity×TP interaction, however, this was mostly driven by differences in volume between the two time points. LC - low-compulsive; MC - medium-compulsive; HC - high-compulsive. (PND63 - LC: n=13; MC: n=16; HC: n=7; PND220-240 - LC: n=14; MC: n=9; HC: n=7.)



 β using an analysis of covariance (ANCOVA) did not find any interactive effects between compulsivity and β on AI volume. A results summary is shown in Fig. 4.10.

Fig. 4.10 Correlation matrix representing the relationship between RL parameters and ROI volumes for all animals.

The colour bar on the right-hand side represents the correlation strength. Alpha – α parameter representing the learning rate from rewarded and non-rewarded trials; Beta – the exploitation/exploration parameter β ; Kappa – κ the stickiness parameter. CPu - caudate putamen; NAcC - nucleus accumbens core; NAcS - nucleus accumbens shell.

4.6 Discussion

In this study, rats that underwent cocaine SA were behaviourally phenotyped prior to cocaine exposure, and brain scans were taken at multiple time points during development. As part of the behavioural testing, animals completed a DRL task. Based on animals' behaviour during cocaine SA, animals were classified into low-, medium- and high-compulsive groups. A conventional analysis of the DRL task highlighted no behavioural differences between the three groups of animals, as measured using trials to criterion, win-stay, lose-shift and perseverative errors. Five reinforcement learning models were fit to the DRL data, with the winning model including the parameters α , β and κ . It was found that the parameter

 κ , which is representative of an animals' response stickiness, was significantly higher in high-compulsive compared to low-compulsive animals. This parameter was positively correlated with win-stay behaviour and negatively correlated with lose-shift behaviour. These correlations between κ and win-stay/lose-shift were also found in chapter 3, suggesting that despite κ being a measure of value-independent stickiness, it is reflected in lower shifting behaviours after both a win and a loss.

The structural MRI results made it possible to link the various computational parameters to neural substrates. Volumetric group differences were tested on seven ROIs, which included the NAcS and NAcC, CPu, AI, IL and PrL and the BLA. No differences between the different compulsivity groups were seen except in the NAcC, which was affected by a compulsivity×TP interaction. Brain volumes in the NAcC were lower in the high-compulsive group than in the other animals. Additionally, the volumetric data was related to the RL parameters, and the AI volume was found to be positively correlated with the RL parameter β . This provides a novel insight into which region mediates this parameter.

One of the RL models that was fitted to the data, the 'forgetting' model, also provided a good fit to the DRL data. This model included three parameters: a learning rate α , reward sensitivity μ_1 and punishment sensitivity μ_2 . It did not include the two parameters β and κ in the softmax function, compared to the other 3-parameter model. This model was not selected for further analysis as it has not yet been published in the context of compulsivity and would therefore be difficult to interpret. However, the good model fit suggests that it was also able to capture the DRL data well and that the β and κ parameters are not necessarily required for explaining behaviour. The inclusion of the 'forgetting' mechanism may be the reason for improved model fit and appears to be an important aspect of behaviour on this task. In short, it may indicate that 'forgetting' of the non-chosen stimulus may be an important component of RL processes. The BL model also provided good fit to the data, however, none of the parameters from this model were found to be affected by different levels of compulsivity. Model fit was improved through increasing memory size, which gives the model access to a greater number of trials when calculating the Bayesian likelihood. Therefore, it appears that many previous trials influence decision-making. Incorporating memory size into RL models may thus also improve model fit and could capture behavioural dynamics even better.

Increased perseveration following illicit substance abuse has been reported across species. In rats, increased κ and heightened perseveration have been found following cocaine exposure, effects that have also been observed in humans (Allen and Leri, 2014; Ersche et al., 2008; Kanen et al., 2019; Zhukovsky et al., 2019). Here, we show that higher levels of κ are observed prior to the emergence of compulsive drug-taking, a finding that has not been made

in rats before and would be difficult to determine in humans. This suggests that increased κ may *predict* compulsive drug use. In marmosets, it has been shown that excitotoxic lesions of the medial striatum or OFC, but not of the amygdala, result in perseverative impairments on reversal learning paradigms (Clarke et al., 2008). Similar observations have been made in rats, with increased perseverative responding following lesions of the DMS (Castañé et al., 2010). These studies indicate which areas might be underlying κ . Additionally, striatal and OFC regions have been shown to be affected in SUD in humans (Volkow et al., 1991, 1999, 2001). Specifically, activity in the OFC and its connections to brain areas including subcortical structures and the PFC is thought to be involved in impaired decision-making and drug craving in SUD (London et al., 2000). Therefore, higher κ values may be an early indicator and precursor of compulsive drug abuse.

The NAc is a key part of the brain's reward system and is involved in drug addiction (Koob and Volkow, 2009). In this chapter, we found a lower volume of the NAcC in high-compulsive compared to low- and medium-compulsive animals. In heroin-dependent individuals, reduced NAc volumes have been reported, assessed using MRI methods as well as post-mortem analyses (Müller and Homberg, 2015; Seifert et al., 2015). NAc volume was up to 16% lower in individuals with heroin-use disorder. NAc volume was additionally found to be negatively correlated with addiction severity. Although our findings in rats indicate volumetric differences before drug exposure, these findings provide support for differences in brain structure prior to the onset of SUD. Even though it would be difficult to investigate changes in brain volume prior to SUD onset in humans, findings from rodents are helpful in investigating neural and behavioural markers that may predispose to and predict substance abuse.

Only few studies have aimed to identify the neural substrates of RL parameters. Verharen and colleagues pharmacologically inactivated five PFC regions, namely the PrL, IL, IOFC, mOFC and ACC and modelled reversal learning data using RL models to explore how parameters were affected by reversible lesions (Verharen et al., 2020). Even though some of these inactivations lead to changes in RL parameters such as α_{rew} , α_{pun} and κ , none of these areas seemed to modulate the β parameter. Amygdala and OFC 5-HT depletion do not affect the β parameter; but instead alter the other model parameters such as the learning rate and stickiness (Rygula et al., 2015). Thus, the finding that AI volume is correlated with the β parameter provides a new, unexpected insight into the neural correlates of RL. However, it should be considered that this region does not have clearly defined boundaries in the rodent brain, raising the question of how reliable the use of this ROI is. The area was based on the Paxinos and Watson rodent atlas, therefore, it should to a certain extent truly reflect this region. Nonetheless, future studies should aim to verify this and confirm whether the boundaries are not altered as a result of warping, an essential component of MRI pre-processing. The key roles of the AI are in its involvement in affective processing, including pain responses, expressions to disgust, fear, anxiety, and happiness (Uddin et al., 2017). Insular dysfunction has been implicated in many clinical conditions including SUD (Naqvi and Bechara, 2009). Since β can be interpreted as sensitivity to reinforcement and due to the newly found link to the AI, β may be interpreted as the affective component of RL. The following chapters aim to replicate the findings made in this and the previous chapter in humans. This will help to identify how translatable the behavioural results are, and whether the brain regions implicated in these measures are homologous across species.

Chapter 5

Reinforcement learning in patients with gambling disorder and substance use disorder

5.1 Introduction

The diagnoses for substance use disorder and gambling disorder both include criteria for unsuccessful attempts to stop substance abuse or gambling, jeopardising relationships and educational/career opportunities and financial troubles as a consequence of the disorder (APA, 2013). Since the DSM-5, GD has been categorised as a behavioural addiction. Compulsivity is defined as persistent actions inappropriate to a given situation, which have no clear relationship to the overall goal and frequently result in undesirable consequences (Dalley et al., 2011). Both GD and SUD are considered to be disorders of compulsivity, sharing similarities but also some differences (Leeman and Potenza, 2012; Robbins, 2012). The overlapping and diverging aspects of these two disorders of compulsivity, behaviourally and neurally, have yet to be clearly defined. Understanding of this could better inform development of new treatment strategies for compulsive disorders.

Behavioural inflexibility, a deficit in adapting to contingency changes in the environment, is a key feature of SUD (Smith et al., 2020). On the PRL task, stimulant-dependent individuals reportedly show increased perseverative responses to the previously correct stimulus after a contingency change (Ersche et al., 2011). This deficit can be normalised using the DA $D_{2/3}$ receptor agonist pramipexole. However, increased perseverative responding is not found in subjects with OCD, suggesting that this is not a behaviour seen across compulsive

Reinforcement learning in patients with gambling disorder and substance use disorder

disorders generally (Ersche et al., 2011; Smith et al., 2015). Interestingly, cocaine-dependent individuals, but not amphetamine-dependent individuals, perseverate on reversal learning tasks (Ersche et al., 2008). In rats, deficits in reversal learning as a result of cocaine exposure over a period of 2 weeks have been observed (Jentsch et al., 2002). Methamphetamine exposure in rats resulted in similar impairments (Cheng et al., 2007; Izquierdo et al., 2010). In monkeys, protracted exposure to cocaine resulted in long-lasting deficits in reversal learning (Jentsch et al., 2002). Overall, evidence from several species indicates long-term reversal learning deficits in the form of excessive perseveration arising from cocaine exposure. In humans, cocaine use in particular has been associated with maladaptive perseveration (Ersche et al., 2008).

Inflexible responding on cognitive flexibility tasks has also been observed in pathological gamblers (PGs) (Jara-Rizzo et al., 2020; Perandrés-Gómez et al., 2021). Specifically, patients with GD struggle to learn new stimulus-outcome associations following contingency reversals. In the face of repeated negative outcomes, PGs are more likely to stay rather than switch their response on a PRL task or switch prematurely following little or no negative feedback. This suggests maladaptive responding in GD after feedback, as well as behavioural inflexibility when there are changes in the environment (Perandrés-Gómez et al., 2021). An additional study investigated behavioural differences on the IED task, which is also used to measure cognitive flexibility, in PGs between the ages of 18-29 (Leppink et al., 2016). Similarly to the PRL task, participants with GD performed worse on this paradigm and a positive correlation between IED errors and gambling severity was found, as well as a positive correlation with self-reported measures of compulsivity. A meta-analysis including nine studies employing the WCST showed that individuals with GD made more perseverative errors than control participants (van Timmeren et al., 2018). Thus, patients with GD are generally impaired on tasks of cognitive flexibility and show increased perseverative tendencies, as are patients with SUD.

Neuroimaging studies in GD report that prefrontal cortical areas are differentially recruited in reward processing (Leeman and Potenza, 2012). Perseveration on a PRL task has been linked to lower activity in the right vlPFC after positive and negative feedback (Ruiter et al., 2008). The vmPFC, which is active in healthy participants on monetary reward tasks such as the IGT and loss-chasing tasks, is important for tracking expected reward outcomes and reward processing (Campbell-Meiklejohn et al., 2008; Li et al., 2010). In GD, however, vmPFC activation is lower than in healthy individuals on such tasks and activity has been shown to be negatively correlated with gambling severity (Habib and Dixon, 2010; Reuter et al., 2005). On the IGT, PGs showed greater activity in the right caudate, OFC, vmPFC, superior frontal gyrus, amygdala and hippocampus during high-risk choices (Power et al., 2012). These findings point to altered reward processing in GD and suggest the involvement of cortical areas such as the vmPFC and OFC, as well as subcortical structures.

Similarly to PGs, cocaine abusers reportedly exhibit lower BOLD signals in response to monetary gain or loss in the OFC compared to healthy controls (Goldstein et al., 2007). On the IGT, on the other hand, the right OFC shows *stronger* signals, whereas the dorsomedial and medial PFC have weakened activity than in control participants (Bolla et al., 2003). Reversal learning is known to be mediated by circuits involving the dmPFC, dlPFC and vlPFC, amygdala and striatum (Cools et al., 2002). Altered PFC and striatal activity has been found during reversal learning tasks in cocaine abusers. For example, it has been reported that SDIs have reduced activity in the caudate nucleus and MFG following perseverative errors (Ersche et al., 2011). Additionally, resting-state functional connectivity in the dlPFC and SFG was greater in participants with CUD, and was positively correlated with reversal learning performance (Camchong et al., 2011). A meta-analysis including fifty-two studies found that OFC activity is consistently altered in cocaine abusers compared to healthy individuals across a variety of behavioural tasks (Dom et al., 2005). Thus, PFC and striatal areas seem to be affected in CUD, with some of the highlighted areas overlapping with those found to be altered in GD.

In this study, a previously published dataset from (Verdejo-Garcia et al., 2015) including patients with CUD and GD was analysed using novel computational methods. Participants completed the PRL task in an fMRI scanner. In the original publication, conventional PRL measures from the behavioural data were calculated and compared between groups. It was found that only the perseveration error rate was increased in the GD group, and no measures in CUD were different. Neurally, both patient groups showed reduced vIPFC activations on response shifting following a reversal, pointing to a common mechanism in both addictive disorders. Participants from the CUD group had a greater signal in the dmPFC than the GD group during perseveration.

The aim of this chapter was to analyse the behavioural data using RL models, which could highlight differences in the latent behavioural mechanisms due to its more sensitive trial-by-trial approach. Additionally, the imaging data provides an opportunity to relate the computational parameters to neural substrates in humans. To our knowledge, a RL analysis has not been fitted to PRL data from GD patients before, providing novel insights into another disorder of compulsivity. Furthermore, the findings in CUD could be linked to the analysis conducted in rodents in chapter 4. Based on previous publications, it was hypothesised that the winning RL model would contain at least one type of stickiness parameter, which would

Reinforcement learning in patients with gambling disorder and substance use disorder

be increased in GD and CUD. Neurally, it was predicted that the reward learning rate would be linked to activity in the amygdala and OFC, and that stickiness would be reflected in vmPFC and dACC activity.

5.2 Methods

5.2.1 Participants

Fifty-six people participated in this study. Twenty individuals met the criteria for cocaine dependence but no other Axis I or II disorder based on the DSM-IV, eighteen individuals met the DSM-IV-TR criteria for pathological gambling and eighteen control participants that did not meet the criteria for any Axis I or II disorders. The groups did not differ in age, years of education, or IQ as measured by the Kaufman Brief Intelligence Test (Kaufman and Kaufman, 2014) (Table 5.1).

Cocaine users and gamblers were recruited in the outpatient clinic Centro Provincial de Drogodependencias Granada, Spain, and Asociacion Granadina de Jugadores en Rehabilitacion Granada, Spain, respectively. The individuals had to meet the following inclusion criteria: 1) aged between 18-45 years; 2) estimated IQ level above 80; 3) meeting the DSM-IV-TR criteria for cocaine dependence or pathological gambling; 4) having commenced treatment; 5) having been abstinent for more than 15 days. Abstinence was ensured using two urine tests per week plus a test on the scanning day. Gambling abstinence was ensured by self-assessment and confirmation from relatives. The exclusion criteria included: 1) diagnosis of another Axis I or II disorder, except alcohol misuse or nicotine addiction; 2) history of head injury, neurological disease or any other diseases affecting the central nervous system; 3) having undertaken other treatments in the two years prior to the study; 5) court-mandated treatment. Axis I disorders were assessed using Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I-CV) and Axis II disorders with the International Personality Disorders Examination (IPDE) (First, 1997; Loranger, 1994). Controls were recruited from local agencies. All diagnoses were made by a registered clinical psychologist. The study was approved by the Ethics Committee for Research in Humans, University of Granada, Spain. All participants confirmed their voluntary participation by signing an informed consent form. Participants were all equally reimbursed for their time. The data were collected by Prof Verdejo-Garcia and colleagues and were kindly shared with us for this analysis. The data analysis was mostly conducted by myself, with advice from Dr Jon Kanen (Department of Psychology, University of Cambridge) and Dr Rudolf Cardinal (Department of Psychiatry,
University of Cambridge). The scripts used for RL modelling were written by Dr Rudolf Cardinal and adjusted by myself.

5.2.2 Probabilistic Reversal Learning Task

The same PRL task as in (Cools et al., 2002) was used in this study. Individuals were presented with two stimuli, which were abstract, coloured patterns, to the right and left visual fields. The location of the stimuli was randomised. Participants were informed that one stimulus was the 'correct' stimulus (CS+), whereas the other stimulus was the 'incorrect' stimulus (CS-). Participants had to learn via trial-and-error which stimulus was correct, and which was incorrect. After a certain number of trials, the contingencies would reverse, and the individual would have to learn that the other stimulus was now the 'correct' one. Given that this was a probabilistic task, the CS+ was rewarded only 85% of trials, whereas the CS- was rewarded on 15% of trials. Subjects were trained on this task prior to the scan; however, different stimuli were used. In the scanner, they had to complete three consecutive blocks of 11 mins each. Each block included 10 discrimination stages and thus 9 reversals. Contingency reversals occurred after 10 to 15 correct responses.

Stimuli were presented through magnetic-resonance-compatible liquid-crystal display goggles (Resonance Technology, Northridge, CA, USA). Behavioural responses were recorded through a five-button box, Evoke Response Pad System (Resonance Technology Inc.). Stimuli were presented for 2000 ms. Participants would have to respond during this time, otherwise, a 'too late' message was shown. Responses were made using a button box placed on the subject's chest. When a 'correct' response was made, a green smiley face was shown after the response. When an 'incorrect' response was made, a red sad face was presented. The feedback stimuli were shown for 500 ms, during this time the stimulus was still shown. After feedback presentation, there was a variable inter-trial interval. The length was determined by the programme, so that the overall interval between stimuli was 3253 ms.

5.2.3 Correlation analysis between conventional and modelling PRL measures

The relationship between the RL and conventional PRL measures was investigated through a correlation analysis using Spearman's rank coefficient, as the data were non-normally distributed, accounting for multiple comparisons with FDR correction (Benjamini and Hochberg, 1995). The statistical analyses were run in R: A Language and Environment



Fig. 5.1 Diagram of the probabilistic reversal learning task.

Task procedure is shown from bottom to top. Participants are presented with two abstract visual stimuli, and they have to learn through trial-and-error which of the two stimuli is associated with a reward. Following a certain number of correct trials, the contingencies are reversed. Feedback is presented following the subject's response and is either a smiley face or a sad face. From (Cools et al, 2002).

for Statistical Computing (R Core Team, 2020). For all analyses, significance was considered to be below p=0.05.

5.2.4 Imaging data acquisition

Participants were scanned in a 3.0 Tesla clinical MRI scanner with an eight-channel phased-array head coil (Intera Achieva, Philips Medical Systems, Eindhoven, The Netherlands). Three T2*-weighted scans using an EPI sequence were scanned initially (TR = 2000 ms, TE = 35 ms, FOV = 230×230 mm, 96×96 matrix, flip angle = 90° , 21 4-mm axial slices, 1-mm gap, 330 scans each). Next, a sagittal three-dimensional T1-weighted turbo-gradient-echo sequence was obtained (160 slices, TR = 8.3 ms, TE = 3.8 ms, flip angle = 8° , FOV = 240×240 , 1 mm³ voxels). Functional MRI scans were pre-processed as described in section 2.2.3.

5.2.5 First-level models

First-level linear models were fit to fMRI data using FEAT (FSL) (Woolrich et al., 2001). This was done for each run and included the following events: (1) reward Expected Value (EV), (2) positive Reward Prediction Error (RPE), (3) negative RPE, (3) punishment EV, (5) positive Punishment Prediction Error (PPE), (6) negative PPE and (7) feedback/cue presentation. The values for expected value and prediction errors were extracted for each trial from the winning Q-learning model. Explanatory variables 1-6 were based on the extracted values of prediction error and reward or punishment cue values. The model was based on the analysis conducted in (Murray et al., 2019). The RPE is when a reward is expected; it is positive if the reward is received, and negative when it is omitted. The PPE is when the subject expects a punishment and is positive when a reward is received and is negative when the reward is omitted. The model also contained six movement parameters (x, y, z, pitch, roll, yaw) resulting from the image realignment to control for movement artefacts.

5.2.6 Higher-level models

In the second-level models, the first-level models for each subject were averaged across the three runs. Contrasts were subsequently examined in a third-level mixed-effects wholebrain analyses involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05 (Woolrich et al., 2004). The contrasts tested group differences in each of the explanatory variables (control vs gambling disorder; control vs cocaine use disorder;

gambling disorder vs cocaine use disorder). An additional exploratory analysis was run, namely, an ANCOVA. In this analysis, the model parameters from the winning RL model were extracted for each individual and included as a variable for each event type, which are described in the previous subsection. This analysis takes into account the different groups and explores whether there are any group differences in the correlation between activity in a given region and a RL parameter. RL parameters were also correlated with BOLD signal from all subjects, regardless of group. FSLeyes was used to generate the results figures (Smith et al., 2004). In these figures, the right and left sides are inverted from the observer's perspective.

5.3 Results

5.4 Behavioural findings

5.4.1 Demographic information

No statistically significant differences in demographic features such as age, gender, IQ, handedness or years of education between the three groups were found. More detailed information is shown in table 5.1.

	Healthy controls (n=18)	Gamblers (n=18)	Cocaine users (n=20)	Group comparisons
Mean age (SD)	31.2 (4.7)	33.6 (8.0)	34.3 (6.9)	F(2,54) = 1.43, p=0.35
Gender (F)	1	2	1	F(2,54) = 0.93, p=0.77
Verbal IQ (SD)	106.9 (9.0)	102.7 (7.4)	100.9 (7.6)	F(2,54) = 2.31, p=0.082
Years of education (SD)	10.6 (1.9)	10.3 (2.1)	9.8 (1.7)	F(2,54) = 1.37, p=0.47
Handedness (L)	1	1	4	F(2,54) = 1.54, p=0.19

Table 5.1 Demographic information and results.

F - female; L - left; SD - standard deviation. A table based on the same dataset has been published in (Verdejo-Garcia et al., 2015), however, this is an original representation of the data.

5.4.2 Conventional PRL measures

There was no main effect of group on the proportion of correct responses (F(2,53)=1.41, p=0.25). Neither were there any group differences in win-stay or lose-shift behaviour, regardless of the previous trial being correct or incorrect (p=0.66, p=0.25 and p=0.66, respectively). The number of perseverative responses between healthy controls (HCs) and patients also did not differ (F(1,54)=0.26, p=0.77). A summary of the results can be found

in Fig. 5.1. These findings align with the results from the study originally published by (Verdejo-Garcia et al., 2015).

5.4.3 Choosing the winning model

Seven different RL models were fitted to the data and compared with each other in order to select the RL model that fits the behavioural data the best. Model comparison results are presented in table 5.2. All parameters and contrasts had an $\hat{R} < 1.1$, indicating satisfactory model convergence. The maximum \hat{R} in this dataset was 1.006. The best-fitting model was determined using a bridge sampling estimate of the marginal likelihood. This winning model included five parameters: the reward learning rate, which represents how quickly a subject updates its behaviour in response to positive feedback; the punishment learning rate, representing learning from punishment; reinforcement sensitivity, also known as the exploitation vs exploration parameter; stimulus stickiness, which is the tendency to select the same stimulus regardless of outcome, and side stickiness, the tendency to select the same side, irrespective of previous outcome.

Rank Name Parameters Log marginal likelihood Log posterior P 7 Model 1 $\alpha_{rew}, \alpha_{pun}, \beta, \kappa_{side}$ -13168.48 -582.60 Model 2 $\alpha_{rew}, \alpha_{pun}, \beta, \kappa_{side}, \kappa_{stim}$ -11139.35 1 0.000 2 Model 3 $\alpha_{rew}, \alpha_{pun}, \beta, \kappa_{stim}$ -13003.72 -417.85 4 Model 4 $\alpha_{rew}, \alpha_{pun}, \beta$ -13130.08 -544.21 3 Model 5 α , β , κ_{stim} -13018.42 -432.54 6 Model 6 ϕ, ρ, β -575.36 -13161.24 -573.1846 5 Model 7 α, β -13159.06

Table 5.2 Model comparison summary.

Additionally, the posteriors of the parameters from the winning model were visually inspected to ensure that the posterior distributions were properly distributed, i.e., the mean and standard deviations of each group had been properly updated from the initial prior distributions, which were beta distributions for the learning rates, a gamma distribution for the reinforcement sensitivity parameter, and normal distributions for the stickiness parameters (see Table 2.1). The posterior distributions can be seen in Fig. 5.3.

5.4.4 RL analysis

Fig. 5.4. shows the results of the hierarchical Bayesian RL analysis from the winning 5-parameter model. It shows the differences between the parameters from the model for



Fig. 5.2 Key findings of a conventional analysis of the PRL task.

(From previous page). Proportion correct responses; win-stay behaviour; lose-shift behaviour after a correct and an incorrect response and number of perseverative responses was calculated for all subjects from the three groups: healthy controls (HC), Gambling Disorder (GD) and Cocaine Use Disorder (CUD). No significant group differences were found.

the HC, GD and CUD groups. Neither of the learning rate parameters were altered in GD or CUD. The reward learning rate was significantly lower in the CUD than the GD group (difference in parameter per-group mean, posterior 75% highest posterior density interval (HDI)). This was also the case for the reinforcement sensitivity parameter (also known as the exploitation vs exploration parameter). The side stickiness parameter, the tendency to repeat a choice on the same side regardless of outcome of that option on previous trials, was also not significantly changed in either patient group, although a trend towards increased side stickiness in both GD and CUD could be seen. At 75% HDI, a decrease in stimulus stickiness in GD was found, but there was no increase or decrease in the CUD group.



Fig. 5.4 Results of the hierarchical Bayesian Q-learning model.

GD - gambling disorder; SUD - substance use disorder; HC - healthy controls; Reinf sens - reinforcement sensitivity; Stim - stimulus; HDI - highest density interval.

5.4.5 Relating conventional parameters and depression score ratings to RL parameters

Through correlating PRL and RL measures with each other using Spearman's rank coefficient, the relationships between them could be better understood. α_{rew} had a positive



Fig. 5.3 Posterior distributions of parameters from the winning 5-parameter QL model. GD - gambling disorder; CUD - cocaine use disorder; HC - healthy controls; Reinf - reinforcement; Stim - stimulus.

correlation with the proportion of correct responses, win-stay and lose-shift incorrect, as well as a negative relationship with trials required to reach a criterion (Fig. 5.5) $\alpha_{non-rew}$ was correlated with all conventional parameters except the number of perseverative responses. It had a negative relationship with proportion of correct responses and win-stay, but a positive correlation with trials to criterion and lose-shift behaviour. β was also correlated with most measures, except lose-shift after a correct response and perseverative responses. Surprisingly, κ_{side} was not correlated with any conventional measures. However, κ_{stim} had a strong negative correlation with lose-shift after a correct and incorrect response (r(55)=-0.62, p<0.001 and r(55)=-0.58, p<0.001, respectively).

The parameter β was positively correlated with α_{rew} , and κ_{stim} negatively correlated with $\alpha_{non-rew}$ (r(55)=0.46, p=0.002 and r(55)=-0.45, p=0.002, respectively). Some conventional measures were strongly correlated with each other, for example, the proportion of correct responses was negatively correlated with trials to criterion and win-stay behaviour (r(55)=-0.98, p<0.001 and r(55)=0.92, p<0.001, respectively).

5.4.6 Simulations

Simulations of the behavioural data using the values from the winning Q-learning model parameters were run to establish whether the computational models would be able to replicate participant behaviour. Indeed, when the simulated behaviour was analysed using a conventional approach, there were no statistically significant differences between the groups on any of the measures. This indicated that the model was able to capture the behavioural dynamics of three different subject groups on the PRL task.

5.4.7 DDM analysis

A DDM, as described in section 2.1.3, was also applied to the PRL task data to capture further differences in decision-making. Unlike the RL models, this model includes response times to generate parameter estimates, therefore, presenting a different perspective of behaviour on the PRL task. The four parameters in this model are: α , the boundary separation; β , the starting bias; δ , the drift rate and τ , the non-decision time. Prior to running this model, response latencies between the groups were compared to guide the DDM analysis. An LME model including solely the effect of group found no differences (F(1,54)=2.28, p=0.14) (Fig. 5.5). However, when a group×session interaction was included, the interactive factor was found to be significant (F(1,162)=4.47, p=0.011), as was the session factor (F(1,162)=82.9, p<0.001). LME model comparisons using measures such as the BIC and log-likelihood

PropCor	1	0	0	0.002	0.017	0.56	0	0.007	0	0.54	0.39	
Trials-to-crit	-0.98	1	0	0.002	0.019	0.45	0	0.005	0	0.43	0.39	- 1.00
Win-stay-corr	0.92	-0.93	1	0.048	0.063	0.12	0	0.048	0	0.25	0.68	- 0.75
Lose-shift-corr	-0.44	0.44	-0.3	1	0.001	0.72	0.47	0	0.11	0.35	0	- 0.50
Lose-shift-incorr	0.36	-0.35	0.29	0.46	1	0.35	0.007	0.002	0.009	0.089	0	- 0.25
Perseverative_resp	0.088	-0.12	0.24	-0.051	-0.16	1	0.62	0.87	0.088	0.1	0.089	- 0.00
Alpha rew	0.53	-0.5	0.5	-0.11	0.39	-0.075	1	0.49	0.002	0.43	0.22	0.2
Alpha NR	-0.39	0.41	-0.3	0.89	0.45	-0.023	-0.1	1	0.089	0.22	0.002	0.50
Beta	0.89	-0.91	0.94	-0.25	0.38	0.27	0.46	-0.26	1	0.083	0.47	0.7
Kappa side	-0.095	0.13	-0.18	-0.15	-0.26	-0.25	-0.13	-0.2	-0.27	1	0.12	
Kappa stim	0.14	-0.14	0.06	-0.62	-0.58	0.26	-0.19	-0.45	-0.11	0.24	1	
	PropCor	Trials-to-crit	Win-stay-corr	Lose-shift-corr	Lose-shift-incorr	Perseverative_resp	Apha rew	Apha NR	Beta	Kappa side	Kappa stim	

Fig. 5.5 Correlation matrix representing the relationship between conventional PRL measures and RL parameters.

The colour bar on the right-hand side represents the correlation strength using Spearman's rank coefficient. Values in cells in the top right half of the figure indicate p-values, value in bottom left indicate r-values. PropCor - Proportion of correct responses; Trials-to-crit - Trials to criterion; Win-stay-corr - win-stay after a correct reponse; Lose-shift-corr - lose-shift after a correct response; Lose-shift-incorrect - lose-shift after an incorrect response; $Alpha_{rew} - \alpha_{rew}$ parameter representing the learning rate from reward; $Alpha_{NR} - \alpha_{non-rew}$ parameter representing the learning rate from punishment; Beta – the exploitation/exploration parameter β ; Kappa_stim - κ_{side} , side stickiness parameter; Kappa_stim - κ_{stim} , stimulus stickiness parameter.

indicated that the model including the interactive term fit the data better (BIC: -263.22 vs -251.45, log-likelihood: 168.21 vs 128.67). The latency of participants with GD was lower than that of healthy controls (t(35)=3.12, p=0.002), whereas reactions times of the CUD group were higher overall (t(55)=-19.22, p<0.001). Across the three sessions, RTs decreased for all groups, however, the RT increased between the second and the third session for the CUD group only (t(39)=2.45, p=0.014) (Fig. 5.7).

The DDM model results showed that the α parameter was the same across the three groups (p=0.16), indicating that the boundary separation was the same for all participants. This was also the case for the starting bias β (p=0.86), drift rate δ (p=0.90) and τ (p=0.56) (Fig. 5.8). In summary, the DDM model did not highlight any differences in decision-making between HCs and the patient groups, even though differences were observed in the response latencies.

5.4.8 RL-DDM

In addition to the RL and DDM models explored, a novel model combining these two models was tested (the RL-DDM model). Since stimulus stickiness was found to be altered in the GD group and the latencies across sessions were different between the groups, it was hypothesised that the RL-DDM model would be able to capture these changes. The RL-DDM model substitutes the softmax distribution function, which is used to calculate the probability of choosing a stimulus on the next trial, with the WFPT, as is employed in the standard DDM. The four parameters included in this model were *a*, the decision threshold; *v*, the scaling parameter; *t* the non-decision time and *al pha*, the learning rate. Convergence was confirmed using the Gelman-Rubin statistic (\hat{R}), which indicates good convergence below 1.1. Further, it was confirmed that there was no collinearity between the parameters, as can be seen in Fig. 5.9. Next, LMEs were fit to the extracted RL-DDM parameters (*a*: p=0.29, *v*: p=0.54, *t*: p=0.56, *al pha*: p=0.79) (Fig. 5.10).





Latency distributions for all subjects from the three groups: healthy controls (HC), gambling disorder (GD) and cocaine use disorder (CUD). No significant group differences were found.





Latencies across the three runs for healthy controls (HC), gambling disorder (GD) and cocaine use disorder (CUD) groups. Across the three sessions, reaction times decreased for all groups, however, the reaction time increased between the second and the third session for the CUD group only (t(39)=2.45, p=0.014). Group 1 - healthy controls, group 2 - gambling disorder, group 3 - cocaine use disorder.



Fig. 5.8 Results from the drift diffusion model analysis.

DDM parameters from the three groups: healthy controls (HC), gambling disorder (GD) and cocaine use disorder (CUD). The parameters are α , the boundary separation; β , the starting bias; δ , the drift rate; and τ , the non-decision time. No significant group differences were found.



Fig. 5.9 Testing for collinearity between the RL-DDM parameters.

Checking for collinearity between the four RL-DDM parameters, which are: a, the decision threshold; v, the scaling parameter; t, the non-decision time; and α , the learning rate. There is only a low correlation between the parameters, thus, no obvious collinearity can be seen, indicating that the model can be used for further analyses.



Fig. 5.10 Results from the RL-DDM analysis.

RL-DDM parameters for the three groups: healthy controls (HC), gambling disorder (GD) and cocaine use disorder (CUD). The parameters are a, the decision threshold; v, the scaling parameter; t, the non-decision time; and *al pha*, the learning rate. No significant group differences were found.

5.5 Imaging results

5.5.1 Response to feedback/cue presentation

As described in the methods, the model fitted to the task-based fMRI data included seven explanatory variables: (1) reward EV; (2) positive RPE; (3) negative RPE; (3) punishment EV; (5) positive PPE; (6) negative PPE and (7) feedback/cue presentation. Differences in the last explanatory variable, which is the simultaneous presentation of feedback and cue, highlighted that participants with GD have greater activity in the lateral occipital cortex, cingulate gyrus, parahippocampal gyrus, precuneus cortex, middle temporal gyrus and supramarginal gyrus. The CUD group also had significant activations following feedback presentation, however, these were found in the frontal pole, SFG, IFG, precentral gyrus, superior parietal lobule, supramarginal gyrus, precuneus cortex, angular gyrus and lateral occipital cortex. Additionally, the CUD vs GD contrast highlighted greater activity in the insular cortex, IFG and frontal operculum during feedback presentation in the CUD group. The results figures and tables can be found in the Appendix (section B.2).

5.5.2 Brain activity during reward and punishment expected value tracking in Gambling Disorder

Furthermore, it was found that when tracking reward EV, participants with GD showed greater activations in the frontal pole, OFC, SFG, amygdala, hippocampus, parahippocampal gyrus, NAc, lateral occipital cortex, superior, inferior and middle temporal gyri, as well as the precuneus cortex and PCCx than healthy controls. This can be seen in Fig. 5.11, with information on all the activations summarised in table 5.3. The effects were lateralised to the left hemisphere.

The opposite trend was observed for the punishment EV, as control participants showed greater brain activity in the superior parietal lobule, pre- and postcentral gyri and the precuneus cortex, parietal operculum, supramarginal gyrus and angular gyrus (Fig. 5.12, table 5.4). Although the activations were seen in both hemispheres, they were more pronounced in the right hemisphere. These results indicate an imbalanced neural response to tracking expected reward and punishment values in patients with gambling disorder.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm ³)	Mean z-statistic
Middle Temporal Gyrus	21	L	-57, -6, -17	365	5893	3.63
Precuneous Cortex	7	L	-3, -66, 33	224	3617	3.68
Cingulate Gyrus	24, 32	L	-9, -50, 27	182	2939	3.71
Superior Temporal Gyrus	22, 42	L	-46, -14, -8	112	1808	3.49
Lateral Occipital Cortex	19	L	-57, -62, -6	97	1566	3.79
Hippocampus	28	L	-21, -10, -24	81	1308	3.45
Amygdala		L	-23, -5, -17	68	1098	3.41
Parahippocampal gyrus	27	L	-17, -10, -24	59	953	3.38
Inferior Temporal Gyrus	20	L	-57, -57, -13	29	468	3.42

Table 5.3 Reward expected value tracking GD vs controls contrast summary.

Summary of fMRI peak activity for the reward EV GD vs controls contrast (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). The indicated areas show greater activity in participants with GD than healthy controls. BA – Brodmann area; MNI – Montreal Neurological Institute template.



Fig. 5.11 GD vs control contrast for reward EV.

Areas of the contrast highlighting differences in tracking reward EV between participants with gambling disorder and healthy controls (MNI coordinates: X=-16, Y=58, Z=34). Activity was higher in the GD group in areas including the amygdala and hippocampus. Colour bar on the right-hand side represents t-statistic.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm³)	Mean z-statistic
Postcentral Gyrus	1, 2, 3	L	-62, -21, -33	408	6588	3.02
Postcentral Gyrus	1, 2, 3	R	56, -14, -33	1228	19827	2.99
Precentral Gyrus	4	R	43, -14, 45	865	13966	2.95
Superior Parietal Lobule	7	R	28, -44, 59	524	8461	3.04
Precuneous Cortex	7	L	-6, -48, 56	403	6507	2.90
Precuneous Cortex	7	R	10, -51, 56	555	8961	2.90
Lateral Occipital Cortex	19	L	-14, -83, 41	181	2922	2.75
Lateral Occipital Cortex	19	R	17, -79, 41	242	3907	2.88
Supramarginal Gyrus	40	L	-63, -26, 31	258	4166	3.02
Supramarginal Gyrus	40	R	56, -18, 31	254	4101	2.88
Parietal Operculum Cortex	40, 43	L	-51, -33, 21	191	3084	2.98
Parietal Operculum Cortex	40, 43	R	1.5, -33, 21	106	1711	2.83

Table 5.4 Punishment expected value tracking controls vs GD contrast summary.

Summary of fMRI peak activity for the punishment EV controls vs GD contrast (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). The indicated areas show greater activity in controls than participants with GD. BA – Brodmann area; MNI – Montreal Neurological Institute template.



Fig. 5.12 Control vs GD contrast for punishment EV.

Coronal slices of the areas in the contrast highlighting differences in tracking punishment EV between control participants and patients with GD (MNI Y = -24 to -17). Activity was higher in the control group in the highlighted areas. Colour bar on the right-hand side represents t-statistic.

5.5.3 Neural signal to positive and negative punishment prediction errors is affected in Cocaine Use Disorder

The response to positive and negative PPEs highlighted an additional imbalance in neural activity, however, this was observed in participants with CUD. The CUD group showed lower activity in the left SFG and paracingulate gyrus following positive PPE than control participants. In response to the negative PPE, however, the CUD group showed greater activation in the left SFG and left MFG than controls. These results are shown in figures 5.13 and 5.14 and tables 5.5 and 5.6, respectively. Otherwise, no differences in the other explanatory variables such as positive RPE or negative RPE were found between the groups.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm ³)	Mean z-statistic
Frontal Pole	10	L	-7, 47, 46	405	6539	2.91
Frontal Pole	10	R	9, 45, 46	44	710	2.80
Middle Temporal Gyrus	21	L	-57, -6, -17	365	5893	3.63
Superior Frontal Gyrus	8, 9	L	-2, 47, 45	292	4715	2.88
Superior Frontal Gyrus	8, 9	R	6, 44, 49	49	791	2.79
Precuneous Cortex	7	L	-3, -66, 33	224	3617	3.68
Cingulate Gyrus	24, 32	L	-9, -50, 27	182	2939	3.71
Cingulate Gyrus	24, 32	R	1, -52, 27	332	5360	3.27
Frontal Orbital Cortex	11, 12	L	-48, 26, -4	178	2874	3.01
Superior Temporal Gyrus	22, 42	L	-46, -14, -8	112	1808	3.49
Lateral Occipital Cortex	19	L	-57, -62, -6	97	1566	3.79
Hippocampus	28	L	-21, -10, -24	81	1308	3.45
Amygdala		L	-23, -5, -17	68	1098	3.41
Amygdala		R	25, -3, -18	11	178	2.80
Parahippocampal gyrus	27	L	-17, -10, -24	59	953	3.38
Inferior Temporal Gyrus	20	L	-57, -57, -13	29	468	3.42
Nucleus Accumbens		L	-11, 5, -12	21	339	2.91

Table 5.5 Positive PPE controls vs CUD contrast summary.

Summary of fMRI peak activity for the positive PPE CUD vs controls contrast (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). The indicated areas show greater activity in control participants than participants with CUD. BA – Brodmann area; MNI – Montreal Neurological Institute template.



Fig. 5.13 The control vs CUD contrast for positive PPE.

Areas of the contrast highlighting differences in the positive punishment prediction error between participants with CUD and healthy controls (MNI coordinates: X=-5, Y=17, Z=48). Activity was higher in the healthy control group in the superior frontal gyrus and paracingulate gyrus. Colour bar on the right-hand side represents t-statistic.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm³)	Mean z-statistic
Superior Frontal Gyrus	8, 9	L	-57, -6, -17	71	1146	3.41
Middle Frontal Gyrus	8, 9	L	-3, -66, 33	70	1130	3.41

Table 5.6 Negative PPE CUD vs controls contrast summary.

Summary of fMRI peak activity for the negative PPE CUD vs controls contrast (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). The indicated areas show greater activity in participants with CUD than healthy controls. BA – Brodmann area; MNI – Montreal Neurological Institute template.

5.5 Imaging results



Fig. 5.14 The CUD vs control contrast for negative PPE.

Areas of the contrast highlighting differences in the negative punishment prediction error between participants with CUD and healthy controls (MNI coordinates: X=-31, Y=30, Z=56). Activity was higher in the CUD group in the superior frontal gyrus and middle frontal gyrus. Colour bar on the right-hand side represents t-statistic.

5.5.4 Whole-brain connectivity analyses

A whole-brain connectivity analysis with the five RL parameters, which were extracted from the winning model, was run in order to identify if the parameters are correlated with BOLD signal in any of the brain regions. Initially, the parameters were correlated with activity from all subjects. Subsequently, an ANCOVA was run to further account for the three different participant groups. Each RL parameter was correlated with an event type, e.g. reward EV or punishment EV tracking. Across all subjects, α_{rew} negatively correlated with activity in response to reward EV tracking in the cingulate and paracingulate gyrus, IFG, middle and superior temporal gyrus, insular cortex, and OFC. This was also observed in response to the positive PPE. Activity following a positive RPE was negatively correlated with activity in the putamen, OFC and insula.

Moreover, it was discovered that the GD vs control contrast highlighted differential correlations between the SFG, MFG, postcentral gyrus during reward EV brain activity and the RL parameter α_{rew} . More specifically, the results showed that there was a stronger positive correlation between this parameter and the highlighted areas in the GD group, whereas the correlation in the control group was weaker. Additionally, the CUD vs HC contrast for the positive PPE found that α_{rew} was also correlated with the frontal pole, SFG, cingulate and paracingulate gyri, indicating a stronger positive correlation with α_{rew} in the CUD group than in the control group. Stimulus stickiness (κ_{stim}) was found to have a stronger positive correlation with activity in the right MFG and IFG during presentation of the feedback in

both patient groups, the GD and CUD group. No other correlations with RL parameters were found. These results can be seen in Fig. 5.15, as well as the appendix.



Fig. 5.15 The GD vs control contrast - positive correlation between α_{rew} and response to reward EV.

Areas of the contrast highlighting areas that have a stronger correlation with α_{rew} in the GD group than in healthy controls (MNI coordinates: X=-32, Y=12, Z=52). This parameter was positively correlated with activity in areas including the postcentral gyrus, SFG and MFG. Colour bar on the right-hand side represents t-statistic.

5.6 Discussion

In this chapter, computational analyses of behavioural and neuroimaging data from three participant groups were run and the extracted modelling parameters were linked to imaging data. Participants included HCs and individuals diagnosed with GD or CUD. All subjects completed three runs of a PRL task in an fMRI scanner. Performance on the task, as assessed by the proportion of win-stay or lose-shift trials, number of perseverative responses and the proportion of correct responses, did not differ between the groups. Subsequently, multiple computational models were fitted to the data. Besides applying RL models, a DDM and a combined RL-DDM model were employed. Neither the DDM nor the DDM-RL model highlighted group differences on any of the parameters, suggesting that decision-making processes as represented by a WFPT were not affected in the two patient groups. This was surprising, as a preliminary analysis of the response latencies indicated higher RTs in the CUD group that increased between the second and third runs, which was not the case for HCs or PGs.

The winning RL model, on the other hand, did highlight group differences. From the seven models tested, the model with the greatest number of parameters was found to provide

the best fit to the data. This model included five parameters, which were: α_{rew} , $\alpha_{non-rew}$, β , κ_{side} , and κ_{stim} . Comparing the parameters between the groups found that in the GD group, the stimulus stickiness parameter κ_{stim} was significantly decreased, with a non-significant tendency towards increased side stickiness κ_{side} . A comparison between the CUD and GD groups found that SDIs had a lower reward learning rate α_{rew} than PGs, as well as lower reinforcement sensitivity β , but greater stimulus stickiness.

Group differences were also observed on a neural level. When tracking reward and punishment EV, PGs showed an altered balance in brain activity. Specifically, individuals diagnosed with GD had greater activity when tracking reward EV than HCs in areas important for reward processing, such as the OFC, SFG and amygdala. When tracking punishment EV, however, this group showed lower activations in the postcentral gyrus, superior parietal lobule and occipital areas. These results indicate that PGs differentially track EVs of stimuli in their surroundings. In CUD, there was also an altered balance in reinforcement processing, as the neural signal to positive and negative PPE was affected. In the SFG and surrounding areas, the CUD group had a lower response to positive PPEs, and an enhanced response to negative PPEs than control participants. Further, these observations in BOLD activity could be linked to RL parameters, which were also differentially recruited in GD and CUD patients.

Previous studies fitting RL models to PRL data from SDIs have reported reduced learning rates and reinforcement sensitivity or increased stimulus stickiness (Kanen et al., 2019; Lim et al., 2021). In this report, we do not find any differences in RL parameters in cocaine-dependent individuals. Based on findings from conventional analyses of the PRL task, which observed increased perseveration and poorer performance in SUD, increased stickiness and reduced learning rates were expected (Ersche et al., 2008, 2011). However, in this cohort, none of the conventional or RL measures were affected. In (Lim et al., 2021), an instrumental task, rather than a reversal learning task was employed, which could explain the discrepancies. In (Kanen et al., 2019), the sample included patients with any type of SUD, but in this chapter, only cocaine-dependent individuals on the reversal learning task have been observed, which may explain why the results from Kanen et al could not be replicated in this chapter (Ersche et al., 2008). Nonetheless, the sample sizes in all studies were relatively small and in order to generate robust and generalisable results, larger sample sizes need to be used.

To our knowledge, this is the first study investigating how RL parameters in GD are altered. An unexpected finding was made, namely, stimulus stickiness was reduced, whereas side stickiness showed a trend level increase (below 75% HDI). Given that a key feature of GD

is behavioural inflexibility and perseveration, it is unclear why only one stickiness parameter was increased. A possible interpretation of these results is that side stickiness is representative of motor perseveration, whereas stimulus stickiness reflects stimulus perseveration. The increased side stickiness parameter could indicate greater motor perseveration, leading to the behavioural inflexibility commonly seen in GD patients. The reduced stimulus stickiness, on the other hand, could reflect an adaptive behaviour useful for gambling, as lower stimulus stickiness may lead to increased exploration of choices to identify the optimal strategy (Clark, 2010). This disparity in the stimulus stickiness parameter between GD and CUD therefore reveals a novel dissociation between these two compulsive disorders.

Reduced neural prediction error tracking in the left VS and right mOFC on the IGT has been previously reported in SDIs (Tanabe et al., 2007). Additionally, an EEG study found impaired negative RPE signalling in CUD (Parvaz et al., 2015). Here, we reported altered signals in response to positive and negative PPEs, rather than RPEs. Altered prediction error tracking could be a contributing factor to the disadvantageous nature of compulsive drug use, as it persists despite adverse outcomes. The areas involved in PPE tracking, the SFG, MFG, cingulate and paracingulate gyri, were also found to be positively correlated with α_{rew} . Therefore, these regions appear to be of key importance for reward and punishment learning and are likely to be involved in the modulation of RL parameters. This is further supported by a meta-analysis from 35 studies, which reported that these areas are consistently activated when there is a prediction error (Garrison et al., 2013).

Participants with GD in this study also had altered neural responses in RL, although they were impaired in reward and punishment EV tracking, rather than responding to PPEs. This has been reported in a previous study using a card-guessing task. Gamblers had an increased neural signal in the VS and OFC when tracking reward EV (Holst et al., 2012). Together, these results suggest impaired tracking of stimulus EV in GD, which may underlie maladaptive, compulsive behaviours in this group. In summary, this chapter demonstrates that modelling of behaviour on PRL tasks using RL models can help distinguish disorders of compulsivity, which is not possible solely with conventional measures. Moreover, altered brain activity during EV tracking and PPE response in the two patient groups are found, a finding that has not been made in these disorders before. The fMRI data also illuminated the brain regions responsible for RL measures, which thus far was unexplored in humans. In particular, κ was shown to be correlated with activity in the dIPFC and vIPFC, both of which are important for mediating the balance between goal-directed and habitual behaviours (Balleine and O'Doherty, 2009). For example, dIPFC lesions impair motor habits in macaque monkeys and the vIPFC can override habitual behaviours (Badoud et al., 2017; Korponay,

2021). Activity in these regions showed a stronger positive correlation with κ in patients with GD and CUD. The ventro- and dorsolateral PFC are known to be affected in both disorders, providing additional evidence for their involvement (Goldstein and Volkow, 2011; Raimo et al., 2021). Understanding the underlying neural substrates of parameters such as κ further supports their use for diagnosing and distinguishing compulsive disorders.

Chapter 6

Reinforcement learning in patients with Major Depressive Disorder

6.1 Introduction

In previous chapters, the role of RL in MDD was discussed and exemplified in relation to a 'two-hit' model of depression in rats. It was demonstrated that important RL parameters, such as the punishment learning rate $\alpha_{non-rew}$ and stickiness parameter κ , were affected by early-life maternal separation, a second adulthood stress, and by sex. These changes were linked to alterations in resting-state functional connectivity, specifically, connectivity between the amygdala and PFC areas was significantly affected. Additionally, in chapters 4 and 5, RL models were fit to reversal learning data from rodents and humans that have experienced compulsive drug-seeking. The RL parameters in these studies were also linked to functional connectivity as measured via resting-state and task-based fMRI, respectively. An important next step is to translate the findings observed in rodents to humans and find out whether the results are reproducible.

Previous studies have established that patients with MDD have an altered balance between reward and punishment processing, as they show reduced learning from reward and maladaptive punishment learning rates (Eshel and Roiser, 2010; Halahakoon et al., 2020). This imbalance in RL may play a role in the onset and maintenance of MDD (Vrieze et al., 2013). More generally, depression is associated with impairments in executive control and cognitive flexibility. For example, participants with MDD show impaired performance on neutral and emotional Go/No-go tasks, as well as on the WCST and PRL tasks (Murphy et al., 1999, 2012b; Purcell et al., 1997). An additional feature of MDD is negative bias, which is defined as increased sensitivity to negative information (Brolsma et al., 2020a;

Reinforcement learning in patients with Major Depressive Disorder

Robinson et al., 2012). People with acute MDD have a bias to negative emotional stimuli on the Go/No-go task, which is absent in remitted MDD participants (Erickson et al., 2005; Maalouf et al., 2012; Murphy et al., 1999). In some studies, this bias was accompanied by greater impulsive responding, suggesting that inhibitory control is also affected, as well as reduced attention for positive stimuli on an eye-tracking task (Duque and Vázquez, 2015; Maalouf et al., 2012). This negative bias may be related to the altered balance in reward and punishment learning and might manifest itself in the form of changes in reinforcement sensitivity. Anhedonia, another hallmark feature of MDD, has been linked to dysfunctions in reward learning (Vrieze et al., 2013). Therefore, it is important to investigate the RL system in relation to MDD and unravel the underlying processes.

Current medications to treat MDD do not target impairments in reward processing (Admon and Pizzagalli, 2015; Halahakoon et al., 2020). A recent systematic review showed that across forty-eight case-control studies, depressed individuals exhibited impairments in reward processing behaviour compared to a control group, with medication status having no significant effect on RL (Halahakoon et al., 2020). Interestingly, another study found that patients who were not impaired in reward learning prior to pharmacological treatment continued to have a persisting diagnosis of MDD after eight weeks (Vrieze et al., 2013). A cross-sectional cohort study reported that following CBT, RL parameters were normalised in patients with MDD (Brown et al., 2021). This was accompanied by a fMRI study reporting reduced activity in the right striatum and right amygdala in response to reward prediction errors following CBT, providing a potential explanation for the neural mechanisms associated with RL changes and demonstrating that brain activity can be normalised through psychotherapeutic interventions (Queirazza et al., 2019). Gaining a deeper understanding how the RL system is affected by MDD and changed as a result of medication or other therapies is essential to develop future generations of treatments.

Approximately two-thirds of patients diagnosed with depression have a comorbid disorder (Steffen et al., 2020). However, most studies that have investigated RL in people with depressive disorders chose a highly selective sample of patients without other psychiatric diagnoses. Such samples are not easily generalisable; therefore, it is necessary to investigate RL in real-world cohorts. Computational modelling has previously revealed that patients with SUD showed a lower learning rate for reward and a higher learning rate for punishment than HCs, accompanied by increased stimulus stickiness (Kanen et al., 2019). In this study, the best-fitting model contained five parameters, including two separate learning rate terms and two stickiness parameters. In another study, selective for male participants with CUD, the data were best described by a seven-parameter model with four learning rates, two stickiness

terms and one reinforcement sensitivity measure (Lim et al., 2021). Only the punishment learning rate parameter and reinforcement sensitivity were found to be lower in the patient group. In chapter 5, no significant differences in a five-parameter RL model between HCs and individuals with CUD were reported. Therefore, studies including subjects with SUD remain inconclusive, underlining the importance of investigating this further in a naturalistic sample. Moreover, studies have yet to investigate cognitive flexibility in people affected by both MDD and SUD. Based on previous studies, it would be expected that individuals affected by both disorders will have an altered balance in reward and punishment learning, as well as changes in stickiness.

In this chapter, I aim to investigate RL in a large sample including participants with MDD and other diagnoses, such as SUD and ASD. Even though the sample does not include imaging data, it gives an additional insight into the behavioural processes in humans and how these are altered in a variety of psychiatric conditions. Following the results presented in chapter 5, this dataset allows us to extend the findings from SDIs to an even larger patient sample.

6.2 Methods

6.2.1 Participants

Two hundred and seventeen patients and one hundred and one healthy controls completed a PRL task. The study was carried out by Prof Roshan Cools and colleagues at the Department of Psychiatry, Radboud University Medical Centre, Nijmegen, The Netherlands. It was part of the MIND-Set study, an observational, cross-sectional study. Inclusion criteria for adult outpatients were a diagnosis of 1) a current depressive disorder (MDD/dysthymia); 2) an anxiety disorder; 3) an addictive disorder; 4) ADHD and/or 5) ASD. The age range of patients was 18-78, with a mean age of 40. Diagnoses were made using DSM-IV/DSM-5 criteria, as the study took place during the transition period. Depressive disorders, anxiety disorders and psychotic disorders were determined with the SCID-I-CV (First, 1997). ADHD was diagnosed with the Diagnostic Interview for Adult ADHD, second edition (DIVA 2.0) (Kooij and Francken, 2010). For ASD diagnosis, the Dutch Diagnostic Interview for Adult Autism Spectrum Disorders was used (NIDA) (Vuijk et al., 2021) and addictive disorders with the Measurements in the Addictions for Triage and Evaluation and Criminality (MATE-crimi) (Schippers et al., 2010).

Reinforcement learning in patients with Major Depressive Disorder

Exclusion criteria for this study were: 1) a current psychotic disorder according to the SCID section B; 2) an IQ estimation <70; 3) a sensorimotor disability intervening with participation, were mentally incompetent to sign informed consent or had insufficient knowledge of the Dutch language. The study has been approved by the local ethics committee (Commissie Mensgebonden Onderzoek Arnhem-Nijmegen, NL 55618.091.015). Participants gave written informed consent for taking part in this study.

Information on gender, age and level of education was collected. Education could fall into one of four categories: 1) no education (elementary education or education not finished); 2) low (lower vocational and general secondary education); 3) middle (intermediate vocational and higher secondary education) or 4) high (higher vocational education and university) (Ikram et al., 2015). Verbal IQ was measured using the National Adult Reading Test score (NART score, Dutch version) (Schmand et al., 1991). Working memory capacity was used as an additional measure of cognitive ability and was measured using the total number of errors on the spatial working memory task from the Cambridge Neuropsychological Test Automated Battery (CANTAB®; Cambridge Cognition 2019).

For the purpose of the analysis conducted, five groups were selected: 1) HCs (n=81), 2) MDD without any comorbidities (n=27); 3) MDD and ASD (n=17), 4) MDD and SUD (n=24); 5) SUD but not MDD (n=14). Controls underwent the same tests as patients. Data from this study have been previously published in (Brolsma et al., 2020b) and were kindly shared by the authors. In the analysis, only 81 HCs were selected to match the patient sample based on age, gender and education level. The re-analysis of these data was done in collaboration with Dr Jon Kanen (Department of Psychology, Cambridge University) and was conducted with scripts written by Dr Rudolf Cardinal in the Department of Psychiatry at Cambridge University.

6.2.2 Probabilistic reversal learning task

The PRL task used in this study was similar to previously published procedures (Fig. 6.1) (Cools et al, 2001; den Ouden et al, 2013). Subjects were presented with two squares on the left and right side of the visual hemisphere, one of which was 'correct' and one was 'incorrect'. After selecting one, they would receive feedback, which could be either positive or negative. Participants were told to respond to the stimulus that resulted in reward most frequently. They would need to learn through trial-and-error which stimulus is 'correct', and which one is 'incorrect'. Since this was a probabilistic task, the 'correct' stimulus would be rewarded on only 80% of the trials, and the 'incorrect' stimulus would be spuriously

rewarded on 20% of the trials. After 40 trials, the contingencies were reversed. In the instructions, participants were told that the 'correct' stimulus could change, but they were not given information about how frequently or when this would happen.



Fig. 6.1 Diagram of the probabilistic reversal learning task.

a) Two stimuli were presented in two of four possible places, either top, bottom, left, or right. Participants were told to select one of the stimuli by pressing the relevant key. Immediately following the response, feedback was presented, which was either a reward (a green smiley face and a high-pitched sound), or punishment (a red, sad face and a low-pitched sound), both shown for 1500 ms. After 1000 ms, the next stimulus was presented. b) Stimulus contingencies prior to and following a reversal. From (Brolsma et al, 2020).

6.2.3 Correlation analysis between conventional and modelling PRL measures

The relationship between the RL and conventional PRL measures was investigated through a correlation analysis using Spearman's rank coefficient, as the data were non-normally distributed, accounting for multiple comparisons with FDR correction (Benjamini and Hochberg, 1995). The statistical analyses were run in R: A Language and Environment for Statistical Computing (R Core Team, 2020). For all analyses, significance was considered to be below p=0.05.

6.3 Results

6.3.1 Demographic information

There were no statistically significant differences in age, IQ and education level between the five groups. However, there were differences in the gender distributions and depressive symptom severity scores. All patient groups had lower numbers of females than males. The MDD and ASD group had a significantly greater number of males than females compared to the HC group (t(93)=2.45, p=0.016) as did the SUD no MDD group (t(95)=2.50, p=0.014). Additionally, patients with MDD without comorbidities and patients with both MDD with SUD had significantly higher depressive symptom severity scores (t(108)=-4.42, p<0.001 and t(103)=-3.65, p<0.001, respectively).

Table 6.1 Demographic information

	Healthy controls (n=81)	MDD only (n=27)	MDD and ASD (n=17)	MDD and SUD (n=24)	SUD w/o MDD (n=14)	Group comparisons
Mean age (SD)	40.3 (17)	44.74 (15.2)	37.3 (13.7)	38.7 (16.9)	46.1 (14.5)	F(4,157) = 1.09, p=0.36
Gender (F)	40	10	3	9	2	F(4,158) = 2.74, p=0.031
Mean IQ (SD)	101 (12.4)	100 (10.6)	96.5 (9.2)	96.5 (8.2)	97 (11.5)	F(4,156) = 1.28, p=0.28
Education level	None: 0%	3.8 %	0 %	0 %	21.4 %	F(4,158) = 1.77, p=0.13
	Low: 7.4 %	15.4 %	6.3 %	34.7 %	0%	
	Middle: 40.7 %	50%	68.8 %	34.8 %	42.9 %	
	High: 51.9 %	30.7 %	25 %	30.4%	35.7 %	
Mean IDS-SR score (SD)	4.9 (0.4)	37.7 (13.2)	38.2 (9.7)	41.3 (10.5)	30.6 (16.9)	F(4,156) = 130.2, p < 0.001

IDS-SR - Inventory of Depressive Symptomatology, self-report version ; MDD - Major Depressive Disorder; ASD - Autism Spectrum Disorder; SUD - Substance Use Disorder; F - female; L - left; SD standard deviation. A table based on the same dataset has been published in (Brolsma et al, 2020), however, this is an original extraction of the data.

6.3.2 Conventional PRL measures

Fitting LMEs to conventional PRL measures such as total number of errors, win-stay, lose-shift and number of perseverative responses revealed whether there were any significant group differences. There was no effect of group on the number of errors, on both acquisition and reversal errors, as well as total errors (F(4,158)=2.04, p=0.091). Perseverative errors were also not affected by group (F(4,158)=1.00, p=0.41). Total win-stay and lose-shift behaviour did not differ between the groups either (F(4,158)=1.14, p=0.34 and F(4,158)=0.50, p=0.73, respectively). A summary of these results can be seen in Fig. 6.2

6.3 Results



Fig. 6.2 Key findings of a conventional analysis of the PRL task.

Total errors; win-stay behaviour; lose-shift behaviour and number of perseverative responses was calculated for all subjects from the five groups: healthy controls, SUD, SUD and MDD, MDD only and MDD with ASD. No significant group differences were found.

6.3.3 Choosing the winning RL model

To identify which RL model would be able to capture the behavioural dynamics and fit the data the best, five RL models were run and compared with each other. Models including the side stickiness parameter κ_{side} were not included, as the task included four locations which were randomised on each trial in order to avoid side bias on the task. Results from the comparison can be found in table 6.2. Good convergence was seen, with all parameters and contrasts having an $\hat{R} < 1.1$. An \hat{R} value of 1.1 was set as the threshold for model

Reinforcement learning in patients with Major Depressive Disorder

convergence. The maximum \hat{R} was 1.029. The best-fitting model was determined using a bridge sampling estimate of the marginal likelihood. This model contained three parameters: a combined learning rate α , which represents how quickly a subject updates its behaviour in response to feedback; reinforcement sensitivity β , also known as the exploitation vs exploration parameter and stimulus stickiness κ_{stim} , which is the tendency to select the same stimulus regardless of outcome.

Rank	Name	Parameters	Log marginal likelihood	Log posterior P (model)
2	Model 1	$\alpha_{rew}, \alpha_{pun}, \beta, \kappa_{stim}$	-8202.28	-7.71
3	Model 2	$\alpha_{rew}, \alpha_{pun}, \beta$	-8270.00	-75.43
1	Model 3	α , β , κ_{stim}	-8194.57	-4.5x10 ⁻⁴
4	Model 4	ϕ, ρ, β	-8320.71	-126.14
5	Model 5	α, β	-8275.19	-80.61

Table 6.2 Model comparison summary.

6.3.4 Group differences

Fig. 6.3. summarises the results from the hierarchical Bayesian Q-learning model analysis from patients with SUD, SUD and MDD, ASD and MDD and MDD only. For the SUD group, a significant increase in the learning rate and stimulus stickiness rate was found (difference in parameter per-group mean, posterior 95% highest posterior density interval), as well as a decrease in reinforcement sensitivity at 75% HDI. The stimulus stickiness rate assesses stimulus-bound behaviour, irrespective of outcome of that stimulus on previous trials. For the SUD and MDD group, there was a decrease in reinforcement sensitivity and a slight increase in stimulus stickiness (group difference, 75% HDI). The MDD only group showed no differences, whereas patients with MDD and ASD had decreased reinforcement sensitivity at 75% HDI compared to healthy controls.

6.3.5 Relating conventional parameters and depression score ratings to RL parameters

Correlating the various measures with each other using Spearman's rank coefficient revealed some interesting relationships between the parameters. α , the learning rate, was significantly positively correlated with win-stay and negatively correlated with lose-shift behaviour (r(159)=0.36, p<0.001 and r(159)=-0.44, p<0.001, respectively). The reinforcement sensitivity parameter β showed a positive correlation with win-stay behaviour (r(159)=0.41, p<0.001) and a negative relationship with lose-shift (r(159)=-0.23, p<0.001). The same trend


Fig. 6.3 Results of the hierarchical Bayesian Q-learning model.

MDD - Major Depressive Disorder; HC - healthy controls; ASD - Autism Spectrum Disorder; Reinf sens - reinforcement sensitivity; Stim - stimulus; HDI - highest density interval. Red indicates differences between groups at an HDI of 95%, whereas yellow shows differences at an HDI of 75%.

Reinforcement learning in patients with Major Depressive Disorder

was observed for κ_{stim} (r(159)=0.27, p=0.003 and r(159)=-0.58, p<0.001) The Inventory of Depressive Symptomatology (IDS), Addiction Severity Index (ASI), Autism Spectrum Quotient (AQ) and Perseverative Thinking Questionnaire (PTQ) scores did not correlate with any computational RL measures. However, IDS scores were positively correlated with total number of errors, as were the ASI and PTQ scores (r(159)=0.24, p=0.009; r(159)=0.22, p=0.014; r(159)=0.25, p=0.004, respectively). The AQ score was positively correlated with the number of perseverative errors (r(159)=0.21, p=0.021). A summary of these results can be found in Fig. 6.4 and table 6.3.

Table 6.3 Summary table of statistics from the correlation analysis between conventional PRL and modelling parameters.

	totalE	perE	WSall	LSall	IDS_Tot	ASI_Tot	AQ_Total_Brits	ptq_tot	Alpha	Beta	Kappa_stim
totalE	-	0.21	0.172	0.063	0.009	0.014	0.088	0.004	0.949	0.706	0.172
perE	0.128	-	0.001	0.172	0.261	0.999	0.021	0.331	0.146	0.972	0.811
WSall	-0.137	-0.295	-	0.08	0.989	0.951	0.234	0.579	0.000	0.000	0.003
LSall	-0.178	-0.139	-0.169	-	0.972	0.951	0.717	0.602	0.000	0.010	0.000
IDS_Tot	0.236	0.117	0.008	-0.013	-	0.000	0.000	0.000	0.972	0.951	0.989
ASI_Tot	0.224	0.003	0.023	-0.028	0.669	-	0.000	0.000	0.999	0.951	0.999
AQ_Total_Brits	0.165	0.21	-0.122	-0.057	0.681	0.514	-	0.000	0.754	0.602	0.904
ptq_tot	0.255	0.105	0.076	-0.071	0.834	0.709	0.641	-	0.754	0.951	0.999
Alpha	-0.03	-0.147	0.36	-0.438	-0.014	0.0	-0.051	0.051	-	0.001	0.003
Beta	-0.06	-0.014	0.409	-0.231	-0.021	0.026	-0.071	0.021	0.288	-	0.028
Kappa_stim	0.136	0.044	0.265	-0.583	-0.009	-0.002	0.035	0.005	0.267	-0.202	-

Summary of r-values and p-values from the correlation analysis. The top right half of the table represents the p-values, and the bottom left shows the r-values. Cells highlighted in red show p-values < 0.05. TotErrors – total number of errors; persevErrors – number of perseverative responses after a reversal; WS - win-stay; LS - lose-shift; IDS - Inventory of Depressive Symptomatology; ASI - Addiction Severity Index; AQ - Autism Spectrum Quotient; PTQ - Perseverative Thinking Questionnaire; Alpha – α parameter representing the learning rate from feedback; Beta – the exploitation/exploration parameter β ; Kappa_stim - κ_{stim} , stimulus stickiness parameter.

6.3.6 Simulations

Behavioural data were simulated using the extracted group-level Q-learning parameters as described in section 2.2.2. For each group, 100 subjects were simulated. The resulting data was then analysed using a conventional PRL analysis. The proportion of correct responses showed a significant effect of group (F(4,196)=412, p<0.001), which was mainly



Fig. 6.4 Correlation matrix representing the relationship between conventional PRL measures, psychiatric questionnaire scores and RL parameters.

The colour bar on the right-hand side represents the correlation strength using Spearman's rank coefficient. TotErrors – total number of errors; persevErrors – number of perseverative responses after a reversal; WS - win-stay; LS - lose-shift; IDS - Inventory of Depressive Symptomatology; ASI - Addiction Severity Index; AQ - Autism Spectrum Quotient; PTQ - Perseverative Thinking Questionnaire; Alpha – α parameter representing the learning rate from feedback; Beta – the exploitation/exploration parameter β ; Kappa_stim - κ_{stim} , stimulus stickiness parameter.

Reinforcement learning in patients with Major Depressive Disorder

driven by differences between the patient groups. The differences between HCs and patients were in the SUD and MDD/ASD groups; both showed a lower proportion of correct responses (t(196)=21.1, p<0.001 and t(196)=-5.56, p<0.001, respectively). The simulations also highlighted effects of group on win-stay and lose-shift behaviour (F(4,196)=373, p<0.001; F(4,196)=70.5, p<0.001, respectively), as well as perseverative responses (F(4,196)=50.8, p<0.001). Win-stay behaviour was greater in control participants than participants with MDD and ASD (t(196)=20.7, p<0.001), but lower than in patients with MDD only and MDD with SUD (t(196)=6.12, p<0.001 and t(196)=-16.1, p<0.001, respectively). A similar effect on lose-shift behaviour was found, as HCs had lower lose-shift behaviour than participants with ASD and MDD (t(196)=10.5, p<0.001) as well as the SUD/MDD group (t(196)=5.04, p=0.10). Compared to the SUD group, controls were modelled to have greater lose-shift behaviour (t(196)=5.04, p<0.001). Perseverative responding was lower in the MDD and ASD group as well as the MDD and SUD group than controls (t(196)=-9.74, p<0.001; t(196)=3.85, p=0.0015, respectively). Graphical representations of these results can be found in the appendix (section B.3). In summary, the simulations could not successfully capture participants' behaviour on the PRL task, since no group differences were observed in the real data.

6.4 Discussion

In this study, RL modelling was applied to a large cohort of patients with MDD without any comorbidities, as well as participants with MDD and various comorbidities such as ASD and SUD. A group solely consisting of individuals with SUD was also included. Conventional PRL measures such as total errors, win-stay, lose-shift and perseverative responses found no significant group differences. However, a hierarchical Bayesian RL analysis with a winning model containing three parameters, namely the learning rate, reinforcement sensitivity parameter and a stimulus stickiness parameter revealed differences in RL between the groups. Firstly, the RL parameters in the MDD group that did not have any comorbidities were unaffected. In participants with SUD only, the reward rate and stimulus stickiness were increased, and the reinforcement sensitivity parameter was reduced. In the SUD and MDD group, stimulus stickiness was increased, and the reinforcement sensitivity was decreased (both at 75% HDI). ASD and MDD participants showed a significant reduction in reinforcement sensitivity.

A correlation analysis found that α , β and κ showed a negative correlation with lose-shift behaviour and a positive relationship with win-stay. This relationship between β and these two conventional measures is novel; it suggests that individuals that have greater sensitivity to reinforcement tend to stay more after a win and shift less after a loss, which by all accounts seems intuitive. No computational measures were found to be related to psychiatric questionnaires. However, patients with higher IDS, ASI and PTQ scores showed poorer performance on the task, and participants with higher AQ scores perseverated more. Even though the simulations did not entirely capture the behaviour of all patient groups, it did manage to capture that there were no differences between HCs and MDD/SUD patients. The reason for the discrepancy between observed and simulated data could be that some of the patient groups had relatively small sample sizes, whereas in the simulations 100 participants were modelled per group. This approach increases statistical power and could be suggestive of what would be observed in a larger sample.

The finding that RL parameters were not altered in patients with MDD was surprising and contrary to our hypotheses. Multiple studies have reported dysfunctional RL in depression, specifically, patients with MDD have been found to learn less from rewards, an attribute thought to be linked to anhedonia (Admon and Pizzagalli, 2015; Herzallah et al., 2013; Mukherjee et al., 2020). Punishment learning is also thought to be altered, however, some reports show no effect on learning from punishment (Mukherjee et al., 2020; Tavares et al., 2008). An alternative argument is that rather than learning mechanisms being impaired per se, reinforcement sensitivity is altered, which is supported by a meta-analysis including six studies (Huys et al., 2013). Neither the learning rate nor the reinforcement sensitivity rate were altered in MDD in the study presented here. Given that most studies to date have demonstrated that this is not the case, it may be that the results from this study do not reflect RL impairments that are generally observed in depression. The patient's medication status may also play a role. For example, medicated MDD patients receiving the SSRI paroxetine learn less from negative feedback than non-medicated patients and HCs (Herzallah et al., 2013). In rats, antidepressants including citalopram, reboxetine and ketamine differentially affect performance on a PRL task and associated RL parameters, with high-dose ketamine reducing the learning rate and reboxetine reducing the exploration vs exploitation parameter (Wilkinson et al., 2020). Nonetheless, these results may also be an indicator of the fact that not all patients with MDD have altered RL mechanisms, reflecting large variability of these parameters. Further studies should aim to elucidate this variability and how it is reflected neurally.

In the previous chapter, RL in SUD was discussed. Although in chapter 5 null results for this patient group were reported, previous studies have found that participants with SUD have a decreased reward learning rate, an increased punishment learning rate, and

Reinforcement learning in patients with Major Depressive Disorder

most importantly, increased stimulus-bound stickiness (Kanen et al., 2019). However, one study solely including patients with CUD found a reduction in the learning rate as well as reinforcement sensitivity (Lim et al., 2021). In PRL experiments analysed using conventional statistical methods, SUD has been associated with increased perseverative responding in humans, thus, heightened stickiness would be expected in SDIs (Ersche et al., 2008, 2011). In rats, greater exploration and heightened stickiness were demonstrated following high cocaine escalation (Zhukovsky et al., 2019). In this chapter, the learning rate in the SUD group was increased, contrary to previously published data. Given that the winning model had a joint reinforcement learning rate, rather than separate reward and punishment learning rates, the increased learning rate may reflect increased punishment learning, rather than learning from reward. The increased stimulus stickiness reported here parallels the results from (Kanen et al., 2019).

There were a few limitations in this study that need to be highlighted. Firstly, the number of participants across samples was imbalanced, with the HC group having more than twice as many participants as the other groups. Additionally, the female to male ratio between the five groups was also different, with some groups having significantly fewer females than others, although HCs were matched to the entire patient sample. The importance of gaining a better understanding of RL in females has been discussed previously. Future studies would benefit from having more female patients enrolled in the study and could explore sex-dependent differences in RL in humans.

Given that MDD is often comorbid with other psychiatric conditions, such as ADHD, ASD or anxiety, this naturalistic sample of patients allows us to study MDD under more ecologically valid settings. It appears that in people diagnosed with both MDD and SUD, there is higher stimulus stickiness and lower reinforcement sensitivity, which may be attributable to SUD, as these effects were also seen in the SUD only group but not the MDD only group. In the ASD/MDD, SUD and SUD/MDD groups, reinforcement sensitivity was reduced, indicating that these differences may be a general, non-specific psychiatric vulnerability. In (Brolsma et al., 2020b), where these data were initially published, none of the Q-learning parameters were altered in MDD, ASD or SUD. In their study, a three-parameter model was used, which included two separate learning rates and an exploitation/exploration parameter, but no stickiness parameter. Therefore, stickiness was not used to explain behaviour and could not be investigated in relation to these conditions. Here, clear effects of the additional stickiness parameter are reported. In particular, this parameter seems to be of great importance in SUD.

The chapters thus far have focused on RL and its neural substrates in rodents and humans. The stickiness parameter κ has been highlighted as a key component of RL across species, with changes in this parameter seen in different states. Moreover, brain areas that mediate this parameter were identified, demonstrating that is has a biological basis. These are findings that have not been reported before. The next chapter will shine a different light on cognitive flexibility, as a novel task measuring it by assessing volitional decision-making will be presented. The task does not involve trial-and-error learning through reward and punishment; therefore, it is not possible to extract κ from behaviour using RL models. However, it allows the investigation of value-free behaviours, manifested as repeating a previous response regardless of feedback, similarly to κ . Thus, a different form of stickiness will be assessed, and its associated brain regions will be identified. This task can help to isolate value-free behaviours from value-based ones, which is not possible in reversal learning tasks. This also allows the exploration of areas solely associated with value-free responding, something that has not been researched previously and could aid in gaining a deeper understanding of the brain regions underlying κ .

Chapter 7

A novel perspective on cognitive flexibility

7.1 Introduction

Cognitive flexibility is the ability to adapt to changes in the environment by switching task sets, responses, or strategies (Cools, 2015). From a psychological perspective, it is often considered alongside related emergent control processes such as inhibition and updating (Miyake et al., 2000). Flexibility allows changing behaviour in response to environmental and internal cues. Greater cognitive flexibility leads to improved life outcomes, better social functioning and reduced cognitive decline with age (Burke et al., 2019; Diamond and Lee, 2011; Koesten et al., 2009). Difficulties in flexible responding have been observed in various psychopathologies. For example, patients with OCD show impaired performance on cognitive flexibility tasks (Remijnse et al., 2006; Vaghi et al., 2019). In patients with ASD, repetitive behaviours are correlated with reversal deficits, highlighting the importance of research on the neural and psychological bases of flexible behaviour (Yerys et al., 2009).

Gaining a full understanding of the behavioural and neural underpinnings of cognitive flexibility has proven to be difficult, as it is not a unitary construct (Dias et al., 1996b). Various behavioural tasks have been developed to measure cognitive flexibility in humans, each measuring overlapping yet distinct aspects of this process (Dajani and Uddin, 2015; Milner, 1963; Roberts et al., 1988). Frequently studied paradigms include task-switching, attentional set-shifting and reversal learning (Brown et al., 2015). Task-switching is a procedure in which the participant switches between two or more tasks, based on pre-defined, deterministic stimulus-response sets (Manoach, 2009; Monsell, 2003). Set-shifting tasks, such as the WCST, IED or the Dimensional Change Card Sort task (DCCS), on the other

A novel perspective on cognitive flexibility

hand, measure the ease with which an individual transitions between cognitive-attentional sets, typically by learning and using external cues about ongoing performance. In reversal learning tasks, subjects adapt their responding following the reversal of previously learnt reward-related contingencies (Izquierdo et al., 2017). The latter has been the focus of the previous chapters. In this section of my thesis, I will present cognitive flexibility from a different perspective, with the aim of highlighting that it is a multifaceted construct. I will also discuss how value-free habits in the form of stickiness, or repeating a response regardless of outcome, are observed in both reversal learning and the novel task.

The neural substrates mediating cognitive flexibility tasks also have diverging and converging aspects. In studies employing either task-switching or set-shifting paradigms, the salience network, composed of the AI and dACC, and executive control network, which includes IFJ, dlPFC and IFG, have been consistently activated (Braun et al., 2015; Dajani and Uddin, 2015; Shine et al., 2016; Uddin, 2021). The default mode network, which includes the mPFC, PCCx, precuneus and angular cortices, has also been associated with set-shifting (Vatansever et al., 2016). Connectivity between the AI, dACC and thalamus, forming the cingulo-opercular network, and the posterior parietal cortex, show greater involvement in task-switching (Dosenbach et al., 2006; Liston et al., 2006; Worringer et al., 2019). Furthermore, striatal and dorsolateral PFC areas are activated in task-switching and attentional set-shifting (Armbruster et al., 2012; Cools et al., 2004; Wager et al., 2004). Reversal learning, however, activates somewhat different circuits, including the striatum, amygdala, mPFC, mOFC and lOFC (Izquierdo et al., 2017; Remijnse et al., 2005). In summary, taskswitching and attentional set-shifting have some overlapping neural substrates, but the former is associated with greater activity in parietal areas and thalamus. As may be expected from task demands, reversal learning mostly engages areas associated with reward processing, such as the OFC and striatum.

Cognitive control more generally has been increasingly recognised to employ proactive and reactive modes (Braver, 2012). The proactive mode requires active maintenance of task goals and anticipates interference before it occurs, which may be mediated by the lateral PFC (Braver et al., 2009; Mäki-Marttunen et al., 2019). Reactive control detects interference demands only after being triggered by external cues. Ventrolateral PFC and AI node activations are common to both proactive and reactive control modes; however, only proactive control elicits the DMN (Ryman et al., 2019). This distinction can further be extended to cognitive flexibility, with task-switching paradigms characteristically testing reactive flexibility, where switching is triggered by external signals informing participants explicitly when to switch. In contrast, reversal learning and attentional set-shifting paradigms characteristically capture both reactive and proactive flexibility, with shifting being both prepared and elicited by feedback in participants (Rogers et al., 2000). There are additional key procedural differences between the common instantiations of these paradigms. For example, many task-switching procedures include similar proportions of switch and repeat trials, whereas most reversal learning and behavioural set-shifting procedures employ fewer shifts. The SN is elicited by less frequent, task relevant events, as is required in proactive responding (Wager et al., 2004). Additional procedural distinctions involve the extent of ongoing learning throughout the task and the employment of feedback or reward to signal a need to change, both of which are less prominent in task-switching (Kiesel et al., 2010).

In daily life, flexibility often involves a change in behaviour being generated endogenously, but in situations where rule learning or set-shifting is not required. Additionally, individuals often choose to change behaviours based on environmental signals, however, these can be difficult to interpret as they are frequently accompanied by noise, especially in situations with high levels of uncertainty. Here, a novel 'change your mind' task enabling the exploration of proactive switching under uncertainty without the need for ongoing rule-based learning is presented. Participants performed a two-alternative forced choice task and following spurious feedback, were presented with the same stimulus again. They could repeat their previous response or change it, acting of their own volition. No commonly used task to study cognitive flexibility provides participants with the opportunity to repeat their choice and assess whether they subsequently choose the 'road not taken'. We propose that the ability to change one's responses constitutes a form of mental flexibility not fully explored within current conceptual frameworks (Uddin, 2021). This chapter reports findings from forty healthy participants who completed the task whilst undergoing a fMRI scan. Behaviourally, it was hypothesised that subjects would change their second response when their first response was incorrect or when feedback was negative, and that any effects of the spurious feedback would decrease with time. Based on the evidence reviewed above, it was predicted that PFC salience and ECN nodes (i.e., AI, dACC, IFJ, IFG and dIPFC) would be activated when participants volitionally 'changed their minds'.

7.2 Methods

7.2.1 Participants

Forty healthy volunteers (21 females) between the ages of 18 and 60 years (mean=31.88, sd=10.03) were recruited from adverts in the Cambridge community and from the Behavioural

and Clinical Neuroscience Institute volunteer panel. The participants were screened to ensure that they had no history of psychiatric or neurological disorders. All participants provided informed written consent and were reimbursed for their time and travel expenses, with ethical approval granted from the Cambridge Local Research Ethics Committee (08/H0308/65). These data were collected by Dr Sharon Morein-Zamir; the data analyses were conducted by myself with the guidance of Dr Morein-Zamir.

7.2.2 Task procedure

Participants performed a two-alternative forced choice task, during which they had to identify whether a masked target letter was 'T' or 'L' (Figure 7.1). During each trial, the same stimulus was presented twice following a fixation ('1' or '2') denoting whether this was the first or second of the pair. During stimulus presentation, the target letter was embedded in a row of 'X's. After the first stimulus presentation, participants received feedback. The feedback was spurious and orthogonal to performance, so that it was negative on half of accurate trials and positive on half of the incorrect trials. Participants were told in advance that the feedback may not always be 100% accurate. Instructions stated that immediately after the feedback the same trial will be presented again and that they would have a chance to change their mind on the second display should they wish. The task consisted of three runs of 56 pairs of trials each. 48 trials were difficult and 8 were easy to help participants learn about the validity of the feedback.

On each trial, following a 500 ms fixation, the target was briefly presented within a row of eight 'X's in grey in the centre of the screen. Following a predetermined duration, the target was replaced with an 'X'. The target location was counterbalanced, appearing in one of the 6 or 4 central locations on difficult or easy trials, respectively. A staircase determined target duration. On difficult trials, it was initially set to 60 ms and decreased by 10 ms following a correct response and increased by 10 ms following an incorrect response. On easy trials, target duration was fixed at 150 ms. The stimuli remained on screen for 2 s and participants responded by pressing one of two buttons on a custom button-box with response mappings counterbalanced across participants. Following 1 to 3 s (duration randomly selected at 100 ms intervals), feedback appeared for 1 s ('+ correct +' in green or 'x wrong x' in red). After another interval (1-3 s), the second fixation appeared, followed by the same stimulus display. The intertrial interval lasted 1 to 2.5 s. This procedure allowed for an average of 6.2 s between the first and second responses of each pair. Participants completed 12 practice pairs



before entering the scanner. Trial order per block was randomized and the experiment was programmed in Visual Basic .NET.

Fig. 7.1 Diagrammatic representation of the task structure.

In pairs of trials a target letter ('T' or 'L') was presented and masked with an 'X' after a predetermined duration. After responding, participants received feedback that was orthogonal to their performance. This was followed by an identical second stimulus presentation. The first and second of each pair were denoted by fixation appearing as the numeral '1' or '2'. Event durations are shown below the task representation.

7.2.3 Neuroimaging acquisition

Subjects were scanned at the Wolfson Brain Imaging Centre in a 3 Tesla Siemens MAGNETOM Trio scanner. A T1-weighted scan was acquired for each participant, followed by a Siemens standard EPI sequence was used depicting BOLD contrast, with TR 2000 ms, flip angle 78°, TE 30 ms, in an interleaved ascending sequence. The field of view was 192x192 mm, with matrix 64x64, echo spacing 0.47 ms, and bandwidth 2442 Hz/Px. Volume number per run varied from 285 to 336. Each image volume comprised 32 slices

of 3-mm thickness, with in-plane resolution of 3x3 mm, orientated parallel to the anterior commissure–posterior commissure line.

7.2.4 Task performance

The main measure of interest was mean change on response two (R2), which was measured as a proportion (0-1). Response accuracy and reaction time were measures of secondary interest. LME models were fit with feedback (positive or negative), response accuracy of the first of each pair (R1: correct or incorrect), run (1-3) and an interactive term (feedback×R1 accuracy×run) as factors in R (R Core Team, 2020) using the lme4 package (Bates et al., 2015). Subsequent post-hoc pairwise comparisons of estimated marginal means with Tukey adjustment were conducted using the emmeans package (Lenth et al., 2018; Pinheiro et al., 2021). Various LMEs were tested that included different random intercepts and were subsequently compared using the Akaike Information Criterion (AIC), BIC and log likelihood. The LME including a random intercept for each subject and run had the lowest scores across all three measures and is thus reported here. The winning model was the following:

Mean change R2 ~ feedback × R1 accuracy × run + (1 + run | subject)

7.2.5 First-level models

First-level linear models were fit to fMRI data using FEAT (FSL) (Woolrich et al., 2001). This was done for each run and included six events: (1) R1 correct; (2) R1 incorrect; (3) positive feedback; (3) negative feedback; (5) R2 change and (6) R2 repeat. Equivalent event types for easy trials and six movement parameters (x, y, z, pitch, roll, yaw) resulting from the image realignment to control for movement artefacts were also included. Contrasts included only difficult trials and consisted of: (1) R1 incorrect vs correct; (2) negative vs positive feedback and (3) R2 change vs repeat.

7.2.6 Higher-level models

In the first second-level model, the first-level models for each subject were averaged across the three runs. An additional model was tested in which the effect of run was included to assess changes in activity over time. The contrasts above were subsequently examined in third-level mixed-effects whole-brain analyses involving one-sample t-tests with cluster thresholding with a Z-threshold of 3.1 and p<0.05 (Woolrich et al., 2004). FSLeyes was used

to generate the results figures (Smith et al., 2004). In these fMRI figures, the right and left hemispheres of the brain are inverted from the observer's perspective.

7.2.7 Brain behaviour correlations

In order to assess the effect of run/session on signal in the striatum, which arose as a question of interest after the higher-level model results, the relationship between striatal activity and R2 change was assessed. Two ROIs, one comprising bilateral VS and the other bilateral DS, based on the MNI structural atlas were created (Collins et al., 1995). For each subject, the mean signal in each of these ROIs across three runs during positive or negative feedback was calculated. The signal was then correlated with mean R2 change after positive or negative feedback, respectively, using PCC, accounting for multiple comparisons with FDR correction (Benjamini and Hochberg, 1995). The data were tested for normality using the Shapiro-Wilk test prior to deciding which correlation measure to use (Shapiro and Wilk, 1965). The data was found to be normally distributed; therefore, PCC was used.

7.2.8 Psychophysiological interaction analysis

In order to identify which circuits may be involved in changing and repeating an action, a whole-brain psychophysiological interaction (PPI) analysis was run in FSL. Four ROIs were created based on the findings from the first analyses, as well as areas reported to mediate proactive cognitive flexibility in previous studies. The regions that were selected for PPI analysis included the caudate nucleus, AI, OFC and SFG. These areas were selected based on literature suggesting their involvement in shifting and flexibility, as well the results from the first analysis showing that the AI, OFC and SFG are more active during a change than a repeat trial. A 5 mm sphere was placed in the respective region based on the Harvard-Oxford Structural Atlas supplied with FSL. Mean timeseries for each participant and each ROI were extracted and used for subsequent analyses. The purpose of a PPI analysis is to test whether there is an interaction between the timeseries from a prespecified brain region and a cognitive process to account for the neural responses observed in other brain regions (Neufang et al., 2008). The design matrix for this analysis included the extracted time series, the response to the cognitive process of interest (change vs repeat) and a term representing the interaction between the two. Six one-sample t-tests for each seed-region were conducted (for the change-repeat and repeat-change contrasts). Clusters above a z-threshold of 2.5 and p<0.05 were considered significant.

7.2.9 Multivariate pattern analysis

A Multivariate Pattern Analysis (MVPA) was run on non-normalised and unsmoothed data to reduce distortions using nilearn and sklearn in Python3.6 (Abraham et al., 2014; Pedregosa et al., 2011). The MVPA was based on the whole-brain individual beta maps obtained from the univariate analysis, more specifically, the beta maps for feedback, R1, and beta maps for the two categories were combined by concatenation. The individual beta maps were represented as separate points in a high-dimensional space, in which a linear decision boundary was determined through a hyperplane. The decision boundary separated the scans based on the labels assigned to them, in our case, changing or repeating a response on the next trial. The classifier identified a hyperplane that could best separate the provided input space (Li et al., 2014). The SVM applied a linear kernel, rather than a non-linear kernel to minimise overfitting, using a fixed regularization parameter (C = 1). This parameter C controls the trade-off between having no training errors and allowing misclassifications and was set to 1 based on previous publications (Yang et al., 2019; Vilgis et al., 2022). Crossvalidation was performed using the leave-one-out method. This method excludes one subject from each group, trains the classifier on the remaining data, and then tests performance on the excluded participants. This procedure is repeated for each subject. Subsequently, the classification process was repeated 1000 times with a different random permutation of the training labels. The purpose of permutation testing was to determine whether the overall classification accuracy was statistically significant and was used to derive a p-value. performance was evaluated using: (1) the F1 score, (2) the area under the receiver operating characteristic curve (AUROC), and (3) classification accuracy. Maps visualising decoder weights showed which areas contributed to classifying whether the subject would change or repeat their choice on the subsequent stimulus presentation.

7.3 Results

7.3.1 Task performance

It was found that mean R2 change was significantly affected by R1 accuracy (F(1,394)=16.9, p<0.001, $\eta^2 p=0.04$) and feedback (F(1,394)=31.2, p<.001, $\eta^2 p=0.07$). When feedback was negative, mean R2 change was higher than when it was positive (M=0.39, SD=0.24; M=0.23, SD=0.22; respectively). When R1 was incorrect, mean R2 change was also higher compared to when it was correct (M=0.41, SD=0.24; M=0.21, SD=0.19, respectively). Furthermore, there was a significant feedback×run interaction (F(1,394)=3.96, p=0.047, $\eta^2 p=0.01$), driven

by participants changing less after negative feedback with each run, i.e., they learned to disregard negative feedback (t(394)=-2.00, p=0.047). Mean changing after positive feedback did not increase or decrease across the three runs (t(394)=-0.40, p=0.69) (Figure 7.2).



Fig. 7.2 Effect of feedback and run on R2 change.

A linear mixed-effects model for R2 change revealed significant feedback and feedback×run interactive effects. Participants changed their response more following spurious negative feedback compared with spurious positive feedback, although the effect of the negative feedback on R2 change reduced across the three runs. Error bars represent standard error of the mean.

Mean R1 accuracy was above chance at 0.59 (SD=0.071, 95%CI[0.56,0.61]), albeit quite low, suggesting participants found the task challenging to complete. Mean R1 accuracy was significantly affected by run as subjects' performance improved (F(1,79)=4.35, p=0.040). Mean R2 accuracy was higher, at 0.64 (SD=0.071, 95%CI[0.61,0.66]). R1 accuracy and R1 accuracy×feedback significantly affected R2 accuracy (F(1,394)=403, p<0.001, η^2 p=0.51 and F(1,394)=63.3, p<0.001, η^2 p=0.14, respectively). Following correct R1, mean R2 accuracy was higher than after incorrect R1 (M=0.78, SD=0.19; M=0.40, SD=0.25, respectively).

When R1 was correct, R2 accuracy was higher when feedback was positive than when it was negative (t(394)=-5.06, p<0.001). When R1 was incorrect, R2 accuracy was higher when feedback was negative than positive (t(394)=6.19, p<0.001). An analysis of mean RT2 (mRT2) found a decrease in mRT2 across the three runs (F(1,39)=9.88, p=0.0032). There was also a feedback×R1 accuracy interaction (F(1,394)=4.21, p=0.41), with faster responding when R1 was correct and feedback was positive (t(394)=5.23, p<0.001).

7.3.2 Task-related brain activation - Changing versus repeating a response

Brain activity in the in the SFG, paracingulate gyrus, cingulate gyrus, insular cortex, OFC, caudate, putamen, thalamus, frontal pole, intracalcarine and supracalcarine cortices and cerebellum was associated with repeating a response. Similar areas were activated when a participant changed their response. The R2 change-repeat contrast found that there was greater activity in the paracingulate/cingulate gyrus, AI, superior, middle and inferior frontal gyri, superior and middle temporal gyri, OFC, frontal pole, caudate as well as in the lateral occipital cortex and cerebellum when a response was changed, rather than repeated (Fig. 7.3 and Table 7.1).

7.3 Results



Fig. 7.3 The R2 change versus repeat contrast.

a) Axial slices (MNI Z = -10 to -1) of the contrast highlighting key regions activated more when participants changed compared to when they repeated their responses. b) Sagittal view (MNI x=6 to 4) of the contrast highlighting key regions activated more when participants changed compared to when they repeated their responses. Activations detected with a wholebrain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 3.1 and p<0.05. The colour bar represents the t-statistic.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm ³)	Mean z- statistic
Frontal Pole	10	R	30, 64, -4	367	12386	3.49
Frontal Pole	10	L	-35, 59, -5	199	6716	3.60
Anterior Insular Cortex	13,16	R	38, 20, 2	96	3240	3.76
Anterior Insular Cortex	13,16	L	-40, 20, 2	86	2903	3.75
Frontal Orbital Cortex	11,12	R	42, 26, -7	167	5636	3.67
Frontal Orbital Cortex	11,12	L	-50, 25, -8	119	4016	3.73
Superior Frontal Gyrus	6,8	R	3, 36, 44	251	8471	3.64
Superior Frontal Gyrus	6,8	L	-7, 27, 47	168	5670	3.85
Middle Frontal Gyrus	46	R	47, 35, 31	252	8505	3.53
Middle Frontal Gyrus	46	L	-48, 18, 42	93	3139	3.39
Paracingulate Gyrus	32	R	3, 29, 42	168	5680	3.71
Paracingulate Gyrus	32	L	-7, 40, 27	168	5680	3.80
Anterior Cingulate Gyrus	24	R	8, 30, 24	104	3510	3.58
Anterior Cingulate Gyrus	24	L	-6, 28, 27	67	2261	3.46
Inferior Frontal Gyrus	44,45	R	45, 20, 9	182	6142	3.50
Inferior Frontal Gyrus	44,45	L	-48, 18, 5	161	5434	3.55
Middle Temporal Gyrus	21	R	58, -25, -6	92	3105	3.50
Middle Temporal Gyrus	21	L	-59, -25, -6	34	1147	3.58
Angular Gyrus	39	R	51, -53, 43	294	9922	3.72
Angular Gyrus	39	L	-55, -54, 43	363	12251	3.95
Supramarginal Gyrus	40	R	56, -44, 45	225	7594	3.58
Supramarginal Gyrus	40	L	-57, -46, 45	295	13331	3.89
Lateral Occipital Cortex	19	R	38, -63, 45	163	5501	3.88
Lateral Occipital Cortex	19	L	-37, -61, 45	275	9281	3.96
Precuneus Cortex	7	R	6, -68, 45	11	371	3.38
Precuneus Cortex	7	L	-6, -73, 45	38	1282	3.50

A novel perspective on cognitive flexibility

Table 7.1 Change-repeat contrast summary.

Summary of fMRI peak activity for the R2 change vs repeat contrast (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 3.1 and p<0.05). BA – Brodmann area; MNI – Montreal Neurological Institute template.

7.3.3 Activations associated with R1 and feedback

Both incorrect and correct R1 activated the cingulate and paracingulate gyri, superior, middle, and inferior frontal gyri, insular cortex, lateral occipital cortex and superior parietal lobule. The incorrect-correct contrast revealed greater activations in the paracingulate, cingulate and superior frontal gyri and the right angular gyrus during an incorrect response. Positive and negative feedback both lead to activity in the superior and middle temporal gyri, occipital cortex, and right frontal pole. The response was significantly greater for negative feedback than for positive feedback only in the superior frontal and superior temporal gyri.

7.3.4 Dorsal and ventral striatum are associated with learning to disregard feedback

Exploratory analyses were run to determine whether activations during the main events of interest (R1, feedback and R2) changed across the three runs. The response to correct R1 increased across the runs in the frontal pole and vmPFC. For incorrect R1, the signal in the DS decreased with time in both hemispheres. Activity in the left DS increased following both positive and negative feedback (Fig. 7.4). Additionally, an increase in the vmPFC, occipital pole and cerebellum was observed in the negative feedback condition. In the R2 change and repeat conditions, only activity in the occipital pole increased with run.

As a result of the observations made for R2 change over time in response to feedback, it was decided to extract the mean signal in the dorsal and ventral striatum during the presentation of positive or negative feedback using anatomical ROIs for these two regions. Then, the mean signal was correlated with mean R2 change after positive or negative feedback, respectively. There was a significant positive correlation between mean R2 change and VS activation after positive feedback (r(40)=0.39, p=0.039). No other correlations survived multiple comparisons correction.



Fig. 7.4 Change in signal in response to positive and negative feedback across the three runs. a) Voxels that show an increased signal in response to positive feedback across the three runs. The voxels in the left striatum are highlighted (MNI coordinates: X=-13, Y=9, Z=13). b) Voxels positively correlated with run when negative feedback is presented. The voxels can be seen in the left striatum, occipital pole and cerebellum (MNI coordinates: X=-19, Y=17, Z=3). Results were obtained with a mixed-effects whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 3.1 and p<0.05. c) Scatter plot showing the relationship between mean ventral striatal activation across the three runs when positive feedback is presented and mean R2 change after positive feedback.

7.3.5 Neural circuits implicated in the repeat and change conditions

PPI analyses highlighted important interactions between brain areas involved in modulating responses on this novel task. When the AI was used as an ROI in the analysis, it was found that there is greater connectivity between the AI and occipital cortex in the repeat compared to the change condition (Fig.7.5). The areas of the occipital cortex include the lateral occipital cortex and occipital pole, with stronger connections seen with the left hemisphere. When the OFC sphere was used as the seed region, similar findings were made. There was stronger connectivity in the repeat condition between the OFC ROI and occipital areas, including the lingual gyrus, occipital pole, and occipital fusiform gyrus (Fig. 7.6). When the SFG PPI analysis was run, there was enhanced connectivity between this region and the left posterior cingulate gyrus and precuneus cortex when a person was changing a response rather than repeating it (Fig 7.7). No observations were made when the caudate nucleus was the ROI.



Fig. 7.5 Results of the AI psychophysiological interaction analysis.

Brain areas with greater connectivity in the repeat condition when the AI ROI was used in the PPI analysis in the repeat-change contrast (MNI coordinates: X=-28, Y=84, Z=-4). These areas include the lateral occipital cortex and occipital pole. The ROI sphere is represented in white, and the areas highlighted by the PPI analysis are shown in yellow. Results were obtained via a mixed-effects whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05.

A novel perspective on cognitive flexibility

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm³)	Mean z- statistic
Lateral Occipital Cortex	19	L	-30, -82, 24	58	1958	2.87
Occipital Pole	18	L	-30, -95, 19	32	1080	2.85

Table 7.2 AI PPI repeat-change contrast summary.

Summary of fMRI peak activity for the PPI contrast of change vs repeat with AI as the seed-region (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). BA – Brodmann area; MNI – Montreal Neurological Institute template.



Fig. 7.6 Results of the OFC psychophysiological interaction analysis.

Brain areas with greater connectivity in the repeat condition when the OFC ROI was used in the PPI analysis (MNI coordinates: X=-3, Y=-91, Z=-4). Areas include the lingual gyrus, occipital fusiform gyrus and occipital pole. The ROI sphere is represented in white, and the areas highlighted by the PPI analysis are shown in yellow. Results were obtained via a mixed-effects whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05.

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm ³)	Mean z-statistic
Lingual Gyrus	17,18	L	-8,-86,-5	47	1586	2.84
Occipital Pole	18	R	-9, 93, -4	18	608	2.82
Occipital Pole	18	L	-5, -93, -5	45	1519	2.89
Occipital Fusiform Gyrus	37	L	-19, -88, -11	33	1114	2.84

Table 7.3 OFC PPI repeat-change contrast summary.

Summary of fMRI peak activity for the PPI contrast of change vs repeat with OFC as the seed-region (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). BA – Brodmann area; MNI – Montreal Neurological Institute template.



Fig. 7.7 Results of the SFG psychophysiological interaction analysis.

Areas that show greater activity with the SFG ROI in the change compared to the repeat condition (MNI coordinates: X=-3, Y=0, Z=24). These areas include the posterior cingulate gyrus and precuneus cortex. The ROI sphere is represented in white, and the areas highlighted by the PPI analysis are shown in yellow. Results were obtained via a mixed-effects whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05.

A novel perspective on cognitive flexibility

Name	BA	Side	MNI coordinates (X, Y, Z)	Number of voxels	Volume (mm ³)	Mean z- statistic
Cingulate Gyrus, posterior	23,31	L	-4, -47, 26	42	1418	2.81
Precuneus Cortex	7	R	4, 61, 38	8	280	2.66
Precuneus Cortex	7	L	-3, -61, 28	44	1485	2.82

Table 7.4 SFG PPI change-repeat contrast summary.

Summary of fMRI peak activity for the PPI contrast of change vs repeat with the SFG as the seed-region (whole-brain analysis involving one-sample t-tests with cluster thresholding with a Z-threshold of 2.5 and p<0.05). BA – Brodmann area; MNI – Montreal Neurological Institute template.

7.3.6 Multivariate Pattern Analysis: changing one's mind

The MVPA analysis enabled us to test if it is possible to predict whether a participant would change or repeat their response on the second pair of a trial. Using this method, it was possible to predict R2 change from brain activity during R1, feedback presentation and the combination of the two. Accuracy reached 0.64 when the beta coefficient maps for R1 accuracy were used. When the feedback beta coefficient maps were tested, predictive accuracy reached 0.74. When R1 and feedback were combined, performance achieved was superior with an accuracy of 0.77. Using fMRI data improved predictive accuracy compared to solely fitting an LME model to behaviour, as accuracy maximally reached 0.70 in that case. A summary of performance metrics, including the AUROC and F1 scores, for the different models can be found in Table 7.5.

Model								
Response 1								
Measure	Response 1	Feedback	+ Feedback					
Accuracy	0.64	0.74	0.77					
AUROC	0.57	0.83	0.86					
F1	0.66	0.74	0.78					

Table 7.5 MVPA analysis summary of performance metrics.

Summary of MVPA analysis performance metrics. In the first analysis, classification was based purely on the beta maps from R1. The second analysis uses the beta maps from the feedback conditions, and the last analysis combined the two by concatenating the beta maps. Performance was measured using the accuracy, AUROC and F1 scores. AUROC – Area Under the Receiver Operating Characteristic Curve.

Maps visualising decoder weights showed which areas helped classify whether participants would repeat or change their response. Areas that contributed to classifying R2 responses based on activity during R1 included the cingulate cortex, caudate, putamen, inferior and middle frontal gyri, frontal pole, occipital cortex, with additional voxels distributed across the brain. When the feedback beta maps were used, the areas involved in classification were the AI and OFC, as well as the SFG, paracingulate gyrus, cingulate gyrus, caudate and frontal pole (see Fig. 7.8).



Fig. 7.8 Results of the multivariate pattern analysis.

MVPA decoder weights for change and repeat conditions based on beta maps during the feedback presentation. The voxels shown contributed to the decoder deciding whether the participant would change or repeat their response on the next stimulus presentation. Blue voxels aided in the classification of a repeat trial, orange voxels helped classify a change trial. MNI coordinates: X=-5, Y=15, Z=4.

7.4 Discussion

In this study, I present a novel task enabling the assessment of proactive cognitive flexibility which does not involve reward or rule-based learning, but instead requires the subject to adjust their behaviour voluntarily. Behaviourally, it was observed that changing a response on the second stimulus presentation was affected by feedback, accuracy on the first response, as well as run. In particular, subjects were more likely to shift if their first response was incorrect and if feedback was negative. Additionally, changing a response after negative feedback decreased with each run, indicating that participants learnt that feedback was unreliable. The reaction time on the second response also decreased across the three runs, suggesting that participants were becoming more confident in their decisions (Desender et al., 2019).

Previous studies have identified the ACC and SFG to be two key structures for error and action monitoring (Bonini et al., 2014; Botvinick et al., 1999; Carter et al., 1998; Veen et al., 2001). Therefore, the activity in these two areas when a response was incorrect likely reflects their involvement in error signalling. Electrophysiological evidence from primates has shown that ACC neurons not only respond to errors, but also signal prior to adjustments following an error, thereby contributing to selection of the next movement (Shima and Tanji, 1998). Activity in the ACC has also been found to reflect a behavioural shift on the subsequent trial on reversal learning tasks (Kawai et al., 2015). Previous studies in humans have investigated the neural substrates of error signalling and error detection (Ullsperger et al., 2014; Veen and Carter, 2002). The findings presented in this chapter complement these by focusing on the subsequent behavioural adjustment, i.e., on the instantiation of flexibility through a change on the second response. This reflects the importance of the ACC when behavioural adjustments are implemented, which has thus far received less attention in humans (Ridderinkhof et al., 2004).

In line with findings from other studies employing tasks of behavioural shifting, it is found that the IFJ, IFG, AI, ACC and dlPFC play a key role in mediating response switching in our task (Braun et al., 2015; Dajani et al., 2020; Shine et al., 2016). The IFJ is thought to mediate response shifting via the integration of multiple behavioural control processes and interactions with the dlPFC, vlPFC, AI and putamen (Cole and Schneider, 2007; Levy and Wagner, 2011; Sundermann and Pfleiderer, 2012). The AI and IFG, on the other hand, are important for response inhibition, whilst having dissociable roles (Cai et al., 2014). The former is intrinsically functionally connected to the ACC and important for salience detection of unexpected or infrequent events. AI involvement here is unlikely due to detection of an

unexpected stimulus as in the case of inhibition in stop signal or Go/No-go tasks. Instead, given changing was more infrequent than staying, it may be associated with the salience of an internally generated infrequent action. The IFG has stronger connectivity with the dlPFC as part of the ECN and facilitates appropriate behavioural control (Cai et al., 2014). The recruitment of both the AI and IFG may be attributable to inhibition of the first response, highlighting the intrinsic role of inhibition in purposeful flexible adjustment of behaviour.

The PPI analysis gave an insight into which brain circuits give rise to repeating or changing a response. When the AI was used as the seed region, stronger connectivity to the lateral occipital cortex and occipital pole in the repeat compared to the change condition was found. This was also the case for the connectivity from the OFC to the occipital pole, lingual gyrus, and occipital fusiform gyrus. A recent publication reported results from an automated meta-analysis of cognitive shifting, updating and inhibition, which highlighted not only the importance of the IFJ, AI and ACC, but also the angular gyrus and visual cortex, areas whose role in cognitive flexibility is oftentimes overlooked (Uddin, 2021). Reports from multiple task-switching studies further confirm the involvement of occipital regions such as the lateral and medial occipital cortices (Armbruster et al., 2012; Dajani et al., 2020). In mice, ACC projection neurons to the visual cortex important for error monitoring have been directly linked to post-error behavioural adaptation through fibre photometry and optogenetic manipulations (Norman et al., 2021). In humans, a causal dependency between the frontal operculum and three occipitotemporal regions in cognitive control has also been demonstrated using transcranial magnetic stimulation, suggesting top-down regulation of visual areas by the AI and frontal operculum (Higo et al., 2011). Whether occipital areas in our task were solely activated due to the stimuli being visual is unclear. Nonetheless, these findings support an involvement of the AI/OFC and connections to occipital regions in cognitive flexibility.

An unexpected finding was the increased activity in the left caudate and putamen during feedback presentation across the three runs. This change in activity may reflect the decrease in response shifting after negative feedback with each run. The striatum is a critical node in response inhibition, particularly proactive inhibitory control. This is in contrast to circuits passing through the subthalamic nucleus, which are involved in reactive inhibition, as cortical signals via the hyperdirect pathway reach the basal ganglia faster than if they would travel through the striatum (Aron and Poldrack, 2006; Vink et al., 2005; Zandbelt and Vink, 2010). Chemogenetic inactivation of frontostriatal neurons in rats is known to impair behavioural inhibition, electrophysiological striatal recordings in macaques reflect future events and the striatum has been found to be activated during a task requiring proactive adjustments of

A novel perspective on cognitive flexibility

response strategies (Blazquez et al., 2002; Terra et al., 2020; Vink et al., 2005). Moreover, these findings are compatible with the established role of the DS in behavioural and cognitive control, decision-making under uncertainty and signalling ambiguity (Hsu et al., 2005; Lopez-Paniagua and Seger, 2013; Mestres-Missé et al.; 2017). Therefore, it seems plausible that the increased activity in the DS observed is representative of processes inhibiting an upcoming response tendency arising from the feedback presented and reflect its contribution to decision-making and response selection (Zandbelt and Vink, 2010).

Using the machine learning MVPA approach, the amount of information gained from fMRI data can be increased, as it considers the response patterns of all voxels that may otherwise not be taken into account in traditional univariate methods (Weaverdyck et al., 2020). Indeed, this approach has provided additional information about the nature of repeating versus changing a response on this task. When feedback was presented, the algorithm could predict the participants next action with an accuracy of 77%. As the prediction of behaviour in this tightly controlled situation from performance data was 70%, the BOLD voxel-wise approach enhanced the predictive properties of the data. Decoder weight maps highlighted that voxels in the AI, OFC, frontal pole, SFG, cingulate gyrus and caudate strongly contributed to predicting a change event on the subsequent stimulus presentation. The frontal pole, AI and OFC were also important for predicting a repeat trial. These results suggest that brain activity prior to an action can encode contributors to the decision that will be made, in this case, representing antecedents to changing one's mind.

The novel task presented in this chapter created high uncertainty via brief stimulus presentation times and spurious feedback. It differs from previous tasks of cognitive flexibility, as it does not require ongoing rule-based learning and allows subjects to repeat or change their response of their own volition. Despite the differences in task implementation, parallels with the reversal learning task can be drawn. Participants change or repeat a response regardless of feedback on the CYM task, reminiscent of behaviour represented by the κ parameter discussed previously. In RL, the κ parameter may be representative of a volitional component, based on internally generated signals. This behaviour has also been described as value-free (Miller et al., 2019). The CYM task makes it possible to separate this value-free component from value-based components. However, behaviour on the CYM task is not free operant, i.e., it is not modifiable by its consequences (Staddon and Cerutti, 2003). Nonetheless, the results presented here may provide an additional insight into which areas mediate value-free behaviours in RL, such as the SFG, cingulate gyrus, insular cortex and DS, as these were areas active when a response was repeated. The involvement of the insular cortex and DS

in habitual responding has also been demonstrated in chapter 3, suggestive of the fact that behaviour on the RL and CYM task reflect similar mechanisms.

Though effective in eliciting participants to change their minds, the task could not dissociate specific contributions of internal versus external sources of uncertainty. Future studies may disentangle these or explore alternate contributors to shifting representative of everyday situations. The task does assist with some difficulties encountered by the impurity problem, which makes it difficult to untangle joint and separable aspects of executive processes (Miyake et al., 2000; Uddin, 2021). Namely, a challenge in studying cognitive flexibility has been to isolate it from working memory and attention demands. These processes were likely kept constant in the change vs repeat contrast, with the task having minimal spatial and rule learning requirements. Future studies could employ the same task with auditory stimuli to verify modality independent involvement of key PFC nodes. Moreover, this could reveal whether occipital region activation is due to the visual nature of most flexibility tasks (Niendam et al., 2012). Ultimately, the present findings widen our view of flexibility in everyday life by allowing a targeted assessment of volitional decision-making.

Chapter 8

Discussion

This PhD thesis set out to investigate how behavioural RL and cognitive flexibility processes are altered in multiple psychiatric disorders, with a focus on major depressive and substance use disorder. Additionally, the neural substrates underlying the observed behavioural differences were explored using neuroimaging techniques. Besides highlighting behavioural and neural RL changes in different patient groups, a new insight into cognitive flexibility was provided via a novel task. Behaviour on this task was value-free, making it possible to draw parallels with the κ RL parameter investigated in the other chapters. Computational modelling and neuroimaging approaches were used across two species, rats and humans, making it possible to translate findings between them. This last chapter of the thesis will aim to bring together the findings from the different chapters, highlight the most important results and build a translational bridge on a behavioural and neural level between species.

8.1 Reinforcement Learning in Depression

8.1.1 RL in rodents exposed to REMS and adulthood stress

In chapter 3, behavioural and neural RL processes were investigated in rats exposed to repeated early-life maternal separation and adulthood stress for the first time. The MS paradigm is commonly used to induce depression-like phenotypes in rats and to investigate the effects of early-life stress. ELS is known to predispose to psychiatric conditions such as MDD and posttraumatic stress disorder in humans (Bölükbas et al., 2020; Syed and Nemeroff, 2017). It has been shown that ELS in rodents and humans leads to reduced cognitive flexibility and decision-making (Goodwill et al., 2018; Harms et al., 2018; Hyer

Discussion

et al., 2021). However, the effects of a 'two-hit' model involving ELS and adulthood stress on RL in rodents have not been investigated before.

Control males showed an increase in lose-shift behaviour after a correct response following adulthood stress, accompanied by a decrease in the learning rate for non-rewarded trials $\alpha_{non-rew}$. MS males were unaffected by the second stressor but showed fewer perseverative responses than control males beforehand. In females, only effects on the lose-shift probability after an incorrect response were observed; this measure was significantly lower in MS females prior to the stressor but increased after secondary stress exposure. This outcome aligned with greater values of the stickiness parameter κ in MS females prior to the stressor. κ increased in control females after adulthood stress. These results suggest that in males, the REMS procedure enhanced resilience to stress later in life, whereas in females, REMS and the secondary stress had a non-additive effect on the animals' behavioural profile. In our study, the latter was reflected by an increased tendency to repeat choices regardless of previous outcomes. In other words, MS females' choices were less driven by reward and punishment. This suggests blunted reward processing, an effect previously observed in humans, which has been related to anhedonia (Safra et al., 2019; Weinberg and Shankman, 2017). Stress is known to disproportionately affect females, increasing their risk for stress-related pathologies that lead to cognitive impairment, an effect that was also shown by the findings reported in this chapter (and see Goodwill et al., 2018).

Analysis of rs-FC revealed that changes in the connectivity of the BLA may underlie the behavioural observations. In control males and MS females, lower strength connectivity of the BLA to the cingulate, IL, AI and DS was observed compared to MS males and control females. These findings are consistent with other studies where reduced amydalar connectivity following ELS in rodents was reported (Cohen et al., 2013; Johnson et al., 2018). Additionally, the RL parameter κ , which was significantly affected in MS females, was linked to connectivity between mOFC and the cingulate, IL, PrL, and insular cortex. This is a unique insight into the circuits giving rise to this RL parameter, which may help inform our understanding of which circuits may mediate it in humans. Pharmacological inactivation of the IL in rats has previously been shown to decrease κ (Verharen et al., 2020). Moreover, lesioning this area impairs habitual behaviour, suggesting the IL is an important component of the network mediating habitual responding (Killcross and Coutureau, 2003). Given that both 'sticky' and habitual behaviour are value-free, the results suggest a potential common involvement of the IL.

8.1.2 RL in patients with MDD and/or SUD

In chapter 6, PRL data from a naturalistic sample of patients with a diagnosis of MDD, either on its own or with a comorbid diagnosis, were analysed using RL models. The additional diagnoses included SUD or ASD. Surprisingly, none of the conventional measures differed between these groups, which was also found in the original publication of this study (Brolsma et al., 2020b). From the five models tested, the best-fitting model included three parameters: a joint learning rate, α ; reinforcement sensitivity, β ; and stimulus stickiness, κ_{stim} . The α parameter was solely affected in SUD, with patients in this group having higher values. The reinforcement sensitivity parameter β was decreased in multiple patient groups, including the SUD, SUD/MDD and SUD/ASD group. The stimulus stickiness rate was increased in both SUD and SUD/MDD groups. MDD only patients showed no differences on any of the parameters. Surprisingly, none of these parameters were found to be correlated with clinical measures, such as the IDS, ASI or AQ.

Although these results were unexpected, a consensus on how RL parameters are altered in MDD has not yet been reached. Some studies report reduced learning rates from rewarded trials, α_{rew} , and/or increases in the learning rate from non-rewarded trials, $\alpha_{non-rew}$ (Admon and Pizzagalli, 2015; Huys et al., 2013). However, null results have also been reported (Gradin et al., 2011). A prominent theory of MDD suggests that reward processing is significantly altered in MDD and that MDD patients exhibit a negative bias on tasks of perception, attention and memory (Eshel and Roiser, 2010; Gotlib and Joormann, 2010). In another study, it was shown that reinforcement sensitivity, rather than learning rates, are correlated with anhedonia severity (Huys et al., 2013). In the study presented here, measures of anhedonia were not collected. Therefore, it was not possible to investigate whether RL parameters correlate with anhedonia, which may explain the discrepancy in the results. Further, the task employed here was different to PRL tasks used in other studies, as in this study, the stimuli could appear in one of four locations to prevent a side bias, which has not been the case in previous reports (Brown et al., 2021; Dombrovski et al., 2013; Mukherjee et al., 2020). Finally, the findings demonstrate that the diagnosis of a comorbidity differentially affects RL measures, underlining the importance of diversity within patient samples.

8.1.3 Building bridges between human and rodent research studies

One of the aims of this thesis was to build translational bridges between the behavioural and neural markers of RL in psychiatric disorders between rats and humans. However, the

Discussion

behavioural profiles arising from MS and the adulthood stressor in rats were not replicated in MDD patients. Whereas the $\alpha_{non-rew}$ and κ parameters were affected by both interventions in rats, patients with MDD showed no differences on either conventional or RL measures. There may be multiple explanations for this. Firstly, the 'two-hit' model in rodents is not the equivalent of MDD in humans. Although ELS is a factor that increases the risk of developing a neuropsychiatric disorder later in life, not all individuals that have experienced ELS develop MDD or other mental health conditions (Kessler et al., 2010; Targum and Nemeroff, 2019). Individuals that have experienced ELS show decreased positive feedback sensitivity and a reduced learning rate (Wilkinson et al., 2021). The learning rate is also negatively impacted by acute stress (Raio et al., 2017). Moreover, ELS has previously been shown to increase perseverative tendencies in female rats, and humans that have experienced stress in childhood show greater habitual responding than control participants (Brydges et al., 2015; Gordon et al., 2020). These findings more closely reflect the observations made in rodents that experienced MS and adulthood stress (chapter 3). Therefore, a better comparison can be drawn between MS in rats and individuals that have experienced ELS.

Another explanation for the diverging results may be that MS animals did not differ from control animals on the SPT, a measure of anhedonia. Although it has been questioned whether the SPT truly provides a measure of anhedonia in rats, together with the PRL findings in males, reflecting non-depressive behaviours, it raises the question of whether the REMS paradigm causes depressive-like phenotypes (Scheggi et al., 2018). Additionally, sucrose intake may not be an appropriate measure for females, as they tend to drink more sucrose than males and show erratic increases in consumption (Dalla et al., 2005). Null effects of REMS are not uncommon, and it is known that this procedure can result in enhanced resilience (Santarelli et al., 2017; Tan et al., 2017). Nonetheless, given the behavioural profiles of MS females and the associated connectivity changes, it does appear that REMS resulted in behavioural and neural effects in this sex only. It has also been reported that the MS paradigm in rats enhances vulnerability to drug abuse, suggesting that this intervention results in a profile reflecting addiction, rather than depression vulnerability (Alves et al., 2020). This would align with the increased κ observed in MS females and MS-induced compulsivity in previous studies (Brydges et al., 2015).

The BLA and mOFC, which are areas associated with goal-directed behaviour, were key regions in rats whose connectivity to other areas was altered following REMS. In humans with low chronic job stress, the OFC, vIPFC, anterior caudate, and insula are active during reversal learning tasks, which is not observed in men with high chronic job stress (Ohira et al., 2011). Altered amygdala connectivity following ELS in rodents and humans and in
patients with MDD has been frequently reported, suggesting that this area is involved in depressive phenotypes across species (Guadagno et al., 2021; He et al., 2019; Qin et al., 2019; Wackerhagen et al., 2020). The mOFC has also been previously reported to show altered connectivity in these groups (Cheng et al., 2016; Drevets, 2005). Further, the putamen and insular cortex have been found to be involved in habitual responding in humans (Eryilmaz et al., 2017; Tricomi et al., 2009). These areas were found to be correlated with the RL parameter κ in chapter 3, therefore, providing support for the fact that they may also be involved in mediating κ in humans.

8.1.4 Implications for the use of the maternal separation procedure to study depression

In the third chapter of this thesis, the 'two-hit' model of depression was used to elicit a depression-like phenotype in rats. Shortly after birth, pups were separated from their mothers for six hours a day between PND5 and PND19. In adulthood, at around PND260-280, they received the second 'hit' - repeated shock-stressors. Although MS is one of the most commonly used procedures to simulate the effects of ELS in rats, there is a large variability in the results, with some studies reporting increased anxiety- and depressive-like behaviours in adulthood, and other studies finding no effects (Masrour et al., 2018; Millstein and Holmes, 2007; Kalinichev et al., 2002; Savignac et al., 2011). Sex-specific effects have also been reported, but there is no clear consensus on which rodent sex is more vulnerable (Eklund and Arborelius, 2006; Lehmann et al., 1999; White et al., 2020). Not only have the inconsistent findings resulting from MS been criticised, but questions have been raised regarding whether MS reflects neglect, abuse or a combination of both (Murthy and Gould, 2018). In some studies, it has been observed that dams increased maternal care following separation, dampening the effects of the separation (Millstein and Holmes, 2007). In the study presented in chapter 3, MS did result in different behavioural profiles, mostly observed in females, and changes in behaviour and connectivity following the adulthood stressor. However, it did not result in increased symptoms of anhedonia, which would be expected. Moreover, it seemed that males were resistant to the separation. This raises the question how much translational value does this paradigm bring?

The great variability in the results of MS studies come down to differences in study design. Firstly, the genetic background of the animal is important, as some mouse or rat strains are more resistant to stress and others are more anxious (Millstein and Holmes, 2007; Savignac et al., 2011). Moreover, the timing and duration of MS may play a role. In studies during

Discussion

which the MS occurs at the same time every day, the intervention may become predictable and may therefore be less stressful for pups. It has also been suggested that there is a sensitive period during which rats are most vulnerable, an effect that does not exist in humans (Dunn et al., 2018; van der Kooij et al., 2015). The 'two-hit' model of depression was developed to mimic 'cumulative' stress related to mental health in humans, but null effects of this model have been presented (Santarelli et al., 2017; Tan et al., 2017). It has also been questioned whether the existing measures of anxiety or depression are sensitive to behavioural changes in females, as these were developed for use in males (Murthy and Gould, 2018; Verharen et al., 2019). Nonetheless, it is important to consider that there may exist individual variability in response to ELS in rodent populations, which may not be statistically significant due to the small sample sizes in rodent studies. In summary, the 'two-hit' model of depression in rats has provided insights into the consequences of ELS, but the factors influencing outcomes following ELS need to be investigated further and understood better.

8.1.5 Implications for investigating sex-dependent differences in depression

In preclinical pharmacological studies, male rodents were routinely used in the past, thus, most animal models in psychiatry were developed and validated in males (Beery and Zucker, 2011). It was thought that there is greater variability in female behaviour due to the oestrous cycle, requiring increased sample sizes and resulting in few studies including female rats (Beery and Zucker, 2011). Although it has been shown that females do not exhibit greater behavioural variability than males, and some funding bodies such as the NIH and MRC now requiring the inclusion of both sexes in rodent studies, many years of research solely on male animals need to be caught up on (Mogil and Chanda, 2005). Until 2020, only two studies investigating the effects of ELS in rodents using imaging techniques involved both females and males, even though considerable sex differences in mouse neuroanatomy have been reported (Honeycutt et al., 2020; Qiu et al., 2018; White et al., 2020). Thus, the findings from chapter 3 contribute significantly to our understanding of sex-specific functional connectivity differences arising from ELS. The inclusion of both sexes in animal studies will increase the validity of animal models in psychiatry and lead to improvements in prevention, diagnosis and treatment of psychiatric disorders (Kokras et al., 2015).

Clear sex-specific differences in behaviour and brain function following MS and adulthood stress were observed in the third chapter. This is not a novel observation; for example, ELS has been found to result in fronto-limbic hyperconnectivity in males, whereas there was either hypoconnectivity or no change in females (White et al., 2020). In another study, ELS impaired reversal learning in females only, which was associated with decreased expression and density of parvalbumin expressing interneurons in the OFC (Goodwill et al., 2018). In humans, sex-differences in the transcriptional patterns associated with MDD have been reported (Labonté et al., 2017). ELS has shown to reduce hippocampal and corpus callosum volume, with effects more pronounced in males than females, whereas another study has found reduced caudate size in females exposed to ELS that was not present in males (Frodl et al., 2017; Teicher and Samson, 2016). The oestrous cycle also affects depressive symptoms in humans, such as irritability, appetite, and insomnia (Kornstein et al., 2010). These results underline the importance of including both sexes in animal models of psychiatric disorders and understand sex differences in relation to these disorders better.

8.1.6 Implications for the study of reinforcement learning and decisionmaking in stress and depression

The findings in chapters 3 and 6 show that a variety of changes in decision-making and RL arise from early life and adulthood stress, as well as MDD. Not only were the learning rates α_{rew} and $\alpha_{non-rew}$ altered following repeated shock stressors in male rats, but all the parameters from the DDM were significantly altered. These parameters included the boundary separation α , bias β , drift rate δ and the non-decision time τ . This suggests that decision-making is greatly affected following chronic stress. Stress is argued to lead to a perceptual bias of the environment being very harsh, resulting in tendencies to exploit information, rather than explore other options, and has been linked to reduced cognitive flexibility (Kim and Lee, 2011; Lenow et al., 2017). Moreover, stress has been found to shift an individuals' cognitive focus to the present over the future (Frankenhuis et al., 2016). Although no changes in the RL exploit vs explore parameter were observed following MS or chronic stress in chapter 3, results from the DDM support these theories; the boundary separation increased following the repeated shock-stressor, indicating that these animals crossed a higher threshold before making a choice, which also reflects a higher threshold required to move on and explore other choices. Additionally, the bias β was altered, reflecting the changes in bias observed in humans following stress.

Acute and chronic stress have been shown to impair function of the PFC and shift activity to limbic and striatal structures to guide decision-making processes (Harms, 2017). The PFC is highly sensitive to stress, for example, through the effect of glucocorticoid release, which disrupts intracellular signalling pathways that impair PFC function (Arnsten, 2009). Chronic

Discussion

stress reduces spine number and causes dendritic retraction, also leading to impairments in PFC function (Dias-Ferreira et al., 2009; Joëls et al., 2007). Additionally, both acute and chronic stress increase amygdala and striatal function, facilitating habit-based learning and perseveration (Harms, 2017; Hermans et al., 2011; Schwabe and Wolf, 2009). Rats and humans exposed to ELS show reduced exploration, information sampling as well as reduced responsivity of the striatum to reward (Dillon et al., 2009; Humphreys et al., 2015). Adolescents that experienced ELS and showed reduced reward activity in the striatum had higher levels of depression; an inverse correlation between NAc activity and depression scores was observed (Goff et al., 2013).

The summarised results align well with our findings. We observe an increase in BLA activity following adulthood stress in a subset of animals, as well as strengthened connectivity from the mOFC to the VS. Moreover, there is a clear increase in habitual behaviour following maternal separation and adulthood stress as measured by the parameter κ in females, providing additional evidence for decision-making changes following ELS and chronic stress. Investigating these differences in more detail and understanding how ELS/chronic stress predisposes to psychiatric conditions such as MDD is important in order to prevent the onset of disease, particularly in individuals that have experienced childhood adversity. These results also provide support for the use of RL models in psychiatric research.

8.2 Reinforcement Learning in Substance Use Disorder

8.2.1 RL in rodents exposed to a cocaine self-administration paradigm

In the fourth chapter, behavioural and neural markers predisposing to high-compulsive drug-seeking in rats were identified. Although conventionally, rats showed no significant differences on any of the PRL measures, computational modelling could extract differences between low-, medium- and high-compulsive animals. Specifically, high-compulsive animals had greater values of the stickiness parameter κ than animals that showed lower levels of drug-seeking. This is in line with previous rodent and human studies, as increased levels of κ have been reported after intravenous cocaine SA in high-compulsive rats as well as in patients with SUD (Kanen et al., 2019; Zhukovsky et al., 2019). Increased perseverative responses, as measured via conventional PRL analyses, are also a known feature of compulsivity across species (Allen and Leri, 2014; Ersche et al., 2008). After cocaine SA in rats, the exploration vs exploitation parameter β has also been found to be increased in high-compulsive animals, indicating lower exploitation and higher exploration (Zhukovsky et al., 2019). This may

reflect a lower threshold for the animal to explore other options rather than exploit the information they are presented with. It is known that high impulsivity levels may predispose an individual to compulsive behaviours and is a feature present following drug abuse (Dalley et al., 2007; de Wit, 2009). Higher impulsivity levels may be linked to increased exploration, as impulsivity is the tendency to act prematurely and leads to rash actions without forethought (Dalley and Robbins, 2017). This is supported by a recent study in humans, which found that participants with methamphetamine use disorder (MUD) are impaired in exploitative decision-making and showed behaviour consistent with random exploration on the IGT (Robinson et al., 2022). In chapter 4, it was found that the exploitation vs exploration parameter, β , was positively correlated with the volume of the AI. In humans, the AI and ACC have been implicated in exploration (Addicott et al., 2014; Blanchard and Gershman, 2018; Chakroun et al., 2020). Their exact role is unclear, but it has been proposed that these regions reallocate attentional mechanisms to alternative salient options in the presence of increasing uncertainty (Laureiro-Martínez et al., 2015). Exploitation, however, has mostly been associated with the vmPFC, OFC, VS and hippocampus, which form a 'valuation' network (Daw et al., 2006; Laureiro-Martínez et al., 2013). Based on these studies, the link between AI volume and β is strengthened; increased β values suggest increased exploration, rather than exploitation, which is linked to an area of the brain that has previously been shown to mediate exploration in humans (the AI). This is a novel finding and provides insight into the neural mechanisms of this RL measure.

8.2.2 RL in patients with Cocaine Use Disorder and Gambling Disorder

The reward system is one of the main processes hijacked in SUDs such as CUD (Luijten et al., 2017). In another disorder of compulsivity, namely GD, reward-related circuitries are also dysregulated. However, in both disorders the exact effects are not entirely understood and the similarities and differences between the two still need to be identified (Clark and Goudriaan, 2018). In chapter 5, the behavioural and neural underpinnings of these two disorders of compulsivity were investigated, with the aim of elucidating converging and diverging aspects. No differences were found in any of the RL parameters in the CUD group, which was unexpected, given that both rodent and human studies have suggested increased perseveration and greater values of the RL parameter κ in SUD. Most reports in humans investigated reversal learning across all SUDs, rather than focussing on one type of SUD, which may explain the observed discrepancies. The side stickiness parameter

Discussion

was non-significantly increased in both GD and SUD, whereas stimulus stickiness was significantly *reduced* in GD. The latter finding was unexpected, but points at a separation of the two parameters into motor and stimulus perseveration, respectively. Increased motor perseveration in PGs has been reported before and was linked to reduced activity in the vIPFC (Ruiter et al., 2008). Although reduced stickiness towards stimuli has not been found in PGs before, it is known that they have higher levels of impulsivity compared to healthy individuals (Odlaug et al., 2011). This represents less 'sticky' behaviour, which may be an adaptive benefit of gambling, as it reflects quicker adaptation to strategies. This was also demonstrated by the lower RTs of PGs compared to CUD and HC groups.

Tracking of reward and punishment EV was reported to be altered in GD in chapter 5, as were positive and negative PPEs in CUD. Reward EV tracking in PGs showed higher activity in the cingulate gyrus, hippocampus and amygdala, which are important areas for reward processing (Haber and Knutson, 2010). When tracking punishment EV, the postcentral gyrus, superior parietal lobule, precuneus cortex and lateral occipital cortex showed reduced activity compared to controls. In CUD, BOLD response to positive PPE was reduced in the frontal pole, SFG, cingulate gyrus and OFC. The negative PPE response was increased in the superior and middle frontal gyrus. Stronger activity in the reward system during reward expectation has been previously shown in PGs, although no differences in punishment tracking have been found before (Holst et al., 2012). Impaired positive and negative RPEs have been observed in CUD in previous studies, with no effects on PPE being reported (Parvaz et al., 2015). In women with MUD, stronger reward expectations and stronger response to harm avoidance have been found (Wei et al., 2021). Overall, it is clear that there is an imbalance of reward and punishment processing in SUDs, which is supported by previous studies. The findings that PGs have reduced neural responses to punishment EV tracking and cocaine users have altered PPE processing are novel and may contribute to our understanding of how reward mechanisms in these groups are altered.

8.2.3 Building bridges between rodent and human studies

The results from the RL analysis from rodent and human data in this thesis did not fully converge, as higher levels of κ in high-compulsive rats were seen, but there were no changes in this parameter in CUD. A slight increase in the side stickiness parameter was seen in the CUD group, but this did not reach significance. The discrepancies may be due to rats completing a deterministic reversal learning task, whereas CUD patients completed the probabilistic version of this paradigm. Furthermore, rats were tested on this task prior to

cocaine SA, whereas the human data consisted of participants that already had a diagnosis of CUD. Moreover, the sample size of the CUD patient group was relatively small, as it consisted of 20 participants, which may be why no significant effects were found. As mentioned previously, the rodent and human literature on PRL has converged thus far; both rodent and human studies reported increased perseverative tendencies and greater κ values following drug use. In chapter 6, increased stickiness in SUD was also observed. Compulsivity is defined as perseverative behaviour in the face of adverse consequences, indicating that in high-compulsive individuals, perseverative tendencies are present (Wolffgramm and Heyne, 1995). Therefore, higher stickiness is expected in cocaine-dependent individuals.

Structurally, reduced brain volumes were found in the NAcC in high-compulsive animals, compared to low-compulsive ones. This effect has been observed in humans, such as in heroin use disorder and crack-cocaine abusers (Schuch-Goi et al., 2017; Seifert et al., 2015). The NAc plays a vital role in the brain's reward system, and it is known that the rewarding properties of psychostimulants depend on the mesolimbic dopamine system, which the NAc is part of (Koob and Volkow, 2009). In high-drinker compulsive rats, increased amygdala and decreased hippocampal volumes have been reported, without any effects on NAc volume (Mora et al., 2020). Only few MRI studies in rodents have investigated differences in brain structure prior to cocaine SA, therefore, no major comparisons can be drawn. Given that this volume difference is observed prior to the SA paradigm and is observed at a young age, it may be a useful biomarker for the early prediction of drug abuse. Although it is difficult to investigate whether this finding can be replicated humans, one could investigate whether relatives of individuals diagnosed with CUD have different NAc volumes compared to the general population.

Even though no fMRI analyses were included in chapter 4, some conclusions can be drawn from the whole-brain connectivity analyses conducted in chapters 3 and 5. In rodents, the parameter κ was(93)=2.45 positively correlated with rs-FC connectivity between the mOFC and the cingulate cortex, IL, PrL, DS and insular cortex across all animals, regardless of their group. This relationship between κ and functional connectivity would also be expected in the study presented in chapter 4, with stronger functional connectivity between these areas expected in high-compulsive compared to low-compulsive rats. In chapter 5, it is reported that the learning rate parameter α_{rew} is positively correlated with reward EV activity in the cingulate and paracingulate gyri, as well as the insular cortex and OFC. The cingulate and paracingulate gyri and occipital areas are also negatively correlated with this parameter during positive PPE activity. Interestingly, κ_{stim} had a positive correlation with activity in the MFG and IFG, which was strengthened in both patient groups, CUD and GD.

Discussion

Thus, the neural substrates of κ do not align between the two species. In rats, this measure is mediated by medial cortical structures, such as the medial PFC areas PrL and IL, and subcortical structures, such as the DS. In humans, it seems that lateral PFC structures give rise to this parameter and are altered in disorders of compulsivity. Regardless, both studies have identified the regions involved in κ in both species, which is of great importance for future studies researching stickiness in relation to MDD and SUD.

8.2.4 Relating conventional parameters to RL parameters

One important question that remains unanswered is how the conventional measures extracted from the PRL task are related to computational RL parameters. Gaining an understanding of this would help the interpretation of RL parameter results. In chapter 3, a positive correlation between the parameter κ and win-stay behaviour after a correct or an incorrect trial was found. Moreover, there was a positive correlation between α_{rew} and win-stay correct and the percentage of correct responses. In the next chapter, a positive correlation between the learning rate α and win-stay was found. This RL parameter was also negatively correlated with the total trials to criterion and the number of perseverative errors. β was positively correlated with lose-shift, and κ had a positive relationship with win-stay and a negative relationship with lose-shift behaviour. The correlation analysis in chapter 5 highlighted a positive relationship between α_{rew} and the proportion of correct responses, as well as a negative relationship with trials to criterion. This parameter also had a positive correlation with lose-shift incorrect and win-stay. $\alpha_{non-rew}$ had the opposite relationship to proportion of correct responses and trials to criterion compared to α_{rew} and had a positive correlation with lose-shift generally. β was inversely related to trials to criterion but positively correlated with proportion of correct responses, win-stay correct and lose-shift incorrect. κ_{stim} had a significant negative relationship with lose-shift behaviour. In the second chapter involving analysis of human data, chapter 6, the learning rate α was positively related to win-stay and negatively correlated with lose-shift behaviour. The exploitation vs exploration parameter β was positively related to win-stay behaviour and had an inverse relationship with lose-shift. This was also found for κ_{stim} . These results are summarised in table 9.1.

	Proportion correct	Trials to criterion	Win-stay	Lose-shift	Perseverative errors
0/ _{rew}	Chapter 3: + Chapter 5: +	Chapter 5: –	Chapter 3: + (after correct response) Chapter 5: +	Chapter 5: + (after incorrect response)	
α _{non-rew}	Chapter 5: –	Chapter 5: +	Chapter 5: –	Chapter 5: +	
α		Chapter 4: –	Chapter 4: + Chapter 6: +	Chapter 4: + Chapter 6: -	Chapter 4: –
β	Chapter 5: +	Chapter 5: –	Chapter 5: + (after correct response) Chapter 6: +	Chapter 4: + Chapter 5: + (after incorrect response) Chapter 6: -	Chapter 4: +
K _{side}		Chapter 4: –	Chapter 3: + Chapter 4: +	Chapter 4: –	
K _{stim}			Chapter 6: +	Chapter 5: – Chapter 6: –	

8.2 Reinforcement Learning in Substance Use Disorder

Table 8.1 Summary of correlations between PRL and RL measures across all chapters. A table summarising the correlation results between conventional PRL measures and RL parameters. α_{rew} - the learning rate alpha_{rew} from rewarded trials; $\alpha_{non-rew}$ - the learning rate alpha_{non-rew} from non-rewarded trials; β - the exploitation/exploration parameters, otherwise known as the reinforcement sensitivity; κ_{side} , the side stickiness parameter; κ_{stim} , the stimulus stickiness parameter.

In summary, across most of the studies, rodent and human, α_{rew} was positively correlated with win-stay behaviour. This suggests that the higher the subject's learning from reward, the more likely they are to stay after a win. Moreover, it was positively associated with performance on the task, with higher values of this parameter leading to improved performance. β had a positive relationship with lose-shift behaviour, as the greater the exploration, the more likely they will shift their response. κ had a positive relationship with win-stay behaviour, which can be explained by higher stickiness leading to higher levels of staying, regardless of the outcome of the previous trial. These results provide explanations for the trends in RL parameters and how these relate to PRL measures. As the trends are observed across multiple studies, this increases the validity of the use of RL parameters for future studies.

8.2.5 Implications for the study of reinforcement learning and decisionmaking in disorders of compulsivity

Decision-making is characterised by two parallel systems: the goal-directed and habitual systems (Dolan and Dayan, 2013). These systems are thought to reflect model-based and model-free computational learning mechanisms (Daw et al., 2005). When choices based on habitual control are made, actions that have been previously rewarded are repeated. In goal-directed control, choices are made based on an internal model of the environment, which predicts outcomes. It has been shown that in disorders of compulsivity, including binge eating disorder, MUD and OCD, there is a bias towards the model-free, habit-based system (Voon et al., 2015). This bias towards the habit-based system has been associated with reduced grey matter volumes in the caudate, lateral PFC and mOFC. Both preclinical and clinical studies have shown that addiction to stimulants is associated with increased habitual behaviour (Everitt et al., 2001, 2005; Tiffany, 1990). Moreover, in a study involving a two-step sequential choice task, it was found that PGs had impaired model-based, i.e., goal-directed decision-making, particularly after unrewarded trials (Wyckmans et al., 2019). In rodents, lesions of the posterior DMS or PrL impair goal-directed learning (Balleine and Dickinson, 1998). Through human fMRI studies, it is known that reward value is encoded in the mOFC and mPFC, and the caudate computes action-outcome contingencies and is vital for goal-directed behaviour (Daw et al., 2006; Tanaka et al., 2008). In disorders of compulsivity, the structure and function of these areas is altered, resulting in impaired goaldirected function. Specifically, it has been shown that abstinent SDIs have reduced mOFC volumes, as do individuals with MUD (Franklin et al., 2002; Nakama et al., 2011). In OCD, reductions in OFC and caudate volumes are also present, providing a common biomarker of compulsive disorders that is likely associated with altered habit formation (Maia et al., 2008; Radua and Mataix-Cols, 2009).

The summarised literature clearly indicates that disorders of compulsivity, such as OCD, SUD and GD, have overlapping behavioural profiles as well as neural substrates. This leads to the question: how can these disorders be distinguished? A distinction of the behavioural and neurobiological mechanisms is necessary for targeted treatments. This thesis has contributed to differentiation and understanding of a subset of these disorders. It was shown that RL models can help to distinguish the behavioural profiles of individuals with GD and CUD. It is known that both of these disorders lead to increased habitual behaviour, but a conventional analysis of the PRL task cannot extract measures that reflect the behavioural mechanisms well enough to separate them. The RL models that were employed made it possible to

8.3 'Change Your Mind' - a novel task to investigate cognitive flexibility

quantify habitual behaviour in the form of the stickiness parameter κ , which could be further separated into side and stimulus stickiness. Through this differentiation, it was found that κ_{side} and κ_{stim} can be used to identify each of the two disorders, as individuals with GD show reduced κ_{stim} values compared to HCs, an effect not observed in CUD. Moreover, the RL model was able to uncover significant differences in parameters between GD and CUD, as participants with GD had higher reward learning rates than participants with CUD, and they also had greater reinforcement sensitivity values. Conventionally analysing the PRL data did not highlight such differences. In rats, the κ_{side} parameter was reported to be increased in high-compulsive animals, demonstrating that it is a marker of compulsivity across species. It was also shown that this parameter is positively correlated with connectivity between the mOFC and DS. In rodents and humans, the DS is known to be a component of the habit system (Graybiel, 2008). Lesioning the DS in rats has resulted in reduced perseverative responding and improved flexible behaviour (Yin and Knowlton, 2006). Therefore, it is plausible that the κ parameter is also modulated by the connectivity between the mOFC and DS in humans, with increased functional connectivity leading to increased κ values and thus habitual behaviour. This is further supported by the fact that the mOFC and its involvement in probability estimation and goal-formation is disrupted in disorders of compulsivity, and that the DS plays a key role in the development of compulsions (Goldstein and Volkow, 2011; Lipton et al., 2019). Across all studies, it was shown that the stickiness parameter κ improved model fit, suggesting that it reflects an important component of RL and can be used to investigate differences between psychiatric conditions. These findings provide a positive outlook for the study of RL in disorders of compulsivity and the use of RL models to distinguish these disorders more clearly.

8.3 'Change Your Mind' - a novel task to investigate cognitive flexibility

The final results chapter presents and validates a novel behavioural task for the study of cognitive flexibility. Existing tasks, such as the WCST or reversal learning task, investigate set-shifting and rule learning, respectively. However, in daily life, behavioural flexibility is frequently elicited via internal signals, and does not require the use of rule learning or set-shifting. The 'change your mind' (CYM) task is a two-alternative choice task, during which a stimulus is presented, followed by spurious feedback and the same stimulus being presented again. This gives subjects the opportunity to repeat their choice or change it, acting

Discussion

of their own volition. Existing paradigms do not give subjects the opportunity to repeat or change their choice following unreliable feedback, thus, ecological validity is increased in the presented task. The results suggest that the highlighted processes present a form of cognitive flexibility that have not yet been explored within existing frameworks (Dias et al., 1996a).

It was found that participants changed their response more frequently following negative, rather than positive feedback, and when their response was incorrect on the first stimulus presentation. Participants learnt to disregard negative feedback, as response shifting following negative feedback presentation decreased across the three runs. Increased switching after negative feedback is a common observation across tasks of flexibility; on reversal learning tasks, subjects will frequently shift after a loss and stay after a win. Whether a participant shifts behaviour or not following presentation of feedback is dependent on a variety of factors, e.g., neurochemical modulators such as noradrenaline and 5-HT (Evers et al., 2005; Swanson et al., 2021). Individuals diagnosed with disorders such as OCD or schizophrenia also show altered shifting behaviour on tasks of cognitive flexibility (Becker et al., 2014; Huddy et al., 2011). Therefore, the CYM task may be relevant for investigating decision-making under uncertainty in various psychopathologies. Greater activations in areas including the SFG, paracingulate gyrus, cingulate gyrus, AI, OFC, putamen, and frontal pole were associated with changing, compared to repeating a response. The AI, dACC IFJ, dlPFC and IFG are areas frequently found to be activated in both task-switching and set-shifting paradigms, therefore, there appears to be some overlap with the areas activated on the CYM task. However, the task presented here leads to greater activity in the frontal pole, as well as more caudal regions such as the middle temporal gyrus and angular gyrus. PPI analyses further implicated the connection between the AI and lateral occipital cortex as well as occipital pole, and between the OFC and the lingual gyrus/occipital pole in repeat trials. In summary, the fMRI data support the idea that a previously unexplored subtype of cognitive flexibility can be investigated using the CYM task, as it has overlapping, yet distinct neural circuits compared to those found in other paradigms.

The CYM task created a highly uncertain environment through the short presentation of stimuli and unreliable feedback. Unlike other tasks to measure cognitive flexibility, such as the reversal learning task, it does not require ongoing rule-based learning, and enables subjects to act of their own volition. Although there are clear differences between the reversal learning and CYM tasks, some parallels can be drawn. The κ parameter, as extracted from the PRL task, represents a subject repeating a response regardless of outcome. This is similar to the CYM task, as responding is independent of feedback.

8.3 'Change Your Mind' - a novel task to investigate cognitive flexibility

rather than value-based, behaviour (Miller et al., 2019). Therefore, the task presents the opportunity to separate value-free components of behaviour, particularly stickiness. In RL, the κ parameter may also represent a volitional component, based on internally generated signals. However, behaviour on the CYM task is not free operant, i.e., it is not modifiable by its consequences (Staddon and Cerutti, 2003). Nonetheless, the results presented here may provide an additional insight into which areas may contribute to value-free behaviours in RL. Additionally, it allows the investigation of volitional behaviour under uncertain circumstances. As was seen in chapter 3, the behavioural consequences of ELS in rats manifested themselves only under uncertainty. In humans with a psychiatric diagnosis, particularly OCD or anxiety disorders, patients have difficulty tolerating ambiguous situations (Tolin et al., 2003). In the case of OCD, this has been associated with checking compulsions (van den Hout and Kindt, 2004). Therefore, the CYM task also provides an opportunity to investigate volitional, value-free behaviour under uncertainty across mental health disorders, an important aspect of cognition that has the potential for further research.

8.3.1 Implications for the study of cognitive flexibility

In the first chapters of this thesis, the circuits underlying reversal learning and RL in rodents and humans were investigated. The circuits implicated in this type of cognitive flexibility are mostly those underlying reward processing, such as the OFC, amygdala and striatum. Alongside set-shifting and task-switching paradigms, reversal learning processes have been greatly focused on in the cognitive flexibility literature. The results from the CYM task show that these three processes do not explain the entire construct of mental flexibility; this space is far from being saturated. Rather than separating the construct of flexibility into categorisations, such as reversal learning, attentional set-shifting, task-switching and 'changing your mind', it may be more useful considering it as a continuum, with both converging and diverging aspects. Furthermore, the importance of greater ecological validity is highlighted. Understanding how circuits mediating flexibility develop in adolescence, stabilise in adulthood and break down in ageing is vital for identifying interventions for flexibility deficits in psychiatric and neurological disorders and to advance precision medicine in this field (Uddin, 2021).

8.4 Future research directions

Fitting computational models, such as RL models, to behavioural data is imperative for a better understanding of the latent processes underlying animal behaviours. In this thesis, RL models helped uncover observations that were not clear with the use of conventional analyses. This shows that the use of conventional analyses may be somewhat simplistic; only the use of higher-level mathematical models can truly uncover the processes that mediate behaviour (Daw, 2009). Further, parameters extracted from such models can aid the differentiation and diagnosis of neuropsychiatric disorders. They can also help to distinguish disorders that seemingly have similar profiles. For example, for distinguishing disorders of compulsivity, e.g., GD and OCD, as well as different types of SUDs. This could contribute to the advancement of precision medicine in neurology and psychiatry.

Despite the disadvantages of using translational animal models, such as the fact that they are not able to replicate the full extent of a disease state and that the functional homology and analogy do not always align, their use in neuroscientific research remains vital for the advancement of our understanding of the human brain. Dissecting a disorder into more tractable dimensions through the use of endophenotypes makes it possible to better integrate preclinical and clinical findings (Pizzagalli and Roberts, 2021). Another issue regarding the use of animals in translational research, particularly with neuroimaging, is the use of different atlases for defining brain regions. In rats, there are multiple atlases defining regions differently, without there being a standardised approach. Additionally, given that the boundaries in the rodent brain are not always clearly defined, it raises the question how reliable the atlases are and how confident one can be about the ROI used, especially after warping them into a study-specific template. Moreover, there is an unmet gap in our understanding how sex affects behaviour and cognition in preclinical models as well as clinically (Buoncervello et al., 2017; Riecher-Rössler, 2017). Studying both sexes preclinically is vital to reduce this gap and improve treatments for women. Future studies should aim to consider sex when planning their study design and contribute to a better understanding of sexual dimorphisms.

Finally, future directions for research on cognitive flexibility include maximising ecological and construct validity (Uddin, 2021). It is important to address the question of how laboratory-based measures translate to real-world behaviour. One method that can help achieve this is the development of novel behavioural tasks, as was presented in chapter 7. Additionally, it has been suggested that self-report of informant-report measures could better predict real-life inter-individual differences (Dang et al., 2020). The consistent use of questionnaires such as the Cognitive Control and Flexibility Questionnaire, the Cognitive Flexibility Scale and the Psychological Flexibility Questionnaires can help to achieve a better understanding of flexibility deficits (Ben-Itzchak et al., 2014; Gabrys et al., 2018; Martin and Rubin, 1995).

8.5 Conclusions

This thesis aimed to gain a better understanding of the behavioural and neural mechanisms involved in RL in rats and humans and translate the findings between the two species. Moreover, the role of RL in MDD and SUD was investigated. To extend our current knowledge of cognitive flexibility beyond the existing constructs, a novel task was presented in the final results chapter. Through the use of four datasets from rats and humans that have completed the reversal learning task, I was able to improve our understanding of how different states affect RL parameters and how this is implemented in the brain. Through a unique study exploring the 'two-hit' model of depression in rats, I found that both maternal separation and chronic stress in adulthood altered reward and punishment learning rates, as well as stickiness. An important finding was the sexually dimorphic effects of these interventions; females showed greater vulnerability to MS than males, the behavioural and neural consequences between them were distinct and showed opposing trends. The heightened vulnerability of females aligns with human studies, demonstrating the translational importance of rodent research. Additionally, I identified functional circuits associated with the RL parameter κ , which has not been shown previously and improves our understanding of how this parameters is implemented in the brain. Using the second rodent dataset it was possible to demonstrate the parameter κ is a behavioural marker which can identify high-compulsive animals prior to cocaine exposure. This may be of translational importance, as this measure could be used to identify individuals at risk of becoming stimulant-dependent. Slightly increased side stickiness was also found to be a consequence of CUD, as well as GD. The latter group was also found to have reduced stimulus stickiness, making it possible to distinguish the two compulsive disorders, which could be of clinical diagnostic relevance in the future. The two disorders could also be distinguished based on differences in functional connectivity underlying RL. The CUD group had an altered balance in processing positive and negative PPEs compared to control participants, whereas the GD group showed differences in tracking EVs of reward and punishments. Finally, I introduced a new method of measuring cognitive flexibility, which is of importance for extending our understanding of cognitive control and may lead to new developments in improving flexibility associated with a number of disorders

Discussion

as well as aging. In summary, I have been able to contribute to our understanding of RL and cognitive flexibility generally and how they are putatively altered in psychiatric conditions.

Appendix A

Appendix for rodent chapters (chapters 3-4)

A.1 Reinforcement Learning in rodents exposed to maternal separation and adulthood stress

A.1.1 Additional conventional PRL measures

The proportion of correct responses on the PRL task was significantly affected by sex and had a near-significant sex×stress interaction (F(1,46)=7.31, p=0.010 and F(1,46)=3.72, p=0.060, respectively). Females had a lower proportion of correct responses than males (t(44)=3.38, p=0.031). The trials required to reach criterion did not differ between the groups. Win-stay behaviour after both a correct and an incorrect response was significantly affected by sex, with males having higher values (t(89)=3.49, p<0.001 and t(89)=2.68, p<0.001, respectively). The results can be found in figure A1 and A2.



Appendix for rodent chapters (chapters 3-4)



Control (con) and maternally separated (MS) males and females prior to (pre) and after (post) adulthood stressors. None of the behavioural measures presented in this figure were affected as a result of MS or adulthood stress. A) Proportion of correct responses; B) Trials to criterion.



A.1 Reinforcement Learning in rodents exposed to maternal separation and adulthood stress



Control (con) and maternally separated (MS) males and females prior to (pre) and after (post) adulthood stressors. None of the behavioural measures presented in this figure were affected as a result of MS or adulthood stress. C) Win-stay behaviour after a correct response; D) Win-stay behaviour after an incorrect response.

Appendix for rodent chapters (chapters 3-4)



Fig. A.3 Figure reflecting the ability of the model to capture animal behaviour. The yellow line represents the animal's true behaviour. When this value is 1, it chooses the left stimulus, and when it is 2, it selects the right stimulus. The blue line demonstrates the probability of choosing the left stimulus and the orange line reflects the probability of choosing the right stimulus.



A.1.2 Reaction time data



Reaction times (RTs) of control (con) and maternally separated (MS) males (left) and females (right) on the PRL task prior to (pre) and after (post) adulthood stress. RTs increased in males but decreased in females following adulthood stress. Linear mixed-effects models were fitted to the data. * - p < 0.05.





Fig. A.5 Example of quality checks conducted on rodent imaging scans. Quality checks conducted on all the functional images to ensure successful pre-processing. A) absolute and relative estimated mean displacement (mm), B) estimated rotations (radians), C) estimated translations (mm), D) example registration of a functional scan to standard space, E) example registration of a structural scan to standard space, F) functional connectivity matrix of a subset of regions. These data were generated using FSL (Smith et al., 2004).



A.1.4 Additional model comparison figures

Fig. A.6 Additional model comparison figures.

A) Bayesian Information Criterion (BIC) comparison of the 4-parameter and 3-parameter 'forgetting' model, B) BIC comparison of the 4-parameter and 4-parameter 'differential forgetting' model, C) the log likelihood ratio test for the 4- and 3-parameter 'forgetting models, D) the log likelihood ratio test for 4- and 4-parameter 'differential forgetting' models.

Appendix for rodent chapters (chapters 3-4)

A.1.5 Separate t-tests - seed-based analysis



В	Seed-region	Number of voxels in cluster(s)	Volume of cluster(s) (mm ³)	Degrees of Freedom	Range of t- statistics
	BLA	15507	52336	16	2.24-3.97

Fig. A.7 BLA seed-based analysis results - females.

A) Areas to which the rs-FC from the BLA is greater in control females than MS females. These areas are: the cingulate cortex, infralimbic cortex, dorsal striatum and anterior insular cortex. The colour bar represents the t-statistics of the respective voxels. B) A table summarising the characteristics of the highlighted clusters.

A.1 Reinforcement Learning in rodents exposed to maternal separation and adulthood stress



В	Seed-region	Number of voxels in cluster(s)	Volume of cluster(s) (mm³)	Degrees of Freedom	Range of t- statistics
	BLA	1765	5957	26	2.82-4.64

Fig. A.8 BLA seed-based analysis results - males.

A) Areas to which the rs-FC from the BLA is greater in MS males than control males. These areas include the dorsal striatum and anterior insular cortex. The colour bar represents the t-statistics of the respective voxels. B) A table summarising the characteristics of the highlighted clusters.

A.2 RL in rodents exposed to a cocaine self-administration paradigm

A.2.1 Additional RL model comparison measures



Fig. A.9 Additional model comparison measures for RL models.

A) Log likelihood comparison of the 3-parameter vs 2-parameter model, red line indicated the threshold for model comparison which is 3.842; B) Log likelihood comparison of the 3-parameters vs 3parameter 'forgetting' model; C) Rho squared difference of the 3-parameter vs the 2-parameter model; D) Rho squared difference of the 3-parameter vs the 3-parameter 'forgetting' model.



A.2.2 Structural analyses

Fig. A.10 Volume differences across six brain regions.

Volumetric differences of six brain structures between low-, medium- and high-compulsive (LC, MC and HC, respectively) animals at two time points. Brain volumes for low-, mediumand high-compulsive animals, at PND63 and PND220-240. A) Nucleus accumbens shell; B) Caudate putamen; C) Amygdala; D) Infralimbic cortex; E) Prelimbic cortex; F) Anterior insular cortex. The brain volumes were solely affected by time, but showed no effects of compulsivity. Linear mixed-effects models were fitted to the data. * - p<0.05.

Appendix B

Additional figures from human chapters (chapters 5-7)

B.1 fMRI quality checks



Fig. B.1 Brain mask and brain tissue segmentation of T1w image. This panel shows the template T1-weighted image (if several T1w images were found), with contours delineating the detected brain mask and brain tissue segmentations.



Additional figures from human chapters (chapters 5-7)

Fig. B.2 Alignment of functional and anatomical MRI scans.



Fig. B.3 BOLD summary.

Summary statistics of fMRI scans. Global signals calculated within the whole-brain (GS), within the white-matter (WM) and within cerebro-spinal fluid (CSF) show the mean BOLD signal in their corresponding masks. DVARS and FD show the standardized DVARS and framewise-displacement measures for each time point.



Fig. B.4 Correlations among nuisance regressors.

Left: Heatmap summarizing the correlation structure among confound variables. Right: magnitude of the correlation between each confound time series and the mean global signal.

B.2 RL in patients with GD and SUD



Fig. B.5 The GD vs control contrast for feedback/cue presentation.

Areas of the contrast highlighting differences in response to feedback/cue presentation between participants with gambling disorder and healthy controls (MNI coordinates: X=31, Y=-68, Z=27). Activity was higher in the GD group in areas including the precuneus cortex, middle temporal gyrus and supramarginal gyrus. Colour bar on the right-hand side represents t-statistic.

Additional figures from human chapters (chapters 5-7)



Fig. B.6 The SUD vs control contrast for feedback/cue presentation.

Areas of the contrast highlighting differences in response to feedback/cue presentation between participants with SUD and healthy controls (MNI coordinates: X=39, Y=34, Z=17). Activity was higher in the SUD group in areas including the frontal pole, SFG and IFG. Colour bar on the right-hand side represents t-statistic.



Fig. B.7 The SUD vs control contrast - positive correlation between α_{rew} and response to positive PPE.

Areas of the contrast highlighting areas that have a stronger correlation with α_{rew} in the SUD group than in healthy controls (MNI coordinates: X=-5, Y=37, Z=22). This parameter was more strongly correlated with activity in areas including the frontal pole, SFG and cingulate and paracingulate gyri. Colour bar on the right-hand side represents t-statistic.



Fig. B.8 The GD vs control contrast - positive correlation between κ_{stim} and response to cue/feedback presentation.

Areas of the contrast highlighting areas that have a stronger positive correlation with κ_{stim} in the GD group than in healthy controls (MNI coordinates: X=48, Y=29, Z=22). This parameter was more strongly correlated with activity in areas including the right IFG, and right MFG. Colour bar on the right-hand side represents t-statistic.



Fig. B.9 The SUD vs control contrast - positive correlation between κ_{stim} and response to cue/feedback presentation.

Areas of the contrast highlighting areas that have a stronger positive correlation with κ_{stim} in the SUD group than in healthy controls (MNI coordinates: X=48, Y=29, Z=20). This parameter was more strongly correlated with activity in areas including the right IFG, and right MFG. Colour bar on the right-hand side represents t-statistic.



B.3 RL in patients with MDD



A) Proportion of correct responses; B) win-stay behaviour; C) lose-shift behaviour; D) number of perseverative responses were calculated for all simulated subjects from the five groups: healthy controls (HC), substance use disorder (SUD), major depressive disorder (MDD) only, SUD and MDD (SUD/MDD) and MDD with autism spectrum disorder (ASD). * - p<0.05.
Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 0:14.

Acosta, J. I., Thiel, K. J., Sanabria, F., Browning, J. R., and Neisewander, J. L. (2008). Effect of schedule of reinforcement on cue-elicited reinstatement of cocaine-seeking behavior. *Behavioural pharmacology*, 19(2):129–136.

Adamantidis, A. R., Tsai, H. C., Boutrel, B., Zhang, F., Stuber, G. D., Budygin, E. A., Touriño, C., Bonci, A., Deisseroth, K., and de Lecea, L. (2011). Optogenetic Interrogation of Dopaminergic Modulation of the Multiple Phases of Reward-Seeking Behavior. *Journal of Neuroscience*, 31(30):10829–10835.

Addicott, M. A., Pearson, J. M., Froeliger, B., Platt, M. L., and McClernon, F. J. (2014). Smoking automaticity and tolerance moderate brain activation during explore-exploit behavior. *Psychiatry research*, 224:254–261.

Adler, L. E., Pachtman, E., Franks, R. D., Pecevich, M., Waldo, M. C., and Freedman, R. (1982). Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia. *Biological Psychiatry*, 17(6).

Admon, R. and Pizzagalli, D. A. (2015). Dysfunctional reward processing in depression. *Current Opinion in Psychology*, 4:114–118.

Afyouni, S. and Nichols, T. E. (2018). Insight and inference for dvars. *Neuroimage*, 172:291.

Albert, P. R. (2015). Why is depression more prevalent in women? *Journal of Psychiatry Neuroscience : JPN*, 40:219.

Alexopoulos, G. S., Hoptman, M. J., Kanellopoulos, D., Murphy, C. F., Lim, K. O., and Gunning, F. M. (2012). Functional connectivity in the cognitive control network and the default mode network in late-life depression. *Journal of affective disorders*, 139(1):56–65.

Allen, C. P. and Leri, F. (2014). Perseveration in the presence of punishment: The effects of chronic cocaine exposure and lesions to the prefrontal cortex. *Behavioural Brain Research*, 261:185–192.

Altshuler, L. L., Abulseoud, O. A., Foland-Ross, L., Bartzokis, G., Chang, S., Mintz, J., Hellemann, G., and Vinters, H. V. (2010). Amygdala astrocyte reduction in subjects with major depressive disorder but not bipolar disorder. *Bipolar disorders*, 12(5):541–549.

Alves, R. L., Oliveira, P., Lopes, I. M., Portugal, C. C., Alves, C. J., Barbosa, F., Summavielle, T., and Magalhães, A. (2020). Early-life stress affects drug abuse susceptibility in adolescent rat model independently of depression vulnerability. *Scientific Reports 2020 10:1*, 10:1–11.

Ames, S. L., Wong, S. W., Bechara, A., Cappelli, C., Dust, M., Grenard, J. L., and Stacy, A. W. (2014). Neural Correlates of a Go/NoGo Task with Alcohol Stimuli in Light and Heavy Young Drinkers. *Behavioural brain research*, 0:382.

Anastasio, N. C., Stutz, S. J., Fox, R. G., Sears, R. M., Emeson, R. B., Dileone, R. J., O'Neil, R. T., Fink, L. H., Li, D., Green, T. A., Gerard Moeller, F., and Cunningham, K. A. (2014). Functional status of the serotonin 5-HT 2C receptor (5-HT 2C R) drives interlocked phenotypes that precipitate relapse-like behaviors in cocaine dependence. *Neuropsychopharmacology*, 39(2):360–372.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*, 65(4):550.

Andrews-Hanna, J. R., Smallwood, J., and Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, 1316(1):29–52.

Angst, J., Gamma, A., Gastpar, M., Lépine, J. P., Mendlewicz, J., and Tylee, A. (2002). Gender differences in depression. Epidemiological findings from the European DEPRES I and II studies. *European archives of psychiatry and clinical neuroscience*, 252(5):201–209.

APA (2013). American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Association.

Armbruster, D. J., Ueltzhöffer, K., Basten, U., and Fiebach, C. J. (2012). Prefrontal cortical mechanisms underlying individual differences in cognitive flexibility and stability. *Journal of Cognitive Neuroscience*, 24.

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature reviews. Neuroscience*, 10:410.

Aron, A. R. and Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: Role of the subthalamic nucleus. *Journal of Neuroscience*, 26:2424–2433.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, 38:95–113.

Augier, E., Barbier, E., Dulman, R. S., Licheri, V., Augier, G., Domi, E., Barchiesi, R., Farris, S., Nätt, D., Dayne Mayfield, R., Adermark, L., and Heilig, M. (2018). A molecular mechanism for choosing alcohol over an alternative reward. *Science*, 360(6395):1321–1326.

Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12:26–41.

Avery, J. A., Drevets, W. C., Moseman, S. E., Bodurka, J., Barcalow, J. C., and Simmons, W. K. (2014). Major depressive disorder is associated with abnormal interoceptive activity and functional connectivity in the insula. *Biological psychiatry*, 76(3):258–266.

Badoud, S., Borgognon, S., Cottet, J., Chatagny, P., Moret, V., Fregosi, M., Kaeser, M., Fortis, E., Schmidlin, E., Bloch, J., Brunet, J. F., and Rouiller, E. M. (2017). Effects of dorsolateral prefrontal cortex lesion on motor habit and performance assessed with manual grasping and control of force in macaque monkeys. *Brain Structure & Function*, 222:1193.

Bajic, D., Craig, M. M., Mongerson, C. R. L., Borsook, D., and Becerra, L. (2017). Identifying rodent resting-state brain networks with independent component analysis. *Front. Neurosci*, 11:685.

Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, 27:8161–8165.

Balleine, B. W. and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37:407–419.

Balleine, B. W. and O'Doherty, J. P. (2009). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 2010 35:1, 35:48–69.

Banich, M. T. (2009). Executive Function: The Search for an Integrated Account. *https://doi.org/10.1111/j.1467-8721.2009.01615.x*, 18(2):89–94.

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(6):1290–1301.

Barraclough, D. J., Conroy, M. L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, 7(4):404–410.

Baskin-Sommers, A. R. and Foti, D. (2015). Abnormal reward functioning across substance use disorders and major depressive disorder: Considering reward as a transdiagnostic mechanism. *International Journal of Psychophysiology*, 98:227–239.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67:1–48.

Bath, K. G., Nitenson, A. S., Lichtman, E., Lopez, C., Chen, W., Gallo, M., Goodwill, H., and Manzano-Nieves, G. (2017). Early life stress leads to developmental and sex selective effects on performance in a novel object placement task. *Neurobiology of Stress*, 7:57–67.

Becker, M. P., Nitsch, A. M., Schlösser, R., Koch, K., Schachtzabel, C., Wagner, G., Miltner, W. H., and Straube, T. (2014). Altered emotional and bold responses to negative, positive and ambiguous performance feedback in ocd. *Social Cognitive and Affective Neuroscience*, 9:1127.

Beckmann, C., Mackay, C., Filippini, N., and Smith, S. (2009). Group comparison of resting-state fmri data using multi-subject ica and dual regression. *NeuroImage*, 47.

Beery, A. K. and Zucker, I. (2011). Sex bias in neuroscience and biomedical research. *Neuroscience and biobehavioral reviews*, 35:565.

Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10.

Belova, M. A., Paton, J. J., and Salzman, C. D. (2008). Moment-to-Moment Tracking of State Value in the Amygdala. *Journal of Neuroscience*, 28(40):10023–10030.

Belsky, D. W., Moffitt, T. E., Baker, T. B., Biddle, A. K., Evans, J. P., Harrington, H. L., Houts, R., Meier, M., Sugden, K., Williams, B., Poulton, R., and Caspi, A. (2013). Polygenic risk and the developmental progression to heavy, persistent smoking and nicotine dependence: evidence from a 4-decade longitudinal study. *JAMA psychiatry*, 70(5):534–542.

Ben-Itzchak, E., Watson, L. R., and Zachor, D. A. (2014). Cognitive ability is associated with different outcome trajectories in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 44.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57.

Benzina, N., Mondragon, S., Ouarti, N., Mallet, L., and Burguiere, E. (2014). Translational approach to study flexibility as an endophenotype of obsessive compulsive disorders. *European Psychiatry*, 29(S3):546–546.

Bergeron, Y., Bureau, G., Laurier-Laurin, M. É., Asselin, E., Massicotte, G., and Cyr, M. (2017). Genetic deletion of akt3 induces an endophenotype reminiscent of psychiatric manifestations in mice. *Frontiers in Molecular Neuroscience*, 10.

Berton, O. and Nestler, E. J. (2006). New approaches to antidepressant drug discovery: beyond monoamines. *Nature reviews. Neuroscience*, 7(2):137–151.

Bickel, W. K., Odum, A. L., and Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146(4):447–454.

Binder, E. B., Owens, M. J., Liu, W., Deveau, T. C., Rush, A. J., Trivedi, M. H., Fava, M., Bradley, B., Ressler, K. J., and Nemeroff, C. B. (2010). Association of polymorphisms in genes regulating the corticotropin-releasing factor system with antidepressant treatment response. *Archives of general psychiatry*, 67(4):369–379.

Bissonette, G. B., Martins, G. J., Franz, T. M., Harper, E. S., Schoenbaum, G., and Powell, E. M. (2008). Double Dissociation of the Effects of Medial and Orbital Prefrontal Cortical Lesions on Attentional and Affective Shifts in Mice. *The Journal of Neuroscience*, 28(44):11124.

Blanchard, T. C. and Gershman, S. J. (2018). Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective and Behavioral Neuroscience*, 18:117–126.

Blanco, C., Vesga-López, O., Stewart, J. W., Liu, S. M., Grant, B. F., and Hasin, D. S. (2012). Epidemiology of major depression with atypical features: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *The Journal of clinical psychiatry*, 73(2):224–232.

Blazquez, P. M., Fujii, N., Kojima, J., and Graybiel, A. M. (2002). A network representation of response probability in the striatum. *Neuron*, 33:973–982.

Bogren, M., Brådvik, L., Holmstrand, C., Nöbbelin, L., and Mattisson, C. (2018). Gender differences in subtypes of depression by first incidence and age of onset: a follow-up of the Lundby population. *European Archives of Psychiatry and Clinical Neuroscience*, 268(2):179–189.

Bohn, I., Giertler, C., and Hauber, W. (2003). Orbital prefrontal cortex and guidance of instrumental behaviour in rats under reversal conditions. *Behavioural brain research*, 143(1):49–56.

Boileau, I., Dagher, A., Leyton, M., Gunn, R. N., Baker, G. B., Diksic, M., and Benkelfat, C. (2006). Modeling sensitization to stimulants in humans: an [11c]raclopride/positron emission tomography study in healthy men. *Archives of general psychiatry*, 63:1386–1395.

Bolla, K. I., Eldreth, D. A., London, E. D., Kiehl, K. A., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet, J. L., Kimes, A. S., Funderburk, F. R., and Ernst, M. (2003). Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *NeuroImage*, 19:1085–1094.

Bonini, F., Burle, B., Liégeois-Chauvel, C., Régis, J., Chauvel, P., and Vidal, F. (2014). Action monitoring and medial frontal cortex: Leading role of supplementary motor area. *Science*, 343:888–891.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., and Cohen, J. D. (1999). Conflict monitoring versus selection for-action in anterior cingulate cortex. *Nature*, 402:179–181.

Bowley, M. P., Drevets, W. C., Öngür, D., and Price, J. L. (2002). Low glial numbers in the amygdala in major depressive disorder. *Biological psychiatry*, 52(5):404–412.

Braff, D. L. and Freedman, R. (2002). Endophenotypes in studies of the genetics of schizophrenia. *Neuropsychopharmacology*, pages 703–717.

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J. I., Grimm, O., Heinz, A., Tost, H., Meyer-Lindenberg, A., and Bassett, D. S. (2015). Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences*, 112:11678–11683.

Braver, T. S. (2012). The variable nature of cognitive control: A dual-mechanisms framework. *Trends in Cognitive Sciences*, 16:106.

Braver, T. S., Paxton, J. L., Locke, H. S., and Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106:7351–7356.

Brigman, J. L., Feyder, M., Saksida, L. M., Bussey, T. J., Mishina, M., and Holmes, A. (2008). Impaired discrimination learning in mice lacking the NMDA receptor NR2A subunit. *Learning & memory (Cold Spring Harbor, N.Y.)*, 15(2):50.

Brigman, J. L., Graybeal, C., and Holmes, A. (2010). Predictably irrational: Assaying cognitive inflexibility in mouse models of schizophrenia. *Frontiers in Neuroscience*, 4.

Brivio, E., Lopez, J. P., and Chen, A. (2020). Sex differences: Transcriptional signatures of stress exposure in male and female brains. *Genes, brain, and behavior*, 19(3).

Brolsma, S. C., Vassena, E., Vrijsen, J. N., Sescousse, G., Collard, R. M., van Eijndhoven, P. F., Schene, A. H., and Cools, R. (2020a). Negative learning bias in depression revisited: Enhanced neural response to surprising reward across psychiatric disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.

Brolsma, S. C., Vrijsen, J. N., Vassena, E., Kandroodi, M. R., Bergman, M. A., Eijndhoven, P. F. V., Collard, R. M., Ouden, H. E. D., Schene, A. H., and Cools, R. (2020b). Challenging the negative learning bias hypothesis of depression: Reversal learning in a naturalistic psychiatric sample. *Psychological Medicine*, 52.

Bromberger, J. T. and Epperson, C. N. (2018). Depression During and After the Perimenopause: Impact of Hormones, Genetics, and Environmental Determinants of Disease. *Obstetrics and gynecology clinics of North America*, 45(4):663.

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., de Graaf, R., Demyttenaere, K., Hu, C., Iwata, N., Karam, A. N., Kaur, J., Kostyuchenko, S., Lépine, J. P., Levinson, D., Matschinger, H., Mora, M. E., Browne, M. O., Posada-Villa, J., Viana, M. C., Williams, D. R., and Kessler, R. C. (2011). Cross-national epidemiology of dsm-iv major depressive episode. *BMC Medicine*, 9:90.

Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations)? *Journal of Computational and Graphical Statistics*, 7.

Brown, V. J., Tait, D. S., Brown, V. J., and Tait, D. S. (2015). Attentional set-shifting across species. *Current Topics in Behavioral Neurosciences*, 28:363–395.

Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., and Chiu, P. H. (2021). Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. *JAMA Psychiatry*, 78:1113–1122.

Bryce, C. A. and Howland, J. G. (2015). Stress facilitates late reversal learning using a touchscreen-based visual discrimination procedure in male long evans rats. *Behavioural brain research*, 278:21.

Brydges, N. M., Holmes, M. C., Harris, A. P., Cardinal, R. N., and Hall, J. (2015). Early life stress produces compulsive-like, but not impulsive, behavior in females. *Behavioral neuroscience*, 129:300–308.

Buoncervello, M., Marconi, M., Carè, A., Piscopo, P., Malorni, W., and Matarrese, P. (2017). Preclinical models in the study of sex differences. *Clinical science (London, England : 1979)*, 131:449–469.

Burke, S. N., Mormino, E. C., Rogalski, E. J., Kawas, C. H., Willis, R. J., and Park, D. C. (2019). What are the later life contributions to reserve, resilience, and compensation? *Neurobiology of Aging*, 83:140–144.

Burrows, C. A., Timpano, K. R., and Uddin, L. Q. (2017). Putative Brain Networks Underlying Repetitive Negative Thinking and Comorbid Internalizing Problems in Autism. *Clinical psychological science : a journal of the Association for Psychological Science*, 5(3):522–536.

Bölükbas, I., Mundorf, A., and Freund, N. (2020). Maternal separation in rats induces neurobiological and behavioral changes on the maternal side. *Scientific Reports 2020 10:1*, 10:1–9.

Cai, W., Ryali, S., Chen, T., Li, C. S. R., and Menon, V. (2014). Dissociable roles of right inferior frontal cortex and anterior insula in inhibitory control: Evidence from intrinsic and task-related functional parcellation, connectivity, and response profile analyses across multiple datasets. *Journal of Neuroscience*, 34.

Cai, X., Kim, S., and Lee, D. (2011). Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron*, 69(1):170–182.

Calu, D. J., Stalnaker, T. A., Franz, T. M., Singh, T., Shaham, Y., and Schoenbaum, G. (2007). Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. *Learning & Memory*, 14:325–328.

Camchong, J., MacDonald, A. W., Nelson, B., Bell, C., Mueller, B. A., Specker, S., and Lim, K. O. (2011). Frontal hyperconnectivity related to discounting and reversal learning in cocaine subjects. *Biological psychiatry*, 69:1117.

Camerer, C. and Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67.

Cameron, A. C. and Windmeijer, F. A. (1997). An r-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, 77:329–342.

Campbell-Meiklejohn, D. K., Woolrich, M. W., Passingham, R. E., and Rogers, R. D. (2008). Knowing when to stop: The brain mechanisms of chasing losses. *Biological Psychiatry*, 63:293–300.

Carelli, R. M. and Wondolowski, J. (2003). Selective encoding of cocaine versus natural rewards by nucleus accumbens neurons is not related to chronic drug exposure. *Journal of Neuroscience*, 23.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76:1–32.

Carrier, N. and Kabbaj, M. (2012). Testosterone and imipramine have antidepressant effects in socially isolated male but not female rats. *Hormones and behavior*, 61(5):678–685.

Carrier, N. and Kabbaj, M. (2013). Sex differences in the antidepressant-like effects of ketamine. *Neuropharmacology*, 70:27–34.

Carroll, K. M. and Onken, L. S. (2005). Behavioral Therapies for Drug Abuse. *The American journal of psychiatry*, 162(8):1452.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280:747–749.

Caspi, A., Hariri, A. R., Andrew, H., Uher, R., and Moffitt, T. E. (2010). Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *The American journal of psychiatry*, 167(5):509–527.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H. L., McClay, J., Mill, J., Martin, J., Braithwaite, A., and Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science (New York, N.Y.)*, 301(5631):386–389.

Castañé, A., Theobald, D. E., and Robbins, T. W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behavioural brain research*, 210(1):74–83.

Ceceli, A. O., Parvaz, M. A., King, S., Schafer, M., Malaker, P., Sharma, A., Alia-Klein, N., and Goldstein, R. Z. (2022). Altered prefrontal signaling during inhibitory control in a salient drug context in cocaine use disorder. *Cerebral Cortex*.

Cella, M., Dymond, S., and Cooper, A. (2010). Impaired flexible decision-making in Major Depressive Disorder. *Journal of affective disorders*, 124(1-2):207–210.

Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., and Peters, J. (2020). Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *eLife*, 9:1–44.

Chamberlain, S. R., Stochl, J., Redden, S. A., and Grant, J. E. (2018). Latent traits of impulsivity and compulsivity: Toward dimensional psychiatry. *Psychological Medicine*, 48(5):810–821.

Chase, H. W., Eickhoff, S. B., Laird, A. R., and Hogarth, L. (2011). The neural basis of drug stimulus processing and craving: an activation likelihood estimation meta-analysis. *Biological psychiatry*, 70(8):785–793.

Chase, H. W., Frank, M. J., Michael, A., Bullmore, E. T., Sahakian, B. J., and Robbins, T. W. (2010). Approach and avoidance learning in patients with major depression and healthy controls: Relation to anhedonia. *Psychological Medicine*, 40:433–440.

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., and Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience and Biobehavioral Reviews*, 55:247–267.

Cheng, R. K., Etchegaray, M., and Meck, W. H. (2007). Impairments in timing, temporal memory, and reversal learning linked to neurotoxic regimens of methamphetamine intoxication. *Brain research*, 1186:255–266.

Cheng, W., Rolls, E. T., Qiu, J., Liu, W., Tang, Y., Huang, C. C., Wang, X. F., Zhang, J., Lin, W., Zheng, L., Pu, J. C., Tsai, S. J., Yang, A. C., Lin, C. P., Wang, F., Xie, P., and Feng, J. (2016). Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression. *Brain : a journal of neurology*, 139:3296–3309.

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., and Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. *Journal of cognitive neuroscience*, 25:1807–1823.

Chudasama, Y., Daniels, T. E., Gorrin, D. P., Rhodes, S. E., Rudebeck, P. H., and Murray, E. A. (2013). The role of the anterior cingulate cortex in choices based on reward value and reward contingency. *Cerebral cortex (New York, N.Y. : 1991)*, 23(12):2884–2898.

Clark, L. (2010). Decision-making during gambling: an integration of cognitive and psychobiological approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365:319.

Clark, L. and Goudriaan, A. E. (2018). The neuroscience and neuropsychology of gambling and gambling addiction: an introduction to the special issue. *International Gambling Studies*, 18.

Clark, S. L., Aberg, K. A., Nerella, S., Kumar, G., Mcclay, J. L., Chen, W., Xie, L. Y., Harada, A., Shabalin, A. A., Gao, G., Bergen, S. E., Hultman, C. M., Magnusson, P. K., Sullivan, P. F., and van den Oord, E. J. (2015). Combined Whole Methylome and Genomewide Association Study Implicates CNTN4 in Alcohol Use. *Alcoholism, clinical and experimental research*, 39(8):1396–1405.

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. *The Journal of Neuroscience*, 31(11):4290.

Clarke, H. F., Horst, N. K., and Roberts, A. C. (2015). Regional inactivations of primate ventral prefrontal cortex reveal two distinct mechanisms underlying negative bias in decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13):4176–4181.

Clarke, H. F., Robbins, T. W., and Roberts, A. C. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *Journal of Neuroscience*, 28(43):10972–10982.

Clarke, H. F., Walker, S. C., Crofts, H. S., Dalley, J. W., Robbins, T. W., and Roberts, A. C. (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(2):532–538.

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. *Cerebral cortex (New York, N.Y. : 1991)*, 17(1):18–27.

Clatworthy, P. L., Lewis, S. J., Brichard, L., Hong, Y. T., Izquierdo, D., Clark, L., Cools, R., Aigbirhio, F. I., Baron, J. C., Fryer, T. D., and Robbins, T. W. (2009). Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(15):4690–4696.

Claus, E. D., Kiehl, K. A., and Hutchison, K. E. (2011). Neural and Behavioral Mechanisms of Impulsive Choice in Alcohol Use Disorder. *Alcoholism, clinical and experimental research*, 35(7):1209.

Cohen, M. M., Jing, D., Yang, R. R., Tottenham, N., Lee, F. S., and Casey, B. J. (2013). Early-life stress has persistent effects on amygdala function and development in mice and humans. *Proceedings of the National Academy of Sciences of the United States of America*, 110:18274–18278.

Colasanti, A., Searle, G. E., Long, C. J., Hill, S. P., Reiley, R. R., Quelch, D., Erritzoe, D., Tziortzi, A. C., Reed, L. J., Lingford-Hughes, A. R., Waldman, A. D., Schruers, K. R., Matthews, P. M., Gunn, R. N., Nutt, D. J., and Rabiner, E. A. (2012). Endogenous opioid release in the human brain reward system induced by acute amphetamine administration. *Biological psychiatry*, 72(5):371–377.

Cole, M. W. and Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, 37:343–360.

Colich, N. L., Williams, E. S., Ho, T. C., King, L. S., Humphreys, K. L., Price, A. N., Ordaz, S. J., and Gotlib, I. H. (2017). The association between early life stress and prefrontal cortex activation during implicit emotion regulation is moderated by sex in early adolescence. *Development and Psychopathology*, 29:1851–1864.

Collins, D. L., Holmes, C. J., Peters, T. M., and Evans, A. C. (1995). Automatic 3-d model-based neuroanatomical segmentation. *Human Brain Mapping*, 3.

Colzato, L. S., van den Wildenberg, W. P., and Hommel, B. (2007). Impaired Inhibitory Control in Recreational Cocaine Users. *PLOS ONE*, 2(11):e1143.

Cools, R. (2015). Neuropsychopharmacology of Cognitive Flexibility. *Brain Mapping: An Encyclopedic Reference*, 3:349–353.

Cools, R., Clark, L., Owen, A. M., and Robbins, T. W. (2002). Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, 22(11):4563–4567.

Cools, R., Clark, L., and Robbins, T. W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience*.

Cope, E. C., Morris, D. R., Scrimgeour, A. G., and Levenson, C. W. (2012). Use of zinc as a treatment for traumatic brain injury in the rat: Effects on cognitive and behavioral outcomes. *Neurorehabilitation and Neural Repair*, 26:907–913.

Corlett, P. R., Honey, G. D., and Fletcher, P. C. (2007). From prediction error to psychosis: ketamine as a pharmacological model of delusions. *Journal of psychopharmacology (Oxford, England)*, 21(3):238–252.

Corriger, A. and Pickering, G. (2019). Ketamine and depression: a narrative review. *Drug Design, Development and Therapy*, 13:3051.

Costa, V. D., Tran, V. L., Turchi, J., and Averbeck, B. B. (2015). Reversal Learning and Dopamine: A Bayesian Perspective. *Journal of Neuroscience*, 35(6):2407–2416.

Couto, F. S. D., Batalha, V. L., Valadas, J. S., Data-Franca, J., Ribeiro, J. A., and Lopes, L. V. (2012). Escitalopram improves memory deficits induced by maternal separation in the rat. *European Journal of Pharmacology*, 695:71–75.

Cowen, P. J. (2016). Neuroendocrine and Neurochemical Processes in Depression. *https://doi.org/10.5127/pr.034513*, a3(1):3–15.

Cox, R. W. (1996). Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29.

Cui, M., Yang, Y., Yang, J., Zhang, J., Han, H., Ma, W., Li, H., Mao, R., Xu, L., Hao, W., and Cao, J. (2006). Enriched environment experience overcomes the memory deficits and depressive-like behavior induced by early life stress. *Neuroscience Letters*, 404:208–212.

Cui, Y., Cao, K., Lin, H., Cui, S., Shen, C., Wen, W., Mo, H., Dong, Z., Bai, S., Yang, L., Shi, Y., and Zhang, R. (2020). Early-life stress induces depression-like behavior and synaptic-plasticity changes in a maternal separation rat model: Gender difference and metabolomics study. *Frontiers in Pharmacology*, 11:1.

Culverhouse, R. C., Saccone, N. L., Horton, A. C., Ma, Y., Anstey, K. J., Banaschewski, T., Burmeister, M., Cohen-Woods, S., Etain, B., Fisher, H. L., Goldman, N., Guillaume, S., Horwood, J., Juhasz, G., Lester, K. J., Mandelli, L., Middeldorp, C. M., Olié, E., Villafuerte, S., Air, T. M., Araya, R., Bowes, L., Burns, R., Byrne, E. M., Coffey, C., Coventry, W. L., Gawronski, K. A., Glei, D., Hatzimanolis, A., Hottenga, J. J., Jaussent, I., Jawahar, C., Jennen-Steinmetz, C., Kramer, J. R., Lajnef, M., Little, K., Schwabedissen, H. M. Z., Nauck, M., Nederhof, E., Petschner, P., Peyrot, W. J., Schwahn, C., Sinnamon, G., Stacey, D., Tian, Y., Toben, C., Auwera, S. V. D., Wainwright, N., Wang, J. C., Willemsen, G., Anderson, I. M., Arolt, V., Aslund, C., Bagdy, G., Baune, B. T., Bellivier, F., Boomsma, D. I., Courtet, P., Dannlowski, U., Geus, E. J. D., Deakin, J. F., Easteal, S., Eley, T., Fergusson, D. M., Goate, A. M., Gonda, X., Grabe, H. J., Holzman, C., Johnson, E. O., Kennedy, M., Laucht, M., Martin, N. G., Munafò, M. R., Nilsson, K. W., Oldehinkel, A. J., Olsson, C. A., Ormel, J., Otte, C., Patton, G. C., Penninx, B. W., Ritchie, K., Sarchiapone, M., Scheid, J. M., Serretti, A., Smit, J. H., Stefanis, N. C., Surtees, P. G., Völzke, H., Weinstein, M., Whooley, M., Nurnberger, J. I., Breslau, N., and Bierut, L. J. (2018). Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-httlpr genotype contributing to the development of depression. *Molecular psychiatry*, 23:133–142.

Dajani, D. R., Odriozola, P., Winters, M., Voorhies, W., Marcano, S., Baez, A., Gates, K. M., Dick, A. S., and Uddin, L. Q. (2020). Measuring cognitive flexibility with the flexible item selection task: From fmri adaptation to individual connectome mapping. *Journal of Cognitive Neuroscience*, 32:1026–1045.

Dajani, D. R. and Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience.

Dakwar, E. and Kleber, H. D. (2022). Minding the brain: the role of pharmacotherapy in substance-use disorder treatment. *https://doi.org/10.31887/DCNS.2017.19.3/hkleber*, 19(3):289–297.

Dalla, C., Antoniou, K., Drossopoulou, G., Xagoraris, M., Kokras, N., Sfikakis, A., and Papadopoulou-Daifoti, Z. (2005). Chronic mild stress impact: are females more vulnerable? *Neuroscience*, 135:703–714.

Dalley, J. W., Everitt, B. J., and Robbins, T. W. (2011). Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron*, 69(4):680–694.

Dalley, J. W., Fryer, T. D., Brichard, L., Robinson, E. S., Theobald, D. E., Lääne, K., Peña, Y., Murphy, E. R., Shah, Y., Probst, K., Abakumova, I., Aigbirhio, F. I., Richards, H. K., Hong, Y., Baron, J. C., Everitt, B. J., and Robbins, T. W. (2007). Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science (New York, N.Y.)*, 315(5816):1267–1270.

Dalley, J. W. and Robbins, T. W. (2017). Fractionating impulsivity: Neuropsychiatric implications. *Nature Reviews Neuroscience*, 18(3):158–171.

Dalton, G. L., Ma, L. M., Phillips, A. G., and Floresco, S. B. (2011). Blockade of NMDA GluN2B receptors selectively impairs behavioral flexibility but not initial discrimination learning. *Psychopharmacology*, 216(4):525–535.

Dalton, G. L., Phillips, A. G., and Floresco, S. B. (2014). Preferential involvement by nucleus accumbens shell in mediating probabilistic learning and reversal shifts. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 34(13):4618–4626.

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. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 36:1996–2006.

Dang, J., King, K. M., and Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in cognitive sciences*, 24:267–269.

David, D. J., Dhonnchadha, B. A., Jolliet, P., Hascoët, M., and Bourin, M. (2001). Are there gender differences in the temperature profile of mice after acute antidepressant administration and exposure to two animal models of depression? *Behavioural brain research*, 119(2):203–211.

Daw, N. D. (2009). Trial-by-trial data analysis using computational models. *Attention & Performance XXIII*.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience* 2005 8:12, 8:1704–1711.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature 2006 441:7095*, 441:876–879.

Daws, R. E., Timmermann, C., Giribaldi, B., Sexton, J. D., Wall, M. B., Erritzoe, D., Roseman, L., Nutt, D., and Carhart-Harris, R. (2022). Increased global integration in the brain after psilocybin therapy for depression. *Nature Medicine* 2022 28:4, 28(4):844–851.

de Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addiction biology*, 14:22–31.

Deak, J. D. and Johnson, E. C. (2021). Genetics of substance use disorders: a review. *Psychological Medicine*, 51(13):2189.

Deak, J. D., Miller, A. P., and Gizer, I. R. (2019). Genetics of alcohol use disorder: a review. *Current Opinion in Psychology*, 27:56–61.

Deak, J. D., Zhou, H., Galimberti, M., Levey, D. F., Wendt, F. R., Sanchez-Roige, S., Hatoum, A. S., Johnson, E. C., Nunez, Y. Z., Demontis, D., Børglum, A. D., Rajagopal, V. M., Jennings, M. V., Kember, R. L., Justice, A. C., Edenberg, H. J., Agrawal, A., Polimanti, R., Kranzler, H. R., and Gelernter, J. (2022). Genome-wide association study in individuals of european and african ancestry and multi-trait analysis of opioid use disorder identifies 19 independent genome-wide significant risk loci. *Molecular psychiatry*.

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J., Pye, J. E., Hickie, I., and Guastella, A. J. (2018). Autism spectrum disorders: a meta-analysis of executive function. *Molecular psychiatry*, 23(5):1198–1204.

den Ouden, H. and O'Reilly, J. (2015). Models of learning.

den Ouden, H. E., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., Franke, B., and Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*, 80:1090–1100.

Der-Avakian, A. and Pizzagalli, D. A. (2018). Translational Assessments of Reward and Anhedonia: A Tribute to Athina Markou. *Biological psychiatry*, 83(11):932–939.

DeRubeis, R. J., Brotman, M. A., and Gibbons, C. J. (2005). A Conceptual and Methodological Analysis of the Nonspecifics Argument. *Clinical Psychology: Science and Practice*, 12(2):174–183.

Desender, K., Boldt, A., Verguts, T., and Donner, T. H. (2019). Confidence predicts speed-accuracy tradeoff for subsequent decisions. *eLife*, 8.

Di Ciano, P. and Everitt, B. J. (2001). Dissociable effects of antagonism of NMDA and AMPA/KA receptors in the nucleus accumbens core and shell on cocaine-seeking behavior. *Neuropsychopharmacology*, 25(3):341–360.

Diamond, A. and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old.

Dias, R., Robbins, T. W., and Roberts, A. C. (1996a). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, 380.

Dias, R., Robbins, T. W., and Roberts, A. C. (1996b). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral neuroscience*, 110(5):872–886.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M., and Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, 325:621–625.

Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., and Garbutt, J. C. (2010). Unipolar Depression Does Not Moderate Responses to the Sweet Taste Test. *Depression and anxiety*, 27(9):859.

Dick, D. M., Smith, G., Olausson, P., Mitchell, S. H., Leeman, R. F., O'Malley, S. S., and Sher, K. (2010). Understanding the construct of impulsivity and its relationship to alcohol use disorders. *Addiction biology*, 15(2):217.

Diefendorf, A. R. and Dodge, R. (1908). AN EXPERIMENTAL STUDY OF THE OC-ULAR REACTIONS OF THE INSANE FROM PHOTOGRAPHIC RECORDS. *Brain*, 31(3):451–489.

Diergaarde, L., Pattij, T., Poortvliet, I., Hogenboom, F., de Vries, W., Schoffelmeer, A. N., and De Vries, T. J. (2008). Impulsive Choice and Impulsive Action Predict Vulnerability to Distinct Stages of Nicotine Seeking in Rats. *Biological Psychiatry*, 63(3):301–308.

Dillon, D. G., Holmes, A. J., Birk, J. L., Brooks, N., Lyons-Ruth, K., and Pizzagalli, D. A. (2009). Childhood adversity is associated with left basal ganglia dysfunction during reward anticipation in adulthood. *Biological psychiatry*, 66:206–213.

Dolan, R. J. and Dayan, P. (2013). Goals and habits in the brain. Neuron, 80:312.

Dom, G., Sabbe, B., Hulstijn, W., and Brink, W. V. D. (2005). Substance use disorders and the orbitofrontal cortex: Systematic review of behavioural decision-making and neuroimaging studies. *The British Journal of Psychiatry*, 187:209–220.

Dombrovski, A. Y., Szanto, K., Clark, L., Reynolds, C. F., and Siegle, G. J. (2013). Reward signals, attempted suicide, and impulsivity in late-life depression. *JAMA Psychiatry*, 70:1020–1030.

Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., and Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, 50:799–812.

Draper, N. and Smith, H. (1981). Applied regression analysis. series in probability and mathematical statistics. *Wiley*.

Drevets, W. C. (2007). Orbitofrontal cortex function and structure in depression. *Annals of the New York Academy of Sciences*, 1121:499–527.

Dunn, E. C., Soare, T. W., Raffeld, M. R., Busso, D. S., Crawford, K. M., Davis, K. A., Fisher, V. A., Slopen, N., Smith, A. D., Tiemeier, H., and Susser, E. S. (2018). What life course theoretical models best explain the relationship between exposure to childhood adversity and psychopathology symptoms: recency, accumulation, or sensitive periods? *Psychological medicine*, 48:2550–2561.

Duque, A. and Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-tracking study. *Journal of Behavior Therapy and Experimental Psychiatry*, 46.

Edenberg, H. J., Dick, D. M., Xuei, X., Tian, H., Almasy, L., Bauer, L. O., Crowe, R. R., Goate, A., Hesselbrock, V., Jones, K., Kwon, J., Li, T. K., Nurnberger, J. I., O'Connor, S. J., Reich, T., Rice, J., Schuckit, M. A., Porjesz, B., Foroud, T., and Begleiter, H. (2004). Variations in GABRA2, Encoding the α 2 Subunit of the GABAA Receptor, Are Associated with Alcohol Dependence and with Brain Oscillations. *American Journal of Human Genetics*, 74(4):705.

Edenberg, H. J. and Foroud, T. (2006). The genetics of alcoholism: identifying specific genes through family studies. *Addiction biology*, 11(3-4):386–396.

Ehrich, E., Turncliff, R., Du, Y., Leigh-Pemberton, R., Fernandez, E., Jones, R., and Fava, M. (2015). Evaluation of opioid modulation in major depressive disorder. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 40(6):1448–1455.

Eid, R. S., Gobinath, A. R., and Galea, L. A. (2019). Sex differences in depression: Insights from clinical and preclinical studies. *Progress in Neurobiology*, 176:86–102.

Eklund, M. B. and Arborelius, L. (2006). Twice daily long maternal separations in wistar rats decreases anxiety-like behaviour in females but does not affect males. *Behavioural Brain Research*, 172:278–285.

Erickson, K., Drevets, W. C., Clark, L., Cannon, D. M., Bain, E. E., Zarate, C. A., Charney, D. S., and Sahakian, B. J. (2005). Mood-congruent bias in affective Go/No-Go performance of unmedicated patients with major depressive disorder. *American Journal of Psychiatry*, 162(11):2171–2173.

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Mller, U., Suckling, J., Ooi, C., Shabbir, S. S., Clark, L., Sahakian, B. J., Fineberg, N. A., Merlo-Pich, E. V., Robbins, T. W., and Bullmore, E. T. (2011). Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a d2/3 receptor agonist. *Biological Psychiatry*, 70:754–762.

Ersche, K. D., Roiser, J. P., Robbins, T. W., and Sahakian, B. J. (2008). Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. *Psychopharmacology*, 197.

Ersche, K. D., Turton, A. J., Pradhan, S., Bullmore, E. T., and Robbins, T. W. (2010). Drug addiction endophenotypes: Impulsive versus sensation-seeking personality traits. *Biological Psychiatry*, 68(8):770–773.

Eryilmaz, H., Rodriguez-Thompson, A., Tanner, A. S., Giegold, M., Huntington, F. C., and Roffman, J. L. (2017). Neural determinants of human goal-directed vs. habitual action control and their relation to trait motivation. *Scientific Reports*, 7:1–11.

Eshel, N. and Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, 68:118–124.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., and Gorgolewski, K. J. (2018). fmriprep: a robust preprocessing pipeline for functional mri. *Nature Methods 2018 16:1*, 16:111–116.

Evenden, J. L. (1999). Varieties of impulsivity. Psychopharmacology, 146(4):348-361.

Everitt, B. J., Dickinson, A., and Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain research. Brain research reviews*, 36:129–138.

Everitt, B. J. and Robbins, T. W. (2000). Second-order schedules of drug reinforcement in rats and monkeys: measurement of reinforcing efficacy and drug-seeking behaviour. *Psychopharmacology 2000 153:1*, 153:17–30.

Everitt, B. J. and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience 2005 8:11*, 8:1481–1489.

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.

Evers, E. A., Cools, R., Clark, L., Veen, F. M. V. D., Jolles, J., Sahakian, B. J., and Robbins, T. W. (2005). Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacology 2005 30:6*, 30:1138–1147.

Fales, C. L., Barch, D. M., Rundle, M. M., Mintun, M. A., Mathews, J., Snyder, A. Z., and Sheline, Y. I. (2009). Antidepressant treatment normalizes hypoactivity in dorsolateral prefrontal cortex during emotional interference processing in major depression. *Journal of Affective Disorders*, 112:206–211.

Fan, M., Liu, B., Jiang, T., Jiang, X., Zhao, H., and Zhang, J. (2010). Meta-analysis of the association between the monoamine oxidase-A gene and mood disorders. *Psychiatric genetics*, 20(1):1–7.

Fan, Y., Herrera-Melendez, A. L., Pestke, K., Feeser, M., Aust, S., Otte, C., Pruessner, J. C., Böker, H., Bajbouj, M., and Grimm, S. (2014). Early life stress modulates amygdalaprefrontal functional connectivity: Implications for oxytocin effects. *Human Brain Mapping*, 35:5328–5339.

Fattore, L. (2015). Reward processing and drug addiction: Does sex matter? *Frontiers in Neuroscience*, 9:329.

Fecteau, J. H. and Munoz, D. P. (2003). Exploring the consequences of the previous trial. *Nature Reviews Neuroscience*, 4.

Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A. M., Nigg, J. T., and Fair, D. A. (2019). The Heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends in cognitive sciences*, 23(7):584.

Fellows, L. K. and Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain*, 126(8):1830–1837.

First, M. B. (1997). Structured clinical interview for dsm-iv axis i disorders. *Biometrics Research Department*.

Fisher, P. M. and Hariri, A. R. (2013). Identifying serotonergic mechanisms underlying the corticolimbic response to threat in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1615).

Floresco, S. B., Tse, M. T., and Ghods-Sharifi, S. (2008). Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. *Neuropsychopharmacology*, 33(8):1966–1979.

Fontanesi, L., Gluth, S., Spektor, M. S., and Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin and Review*, 26.

Frank, M. J., Santamaria, A., O'Reilly, R. C., and Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 32(7):1583–1599.

Frankenhuis, W. E., Panchanathan, K., Nettle, D., Gangestad, S. W., and Tybur, J. M. (2016). Cognition in harsh and unpredictable environments. *Current Opinion in Psychology*.

Franklin, T. R., Acton, P. D., Maldjian, J. A., Gray, J. D., Croft, J. R., Dackis, C. A., O'Brien, C. P., and Childress, A. R. (2002). Decreased gray matter concentration in the insular, orbitofrontal, cingulate, and temporal cortices of cocaine patients. *Biological psychiatry*, 51:134–142.

Freedman, R., Coon, H., Myles-Worsley, M., Orr-Urtreger, A., Olincy, A., Davis, A., Polymeropoulos, M., Holik, J., Hopkins, J., Hoff, M., Rosenthal, J., Waldo, M. C., Reimherr, F., Wender, P., Yaw, J., Young, D. A., Breese, C. R., Adams, C., Patterson, D., Adler, L. E., Kruglyak, L., Leonard, S., and Byerley, W. (1997). Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proceedings of the National Academy of Sciences of the United States of America*, 94(2):587–592.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. Frith, C. D., and Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2.

Frodl, T., Janowitz, D., Schmaal, L., Tozzi, L., Dobrowolny, H., Stein, D. J., Veltman, D. J., Wittfeld, K., van Erp, T. G., Jahanshad, N., Block, A., Hegenscheid, K., Völzke, H., Lagopoulos, J., Hatton, S. N., Hickie, I. B., Frey, E. M., Carballedo, A., Brooks, S. J., Vuletic, D., Uhlmann, A., Veer, I. M., Walter, H., Schnell, K., Grotegerd, D., Arolt, V., Kugel, H., Schramm, E., Konrad, C., Zurowski, B., Baune, B. T., van der Wee, N. J., van Tol, M. J., Penninx, B. W., Thompson, P. M., Hibar, D. P., Dannlowski, U., and Grabe, H. J. (2017). Childhood adversity impacts on brain subcortical structures relevant to depression. *Journal of psychiatric research*, 86:58–65.

Frodl, T., Meisenzahl, E., Zetzsche, T., Bottlender, R., Born, C., Groll, C., Jäger, M., Leinsinger, G., Hahn, K., and Möller, H. J. (2002). Enlargement of the amygdala in patients with a first episode of major depression. *Biological psychiatry*, 51(9):708–714.

Fudenberg, D., Newey, W., Strack, P., and Strzalecki, T. (2020). Testing the drift-diffusion model. *Proceedings of the National Academy of Sciences of the United States of America*, 117:33141–33148.

Furman, D. J., Hamilton, J. P., and Gotlib, I. H. (2011). Frontostriatal functional connectivity in major depressive disorder. *Biology of Mood & Anxiety Disorders 2011 1:1*, 1(1):1–11.

Gabrys, R. L., Tabri, N., Anisman, H., and Matheson, K. (2018). Cognitive control and flexibility in the context of stress and depressive symptoms: The cognitive control and flexibility questionnaire. *Frontiers in Psychology*, 9.

Gallo, E. A. G., Munhoz, T. N., de Mola, C. L., and Murray, J. (2018). Gender differences in the effects of childhood maltreatment on adult depression and anxiety: A systematic review and meta-analysis. *Child abuse neglect*, 79:107–114.

García-Marchena, N., de Guevara-Miranda, D. L., Pedraz, M., Araos, P. F., Rubio, G., Ruiz, J. J., Pavón, F. J., Serrano, A., Castilla-Ortega, E., Santín, L. J., and de Fonseca, F. R. (2018). Higher impulsivity as a distinctive trait of severe cocaine addiction among individuals treated for cocaine or alcohol use disorders. *Frontiers in Psychiatry*, 9(FEB).

Garcia-Toro, M., Medina, E., Galan, J. L., Gonzalez, M. A., and Maurino, J. (2012). Treatment patterns in major depressive disorder after an inadequate response to first-line antidepressant treatment. *BMC Psychiatry*, 12(1):143.

Garrison, J., Erdeniz, B., and Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 37:1297–1310.

Garvert, L., Kirchner, K., Grabe, H. J., and der Auwera, S. V. (2022). Genome-wide gene-gene interaction of the 5-httlpr promoter polymorphism emphasizes the important role of neuroplasticity in depression. *Progress in neuro-psychopharmacology biological psychiatry*, 119.

Gelernter, J. and Polimanti, R. (2021). Genetics of substance use disorders in the era of big data. *Nature reviews. Genetics*, 22(11):712–729.

Gelernter, J., Sherva, R., Koesterer, R., Almasy, L., Zhao, H., Kranzler, H. R., and Farrer, L. (2014). Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Molecular psychiatry*, 19(6):717–723.

Gelman, A. (2013). Bayesian data analysis, volume 53. Chapman & Hall/CRC.

Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, 71:1–6.

Gershon, E. S. and Goldin, L. R. (1986). Clinical methods in psychiatric genetics. I. Robustness of genetic marker investigative strategies. *Acta psychiatrica Scandinavica*, 74(2):113–118.

Geschwind, D. H. and Flint, J. (2015). Genetics and genomics of psychiatric disease. *Science (New York, N.Y.)*, 349(6255):1489–1494.

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. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological psychiatry*, 75(8):631–638.

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. *American Journal of Psychiatry*, 168(7):718–726.

Gobinath, A. R., Richardson, R. J., Chow, C., Workman, J. L., Lieblich, S. E., Barr, A. M., and Galea, L. A. (2018). Voluntary running influences the efficacy of fluoxetine in a model of postpartum depression. *Neuropharmacology*, 128:106–118.

Goff, B., Gee, D. G., Telzer, E. H., Humphreys, K. L., Gabard-Durnam, L., Flannery, J., and Tottenham, N. (2013). Reduced nucleus accumbens reactivity and adolescent depression following early-life stress. *Neuroscience*, 249:129–138.

Goldberg, T. and Green, M. F. (2002). Neurocognitive functioning in patients with schizophrenia: an overview. *Neuropsychopharmacology: The Fifth Generation of Progress*.

Goldstein, B. L. and Klein, D. N. (2014). A Review of Selected Candidate Endophenotypes for Depression. *Clinical psychology review*, 34(5):417.

Goldstein, R. Z., Alia-Klein, N., Tomasi, D., Zhang, L., Cottone, L. A., Maloney, T., Telang, F., Caparelli, E. C., Chang, L., Ernst, T., Samaras, D., Squires, N. K., and Volkow, N. D. (2007). Is decreased prefrontal cortical sensitivity to monetary reward associated with impaired motivation and self-control in cocaine addiction? *American Journal of Psychiatry*, 164:43–51.

Goldstein, R. Z. and Volkow, N. D. (2011). Dysfunction of the prefrontal cortex in addiction: Neuroimaging findings and clinical implications. *Nature Reviews Neuroscience*, 12.

Gómez, M. L., Martínez-Mota, L., Estrada-Camarena, E., and Fernández-Guasti, A. (2014). Influence of the brain sexual differentiation process on despair and antidepressant-like effect of fluoxetine in the rat forced swim test. *Neuroscience*, 261:11–22.

Goodwill, H. L., Manzano-Nieves, G., Gallo, M., Lee, H. I., Oyerinde, E., Serre, T., and Bath, K. G. (2019). Early life stress leads to sex differences in development of depressive-like outcomes in a mouse model. *Neuropsychopharmacology*, 44:711–720.

Goodwill, H. L., Manzano-Nieves, G., LaChance, P., Teramoto, S., Lin, S., Lopez, C., Stevenson, R. J., Theyel, B. B., Moore, C. I., Connors, B. W., and Bath, K. G. (2018). Early life stress drives sex-selective impairment in reversal learning by affecting parvalbumin interneurons in orbitofrontal cortex of mice. *Cell reports*, 25:2299.

Gordon, A. L., Patterson, T. K., and Knowlton, B. J. (2020). Early-life stress is associated with a preponderance of habitual responding in a novel instrumental avoidance learning paradigm. *Neurobiology of Learning and Memory*, 175:107316.

Goswami, D. B., May, W. L., Stockmeier, C. A., and Austin, M. C. (2010). Transcriptional expression of serotonergic regulators in laser-captured microdissected dorsal raphe neurons of subjects with major depressive disorder: sex-specific differences. *Journal of neurochemistry*, 112(2):397–409.

Gotlib, I. H. and Joormann, J. (2010). Cognition and Depression: Current Status and Future Directions. *Annual review of clinical psychology*, 6:285.

Goto, M., Abe, O., Miyati, T., Inano, S., Hayashi, N., Aoki, S., Mori, H., Kabasawa, H., Ino, K., Yano, K., Iida, K., Mima, K., and Ohtomo, K. (2011). 3 tesla mri detects accelerated hippocampal volume reduction in postmenopausal women. *Journal of magnetic resonance imaging : JMRI*, 33:48–53.

Gottesman, I. I. and Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions.

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain : a journal of neurology*, 134:1751–1764.

Grant, B. F., Saha, T. D., June Ruan, W., Goldstein, R. B., Patricia Chou, S., Jung, J., Zhang, H., Smith, S. M., Pickering, R. P., Huang, B., and Hasin, D. S. (2016). Epidemiology of DSM-5 Drug Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions-III. *JAMA psychiatry*, 73(1):39–47.

Gray, A. L., Hyde, T. M., Deep-Soboslay, A., Kleinman, J. E., and Sodhi, M. S. (2015). Sex differences in glutamate receptor gene expression in major depression and suicide. *Molecular psychiatry*, 20(9):1057–1068.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31:359–387.

Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., Reiss, A. L., and Schatzberg, A. F. (2007). Resting-State Functional Connectivity in Major Depression: Abnormally Increased Contributions from Subgenual Cingulate Cortex and Thalamus. *Biological psychiatry*, 62(5):429.

Greve, D. N. and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48:63–72.

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. *Biological psychiatry*, 73(8):756–762.

Groman, S. M., Rich, K. M., Smith, N. J., Lee, D., and Taylor, J. R. (2017). Chronic exposure to methamphetamine disrupts reinforcement-based decision making in rats. *Neuropsychopharmacology* 2018 43:4, 43:770–780.

Gronau, Q. F., Erp, S. V., Heck, D. W., Cesario, J., Jonas, K. J., and Wagenmakers, E. J. (2017). A bayesian model-averaged meta-analysis of the power pose effect with informed and default priors: the case of felt power. *Comprehensive Results in Social Psychology*, 2.

Gronau, Q. F., Singmann, H., and Wagenmakers, E. J. (2020). Bridgesampling: An r package for estimating normalizing constants. *Journal of Statistical Software*, 92.

Guadagno, A., Belliveau, C., Mechawar, N., and Walker, C. D. (2021). Effects of early life stress on the developing basolateral amygdala-prefrontal cortex circuit: The emerging role of local inhibition and perineuronal nets. *Frontiers in Human Neuroscience*, 15:484.

Guillin, O., Abi-Dargham, A., and Laruelle, M. (2007). Neurobiology of dopamine in schizophrenia. *International review of neurobiology*, 78:1–39.

Haber, S. N. and Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35.

Habib, R. and Dixon, M. R. (2010). Neurobehavioral evidence for the "near-miss" effect in pathological gamblers. *Journal of the experimental analysis of behavior*, 93:313–328.

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., and Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic review and meta-analysis. *JAMA Psychiatry*, 77:1286–1295.

Hales, C. A., Robinson, E. S., and Houghton, C. J. (2016). Diffusion modelling reveals the decision making processes underlying negative judgement bias in rats. *PLOS ONE*, 11:e0152592.

Han, G., Klimes-Dougan, B., Jepsen, S., Ballard, K., Nelson, M., Houri, A., Kumra, S., and Cullen, K. (2012). Selective neurocognitive impairments in adolescents with major depressive disorder. *Journal of adolescence*, 35(1):11.

Harari, O., Wang, J. C., Bucholz, K., Edenberg, H. J., Heath, A., Martin, N. G., Pergadia, M. L., Montgomery, G., Schrage, A., Bierut, L. J., Madden, P. F., and Goate, A. M. (2012). Pathway Analysis of Smoking Quantity in Multiple GWAS Identifies Cholinergic and Sensory Pathways. *PLOS ONE*, 7(12):e50913.

Harmer, C. J., Duman, R. S., and Cowen, P. J. (2017). How do antidepressants work? new perspectives for refining future treatment approaches. *The Lancet Psychiatry*, 4.

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., Goodwin, G. M., and Cowen, P. J. (2009). Effect of acute antidepressant administration on negative affective bias in depressed patients. *American Journal of Psychiatry*, 166(10):1178–1184.

Harms, M. B. (2017). Stress and exploitative decision-making. *Journal of Neuroscience*, 37:10035–10037.

Harms, M. B., Bowen, K. E. S., Hanson, J. L., and Pollak, S. D. (2018). Instrumental learning and cognitive flexibility processes are impaired in children exposed to early life stress. *Developmental Science*, 21:e12596.

Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., Compton, W. M., Crowley, T., Ling, W., Petry, N. M., Schuckit, M., and Grant, B. F. (2013). DSM-5 criteria for substance use disorders: Recommendations and rationale. *American Journal of Psychiatry*, 170(8):834–851.

Hasselbalch, B. J., Knorr, U., Hasselbalch, S. G., Gade, A., and Kessing, L. V. (2012). Cognitive deficits in the remitted state of unipolar depressive disorder. *Neuropsychology*, 26(5):642–651.

Hayes, A., Herlinger, K., Paterson, L., and Lingford-Hughes SUMMARY, A. (2020). The neurobiology of substance use and addiction: evidence from neuroimaging and relevance to treatment. *BJPsych Advances*, 26(6):367–378.

Hayes, J. P., Hayes, S. M., and Mikedis, A. M. (2012). Quantitative meta-analysis of neural activity in posttraumatic stress disorder. *Biology of Mood & Anxiety Disorders 2012 2:1*, 2:1–13.

He, C., Gong, L., Yin, Y., Yuan, Y., Zhang, H., Lv, L., Zhang, X., Soares, J. C., Zhang, H., Xie, C., and Zhang, Z. (2019). Amygdala connectivity mediates the association between anxiety and depression in patients with major depressive disorder. *Brain imaging and behavior*, 13:1146–1159.

Hedegaard, H., Miniño, A. M., and Warner, M. (2020). Drug Overdose Deaths in the United States, 1999-2018. *NCHS data brief*, 356.

Heil, S. H., Johnson, M. W., Higgins, S. T., and Bickel, W. K. (2006). Delay discounting in currently using and currently abstinent cocaine-dependent outpatients and non-drug-using matched controls. *Addictive behaviors*, 31(7):1290–1294.

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., and Lesch, K. P. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of neurochemistry*, 66(6):2621–2624.

Hermans, E. J., Henckens, M. J., Joëls, M., and Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in neurosciences*, 37(6):304–314.

Hermans, E. J., Marle, H. J. V., Ossewaarde, L., Henckens, M. J., Qin, S., Kesteren, M. T. V., Schoots, V. C., Cousijn, H., Rijpkema, M., Oostenveld, R., and Fernández, G. (2011). Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science*, 334:1151–1153.

Herringa, R. J., Birn, R. M., Ruttle, P. L., Burghy, C. A., Stodola, D. E., Davidson, R. J., and Essex, M. J. (2013). Childhood maltreatment is associated with altered fear circuitry and increased internalizing symptoms by late adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, 110:19119–19124.

Herzallah, M. M., Moustafa, A. A., Natsheh, J. Y., Abdellatif, S. M., Taha, M. B., Tayem, Y. I., Sehwail, M. A., Amleh, I., Petrides, G., Myers, C. E., and Gluck, M. A. (2013). Learning from negative feedback in patients with major depressive disorder is attenuated by ssri antidepressants. *Frontiers in Integrative Neuroscience*, 7.

Higo, T., Mars, R. B., Boorman, E. D., Buch, E. R., and Rushworth, M. F. (2011). Distributed and causal influence of frontal operculum in task control. *Proceedings of the National Academy of Sciences of the United States of America*, 108:4230–4235.

Holmes, A., Le Guisquet, A. M., Vogel, E., Millstein, R. A., Leman, S., and Belzung, C. (2005). Early life genetic, epigenetic and environmental factors shaping emotionality in rodents. In *Neuroscience and Biobehavioral Reviews*, volume 29, pages 1335–1346. Neurosci Biobehav Rev.

Holst, R. J. V., Veltman, D. J., Bchel, C., Brink, W. V. D., and Goudriaan, A. E. (2012). Distorted expectancy coding in problem gambling: Is the addictive in the anticipation? *Biological Psychiatry*, 71:741–748.

Honeycutt, J. A., Demaestri, C., Peterzell, S., Silveri, M. M., Cai, X., Kulkarni, P., Cunningham, M. G., Ferris, C. F., and Brenhouse, H. C. (2020). Altered corticolimbic connectivity reveals sex-specific adolescent outcomes in a rat model of early life adversity. *eLife*, 9.

Horn, D. I., Yu, C., Steiner, J., Buchmann, J., Kaufmann, J., Osoba, A., Eckert, U., Zierhut, K. C., Schiltz, K., He, H., Biswa, B., Bogerts, B., and Walter, M. (2010). Glutamatergic and resting-state functional connectivity correlates of severity in major depression - the role of pregenual anterior cingulate cortex and anterior insula. *Frontiers in systems neuroscience*, 4.

Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., Hemani, G., Berger, K., Teismann, H., Rawal, R., Arolt, V., Baune, B. T., Dannlowski, U., Domschke, K., Tian, C., Hinds, D. A., Trzaskowski, M., Byrne, E. M., Ripke, S., Smith, D. J., Sullivan, P. F., Wray, N. R., Breen, G., Lewis, C. M., and McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*, 22:343–352.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. - supporting material. *Science (New York, N.Y.)*, 310.

Hu, S., Ide, J. S., Zhang, S., Sinha, R., and Li, C. S. R. (2015). Conflict anticipation in alcohol dependence — A model-based fMRI study of stop signal task. *NeuroImage: Clinical*, 8:39–50.

Huang, Y., Coupland, N. J., Lebel, R. M., Carter, R., Seres, P., Wilman, A. H., and Malykhin, N. V. (2013). Structural Changes in Hippocampal Subfields in Major Depressive Disorder: A High-Field Magnetic Resonance Imaging Study. *Biological Psychiatry*, 74(1):62–68.

Huddy, V. C., Hodgson, T. L., Ron, M. A., Barnes, T. R., and Joyce, E. M. (2011). Abnormal negative feedback processing in first episode schizophrenia: evidence from an oculomotor rule switching task. *Psychological medicine*, 41:1805.

Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., and Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49.

Hultman, R., Mague, S. D., Li, Q., Katz, B. M., Michel, N., Lin, L., Wang, J., David, L. K., Blount, C., Chandy, R., Carlson, D., Ulrich, K., Carin, L., Dunson, D., Kumar, S., Deisseroth, K., Moore, S. D., and Dzirasa, K. (2016). Dysregulation of prefrontal cortex-mediated slow-evolving limbic dynamics drives stress-induced emotional pathology. *Neuron*, 91:439–452.

Humphreys, K. L., Lee, S. S., Telzer, E. H., Gabard-Durnam, L. J., Goff, B., Flannery, J., and Tottenham, N. (2015). Exploration-exploitation strategy is dependent on early experience. *Developmental psychobiology*, 57:313–321.

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., and Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of mood & anxiety disorders*, 3.

Huys, Q. J., Tobler, P. N., Hasler, G., and Flagel, S. B. (2014). The role of learning-related dopamine signals in addiction vulnerability. *Progress in brain research*, 211:31–77.

Hyer, M. M., Shaw, G. A., Goswamee, P., Dyer, S. K., Burns, C. M., Soriano, E., Sanchez, C. S., Rowson, S. A., McQuiston, A. R., and Neigh, G. N. (2021). Chronic adolescent stress causes sustained impairment of cognitive flexibility and hippocampal synaptic strength in female rats. *Neurobiology of Stress*, 14:100303.

Hyman, S. (2014). Mental health: Depression needs large human-genetics studies. *Nature* 2014 515:7526, 515(7526):189–191.

Hyman, S. E. and Nestler, E. J. (1996). Initiation and adaptation: a paradigm for understanding psychotropic drug action. *The American journal of psychiatry*, 153(2):151–162.

Hägele, C., Schlagenhauf, F., Rapp, M., Sterzer, P., Beck, A., Bermpohl, F., Stoy, M., Ströhle, A., Wittchen, H. U., Dolan, R. J., and Heinz, A. (2015). Dimensional psychiatry: Reward dysfunction and depressive mood across psychiatric disorders. *Psychopharmacology*, 232.

Iacono, W. G. (2018). Endophenotypes in psychiatric disease: Prospects and challenges.

Iacono, W. G., Carlson, S. R., Malone, S. M., and McGue, M. (2002). P3 event-related potential amplitude and the risk for disinhibitory disorders in adolescent boys. *Archives of General Psychiatry*, 59.

Ikram, U. Z., Snijder, M. B., Fassaert, T. J., Schene, A. H., Kunst, A. E., and Stronks, K. (2015). The contribution of perceived ethnic discrimination to the prevalence of depression. *European Journal of Public Health*, 25.

Ito, M. and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *Journal of Neuroscience*, 29(31):9861–9874.

Ito, R., Dalley, J. W., Robbins, T. W., and Everitt, B. J. (2002). Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. *Journal of Neuroscience*, 22(14):6247–6253.

Iversen, S. D. and Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, 11.

Izquierdo, A. (2017). Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *Journal of Neuroscience*, 37(44):10529–10540.

Izquierdo, A., Belcher, A. M., Scott, L., Cazares, V. A., Chen, J., O'Dell, S. J., Malvaez, M., Wu, T., and Marshall, J. F. (2010). Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: Possible involvement of striatal dopamine. *Neuropsychopharmacology*, 35.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., and Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, 345:12–26.

Izquierdo, A., Carlos, K., Ostrander, S., Rodriguez, D., McCall-Craddolph, A., Yagnik, G., and Zhou, F. (2012). Impaired reward learning and intact motivation after serotonin depletion in rats. *Behavioural brain research*, 233(2):494–499.

Izquierdo, A., Darling, C., Manos, N., Pozos, H., Kim, C., Ostrander, S., Cazares, V., Stepp, H., and Rudebeck, P. H. (2013). Basolateral Amygdala Lesions Facilitate Reward Choices after Negative Feedback in Rats. *Journal of Neuroscience*, 33(9):4105–4109.

Jara-Rizzo, M. F., Navas, J. F., Rodas, J. A., and Perales, J. C. (2020). Decision-making inflexibility in a reversal learning task is associated with severity of problem gambling symptoms but not with a diagnosis of substance use disorder. *BMC psychology*, 8:120.

Jaworska, N., Macmaster, F. P., Gaxiola, I., Cortese, F., Goodyear, B., and Ramasubbu, R. (2014). A preliminary study of the influence of age of onset and childhood trauma on cortical thickness in major depressive disorder. *BioMed Research International*, 2014.

Jenkinson, M. (2003). Fast, automated, n-dimensional phase-unwrapping algorithm. *Magnetic Resonance in Medicine*, 49:193–197.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17:825–841.

Jenkinson, M. and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5.

Jentsch, J. D., Olausson, P., Garza, R. D. L., and Taylor, J. R. (2002). Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. *Neuropsychopharmacology*, 26:183–190.

Jentsch, J. D. and Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of behavior by reward-related stimuli. *Psychopharmacology*, 146(4):373–390.

Jentsch, J. D. and Taylor, J. R. (2001). Impaired Inhibition of Conditioned Responses Produced by Subchronic Administration of Phencyclidine to Rats. *Neuropsychopharmacology* 2000 24:1, 24(1):66–74.

Jog, M. S., Kubota, Y., Connolly, C. I., and Graybiel, A. M. (1999). Building neural representations of habits. *Science*, 286:1745–1749.

Johnson, E. C., Demontis, D., Thorgeirsson, T. E., Walters, R. K., Polimanti, R., Hatoum, A. S., Sanchez-Roige, S., Paul, S. E., Wendt, F. R., Clarke, T. K., Lai, D., Reginsson, G. W., Zhou, H., He, J., Baranger, D. A., Gudbjartsson, D. F., Wedow, R., Adkins, D. E., Adkins, A. E., Alexander, J., Bacanu, S. A., Bigdeli, T. B., Boden, J., Brown, S. A., Bucholz, K. K., Bybjerg-Grauholm, J., Corley, R. P., Degenhardt, L., Dick, D. M., Domingue, B. W., Fox, L., Goate, A. M., Gordon, S. D., Hack, L. M., Hancock, D. B., Hartz, S. M., Hickie, I. B., Hougaard, D. M., Krauter, K., Lind, P. A., McClintick, J. N., McQueen, M. B., Meyers, J. L., Montgomery, G. W., Mors, O., Mortensen, P. B., Nordentoft, M., Pearson, J. F., Peterson, R. E., Reynolds, M. D., Rice, J. P., Runarsdottir, V., Saccone, N. L., Sherva, R., Silberg, J. L., Tarter, R. E., Tyrfingsson, T., Wall, T. L., Webb, B. T., Werge, T., Wetherill, L., Wright, M. J., Zellers, S., Adams, M. J., Bierut, L. J., Boardman, J. D., Copeland, W. E., Farrer, L. A., Foroud, T. M., Gillespie, N. A., Grucza, R. A., Harris, K. M., Heath, A. C., Hesselbrock, V., Hewitt, J. K., Hopfer, C. J., Horwood, J., Iacono, W. G., Johnson, E. O., Kendler, K. S., Kennedy, M. A., Kranzler, H. R., Madden, P. A., Maes, H. H., Maher, B. S., Martin, N. G., McGue, M., McIntosh, A. M., Medland, S. E., Nelson, E. C., Porjesz, B., Riley, B. P., Stallings, M. C., Vanyukov, M. M., Vrieze, S., Walters, R., Johnson, E., McClintick, J., Hatoum, A., Wendt, F., Adams, M., Adkins, A., Aliev, F., Batzler, A., Bertelsen, S., Biernacka, J., Bigdeli, T., Chen, L. S., Chou, Y. L., Degenhardt, F., Docherty, A., Edwards, A., Fontanillas, P., Foo, J., Frank, J., Giegling, I., Gordon, S., Hack, L., Hartmann, A., Heilmann-Heimbach, S., Herms, S., Hodgkinson, C., Hoffman, P., Hottenga, J., Kennedy, M., Alanne-Kinnunen, M., Konte, B., Lahti, J., Lahti-Pulkkinen, M., Ligthart, L., Loukola, A., Maher, B., Mbarek, H., McQueen, M., Meyers, J., Milaneschi, Y., Palviainen, T., Pearson, J., Peterson, R., Ripatti, S., Ryu, E., Saccone, N., Salvatore, J., Schwandt, M., Streit, F., Strohmaier, J., Thomas, N., Wang, J. C., Webb, B., Wills, A., Boardman, J., Chen, D., Choi, D. S., Copeland, W., Culverhouse, R., Dahmen, N., Domingue, B., Elson, S., Frye, M., Gäbel, W., Hayward, C., Ising, M., Keyes, M., Kiefer, F., Kramer, J., Kuperman, S., Lucae, S., Lynskey, M., Maier, W., Mann, K., Männistö, S., Müller-Myhsok, B., Murray, A., Nurnberger, J., Palotie, A., Preuss, U., Räikkönen, K., Reynolds, M., Ridinger, M., Scherbaum, N., Schuckit, M., Soyka, M., Treutlein, J., Witt, S., Wodarz, N., Zill, P., Adkins, D., Boomsma, D., Brown, S., Cichon, S., Costello, E. J., de Wit, H., Diazgranados, N., Dick, D., Eriksson, J., Farrer, L., Foroud, T., Gillespie, N., Goate, A., Goldman, D., Grucza, R., Hancock, D., Hewitt, J., Hopfer, C., Iacono, W., Johnson, E., Kaprio, J., Karpyak, V., Kranzler, H., Lichtenstein, P., Lind, P., McGue, M., MacKillop, J., Maes, H., Magnusson, P., Martin, N., Montgomery, G., Nelson, E., Nöthen, M., Palmer, A., Pederson, N., Penninx, B., Rice, J., Rietschel, M., Riley, B., Rose, R., Rujescu, D., Shen, P. H., Silberg, J., Tarter, R., Vanyukov, M., Wall, T., Whitfield, J., Zhao, H., Neale, B., Gelernter, J., Edenberg, H., Agrawal, A., Davis, L. K., Bogdan, R., Edenberg, H. J., Stefansson, K., and Børglum, A. D. (2020). A large-scale genome-wide association study meta-analysis of cannabis use disorder. The lancet. Psychiatry, 7:1032–1045.

Johnson, F. K., Delpech, J. C., Thompson, G. J., Wei, L., Hao, J., Herman, P., Hyder, F., and Kaffman, A. (2018). Amygdala hyper-connectivity in a mouse model of unpredictable early life stress. *Translational Psychiatry 2018* 8:1, 8:1–14.

Jokela, M., Räikkönen, K., Lehtimäki, T., Rontu, R., and Keltikangas-Järvinen, L. (2006). Tryptophan hydroxylase 1 gene (TPH1) moderates the influence of social support on depressive symptoms in adults. *Journal of affective disorders*.

Jones, B. and Mishkin, M. (1972). Limbic lesions and the problem of stimulus-reinforcement associations. *Experimental Neurology*, 36.

Jones, J. A., Zuhlsdorff, K., and Dalley, J. W. (2021). Neurochemical substrates linked to impulsive and compulsive phenotypes in addiction: a preclinical perspective. *Journal of Neurochemistry*, page jnc.15380.

Joormann, J. and Gotlib, I. H. (2008). Updating the contents of working memory in depression: interference from irrelevant negative material. *Journal of abnormal psychology*, 117(1):182–192.

Joormann, J. and Gotlib, I. H. (2010). Emotion regulation in depression: relation to cognitive inhibition. *Cognition & emotion*, 24(2):281–298.

Joormann, J. and Siemer, M. (2011). Affective Processing and Emotion Regulation in Dysphoria and Depression: Cognitive Biases and Deficits in Cognitive Control. *Social and Personality Psychology Compass*, 5(1):13–28.

Joormann, J., Siemer, M., and Gotlib, I. H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *Journal of abnormal psychology*, 116(3):484–490.

Joëls, M., Karst, H., Krugers, H. J., and Lucassen, P. J. (2007). Chronic stress: implications for neuronal morphology, function and neurogenesis. *Frontiers in neuroendocrinology*, 28:72–96.

Jupp, B. and Dalley, J. W. (2014). Behavioral endophenotypes of drug addiction: Etiological insights from neuroimaging studies. *Neuropharmacology*, 76 Pt B(PART B):487–497.

Kaffman, A. and Meaney, M. J. (2007). Neurodevelopmental sequelae of postnatal maternal care in rodents: clinical and research implications of molecular insights. *Journal of child psychology and psychiatry, and allied disciplines*, 48(3-4):224–244.

Kalinichev, M., Easterling, K. W., Plotsky, P. M., and Holtzman, S. G. (2002). Long-lasting changes in stress-induced corticosterone response and anxiety-like behaviors as a consequence of neonatal maternal separation in long-evans rats. *Pharmacology, biochemistry, and behavior*, 73:131–140.

Kalivas, P. W. and Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal of Psychiatry*, 162:1403–1413.

Kampman, K. M. (2019). The treatment of cocaine use disorder. Science advances, 5(10).

Kandilarova, S., Stoyanov, D., Kostianev, S., and Specht, K. (2018). Altered resting state effective connectivity of anterior insula in depression. *Frontiers in Psychiatry*, 9.

Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W., and Cardinal, R. N. (2019). Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: remediating effects of dopaminergic d2/3 receptor agents. *Psychopharmacology*, 236:2337–2358.

Kang, H. J., Park, Y., Yoo, K. H., Kim, K. T., Kim, E. S., Kim, J. W., Kim, S. W., Shin, I. S., Yoon, J. S., Kim, J. H., and Kim, J. M. (2020). Sex differences in the genetic architecture of depression. *Scientific Reports 2020 10:1*, 10(1):1–12.

Karrouri, R., Hammani, Z., Otheman, Y., and Benjelloun, R. (2021). Major depressive disorder: Validated treatments and future challenges. *World Journal of Clinical Cases*, 9:9350.

Katz, R. J. (1982). Animal model of depression: pharmacological sensitivity of a hedonic deficit. *Pharmacology, biochemistry, and behavior*, 16(6):965–968.

Kaufman, A. S. and Kaufman, N. L. (2014). *Kaufman Brief Intelligence Test, Second Edition*. Pearson.

Kaufman, J. N., Ross, T. J., Stein, E. A., and Garavan, H. (2003). Cingulate Hypoactivity in Cocaine Users During a GO-NOGO Task as Revealed by Event-Related Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, 23(21):7839–7843.

Kawai, T., Yamada, H., Sato, N., Takada, M., and Matsumoto, M. (2015). Roles of the Lateral Habenula and Anterior Cingulate Cortex in Negative Outcome Monitoring and Behavioral Adjustment in Nonhuman Primates. *Neuron*, 88(4):792–804.

Kelleher, R. T. (1958). Fixed-ratio schedules of conditioned reinforcement with chimpanzees. *Journal of the Experimental Analysis of Behavior*, 1(3):281.

Kellner, C. H., Greenberg, R. M., Murrough, J. W., Bryson, E. O., Briggs, M. C., and Pasculli, R. M. (2012). ECT in Treatment-Resistant Depression. *https://doi.org/10.1176/appi.ajp.2012.12050648*, 169(12):1238–1244.

Kendler, K. S. and Gardner, C. O. (2014). Sex differences in the pathways to major depression: A study of opposite-sex twin pairs. *American Journal of Psychiatry*, 171.

Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual review of psychology*, 48:191–214.

Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders*, 74.

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alhamzawi, A. O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., Girolamo, G. D., Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., Haro, J. M., Hu, C. Y., Karam, E. G., Kawakami, N., Lee, S., Lépine, J. P., Ormel, J., Posada-Villa, J., Sagar, R., Tsang, A., Üstün, T. B., Vassilev, S., Viana, M. C., and Williams, D. R. (2010). Childhood adversities and adult psychopathology in the who world mental health surveys. *The British journal of psychiatry : the journal of mental science*, 197:378–385.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., and Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin*, 136:849–874.

Killcross, S. and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, 13:400–408.

Kim, H., Sul, J. H., Huh, N., Lee, D., and Jung, M. W. (2009). Role of striatum in updating values of chosen actions. *Journal of Neuroscience*, 29:14701–14712.

Kim, S., Cai, X., Hwang, J., and Lee, D. (2012). Prefrontal and Striatal Activity Related to Values of Objects and Locations. *Frontiers in Neuroscience*, 6(JULY):1–13.

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

Kim, Y. B., Huh, N., Lee, H., Baeg, E. H., Lee, D., and Jung, M. W. (2007). Encoding of action history in the rat ventral striatum. *Journal of neurophysiology*, 98:3548–3556.

Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J. C., Pariante, C. M., Pace, T. W., Mercer, K. B., Mayberg, H. S., Bradley, B., Nemeroff, C. B., Holsboer, F., Heim, C. M., Ressler, K. J., Rein, T., and Binder, E. B. (2012). Allele-specific FKBP5 DNA demethylation mediates gene–childhood trauma interactions. *Nature Neuroscience 2012 16*:*1*, 16(1):33–41.

Koesten, J., Schrodt, P., and Ford, D. J. (2009). Cognitive flexibility as a mediator of family communication environments and young adults' well-being. *Health Communication*, 24(1).

Kojima, S., Stewart, R. A., Demas, G. E., and Alberts, J. R. (2012). Maternal Contact Differentially Modulates Central and Peripheral Oxytocin in Rat Pups During a Brief Regime of Mother-Pup Interaction that Induces a Filial Huddling Preference. *Journal of Neuroendocrinology*, 24(5):831–840.

Kokras, N., Antoniou, K., Mikail, H. G., Kafetzopoulos, V., Papadopoulou-Daifoti, Z., and Dalla, C. (2015). Forced swim test: What about females? *Neuropharmacology*, 99:408–421.

Kong, L., Chen, K., Tang, Y., Wu, F., Driesen, N., Womer, F., Fan, G., Ren, L., Jiang, W., Cao, Y., Blumberg, H. P., Xu, K., and Wang, F. (2013). Functional connectivity between the amygdala and prefrontal cortex in medication-naive individuals with major depressive disorder. *Journal of Psychiatry and Neuroscience*, 38:417–422.

Koob, G. F. and Le Moal, M. (2001). Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 24(2):97–129.

Koob, G. F. and Volkow, N. D. (2009). Neurocircuitry of Addiction. *Neuropsychopharmacology 2010 35:1*, 35(1):217–238.

Koob, G. F. and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology*, 35:217.

Kooij, J. and Francken, M. (2010). Diagnostic interview for adhd in adults 2.0 (diva 2.0). *DIVA Foundation*.

Kornstein, S. G., Schatzberg, A. F., Thase, M. E., Yonkers, K. A., McCullough, J. P., Keitner, G. I., Gelenberg, A. J., Davis, S. M., Harrison, W. M., and Keller, M. B. (2000). Gender differences in treatment response to sertraline versus imipramine in chronic depression. *The American journal of psychiatry*, 157(9):1445–1452.

Kornstein, S. G., Young, E. A., Harvey, A. T., Wisniewski, S. R., Barkin, J. L., Thase, M. E., Trivedi, M. H., Nierenberg, A. A., and Rush, A. J. (2010). The influence of menopause status and postmenopausal use of hormone therapy on presentation of major depression in women. *Menopause (New York, N.Y.)*, 17:828–839.

Korponay, C. (2021). Snapping out of autopilot: Overriding habits in real time and the role of ventrolateral prefrontal cortex.

Kotoula, V., Stringaris, A., Mackes, N., Mazibuko, N., Hawkins, P. C., Furey, M., Curran, H. V., and Mehta, M. A. (2022). Ketamine modulates the neural correlates of reward processing in unmedicated patients in remission from depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 7:285–292.

Kotoula, V., Webster, T., Stone, J., and Mehta, M. A. (2021). Resting-state connectivity studies as a marker of the acute and delayed effects of subanaesthetic ketamine administration in healthy and depressed individuals: A systematic review. *Brain and Neuroscience Advances*, 5.

Kranzler, H. R., Zhou, H., Kember, R. L., Smith, R. V., Justice, A. C., Damrauer, S., Tsao, P. S., Klarin, D., Baras, A., Reid, J., Overton, J., Rader, D. J., Cheng, Z., Tate, J. P., Becker, W. C., Concato, J., Xu, K., Polimanti, R., Zhao, H., and Gelernter, J. (2019). Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nature communications*, 10.

Krishnan, V. and Nestler, E. J. (2008). The molecular neurobiology of depression. *Nature* 2008 455:7215, 455(7215):894–902.

Kruschke, J. K. (2014). Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, second edition. Elsevier.

Krystal, J. H., Abdallah, C. G., Sanacora, G., Charney, D. S., and Duman, R. S. (2019). Ketamine: A Paradigm Shift for Depression Research and Treatment. *Neuron*, 101(5):774–778.

Kühn, S. and Gallinat, J. (2011). Common biology of craving across legal and illegal drugs - a quantitative meta-analysis of cue-reactivity brain response. *The European journal of neuroscience*, 33(7):1318–1326.

Labonté, B., Engmann, O., Purushothaman, I., Menard, C., Wang, J., Tan, C., Scarpa, J. R., Moy, G., Loh, Y. H. E., Cahill, M., Lorsch, Z. S., Hamilton, P. J., Calipari, E. S., Hodes, G. E., Issler, O., Kronman, H., Pfau, M., Obradovic, A. L., Dong, Y., Neve, R. L., Russo, S., Kazarskis, A., Tamminga, C., Mechawar, N., Turecki, G., Zhang, B., Shen, L., and Nestler, E. J. (2017). Sex-specific transcriptional signatures in human depression. *Nature Medicine*, 23. Lai, C. L. E., Lau, Z., Lui, S. S., Lok, E., Tam, V., Chan, Q., Cheng, K. M., Lam, S. M., and Cheung, E. F. (2017). Meta-analysis of neuropsychological measures of executive functioning in children and adolescents with high-functioning autism spectrum disorder. *Autism research : official journal of the International Society for Autism Research*, 10(5):911–939.

Lau, B. and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, 58(3):451–463.

Laughlin, R. E., Grant, T. L., Williams, R. W., and Jentsch, J. D. (2011). Genetic dissection of behavioral flexibility: Reversal learning in mice. *BPS*, 69:1109–1116.

Laureiro-Martínez, D., Brusoni, S., Canessa, N., and Zollo, M. (2015). Understanding the exploration–exploitation dilemma: An fmri study of attention control and decision-making performance. *Strategic Management Journal*, 36:319–338.

Laureiro-Martínez, D., Canessa, N., Brusoni, S., Zollo, M., Hare, T., Alemanno, F., and Cappa, S. F. (2013). Frontopolar cortex and decision-making efficiency: comparing brain activity of experts with different professional background during an exploration-exploitation task. *Frontiers in Human Neuroscience*, 7.

Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., and Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, 50:1613–1622.

Lawrence, A. D., Sahakian, B. J., Rogers, R. D., Hodges, J. R., and Robbins, T. W. (1999). Discrimination, reversal, and shift learning in Huntington's disease: mechanisms of impaired response selection. *Neuropsychologia*, 37(12):1359–1374.

Leboyer, M., Bellivier, F., Jouvent, R., Nosten-Bertrand, M., Mallet, J., and Pauls, D. (1998). Psychiatric genetics: search for phenotypes. *Trends in neurosciences*, 21(3):102–105.

LeDoux, J. E. (1992). Brain mechanisms of emotion and emotional learning. *Current Opinion in Neurobiology*, 2.

LeDoux, J. E. and Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *American Journal of Psychiatry*, 173:1083–1093.

Lee, B., Tiefenbacher, S., Platt, D. M., and Spealman, R. D. (2004). Pharmacological blockade of α 2-adrenoceptors induces reinstatement of cocaine-seeking behavior in squirrel monkeys. *Neuropsychopharmacology*, 29(4):686–693.

Lee, D., Seo, H., and Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annual Review of Neuroscience*.

Lee, S. H., Ripke, S., Neale, Benjamin M. ... Smoller, J. W., Kendler, K. S., and Wray, N. R. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics 2013 45:9*, 45(9):984–994.

Leeman, R. F. and Potenza, M. N. (2012). Similarities and differences between pathological gambling and substance use disorders: A focus on impulsivity and compulsivity. *Psychopharmacology*, 219:469–490.

Leeson, V. C., Robbins, T. W., Matheson, E., Hutton, S. B., Ron, M. A., Barnes, T. R., and Joyce, E. M. (2009). Discrimination Learning, Reversal, and Set-Shifting in First-Episode Schizophrenia: Stability Over Six Years and Specific Associations with Medication Type and Disorganization Syndrome. *Biological Psychiatry*, 66(6):586.

Lehmann, J., Pryce, C. R., Bettschen, D., and Feldon, J. (1999). The maternal separation paradigm and adult emotionality and cognition in male and female wistar rats. *Pharmacology, biochemistry, and behavior*, 64:705–715.

Lenow, J. K., Constantino, S. M., Daw, N. D., and Phelps, E. A. (2017). Chronic and acute stress promote overexploitation in serial decision making. *The Journal of Neuroscience*, 37:5681.

Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2018). Package 'emmeans'. estimated marginal measn, aka least-squares means.

Leppink, E. W., Redden, S. A., Chamberlain, S. R., and Grant, J. E. (2016). Cognitive flexibility correlates with gambling severity in young adults. *Journal of Psychiatric Research*, 81:9–15.

Levine, S. (1957). Infantile experience and resistance to physiological stress. *Science*, 126:405.

Levy, B. J. and Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: Reflexive reorienting, motor inhibition, and action updating. *Annals of the New York Academy of Sciences*, 1224:40–62.

Li, B., Liu, L., Friston, K. J., Shen, H., Wang, L., Zeng, L. L., and Hu, D. (2013). A treatment-resistant default mode subnetwork in major depression. *Biological psychiatry*, 74(1):48–54.

Li, F., Huang, X., Tang, W., Yang, Y., Li, B., Kemp, G. J., Mechelli, A., and Gong, Q. (2014). Multivariate pattern analysis of dti reveals differential white matter in individuals with obsessive-compulsive disorder. *Human brain mapping*, 35:2643–51.

Li, M., D'Arcy, C., and Meng, X. (2016). Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions. *Psychological Medicine*, 46(4):717–730.

Li, X., Lu, Z. L., D'Argembeau, A., Ng, M., and Bechara, A. (2010). The iowa gambling task in fmri images. *Human brain mapping*, 31:410–423.

Lim, S. I., Song, K. H., Yoo, C. H., Woo, D. C., and Choe, B. Y. (2017). Decreased glutamatergic activity in the frontal cortex of single prolonged stress model: In vivo and ex vivo proton mr spectroscopy. *Neurochemical Research*, 42:2218–2229.

Lim, T. V., Cardinal, R. N., Bullmore, E. T., Robbins, T. W., and Ersche, K. D. (2021). Impaired learning from negative feedback in stimulant use disorder: Dopaminergic modulation. *International Journal of Neuropsychopharmacology*.

Ling, W., Wesson, D. R., Charuvastra, C., and Klett, C. J. (1996). A controlled trial comparing buprenorphine and methadone maintenance in opioid dependence. *Archives of general psychiatry*, 53(5):401–407.

Lipton, D. M., Gonzales, B. J., and Citri, A. (2019). Dorsal striatal circuits for habits, compulsions and addictions. *Frontiers in Systems Neuroscience*, 13.

Liston, C., Matalon, S., Hare, T. A., Davidson, M. C., and Casey, B. J. (2006). Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron*, 50:643–653.

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., Sharma, S., Pearson, D., Plotsky, P. M., and Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science* (*New York, N.Y.*), 277(5332):1659–1662.

Liu, M., Fan, R., Liu, X., Cheng, F., and Wang, J. (2015). Pathways and Networks-Based Analysis of Candidate Genes Associated with Nicotine Addiction. *PLOS ONE*, 10(5):e0127438.

Liu, Z., Liu, W., Yao, L., Yang, C., Xiao, L., Wan, Q., Gao, K., Wang, H., Zhu, F., Wang, G., and Xiao, Z. (2013). Negative life events and corticotropin-releasing-hormone receptor1 gene in recurrent major depressive disorder. *Scientific Reports*, 3.

Lohoff, F. W. and Ferraro, T. N. (2010). Pharmacogenetic considerations in the treatment of psychiatric disorders. *Expert opinion on pharmacotherapy*, 11(3):423–439.

London, E. D., Ernst, M., Grant, S., Bonson, K., and Weinstein, A. (2000). Orbitofrontal cortex and human drug abuse: Functional imaging. *Cerebral Cortex*, 10:334–342.

Lopez-Paniagua, D. and Seger, C. A. (2013). Coding of level of ambiguity within neural systems mediating choice. *Frontiers in Neuroscience*.

Loranger, A. W. (1994). The international personality disorder examination. *Archives of General Psychiatry*, 51.

Lorenzetti, V., Allen, N. B., Fornito, A., and Yücel, M. (2009). Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *Journal of affective disorders*, 117(1-2):1–17.

Lucas, G., Rymar, V. V., Du, J., Mnie-Filali, O., Bisgaard, C., Manta, S., Lambas-Senas, L., Wiborg, O., Haddjeri, N., Piñeyro, G., Sadikot, A. F., and Debonnel, G. (2007). Serotonin(4) (5-HT(4)) receptor agonists are putative antidepressants with a rapid onset of action. *Neuron*, 55(5):712–725.

Luijten, M., Schellekens, A. F., Kühn, S., MacHielse, M. W., and Sescousse, G. (2017). Disruption of Reward Processing in Addiction : An Image-Based Meta-analysis of Functional Magnetic Resonance Imaging Studies. *JAMA Psychiatry*, 74(4):387–398.

Lüscher, C., Robbins, T. W., and Everitt, B. J. (2020). The transition to compulsion in addiction. *Nature Reviews Neuroscience*, 21:247–263.

Maalouf, F. T., Clark, L., Tavitian, L., Sahakian, B. J., Brent, D., and Phillips, M. L. (2012). Bias to negative emotions: A depression state-dependent marker in adolescent major depressive disorder. *Psychiatry Research*, 198:28–33.

Maia, T. V., Cooney, R. E., and Peterson, B. S. (2008). The neural bases of obsessivecompulsive disorder in children and adults. *Development and psychopathology*, 20:1251.

Maia, T. V. and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, 14(2):154–162.

Malykhin, N. V., Carter, R., Seres, P., and Coupland, N. J. (2010). Structural changes in the hippocampus in major depressive disorder: contributions of disease and treatment. *Journal of Psychiatry and Neuroscience*, 35(5):337–343.

Manoach, D. S. (2009). Cognitive deficits in schizophrenia. *Encyclopedia of Neuroscience*, pages 1101–1109.

Manoliu, A., Meng, C., Brandl, F., Doll, A., Tahmasian, M., Scherr, M., Schwerthöffer, D., Zimmer, C., Förstl, H., Bäuml, J., Riedl, V., Wohlschläger, A. M., and Sorg, C. (2014). Insular dysfunction within the salience network is associated with severity of symptoms and aberrant inter-network connectivity in major depressive disorder. *Frontiers in human neuroscience*, 7(JAN).

Marcus, S. M., Kerber, K. B., Rush, A. J., Wisniewski, S. R., Nierenberg, A., Balasubramani, G. K., Ritz, L., Kornstein, S., Young, E. A., and Trivedi, M. H. (2008). Sex differences in depression symptoms in treatment-seeking adults: confirmatory analyses from the Sequenced Treatment Alternatives to Relieve Depression study. *Comprehensive psychiatry*, 49(3):238–246.

Maremmani, I., Pacini, M., Popovic, D., Romano, A., Maremmani, A. G., Perugi, G., Deltito, J., Akiskal, K., and Akiskal, H. (2009). Affective temperaments in heroin addiction. *Journal of affective disorders*, 117(3):186–192.

Marinelli, M. and White, F. J. (2000). Enhanced vulnerability to cocaine self-administration is associated with elevated impulse activity of midbrain dopamine neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(23):8876–8885.

Marquardt, K., Saha, M., Mishina, M., Young, J. W., and Brigman, J. L. (2014). Loss of GluN2A-Containing NMDA Receptors Impairs Extra-dimensional Set-Shifting. *Genes, brain, and behavior*, 13(7):611.

Martin, M. M. and Rubin, R. B. (1995). A new measure of cognitive flexibility:. http://dx.doi.org/10.2466/pr0.1995.76.2.623, 76:623–626.

Masrour, F. F., Peeri, M., Azarbayjani, M. A., and Hosseini, M. J. (2018). Voluntary exercise during adolescence mitigated negative the effects of maternal separation stress on the depressive-like behaviors of adult male rats: Role of nmda receptors. *Neurochemical research*, 43:1067–1074.

Mathews, A. and MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual review of clinical psychology*, 1:167–195.

Matt, G. E., Vázquez, C., and Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, 12(2):227–255.

Matthews, K., Wilkinson, L. S., and Robbins, T. W. (1996). Repeated maternal separation of preweanling rats attenuates behavioral responses to primary and conditioned incentives in adulthood. *Physiology and Behavior*, 59:99–107.

McDonald, A. P., D'Arcy, R. C., and Song, X. (2018). Functional MRI on executive functioning in aging and dementia: A scoping review of cognitive tasks. *Aging Medicine*, 1(2):209.

McIntyre, R. S., Cha, D. S., Soczynska, J. K., Woldeyohannes, H. O., Gallaugher, L. A., Kudlow, P., Alsuwaidan, M., and Baskaran, A. (2013). Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depression and anxiety*, 30(6):515–527.

Meaney, M. J. (2001). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual review of neuroscience*, 24:1161–1192.

Mestres-Missé, A., Trampel, R., Turner, R., and Kotz, S. A. (2017). Uncertainty and expectancy deviations require cortico-subcortical cooperation. *NeuroImage*, 144.

Meyer-Lindenberg, A. and Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience 2006* 7:10, 7(10):818–827.

Miller, K. J., Shenhav, A., and Ludvig, E. A. (2019). Habits without values. *Psychological Review*, 126:292–311.

Millstein, R. A. and Holmes, A. (2007). Effects of repeated maternal separation on anxiety- and depression-related phenotypes in different mouse strains. *Neuroscience and Biobehavioral Reviews*, 31:3–17.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, 9:90–100.

Milton, A. L., Lee, J. L., Butler, V. J., Gardner, R., and Everitt, B. J. (2008). Intraamygdala and systemic antagonism of nmda receptors prevents the reconsolidation of drug-associated memory and impairs subsequently both novel and previously acquired drug-seeking behaviors. *Journal of Neuroscience*, 28:8230–8237.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1):49–100.

Moeller, G. F., Dougherty, D. M., Barratt, E. S., Oderinde, V., Mathias, C. W., Andrew Harper, R., and Swann, A. C. (2002). Increased impulsivity in cocaine dependent subjects independent of antisocial personality disorder and aggression. *Drug and alcohol dependence*, 68(1):105–111.

Mogil, J. S. and Chanda, M. L. (2005). The case for the inclusion of female subjects in basic science studies of pain. *Pain*, 117.

Monroe, S. M. and Harkness, K. L. (2005). Life stress, the "kindling" hypothesis, and the recurrence of depression: considerations from a life stress perspective. *Psychological review*, 112(2):417–445.

Monroe, S. M. and Reid, M. W. (2008). Gene-environment interactions in depression research: genetic polymorphisms and life-stress polyprocedures. *Psychological science*, 19(10):947–956.

Monsell, S. (2003). Task switching. Trends in Cognitive Sciences, 7:134–140.

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., and Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction*, 96(12).

Mora, S., Merchán, A., Aznar, S., Flores, P., and Moreno, M. (2020). Increased amygdala and decreased hippocampus volume after schedule-induced polydipsia in high drinker compulsive rats. *Behavioural Brain Research*, 390:112592.

Moreno-Fernández, R. D., Pérez-Martín, M., Castilla-Ortega, E., Valle, C. R. D., García-Fernández, M. I., Chun, J., Estivill-Torrús, G., Fonseca, F. R. D., Santín, L. J., and Pedraza, C. (2017). Malpa1-null mice as an endophenotype of anxious depression. *Translational Psychiatry*, 7:e1077.

Mukherjee, D., Filipowicz, A. L., Vo, K., Satterthwaite, T. D., and Kable, J. W. (2020). Reward and punishment reversal-learning in major depressive disorder. *Journal of Abnormal Psychology*, 129.

Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F., and Tendolkar, I. (2015). Resting-state functional connectivity in major depressive disorder: A review. *Neuroscience & Biobehavioral Reviews*, 56:330–344.

Murata, T., Hamada, T., Shimokawa, T., Tanifuji, M., and Yanagida, T. (2014). Stochastic process underlying emergent recognition of visual objects hidden in degraded images. *PLOS ONE*, 9:e115658.

Murphy, E. K., Sved, A. F., and Finlay, J. M. (2003). Corticotropin-releasing hormone receptor blockade fails to alter stress-evoked catecholamine release in prefrontal cortex of control or chronically stressed rats. *Neuroscience*, 116(4):1081–1087.

Murphy, E. R., Fernando, A. B., Urcelay, G. P., Robinson, E. S., Mar, A. C., Theobald, D. E., Dalley, J. W., and Robbins, T. W. (2012a). Impulsive behaviour induced by both NMDA receptor antagonism and GABA A receptor activation in rat ventromedial prefrontal cortex. *Psychopharmacology*, 219(2):401–410.

Murphy, F. C., Michael, A., and Sahakian, B. J. (2012b). Emotion modulates cognitive flexibility in patients with major depression. *Psychological medicine*, 42(7):1373–1382.
Murphy, F. C., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., Paykel, E. S., and Sahakian, B. J. (2001). Decision-making cognition in mania and depression. *Psychological medicine*, 31(4):679–693.

Murphy, F. C., Sahakian, B. J., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., and Paykel, E. S. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychological Medicine*, 29(6):1307–1321.

Murray, G. K., Knolle, F., Ersche, K. D., Craig, K. J., Abbott, S., Shabbir, S. S., Fineberg, N. A., Suckling, J., Sahakian, B. J., Bullmore, E. T., and Robbins, T. W. (2019). Dopaminergic drug treatment remediates exaggerated cingulate prediction error responses in obsessivecompulsive disorder. *Psychopharmacology 2019 236:8*, 236:2325–2336.

Murrough, J. W., Iacoviello, B., Neumeister, A., Charney, D. S., and Iosifescu, D. V. (2011). Cognitive dysfunction in depression: neurocircuitry and new therapeutic strategies. *Neurobiology of learning and memory*, 96(4):553–563.

Murthy, S. and Gould, E. (2018). Early Life Stress in Rodents: Animal Models of Illness or Resilience? *Frontiers in Behavioral Neuroscience*, 12:157.

Must, A., Horvath, S., Nemeth, V. L., Janka, Z., Hayden, B., and Chandrasekhar Pammi, V. S. (2013). The Iowa Gambling Task in depression – what have we learned about sub-optimal decision-making strategies? *Frontiers in Psychology*.

Mäki-Marttunen, V., Hagen, T., and Espeseth, T. (2019). Task context load induces reactive cognitive control: An fmri study on cortical and brain stem activity. *Cognitive, Affective, & Behavioral Neuroscience 2019 19:4*, 19:945–965.

Müller, C. P. and Homberg, J. R. (2015). The role of serotonin in drug use and addiction. *Behavioural Brain Research*, 277:146–192.

Naito, S., Sato, K., Yoshida, K., Higuchi, H., Takahashi, H., Kamata, M., Ito, K., Ohkubo, T., and Shimizu, T. (2007). Gender differences in the clinical effects of fluvoxamine and milnacipran in Japanese major depressive patients. *Psychiatry and clinical neurosciences*, 61(4):421–427.

Nakama, H., Chang, L., Fein, G., Shimotsu, R., Jiang, C. S., and Ernst, T. (2011). Methamphetamine users show greater than normal age-related cortical grey matter loss. *Addiction* (*Abingdon, England*), 106:1474.

Naqvi, N. H. and Bechara, A. (2009). The hidden island of addiction: the insula. *Trends in Neurosciences*, 32:56–67.

Negele, A., Kaufhold, J., Kallenbach, L., and Leuzinger-Bohleber, M. (2015). Childhood Trauma and Its Relation to Chronic Depression in Adulthood. *Depression Research and Treatment*, 2015.

Nelson, E. C., Agrawal, A., Heath, A. C., Bogdan, R., Sherva, R., Zhang, B., Al-Hasani, R., Bruchas, M. R., Chou, Y. L., Demers, C. H., Carey, C. E., Conley, E. D., Fakira, A. K., Farrer, L. A., Goate, A., Gordon, S., Henders, A. K., Hesselbrock, V., Kapoor, M., Lynskey, M. T., Madden, P. A., Moron, J. A., Rice, J. P., Saccone, N. L., Schwab, S. G., Shand, F. L.,

Todorov, A. A., Wallace, L., Wang, T., Wray, N. R., Zhou, X., Degenhardt, L., Martin, N. G., Hariri, A. R., Kranzler, H. R., Gelernter, J., Bierut, L. J., Clark, D. J., and Montgomery, G. W. (2016). Evidence of cnih3 involvement in opioid dependence. *Molecular psychiatry*, 21:608–614.

Nelson, J., Bundoc-Baronia, R., Comiskey, G., and McGovern, T. F. (2017). Facing Addiction in America: The Surgeon General's Report on Alcohol, Drugs, and Health: A Commentary. *Alcoholism Treatment Quarterly*, 35(4).

Nestler, E. J., Barrot, M., DiLeone, R. J., Eisch, A. J., Gold, S. J., and Monteggia, L. M. (2002). Neurobiology of Depression. *Neuron*, 34(1):13–25.

Neufang, S., Fink, G. R., Herpertz-Dahlmann, B., Willmes, K., and Konrad, K. (2008). Developmental changes in neural activation and psychophysiological interaction patterns of brain regions associated with interference control and time perception. *NeuroImage*, 43:399–409.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, affective & behavioral neuroscience*, 12:241.

Nigg, J. T., Wong, M. M., Martel, M. M., Jester, J. M., Puttler, L. I., Glass, J. M., Adams, K. M., Fitzgerald, H. E., and Zucker, R. A. (2006). Poor response inhibition as a predictor of problem drinking and illicit drug use in adolescents at risk for alcoholism and other substance use disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(4):468–475.

NIMH (2022). Major Depression.

Noël, X., Brevers, D., Bechara, A., Hanak, C., Kornreich, C., Verbanck, P., and Le Bon, O. (2011). Neurocognitive Determinants of Novelty and Sensation-Seeking in Individuals with Alcoholism. *Alcohol and Alcoholism*, 46(4):407–415.

Norman, K. J., Riceberg, J. S., Koike, H., Bateh, J., McCraney, S. E., Caro, K., Kato, D., Liang, A., Yamamuro, K., Flanigan, M. E., Kam, K., Falk, E. N., Brady, D. M., Cho, C., Sadahiro, M., Yoshitake, K., Maccario, P., Demars, M. P., Waltrip, L., Varga, A. W., Russo, S. J., Baxter, M. G., Shapiro, M. L., Rudebeck, P. H., and Morishita, H. (2021). Post-error recruitment of frontal sensory cortical projections promotes attention in mice. *Neuron*, 109.

Noto, C., Rizzo, L. B., Mansur, R. B., McIntyre, R. S., Maes, M., and Brietzke, E. (2014). Targeting the inflammatory pathway as a therapeutic tool for major depression. *Neuroim-munomodulation*, 21(2-3):131–139.

Oberlin, B. G. and Grahame, N. J. (2009). High-alcohol preferring mice are more impulsive than low-alcohol preferring mice as measured in the delay discounting task. *Alcoholism, clinical and experimental research*, 33(7):1294–1303.

O'Brien, J. W., Lichenstein, S. D., and Hill, S. Y. (2014). Maladaptive Decision Making and Substance Use Outcomes in High-Risk Individuals: Preliminary Evidence for the Role of 5-HTTLPR Variation. *Journal of Studies on Alcohol and Drugs*, 75(4):643.

Odlaug, B. L., Chamberlain, S. R., Kim, S. W., Schreiber, L. R., and Grant, J. E. (2011). A neurocognitive comparison of cognitive flexibility and response inhibition in gamblers with varying degrees of clinical severity. *Psychological Medicine*, 41.

Ohira, H., Matsunaga, M., Kimura, K., Murakami, H., Osumi, T., Isowa, T., Fukuyama, S., Shinoda, J., and Yamada, J. (2011). Chronic stress modulates neural and cardiovascular responses during reversal learning. *Neuroscience*, 193:193–204.

Oldham, S., Murawski, C., Fornito, A., Youssef, G., Yücel, M., and Lorenzetti, V. (2018). The anticipation and outcome phases of reward and loss processing: A neuroimaging metaanalysis of the monetary incentive delay task. *Human brain mapping*, 39(8):3398–3418.

O'Leary, O. F. and Cryan, J. F. (2013). Towards translational rodent models of depression. *Cell and tissue research*, 354(1):141–153.

Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., Mohr, D. C., and Schatzberg, A. F. (2016). Major depressive disorder. *Nature Reviews Disease Primers* 2016 2:1, 2(1):1–20.

Paddock, S., Laje, G., Charney, D., Rush, A. J., Wilson, A. F., Sorant, A. J., Lipsky, R., Wisniewski, S. R., Manji, H., and McMahon, F. J. (2007). Association of GRIK4 with outcome of antidepressant treatment in the STAR*D cohort. *The American journal of psychiatry*, 164(8):1181–1188.

Papciak, J., Popik, P., Fuchs, E., and Rygula, R. (2013). Chronic psychosocial stress makes rats more 'pessimistic' in the ambiguous-cue interpretation paradigm. *Behavioural brain research*, 256:305–310.

Parvaz, M. A., Konova, A. B., Proudfit, G. H., Dunning, J. P., Malaker, P., Moeller, S. J., Maloney, T., Alia-Klein, N., and Goldstein, R. Z. (2015). Impaired neural response to negative prediction errors in cocaine addiction. *The Journal of Neuroscience*, 35:1872.

Pastor-Bernier, A. and Cisek, P. (2011). Neural Correlates of Biased Competition in Premotor Cortex. *Journal of Neuroscience*, 31(19):7083–7088.

Patterson, T. K., Craske, M. G., and Knowlton, B. J. (2019). Enhanced avoidance habits in relation to history of early-life stress. *Frontiers in Psychology*, 10:1876.

Paulus, M. P. and Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, 16(9):476–483.

Peckham, A. D., McHugh, R. K., and Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and anxiety*, 27(12):1135–1142.

Pedersen, M. L., Frank, M. J., and Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin and Review*, 24.

Pedersen, M. L., Ironside, M., ichi Amemori, K., McGrath, C. L., Kang, M. S., Graybiel, A. M., Pizzagalli, D. A., and Frank, M. J. (2021). Computational phenotyping of brainbehavior dynamics underlying approach-avoidance conflict in major depressive disorder. *PLOS Computational Biology*, 17:e1008955.

Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., Perrot, M., and Édouard Duchesnay (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12.

Pelloux, Y., Dilleen, R., Economidou, D., Theobald, D., and Everitt, B. J. (2012). Reduced forebrain serotonin transmission is causally involved in the development of compulsive cocaine seeking in rats. *Neuropsychopharmacology*, 37(11):2505–2514.

Peng, W., Chen, Z., Yin, L., Jia, Z., and Gong, Q. (2016). Essential brain structural alterations in major depressive disorder: A voxel-wise meta-analysis on first episode, medication-naive patients. *Journal of Affective Disorders*, 199:114–123.

Perandrés-Gómez, A., Navas, J. F., van Timmeren, T., and Perales, J. C. (2021). Decisionmaking (in)flexibility in gambling disorder. *Addictive Behaviors*, 112:106534.

Petry, N. M. (2001). Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology*, 154(3):243–250.

Piccinelli, M. and Wilkinson, G. (2000). Gender differences in depression. Critical review. *British Journal of Psychiatry*, 177(DEC.):486–492.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Authors, E., Heisterkamp, S., Willigen, B. V., Ranke, J., and R-core (2021). Linear and nonlinear mixed effects models contact. *Linear and nonlinear mixed effects models, version*, 3.

Pittenger, C. and Duman, R. S. (2007). Stress, Depression, and Neuroplasticity: A Convergence of Mechanisms. *Neuropsychopharmacology* 2008 33:1, 33(1):88–109.

Pizzagalli, D. A. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual review of clinical psychology*, 10:393.

Pizzagalli, D. A. and Roberts, A. C. (2021). Prefrontal cortex and depression. *Neuropsy-chopharmacology* 2021 47:1, 47(1):225–246.

Platt, M. L. and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature 1999 400:6741*, 400(6741):233–238.

Polich, J. (2007). Updating p300: An integrative theory of p3a and p3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 118:2128.

Porsolt, R. D., Le Pichon, M., and Jalfre, M. (1977). Depression: a new animal model sensitive to antidepressant treatments. *Nature*, 266(5604):730–732.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fmri. *NeuroImage*, 84:320–341.

Power, M. J., Dalgleish, T., Claudio, V., Tata, P., and Kentish, J. (2000). The directed forgetting task: Application to emotionally valent material. *Journal of Affective Disorders*, 57(1-3).

Power, Y., Goodyear, B., and Crockford, D. (2012). Neural correlates of pathological gamblers preference for immediate rewards during the iowa gambling task: An fmri study. *Journal of Gambling Studies*, 28:623–636.

Prom-Wormley, E. C., Ebejer, J., Dick, D. M., and Bowers, M. S. (2017). The genetic epidemiology of substance use disorder: A review. *Drug and Alcohol Dependence*, 180:241–259.

Pulcu, E. and Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *eLife*, 6.

Purcell, R., Maruff, P., Kyrios, M., and Pantelis, C. (1997). Neuropsychological function in young patients with unipolar major depression. *Psychological Medicine*, 27(6):1277–1285.

Qin, X., He, Y., Wang, N., Zou, J. X., Zhang, Y. M., Cao, J. L., Pan, B. X., and Zhang, W. H. (2019). Moderate maternal separation mitigates the altered synaptic transmission and neuronal activation in amygdala by chronic stress in adult mice. *Molecular Brain*, 12:1–12.

Qiu, L. R., Fernandes, D. J., Szulc-Lerch, K. U., Dazai, J., Nieman, B. J., Turnbull, D. H., Foster, J. A., Palmert, M. R., and Lerch, J. P. (2018). Mouse mri shows brain areas relatively larger in males emerge before those larger in females. *Nature Communications 2018 9:1*, 9:1–15.

Queirazza, F., Fouragnan, E., Steele, J. D., Cavanagh, J., and Philiastides, M. G. (2019). Neural correlates of weighted reward prediction error during reinforcement learning classify response to cognitive behavioral therapy in depression. *Science Advances*, 5.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radke, A. K., Nakazawa, K., and Holmes, A. (2015). Cortical GluN2B deletion attenuates punished suppression of food reward-seeking. *Psychopharmacology*, 232(20):3753–3761.

Radua, J. and Mataix-Cols, D. (2009). Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. *The British journal of psychiatry : the journal of mental science*, 195:393–402.

Radwanska, K. and Kaczmarek, L. (2012). Characterization of an alcohol addiction-prone phenotype in mice. *Addiction biology*, 17(3):601–612.

Ragozzino, M. E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Annals of the New York Academy of Sciences*, 1121:355–375.

Raimo, S., Cropano, M., Trojano, L., and Santangelo, G. (2021). The neural basis of gambling disorder: An activation likelihood estimation meta-analysis. *Neuroscience and Biobehavioral Reviews*, 120.

Raineki, C., Cortés, M. R., Belnoue, L., and Sullivan, R. M. (2012). Effects of early-life abuse differ across development: Infant social behavior deficits are followed by adolescent depressive-like behaviors mediated by the amygdala. *Journal of Neuroscience*, 32:7758–7765.

Raio, C. M., Hartley, C. A., Orederu, T. A., Li, J., and Phelps, E. A. (2017). Stress attenuates the flexible updating of aversive value. *Proceedings of the National Academy of Sciences of the United States of America*, 114.

Ramasubbu, R., Konduru, N., Cortese, F., Bray, S., Gaxiola-Valdez, I., and Goodyear, B. (2014). Reduced intrinsic connectivity of amygdala in adults with major depressive disorder. *Frontiers in Psychiatry*, 5(FEB):17.

Ratcliff, R. and McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural computation*, 20:873.

Rawson, R. A., Huber, A., McCann, M., Shoptaw, S., Farabee, D., Reiber, C., and Ling, W. (2002). A comparison of contingency management and cognitive-behavioral approaches during methadone maintenance treatment for cocaine dependence. *Archives of General Psychiatry*, 59.

Ray, L. A. and Hutchison, K. E. (2004). A Polymorphism of the-Opioid Receptor Gene (OPRM1) and Sensitivity to the Effects of Alcohol in Humans. *Alcoholism: Clinical and Experimental Research*.

Reich, T., Edenberg, H. J., Goate, A., Williams, J. T., Rice, J. P., Van Eerdewegh, P., Foroud, T., Hesselbrock, V., Schuckit, M. A., Bucholz, K., Porjesz, B., Li, T. K., Conneally, P. M., Nurnberger, J. I., Tischfield, J. A., Crowe, R. R., Cloninger, C. R., Wu, W., Shears, S., Carr, K., Crose, C., Willig, C., and Begleiter, H. (1998). Genome-wide search for genes affecting the risk for alcohol dependence. *American Journal of Medical Genetics - Neuropsychiatric Genetics*, 81(3).

Remijnse, P. L., Nielen, M. M., Uylings, H. B., and Veltman, D. J. (2005). Neural correlates of a reversal learning task with an affectively neutral baseline: An event-related fMRI study. *NeuroImage*, 26(2):609–618.

Remijnse, P. L., Nielen, M. M., Van Balkom, A. J., Cath, D. C., Van Oppen, P., Uylings, H. B., and Veltman, D. J. (2006). Reduced Orbitofrontal-Striatal Activity on a Reversal Learning Task in Obsessive-Compulsive Disorder. *Archives of General Psychiatry*, 63(11):1225–1236.

Renard, G. M., Rivarola, M. A., and Suárez, M. M. (2007). Sexual dimorphism in rats: effects of early maternal separation and variable chronic stress on pituitary-adrenal axis and behavior. *International Journal of Developmental Neuroscience*, 25:373–379.

Reuter, J., Raedler, T., Rose, M., Hand, I., Gläscher, J., and Büchel, C. (2005). Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nature neuroscience*, 8:147–148.

Rezvani, A. H. (2006). Involvement of the NMDA System in Learning and Memory. *Animal Models of Cognitive Impairment*, pages 37–48.

Rice, C. J., Sandman, C. A., Lenjavi, M. R., and Baram, T. Z. (2008). A novel mouse model for acute and long-lasting consequences of early life stress. *Endocrinology*, 149:4892–4900.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306.

Riecher-Rössler, A. (2017). Sex and gender differences in mental disorders. *The Lancet Psychiatry*, 4:8–9.

Ries, R. K., Fiellin, D. A., Miller, S. C., and Saitz, R. (2014). *The ASAM principles of addiction medicine: Fifth edition*. Lippincott Williams & Wilkins.

Robbins, T. W. (2012). Animal models of neuropsychiatry revisited: A personal tribute to Teitelbaum. *Behavioural Brain Research*, 231(2):337–342.

Robbins, T. W. and Costa, R. M. (2017). Habits. Current Biology, 27:R1200-R1206.

Roberson-Nay, R., McClure, E. B., Monk, C. S., Nelson, E. E., Guyer, A. E., Fromm, S. J., Charney, D. S., Leibenluft, E., Blair, J., Ernst, M., and Pine, D. S. (2006). Increased Amygdala Activity During Successful Memory Encoding in Adolescent Major Depressive Disorder: An fMRI Study. *Biological Psychiatry*, 60(9):966–973.

Roberts, A. C., Robbins, T. W., and Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *https://doi.org/10.1080/14640748808402328*, 40:321–341.

Robinson, A. H., Chong, T. T.-J., and Verdejo-Garcia, A. (2022). Computational models of exploration and exploitation characterise onset and efficacy of treatment in methamphetamine use disorder. *Addiction Biology*, 27:e13172.

Robinson, E. S. and Roiser, J. P. (2016). Affective biases in humans and animals. *Current Topics in Behavioral Neurosciences*, 28:263–286.

Robinson, O. J., Cools, R., Carlisi, C. O., Sahakian, B. J., and Drevets, W. C. (2012). Ventral striatum response during reward and punishment reversal learning in unmedicated major depressive disorder. *American Journal of Psychiatry*, 169:152–159.

Robinson, O. J., Krimsky, M., Lieberman, L., Allen, P., Vytal, K., and Grillon, C. (2014). Towards a mechanistic understanding of pathological anxiety: the dorsal medial prefrontalamygdala 'aversive amplification' circuit in unmedicated generalized and social anxiety disorders. *The lancet. Psychiatry*, 1:294–302.

Robinson, T. E. and Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, 18.

Rock, P. L., Roiser, J. P., Riedel, W. J., and Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, 44(10):2029–2040.

Rogers, M. A., Kasai, K., Koji, M., Fukuda, R., Iwanami, A., Nakagome, K., Fukuda, M., and Kato, N. (2004). Executive and prefrontal dysfunction in unipolar depression: a review of neuropsychological and imaging evidence. *Neuroscience research*, 50(1):1–11.

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., and Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *Journal of cognitive neuroscience*, 12(1):142–162.

Rogers, R. D., Everitt, B. J., Baldacchino, A., Blackshaw, A. J., Swainson, R., Wynne, K., Baker, N. B., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J. F., Sahakian, B. J., and Robbins, T. W. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 20(4):322–339.

Rolls, E. T., Critchley, H. D., Mason, R., and Wakeman, E. A. (1996). Orbitofrontal cortex neurons: role in olfactory and visual association learning. *https://doi.org/10.1152/jn.1996.75.5.1970*, 75:1970–1981.

Rose, E. J. and Ebmeier, K. P. (2006). Pattern of impaired working memory during major depression. *Journal of Affective Disorders*, 90(2-3):149–161.

Rossi, M. A., Sukharnikova, T., Hayrapetyan, V. Y., Yang, L., and Yin, H. H. (2013). Operant self-stimulation of dopamine neurons in the substantia nigra. *PloS one*, 8(6).

Rudebeck, P. H., Buckley, M. J., Walton, M. E., and Rushworth, M. F. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science (New York, N.Y.)*, 313(5791):1310–1312.

Ruhé, H. G., Mason, N. S., and Schene, A. H. (2007). Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. *Molecular psychiatry*, 12(4):331–359.

Ruiter, M. B. D., Veltman, D. J., Goudriaan, A. E., Oosterlaan, J., Sjoerds, Z., and Brink, W. V. D. (2008). Response perseveration and ventral prefrontal sensitivity to reward and punishment in male problem gamblers and smokers. *Neuropsychopharmacology 2009 34:4*, 34:1027–1038.

Rygula, R., Clarke, H. F., Cardinal, R. N., Cockcroft, G. J., Xia, J., Dalley, J. W., Robbins, T. W., and Roberts, A. C. (2015). Role of central serotonin in anticipation of rewarding and punishing outcomes: Effects of selective amygdala or orbitofrontal 5-HT Depletion. *Cerebral Cortex*, 25(9):3064–3076.

Ryman, S. G., Shaikh, A. A. E., Shaff, N. A., Hanlon, F. M., Dodd, A. B., Wertz, C. J., Ling, J. M., Barch, D. M., Stromberg, S. F., Lin, D. S., Abrams, S., and Mayer, A. R. (2019). Proactive and reactive cognitive control rely on flexible use of the ventrolateral prefrontal cortex. *Human Brain Mapping*, 40:955–966.

Sacher, J., Neumann, J., Fünfstück, T., Soliman, A., Villringer, A., and Schroeter, M. L. (2012). Mapping the depressed brain: A meta-analysis of structural and functional alterations in major depressive disorder. *Journal of Affective Disorders*, 140(2):142–148.

Safra, L., Chevallier, C., and Palminteri, S. (2019). Depressive symptoms are associated with blunted reward learning in social contexts. *PLOS Computational Biology*, 15:e1007224.

Sailer, U., Robinson, S., Fischmeister, F. P. S., König, D., Oppenauer, C., Lueger-Schuster, B., Moser, E., Kryspin-Exner, I., and Bauer, H. (2008). Altered reward processing in the nucleus accumbens and mesial prefrontal cortex of patients with posttraumatic stress disorder. *Neuropsychologia*, 46:2836–2844.

Salk, R. H., Hyde, J. S., and Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, 143:783–822.

Sanacora, G., Gueorguieva, R., Epperson, C. N., Wu, Y. T., Appel, M., Rothman, D. L., Krystal, J. H., and Mason, G. F. (2004). Subtype-specific alterations of gamma-aminobutyric acid and glutamate in patients with major depression. *Archives of general psychiatry*, 61(7):705–713.

Sánchez, M. M., Ladd, C. O., and Plotsky, P. M. (2001). Early adverse experience as a developmental risk factor for later psychopathology: Evidence from rodent and primate models. *Development and Psychopathology*, 13(3):419–449.

Santarelli, S., Zimmermann, C., Kalideris, G., Lesuis, S. L., Arloth, J., Uribe, A., Dournes, C., Balsevich, G., Hartmann, J., Masana, M., Binder, E. B., Spengler, D., and Schmidt, M. V. (2017). An adverse early life environment can enhance stress resilience in adulthood. *Psychoneuroendocrinology*, 78:213–221.

Satterthwaite, T. D., Cook, P. A., Bruce, S. E., Conway, C., Mikkelsen, E., Satchell, E., Vandekar, S. N., Durbin, T., Shinohara, R. T., and Sheline, Y. I. (2016). Dimensional depression severity in women with major depression and post-traumatic stress disorder correlates with fronto-amygdalar hypoconnectivity. *Molecular Psychiatry*, 21:894–902.

Saulnier, K. G., Allan, N. P., Raines, A. M., and Schmidt, N. B. (2019). Depression and intolerance of uncertainty: Relations between uncertainty subfactors and depression dimensions. *Psychiatry (New York)*, 82:72–79.

Savignac, H. M., Dinan, T. G., and Cryan, J. F. (2011). Resistance to early-life stress in mice: Effects of genetic background and stress duration. *Frontiers in Behavioral Neuroscience*, 5.

Sawiak, S. J., Wood, N. I., Williams, G. B., Morton, A. J., and Carpenter, T. A. (2013). Voxel-based morphometry with templates and validation in a mouse model of huntington's disease. *Magnetic Resonance Imaging*, 31:1522–1531.

Scheggi, S., Montis, M. G. D., and Gambarana, C. (2018). Making sense of rodent models of anhedonia. *International Journal of Neuropsychopharmacology*, 21:1049.

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., and Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, 11:1141–1152.

Schildkraut, J. J. (1995). The catecholamine hypothesis of affective disorders: a review of supporting evidence. 1965. *The Journal of neuropsychiatry and clinical neurosciences*, 7(4).

Schippers, G. M., Broekman, T. G., Buchholz, A., Koeter, M. W., and Brink, W. V. D. (2010). Measurements in the addictions for triage and evaluation (mate): An instrument based on the world health organization family of international classifications. *Addiction*, 105.

Schmand, B., Bakker, D., Saan, R., and Louman, J. (1991). De nederlandse leestest voor volwassenen: een maat voor het premorbide intelligentieniveau. *Tijdschrift voor Gerontologie en Geriatrie*, 22.

Schoenbaum, G., Chiba, A. A., and Gallagher, M. (1999). Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *Journal of Neuroscience*, 19:1876–1884.

Schoenbaum, G., Nugent, S. L., Saddoris, M. P., and Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*, 13(6):885–890.

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., and Gallagher, M. (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning & memory (Cold Spring Harbor, N.Y.)*, 10(2):129–140.

Schuch, J. J., Roest, A. M., Nolen, W. A., Penninx, B. W., and De Jonge, P. (2014). Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. *Journal of affective disorders*, 156:156–163.

Schuch-Goi, S. B., Goi, P. D., Bermudez, M., Fara, L. S., Kessler, F. P., Pechansky, F., Gama, C. S., Massuda, R., and von Diemen, L. (2017). Accumbens volumes are reduced among crack-cocaine users. *Neuroscience letters*, 645:86–89.

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

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

Schwabe, L. and Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience*, 29:7191–7198.

Schwabe, L. and Wolf, O. T. (2010). Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology*, 35:977–986.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., and Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience*, 27(9):2349–2356.

Seifert, C. L., Magon, S., Sprenger, T., Lang, U. E., Huber, C. G., Denier, N., Vogel, M., Schmidt, A., Radue, E. W., Borgwardt, S., and Walter, M. (2015). Reduced volume of the nucleus accumbens in heroin addiction. *European Archives of Psychiatry and Clinical Neuroscience*, 265.

Semple, B. D., Blomgren, K., Gimlin, K., Ferriero, D. M., and Noble-Haeusslein, L. J. (2013). Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Progress in neurobiology*, 0:1.

Seney, M. L., Huo, Z., Cahill, K., French, L., Puralewski, R., Zhang, J., Logan, R. W., Tseng, G., Lewis, D. A., and Sibille, E. (2018). Opposite Molecular Signatures of Depression in Men and Women. *Biological psychiatry*, 84(1):18–27.

Seo, H. and Lee, D. (2007). Temporal Filtering of Reward Signals in the Dorsal Anterior Cingulate Cortex during a Mixed-Strategy Game. *Journal of Neuroscience*, 27(31):8366–8377.

Seo, M., Lee, E., and Averbeck, B. B. (2012). Action selection and action value in frontalstriatal circuits. *Neuron*, 74(5):947.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52.

Sharp, B. M. (2017). Basolateral amygdala and stress-induced hyperexcitability affect motivated behaviors and addiction. *Translational psychiatry*, 7:e1194.

Sharp, T. (2013). Molecular and cellular mechanisms of antidepressant action. *Current topics in behavioral neurosciences*, 14:309–325.

Sheline, Y. I., Gado, M. H., and Price, J. L. (1998). Amygdala core nuclei volumes are decreased in recurrent major depression. *Neuroreport*, 9(9):2023–2028.

Shima, K. and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, 282.

Shin, L. M., Wright, C. I., Cannistraro, P. A., Wedig, M. M., McMullin, K., Martis, B., Macklin, M. L., Lasko, N. B., Cavanagh, S. R., Krangel, T. S., Orr, S. P., Pitman, R. K., Whalen, P. J., and Rauch, S. L. (2005). A functional magnetic resonance imaging study of amygdala and medial prefrontal cortex responses to overtly presented fearful faces in posttraumatic stress disorder. *Archives of general psychiatry*, 62:273–281.

Shine, J., Bissett, P., Bell, P., Koyejo, O., Balsters, J., Gorgolewski, K., Moodie, C., and Poldrack, R. (2016). The dynamics of functional brain networks: Integrated network states during cognitive task performance. *Neuron*, 92:544–554.

Siegel, C., Waldo, M., Mizner, G., Adler, L. E., and Freedman, R. (1984). Deficits in sensory gating in schizophrenic patients and their relatives. Evidence obtained with auditory evoked responses. *Archives of general psychiatry*, 41(6):607–612.

Simon, N. W., Mendez, I. A., and Setlow, B. (2007). Cocaine exposure causes long-term increases in impulsive choice. *Behavioral neuroscience*, 121(3):543.

Smith, A. J., Becker, S., and Kapur, S. (2005). A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural computation*, 17(2):361–395.

Smith, K. S., Virkud, A., Deisseroth, K., and Graybiel, A. M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 109:18932–18937.

Smith, P., Benzina, N., Vorspan, F., Mallet, L., and N'Diaye, K. (2015). Compulsivity and probabilistic reversal learning in ocd and cocaine addiction. *European Psychiatry*, 30:S110–S111.

Smith, R., Schwartenbeck, P., Stewart, J. L., Kuplicki, R., Ekhtiari, H., and Paulus, M. P. (2020). Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma. *Drug and Alcohol Dependence*, 215.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., Luca, M. D., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., Stefano, N. D., Brady, J. M., and Matthews, P. M. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23.

Smoski, M. J., Felder, J., Bizzell, J., Green, S. R., Ernst, M., Lynch, T. R., and Dichter, G. S. (2009). fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *Journal of Affective Disorders*, 118(1-3):69–78.

Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biological psychiatry*, 57(11):1231–1238.

Staddon, J. E. and Cerutti, D. T. (2003). Operant conditioning. *Annual Review of Psychology*, 54.

Stalnaker, T. A., Cooch, N. K., and Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience 2015 18:5*, 18(5):620–627.

Stalnaker, T. A., Takahashi, Y., Roesch, M. R., and Schoenbaum, G. (2009). Neural substrates of cognitive inflexibility after chronic cocaine exposure. *Neuropharmacology*, 56:63–72.

Stan Development Team (2022). RStan: the R interface to Stan. R package version 2.21.5.

Steele, J. D., Kumar, P., and Ebmeier, K. P. (2007). Blunted response to feedback information in depressive illness. *Brain : a journal of neurology*, 130:2367–2374.

Steffen, A., Nübel, J., Jacobi, F., Bätzing, J., and Holstiege, J. (2020). Mental and somatic comorbidity of depression: A comprehensive cross-sectional analysis of 202 diagnosis groups using german nationwide ambulatory claims data. *BMC Psychiatry*, 20:1–15.

Stephenson, M. T. and Helme, D. W. (2006). Authoritative parenting and sensation seeking as predictors of adolescent cigarette and marijuana use. *Journal of drug education*, 36(3):247–270.

Stern, C. E., Passingham, R. E., Stern, C. E., and Passingham, s. R. E. (1995). The nucleus accumbens in monkeys (Macaca fascicularis). *Experimental Brain Research 1995 106:2*, 106(2):239–247.

Steru, L., Chermat, R., Thierry, B., and Simon, P. (1985). The tail suspension test: a new method for screening antidepressants in mice. *Psychopharmacology*, 85(3):367–370.

Stetler, C. and Miller, G. E. (2011). Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosomatic medicine*, 73(2):114–126.

Stuart, S. A., Butler, P., Munafò, M. R., Nutt, D. J., and Robinson, E. S. (2013). A Translational Rodent Assay of Affective Biases in Depression and Antidepressant Therapy. *Neuropsychopharmacology* 2013 38:9, 38(9):1625–1635.

Sul, J. H., Jo, S., Lee, D., and Jung, M. W. (2011). Role of rodent secondary motor cortex in value-based action selection. *Nature neuroscience*, 14:1202–1210.

Sul, J. H., Kim, H., Huh, N., Lee, D., and Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66.

Sun, J., Kranzler, H. R., Gelernter, J., and Bi, J. (2020). A genome-wide association study of cocaine use disorder accounting for phenotypic heterogeneity and gene–environment interaction. *Journal of psychiatry & neuroscience : JPN*, 45(1):34–44.

Sun, W. L. and Yuill, M. B. (2020). Role of the GABAa and GABAb receptors of the central nucleus of the amygdala in compulsive cocaine-seeking behavior in male rats. *Psychopharmacology*, 237(12):3759–3771.

Sundermann, B. and Pfleiderer, B. (2012). Functional connectivity profile of the human inferior frontal junction: involvement in a cognitive control network. *BMC Neuroscience*, 13:119.

Sutton, R. and Barto, A. (1998). Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 9(5):1054–1054.

Svenningsson, P., Chergui, K., Rachleff, I., Flajolet, M., Zhang, X., El Yacoubi, M., Vaugeois, J. M., Nomikos, G. G., and Greengard, P. (2006). Alterations in 5-HT1B receptor function by p11 in depression-like states. *Science (New York, N.Y.)*, 311(5757):77–80.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., and Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia*, 38(5):596–612.

Swanson, K., Averbeck, B. B., and Laubach, M. (2021). Noradrenergic regulation of two-armed bandit performance. *Behavioral Neuroscience*, 136:84–99.

Syed, S. A. and Nemeroff, C. B. (2017). Early life stress, mood, and anxiety disorders. *Chronic Stress*, 1.

Tan, S., Ho, H. S., Song, A. Y., Low, J., and Je, H. S. (2017). Maternal separation does not produce a significant behavioral change in mice. *Experimental neurobiology*, 26:390–398.

Tanabe, J., Thompson, L., Claus, E., Dalwani, M., Hutchison, K., and Banich, M. T. (2007). Prefrontal cortex activity is reduced in gambling and nongambling substance users during decision-making. *Human Brain Mapping*, 28.

Tanaka, S. C., Balleine, B. W., and O'Doherty, J. P. (2008). Calculating consequences: Brain systems that encode the causal effects of actions. *The Journal of Neuroscience*, 28:6750.

Tang, Y., Kong, L., Wu, F., Womer, F., Jiang, W., Cao, Y., Ren, L., Wang, J., Fan, G., Blumberg, H. P., Xu, K., and Wang, F. (2013). Decreased functional connectivity between the amygdala and the left ventral prefrontal cortex in treatment-naive patients with major depressive disorder: a resting-state functional magnetic resonance imaging study. *Psychological Medicine*, 43(9):1921–1927.

Targum, S. D. and Nemeroff, C. B. (2019). The effect of early life stress on adult psychiatric disorders. *Innovations in Clinical Neuroscience*, 16:35.

Tavares, J. V. T., Clark, L., Furey, M. L., Williams, G. B., Sahakian, B. J., and Drevets, W. C. (2008). Neural basis of abnormal response to negative feedback in unmedicated mood disorders. *NeuroImage*, 42:1118–1126.

Teicher, M. H. and Samson, J. A. (2016). Annual research review: Enduring neurobiological effects of childhood abuse and neglect. *Journal of child psychology and psychiatry, and allied disciplines*, 57:241–266.

Terra, H., Bruinsma, B., de Kloet, S. F., van der Roest, M., Pattij, T., and Mansvelder, H. D. (2020). Prefrontal cortical projection neurons targeting dorsomedial striatum control behavioral inhibition. *Current Biology*, 30:4188–4200.e5.

Thai, C. A., Zhang, Y., and Howland, J. G. (2013). Effects of acute restraint stress on setshifting and reversal learning in male rats. *Cognitive, affective & behavioral neuroscience*, 13:164.

Thomas, A. W., Caporale, N., Wu, C., and Wilbrecht, L. (2016). Early maternal separation impacts cognitive flexibility at the age of first independence in mice. *Developmental Cognitive Neuroscience*, 18:49.

Thomason, M. E., Marusak, H. A., Tocco, M. A., Vila, A. M., McGarragle, O., and Rosenberg, D. R. (2014). Altered amygdala connectivity in urban youth exposed to trauma. *Social Cognitive and Affective Neuroscience*, 10:1460–1468.

Thorpe, S. J., Rolls, E. T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Experimental brain research*, 49(1):93–115.

Théberge, F. R., Milton, A. L., Lee, J. L., and Everitt, B. J. (2010). The basolateral amygdala and nucleus accumbens core mediate dissociable aspects of drug memory reconsolidation. *Learning & Memory*, 17:444–453.

Tiego, J., Chamberlain, S. R., Harrison, B. J., Dawson, A., Albertella, L., Youssef, G. J., Fontenelle, L. F., and Yücel, M. (2020). Heritability of overlapping impulsivity and compulsivity dimensional phenotypes. *Scientific Reports*, 10(1):14378.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychological review*, 97:147–168.

Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., and Foa, E. B. (2003). Intolerance of uncertainty in obsessive-compulsive disorder. *Journal of Anxiety Disorders*, 17:233–242.

Trautmann, S., Rehm, J., and Wittchen, H. (2016). The economic costs of mental disorders. *EMBO reports*, 17(9):1245–1249.

Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, 29:2225–2232.

Tripp, A., Oh, H., Guilloux, J. P., Martinowich, K., Lewis, D. A., and Sibille, E. (2012). Brain-derived neurotrophic factor signaling and subgenual anterior cingulate cortex dysfunction in major depressive disorder. *The American journal of psychiatry*, 169(11):1194–1202.

Trivedi, M. H. and Greer, T. L. (2014). Cognitive dysfunction in unipolar depression: implications for treatment. *Journal of affective disorders*, 152-154(1):19–27.

Tsunoka, T., Kishi, T., Ikeda, M., Kitajima, T., Yamanouchi, Y., Kinoshita, Y., Kawashima, K., Okochi, T., Okumura, T., Inada, T., Ozaki, N., and Iwata, N. (2009). Association analysis of group II metabotropic glutamate receptor genes (GRM2 and GRM3) with mood disorders and fluvoxamine response in a Japanese population. *Progress in neuro-psychopharmacology* & *biological psychiatry*, 33(5):875–879.

Turton, S., Myers, J. F., Mick, I., Colasanti, A., Venkataraman, A., Durant, C., Waldman, A., Brailsford, A., Parkin, M. C., Dawe, G., Rabiner, E. A., Gunn, R. N., Lightman, S. L., Nutt, D. J., and Lingford-Hughes, A. (2018). Blunted endogenous opioid release following an oral dexamphetamine challenge in abstinent alcohol-dependent individuals. *Molecular Psychiatry 2018 25:8*, 25(8):1749–1758.

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. (2010). N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29:1310–1320.

Uddin, L. Q. (2021). Cognitive and behavioural flexibility: neural mechanisms and clinical considerations.

Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J., and Boucher, O. (2017). Structure and function of the human insula. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 34:300.

Uddin, L. Q., Yeo, B. T., and Spreng, R. N. (2019). Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. *Brain Topography*, 32(6):926–942.

Uher, R. and McGuffin, P. (2010). The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. *Molecular psychiatry*, 15(1):18–22.

Ullsperger, M., Danielmeier, C., and Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, 94:35–79.

UN, U. N. (2018). World Drug Report 2018. *United Nations Office on Drugs and Crime*, 7(2).

UNODC (2020). UNODC World Drug Report 2020: Global drug use rising; while COVID-19 has far reaching impact on global drug markets.

Vaghi, M. M., Cardinal, R. N., Apergis-Schoute, A. M., Fineberg, N. A., Sule, A., and Robbins, T. W. (2019). Action-Outcome Knowledge Dissociates From Behavior in Obsessive-Compulsive Disorder Following Contingency Degradation. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 4(2):200–209.

Valdés-Hernández, P. A., Sumiyoshi, A., Nonaka, H., Haga, R., Aubert-Vásquez, E., Ogawa, T., Iturria-Medina, Y., Riera, J. J., and Kawashima, R. (2011). An in vivo mri template set for morphometry, tissue segmentation, and fmri localization in rats. *Frontiers in Neuroinformatics*, 5.

van den Bos, R., Homberg, J., and de Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the iowa gambling task. *Behavioural Brain Research*, 238:95–108.

van den Hout, M. and Kindt, M. (2004). Obsessive-compulsive disorder and the paradoxical effects of perseverative behaviour on experienced uncertainty. *Journal of Behavior Therapy and Experimental Psychiatry*, 35:165–181.

van der Kooij, M. A., Grosse, J., Zanoletti, O., Papilloud, A., and Sandi, C. (2015). The effects of stress during early postnatal periods on behavior and hippocampal neuroplasticity markers in adult male mice. *Neuroscience*, 311:508–518.

van Eijndhoven, P., van Wingen, G., van Oijen, K., Rijpkema, M., Goraj, B., Jan Verkes, R., Oude Voshaar, R., Fernández, G., Buitelaar, J., and Tendolkar, I. (2009). Amygdala volume marks the acute state in the early course of depression. *Biological psychiatry*, 65(9):812–818.

Van Rossum, E. F., Binder, E. B., Majer, M., Koper, J. W., Ising, M., Modell, S., Salyakina, D., Lamberts, S. W., and Holsboer, F. (2006). Polymorphisms of the glucocorticoid receptor gene and major depression. *Biological psychiatry*, 59(8):681–688.

van Timmeren, T., Daams, J. G., van Holst, R. J., and Goudriaan, A. E. (2018). Compulsivityrelated neurocognitive performance deficits in gambling disorder: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 84:204–217.

Vatansever, D., Manktelow, A. E., Sahakian, B. J., Menon, D. K., and Stamatakis, E. A. (2016). Cognitive flexibility: A default network and basal ganglia connectivity perspective. *Brain connectivity*, 6:201–207.

Veen, V. V. and Carter, C. S. (2002). The anterior cingulate as a conflict monitor: Fmri and erp studies. *Physiology and Behavior*, 77:477–482.

Veen, V. V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., and Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage*.

Verdejo-Garcia, A., Benbrook, A., Funderburk, F., David, P., Cadet, J. L., and Bolla, K. I. (2007). THE DIFFERENTIAL RELATIONSHIP BETWEEN COCAINE USE AND MAR-IJUANA USE ON DECISION-MAKING PERFORMANCE OVER REPEAT TESTING WITH THE IOWA GAMBLING TASK. *Drug and alcohol dependence*, 90(1):2. Verdejo-Garcia, A., Clark, L., Verdejo-Román, J., Albein-Urios, N., Martinez-Gonzalez, J. M., Gutierrez, B., and Soriano-Mas, C. (2015). Neural substrates of cognitive flexibility in cocaine and gambling addictions. *British Journal of Psychiatry*, 207:158–164.

Verdejo-García, A., Lawrence, A. J., and Clark, L. (2008). Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. *Neuroscience and biobehavioral reviews*, 32(4):777–810.

Verharen, J. P., den Ouden, H. E., Adan, R. A., and Vanderschuren, L. J. (2020). Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology*, 237:1267–1280.

Verharen, J. P., Kentrop, J., Vanderschuren, L. J., and Adan, R. A. (2019). Reinforcement learning across the rat estrous cycle. *Psychoneuroendocrinology*, 100:27–31.

Vetulani, J. (2013). Early maternal separation: A rodent model of depression and a prevailing human condition. *Pharmacological Reports*, 65(6):1451–1461.

Vilgis, V., Yee, D., Silk, T. J., and Vance, A. (2022). Distinct neural profiles of frontoparietal networks in boys with adhd and boys with persistent depressive disorder. *Cognitive, Affective, Behavioral Neuroscience*, 22:1183–1198.

Vink, J. M., Hottenga, J. J., de Geus, E. J., Willemsen, G., Neale, M. C., Furberg, H., and Boomsma, D. I. (2014). Polygenic risk scores for smoking: predictors for alcohol and cannabis use? *Addiction (Abingdon, England)*, 109(7):1141–1151.

Vink, M., Kahn, R. S., Raemaekers, M., van den Heuvel, M., Boersma, M., and Ramsey, N. F. (2005). Function of striatum beyond inhibition and execution of motor responses. *Human Brain Mapping*, 25:336.

Volkow, N. D., Chang, L., Wang, G. J., Fowler, J. S., Ding, Y. S., Sedler, M., Logan, J., Franceschi, D., Gatley, J., Hitzemann, R., Gifford, A., Wong, C., and Pappas, N. (2001). Low level of brain dopamine d2 receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. *The American journal of psychiatry*, 158:2015–2021.

Volkow, N. D. and Fowler, J. S. (2000). Addiction, a disease of compulsion and drive: Involvement of the orbitofrontal cortex. *Cerebral Cortex*, 10.

Volkow, N. D., Fowler, J. S., Wang, G. J., and Swanson, J. M. (2004). Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Molecular psychiatry*, 9:557–569.

Volkow, N. D., Fowler, J. S., Wolf, A. P., Hitzemann, R., Dewey, S., Bendriem, B., Alpert, R., and Hoff, A. (1991). Changes in brain glucose metabolism in cocaine dependence and withdrawal. *The American journal of psychiatry*, 148:621–626.

Volkow, N. D., Wang, G. J., Fowler, J. S., Logan, J., Gatley, S. J., Gifford, A., Hitzemann, R., Ding, Y. S., and Pappas, N. (1999). Prediction of reinforcing responses to psychostimulants in humans by brain dopamine d2 receptor levels. *American Journal of Psychiatry*, 156.

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., Schreiber, L. R., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., and Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, 20.

Voon, V., Irvine, M. A., Derbyshire, K., Worbe, Y., Lange, I., Abbott, S., Morein-Zamir, S., Dudley, R., Caprioli, D., Harrison, N. A., Wood, J., Dalley, J. W., Bullmore, E. T., Grant, J. E., and Robbins, T. W. (2014). Measuring "waiting" impulsivity in substance addictions and binge eating disorder in a novel analogue of rodent serial reaction time task. *Biological psychiatry*, 75(2):148–155.

Vos, T., Barber, R. M., Bell, Brad ... Zunt, J. R., Salomon, J. A., and Murray, C. J. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386(9995):743–800.

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., Boer, P. D., Schmidt, M., and Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biological psychiatry*, 73:639.

Vuijk, R., Deen, M., Arntz, A., and Geurts, H. M. (2021). First psychometric properties of the dutch interview for diagnostic assessment of autism spectrum disorder in adult males without intellectual disability. *Journal of Autism and Developmental Disorders*.

Wackerhagen, C., Veer, I. M., Erk, S., Mohnke, S., Lett, T. A., Wüstenberg, T., Romanczuk-Seiferth, N. Y., Schwarz, K., Schweiger, J. I., Tost, H., Meyer-Lindenberg, A., Heinz, A., and Walter, H. (2020). Amygdala functional connectivity in major depression - disentangling markers of pathology, risk and resilience. *Psychological medicine*, 50.

Wadgave, U. and Nagesh, L. (2016). Nicotine Replacement Therapy: An Overview. *International Journal of Health Sciences*, 10(3):425.

Wager, T. D., Jonides, J., and Reading, S. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *NeuroImage*, 22:1679–1693.

Wainwright, S. R., Workman, J. L., Tehrani, A., Hamson, D. K., Chow, C., Lieblich, S. E., and Galea, L. A. (2016). Testosterone has antidepressant-like efficacy and facilitates imipramine-induced neuroplasticity in male rats exposed to chronic unpredictable stress. *Hormones and behavior*, 79:58–69.

Walker, E. R., McGee, R. E., and Druss, B. G. (2015). Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. *JAMA psychiatry*, 72(4):334–341.

Walker, S. C., Robbins, T. W., and Roberts, A. C. (2009). Differential contributions of dopamine and serotonin to orbitofrontal cortex function in the marmoset. *Cerebral cortex* (*New York, N.Y. : 1991*), 19(4):889–898.

Wang, Q., Chen, C., Cai, Y., Li, S., Zhao, X., Zheng, L., Zhang, H., Liu, J., Chen, C., and Xue, G. (2016). Dissociated neural substrates underlying impulsive choice and impulsive action. *NeuroImage*, 134:540–549.

Wang, Q., Timberlake, M. A., Prall, K., and Dwivedi, Y. (2017). The recent progress in animal models of depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 77:99–109.

Wang, Y., Toyoshima, O., Kunimatsu, J., Yamada, H., and Matsumoto, M. (2021). Tonic firing mode of midbrain dopamine neurons continuously tracks reward values changing moment-by-moment. *eLife*, 10.

Warren, C. A. and McDonough, B. E. (1999). Event-related brain potentials as indicators of smoking cue-reactivity. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 110:1570–1584.

Wassum, K. M. and Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neuroscience and Biobehavioral Reviews*, 57:271–283.

Weaverdyck, M. E., Lieberman, M. D., and Parkinson, C. (2020). Tools of the trade multivoxel pattern analysis in fmri: A practical introduction for social and affective neuroscientists. *Social Cognitive and Affective Neuroscience*, 15:487–509.

Wei, S., Xue, Z., Sun, W., Han, J., Wu, H., and Liu, X. (2021). Altered neural processing of reward and punishment in women with methamphetamine use disorder. *Frontiers in Psychiatry*, 12:1731.

Weinberg, A. and Shankman, S. A. (2017). Blunted reward processing in remitted melancholic depression. *Clinical Psychological Science*, 5.

Weniger, G., Lange, C., and Irle, E. (2006). Abnormal size of the amygdala predicts impaired emotional memory in major depressive disorder. *Journal of affective disorders*, 94(1-3):219–229.

White, J. D., Arefin, T. M., Pugliese, A., Lee, C. H., Gassen, J., Zhang, J., and Kaffman, A. (2020). Early life stress causes sex-specific changes in adult fronto-limbic connectivity that differentially drive learning. *eLife*, 9:1–29.

WHO (2016). Depression.

WHO (2020). Depression.

Whooley, M. A. and Wong, J. M. (2013). Depression and Cardiovascular Disorders. http://dx.doi.org/10.1146/annurev-clinpsy-050212-185526, 9:327–354.

Wiecki, T. V., Sofer, I., and Frank, M. J. (2013). Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, 7:14.

Wiehler, A. and Peters, J. (2020). Diffusion modeling reveals reinforcement learning impairments in gambling disorder that are linked to attenuated ventromedial prefrontal cortex value representations. *bioRxiv*.

Wilkinson, M. P., Grogan, J. P., Mellor, J. R., and Robinson, E. S. J. (2020). Comparison of conventional and rapid-acting antidepressants in a rodent probabilistic reversal learning task. *Brain and neuroscience advances*, 4:2398212820907177.

Wilkinson, M. P., Slaney, C. L., Mellor, J. R., and Robinson, E. S. J. (2021). Investigation of reward learning and feedback sensitivity in non-clinical participants with a history of early life stress. *PLOS ONE*, 16:e0260444.

Willner, P. (2017). The chronic mild stress (CMS) model of depression: History, evaluation and usage. *Neurobiology of Stress*, 6:78.

Willuhn, I., Burgeno, L. M., Everitt, B. J., and Phillips, P. E. (2012). Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50):20703–20708.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., and Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, 81(2):267–279.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., and Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92.

Wolffgramm, J. and Heyne, A. (1995). From controlled drug intake to loss of control: the irreversible development of drug addiction in the rat. *Behavioural Brain Research*, 70.

Wong, D. F., Kuwabara, H., Schretlen, D. J., Bonson, K. R., Zhou, Y., Nandi, A., Brašić, J. R., Kimes, A. S., Maris, M. A., Kumar, A., Contoreggi, C., Links, J., Ernst, M., Rousset, O., Zukin, S., Grace, A. A., Rohde, C., Jasinski, D. R., Gjedde, A., and London, E. D. (2006). Increased occupancy of dopamine receptors in human striatum during cue-elicited cocaine craving. *Neuropsychopharmacology 2006 31:12*, 31:2716–2727.

Wong, M. L. and Licinio, J. (2001). Research and treatment approaches to depression. *Nature reviews. Neuroscience*, 2(5):343–351.

Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2004). Multilevel linear modelling for fmri group analysis using bayesian inference. *NeuroImage*, 21:1732–1747.

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fmri data. *NeuroImage*, 14:1370–1386.

Worringer, B., Langner, R., Koch, I., Eickhoff, S. B., Eickhoff, C. R., and Binkofski, F. C. (2019). Common and distinct neural correlates of dual-tasking and task-switching: a meta-analytic review and a neuro-cognitive processing model of human multitasking. *Brain Structure and Function 2019 224:5*, 224:1845–1869.

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M., Bacanu, S. A., Bækvad-Hansen, M., Beekman, A. F., Bigdeli, T. B., Binder, E. B., Blackwood, D. R., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., Cai, N., Castelao, E., Christensen, J. H., Clarke, T. K., Coleman, J. I., Colodro-Conde, L., Couvy-Duchesne, B., Craddock, N., Crawford, G. E., Crowley, C. A., Dashti, H. S., Davies, G., Deary, I. J., Degenhardt, F., Derks, E. M., DIrek, N., Dolan, C. V., Dunn, E. C., Eley, T. C., Eriksson, N., Escott-Price, V., Kiadeh, F. H. F., Finucane, H. K., Forstner, A. J., Frank, J., Gaspar, H. A., Gill, M., Giusti-Rodríguez, P., Goes, F. S., Gordon, S. D., Grove, J., Hall, L. S., Hannon, E., Hansen, C. S., Hansen, T. F.,

Herms, S., Hickie, I. B., Hoffmann, P., Homuth, G., Horn, C., Hottenga, J. J., Hougaard, D. M., Hu, M., Hyde, C. L., Ising, M., Jansen, R., Jin, F., Jorgenson, E., Knowles, J. A., Kohane, I. S., Kraft, J., Kretzschmar, W. W., Krogh, J., Kutalik, Z., Lane, J. M., Li, Y., Li, Y., Lind, P. A., Liu, X., Lu, L., MacIntyre, D. J., MacKinnon, D. F., Maier, R. M., Maier, W., Marchini, J., Mbarek, H., McGrath, P., McGuffin, P., Medland, S. E., Mehta, D., Middeldorp, C. M., Mihailov, E., Milaneschi, Y., Milani, L., Mill, J., Mondimore, F. M., Montgomery, G. W., Mostafavi, S., Mullins, N., Nauck, M., Ng, B., Nivard, M. G., Nyholt, D. R., O'Reilly, P. F., Oskarsson, H., Owen, M. J., Painter, J. N., Pedersen, C. B., Pedersen, M. G., Peterson, R. E., Pettersson, E., Peyrot, W. J., Pistis, G., Posthuma, D., Purcell, S. M., Quiroz, J. A., Qvist, P., Rice, J. P., Riley, B. P., Rivera, M., Mirza, S. S., Saxena, R., Schoevers, R., Schulte, E. C., Shen, L., Shi, J., Shyn, S. I., Sigurdsson, E., Sinnamon, G. B., Smit, J. H., Smith, D. J., Stefansson, H., Steinberg, S., Stockmeier, C. A., Streit, F., Strohmaier, J., Tansey, K. E., Teismann, H., Teumer, A., Thompson, W., Thomson, P. A., Thorgeirsson, T. E., Tian, C., Traylor, M., Treutlein, J., Trubetskoy, V., Uitterlinden, A. G., Umbricht, D., Auwera, S. V. D., Hemert, A. M. V., Viktorin, A., Visscher, P. M., Wang, Y., Webb, B. T., Weinsheimer, S. M., Wellmann, J., Willemsen, G., Witt, S. H., Wu, Y., Xi, H. S., Yang, J., Zhang, F., Arolt, V., Baune, B. T., Berger, K., Boomsma, D. I., Cichon, S., Dannlowski, U., Geus, E. C. D., Depaulo, J. R., Domenici, E., Domschke, K., Esko, T., Grabe, H. J., Hamilton, S. P., Hayward, C., Heath, A. C., Hinds, D. A., Kendler, K. S., Kloiber, S., Lewis, G., Li, Q. S., Lucae, S., Madden, P. F., Magnusson, P. K., Martin, N. G., McIntosh, A. M., Metspalu, A., Mors, O., Mortensen, P. B., Müller-Myhsok, B., Nordentoft, M., Nöthen, M. M., O'Donovan, M. C., Paciga, S. A., Pedersen, N. L., Penninx, B. W., Perlis, R. H., Porteous, D. J., Potash, J. B., Preisig, M., Rietschel, M., Schaefer, C., Schulze, T. G., Smoller, J. W., Stefansson, K., Tiemeier, H., Uher, R., Völzke, H., Weissman, M. M., Werge, T., Winslow, A. R., Lewis, C. M., Levinson, D. F., Breen, G., Børglum, A. D., and Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nature Genetics 2018 50:5, 50:668-681.

Wu, C. and Rao, J. N. (2006). Pseudo-empirical likelihood ratio confidence intervals for complex surveys. *Canadian Journal of Statistics*, 34.

Wu, F., Tu, Z., Sun, J., Geng, H., Zhou, Y., Jiang, X., Li, H., and Kong, L. (2020). Abnormal functional and structural connectivity of amygdala-prefrontal circuit in first-episode adolescent depression: A combined fmri and dti study. *Frontiers in Psychiatry*, 10:1.

Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., Kornreich, C., Chatard, A., Jaafari, N., and Noël, X. (2019). Reduced model-based decision-making in gambling disorder. *Scientific Reports 2019 9:1*, 9:1–10.

Yan, C. G., Chen, X., Li, L., Castellanos, F. X., Bai, T. J., Bo, Q. J., Cao, J., Chen, G. M., Chen, N. X., Chen, W., Cheng, C., Cheng, Y. Q., Cui, X. L., Duan, J., Fang, Y. R., Gong, Q. Y., Guo, W. B., Hou, Z. H., Hu, L., Kuang, L., Li, F., Li, K. M., Li, T., Liu, Y. S., Liu, Z. N., Long, Y. C., Luo, Q. H., Meng, H. Q., Peng, D. H., Qiu, H. T., Qiu, J., Shen, Y. D., Shi, Y. S., Wang, C. Y., Wang, F., Wang, K., Wang, L., Wang, X., Wang, Y., Wu, X. P., Wu, X. R., Xie, C. M., Xie, G. R., Xie, H. Y., Xie, P., Xu, X. F., Yang, H., Yang, J., Yao, J. S., Yao, S. Q., Yin, Y. Y., Yuan, Y. G., Zhang, A. X., Zhang, H., Zhang, K. R., Zhang, L., Zhang, Z. J., Zhou, R. B., Zhou, Y. T., Zhu, J. J., Zou, C. J., Si, T. M., Zuo, X. N., Zhao, J. P., and Zang, Y. F. (2019). Reduced default mode network functional connectivity in

patients with recurrent major depressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 116:9078–9083.

Yang, S. J., Kim, S. Y., Stewart, R., Kim, J. M., Shin, I. S., Jung, S. W., Lee, M. S., Jeong, S. H., and Jun, T. Y. (2011). Gender differences in 12-week antidepressant treatment outcomes for a naturalistic secondary care cohort: the CRESCEND study. *Psychiatry research*, 189(1):82–90.

Yang, X., Hu, X., Tang, W., Li, B., Yang, Y., Gong, Q., and Huang, X. (2019). Multivariate classification of drug-naive obsessive-compulsive disorder patients and healthy controls by applying an svm to resting-state functional mri data. *BMC Psychiatry*, 19:210.

Yang, X., Peng, Z., Ma, X., Meng, Y., Li, M., Zhang, J., Song, X., Liu, Y., Fan, H., Zhao, L., Deng, W., Li, T., and Ma, X. (2017). Sex differences in the clinical characteristics and brain gray matter volume alterations in unmedicated patients with major depressive disorder. *Scientific Reports*, 7(1).

Ye, T., Peng, J., Nie, B., Gao, J., Liu, J., Li, Y., Wang, G., Ma, X., Li, K., and Shan, B. (2012). Altered functional connectivity of the dorsolateral prefrontal cortex in first-episode patients with major depressive disorder. *European journal of radiology*, 81(12):4035–4040.

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., and Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders: reversal shifting deficits on the Intradimensional/Extradimensional Shift Test correlate with repetitive behaviors. *Autism : the international journal of research and practice*, 13(5):523–538.

Yin, H. H. and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7:464–476.

Yoon, H. K. and Kim, Y. K. (2010). Association between glycogen synthase kinase-3beta gene polymorphisms and major depression and suicidal behavior in a Korean population. *Progress in neuro-psychopharmacology & biological psychiatry*, 34(2):331–334.

Yu, C. and McClellan, J. (2016). Genetics of Substance Use Disorders. *Child and adolescent psychiatric clinics of North America*, 25(3):377–385.

Zandbelt, B. B. and Vink, M. (2010). On the role of the striatum in response inhibition. *PLoS ONE*, 5.

Zapata, A., Minney, V. L., and Shippenberg, T. S. (2010). Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *The Journal of Neuroscience*, 30:15457.

Zhang, A., Yang, C., Li, G., Wang, Y., Liu, P. H., Liu, Z., Sun, N., and Zhang, K. (2020). Functional connectivity of the prefrontal cortex and amygdala is related to depression status in major depressive disorder. *Journal of Affective Disorders*, 274:897–902.

Zhang, W. N., Chang, S. H., Guo, L. Y., Zhang, K. L., and Wang, J. (2013). The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *Journal of affective disorders*, 151:531–539.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20:45–57.

Zhou, H., Sealock, J. M., Sanchez-Roige, S., Clarke, T. K., Levey, D. F., Cheng, Z., Li, B., Polimanti, R., Kember, R. L., Smith, R. V., Thygesen, J. H., Morgan, M. Y., Atkinson, S. R., Thursz, M. R., Nyegaard, M., Mattheisen, M., Børglum, A. D., Johnson, E. C., Justice, A. C., Palmer, A. A., McQuillin, A., Davis, L. K., Edenberg, H. J., Agrawal, A., Kranzler, H. R., and Gelernter, J. (2020). Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nature neuroscience*, 23:809–818.

Zhukovsky, P. (2015a). Bayesian learner.

Zhukovsky, P. (2015b). Reversal learning.

Zhukovsky, P., Puaud, M., Jupp, B., Sala-Bayo, J., Alsiö, J., Xia, J., Searle, L., Morris, Z., Sabir, A., Giuliano, C., Everitt, B. J., Robbins, T. W., and Dalley, J. W. (2019). Withdrawal from escalated cocaine self-administration impairs reversal learning by disrupting the effects of negative feedback on reward exploitation: a behavioral and computational analysis. *Neuropsychopharmacology*, 44:2163–2173.

Zuckerman, M. and Neeb, M. (1979). Sensation seeking and psychopathology. *Psychiatry Research*, 1(3):255–264.