

The flexibly ordered brain

Richard Elliot Daws, MSc, MRes, BSc.

Imperial College London

*The computational, cognitive and clinical neuroimaging laboratory (C3NL),
Department of Brain Sciences.*

*The engineering physical sciences research council centre for doctoral training in
neurotechnology.*

A thesis submitted for the degree of
Doctor of Philosophy, 2021.

*Dedicated to my mother, Alison Jean Daws,
and to my brother-in-law, James Patrick Connolly.*

Acknowledgements

Thank you to my supervisors Dr Adam Hampshrie and Prof. Robert Leech. I first met Adam and Rob during a job interview in 2014. I am grateful to Adam for choosing me to become the first member of his lab and for providing me with experience in many (if not all) stages of the scientific and academic process. This has included the conception of ideas, empirically testing them, drawing inferences and then disseminating the findings via publications and the media. Adam, thank you for the opportunities, insight and support over the years.

To Rob, I am grateful for your sardonic take on everything. Our discussions on the brain, data science and politics have helped me appreciate the bigger picture of working in science. I am very proud to be joining your PhD alumni as each of them have genuinely inspired me.

To Eyal, my unofficial supervisor. Your bombastic, Israeli, caring, no-nonsense, feedback and data-science - *artistry* - has significantly raised my game. Thank you.

I am the first in my family to enter academia. I am grateful to Bristol UWE for accepting me as an undergraduate (despite not making their grade offer), to Prof Mitul Mehta for persuading me to study a Neuroimaging MSc in 2011, my Father for loaning the MSc tuition fee, and to Prof Mairead MacSweeney for providing my first ever job as a scientist in 2013.

To Jen, my wife. We moved to London in 2012 so that I could pursue a career in brains. From gruelling 7 day weeks of studying and working in pubs and supermarkets, to tolerating my many years of abysmal pay and long hours as an RA and PhD student, you have been beyond patient and supportive. Thank you.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Noncommercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Declaration of originality

Experimental data presented in chapters 3 and 4 were collected by Dr Adam Hampshire. Data presented in chapter 5 were collected by Dr Robin Carhart-Harris. I declare that all other work contained within this thesis is my own unless otherwise stated and referenced, this includes all data analyses, visualizations and the text written regarding hypothesis testing and the inferences made. This thesis conforms to the rules and guidelines set out by Imperial College London.

Statement of publications

The data reported in chapters 3 and 4 [1, 2] are available on the bioRxiv preprint server. The data reported in chapter 5 [3] are available on the Research Square preprint server. The up to date publication list is available here: tinyurl.com/rdawsScholar

1. Daws, R.E., Soreq, E., Li, Y., Sandrone, S. and Hampshire, A., 2020. Contrasting hierarchical and multiple-demand accounts of frontal lobe functional organisation during task-switching. bioRxiv. doi: <https://doi.org/10.1101/2020.06.27.175133>
2. Daws, R.E., Scott, G., Soreq, E., Leech, R., Hellyer, P. and Hampshire, A., 2020. Optimisation of brain states and behavioural strategies when learning complex tasks. bioRxiv. doi: <https://doi.org/10.1101/2020.06.17.156570>
3. Daws, R., Timmermann, C., Giribaldi, B., Sexton, J., Wall, M., Erritzoe, D., Roseman, L., Nutt, D. and Carhart-Harris, R., 2021. Decreased brain modularity after psilocybin therapy for depression. Research Square. doi: <https://doi.org/10.21203/rs.3.rs-513323/v1>

Abbreviations

ND	N-dimensions, e.g., a 3D image.
ANOVA	ANalysis Of VAriance
aPFC	Anterior PreFrontal Cortex
ACC	Anterior Cingulate Cortex
ADHD	Attention Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorder
BDI	Beck's Depression Inventory
BDM	Block Decomposition Method
BOLD	Blood Oxygen Level Dependent
CB	Catchment Basin
CBT	Cognitive Behavioural Therapy
CTM	Coding Theorem Method
DA	Dorsal Attention
DLPFC	DorsoLateral PreFrontal Cortex
DMN	Default Mode Network
EN	Executive Network
EPI	Echo Planar Imaging
FC	Functional Connectivity
FD	Framewise Displacement
FDR	False Discovery Rate
FID	Free Induction Decay
fMRI	Functional Magnetic Resonance Imaging
FMRI	Oxford Centre for FMRI of the Brain
FSL	FMRI Software Library
FWHM	Full-Width at Half-Maximum
GE	Gradient Echo
GLM	General Linear Model

gPPI	Generalised PsychoPhysiological Interaction
HCP	Human Connectome Project
HRF	Haemodynamic Response Function
ICA	Independent Component Analysis
ICC	Independent Connected Components
IDED	Intra-Dimensional Extra-Dimensional
IFG	Inferior Frontal Gyrus
IFS	Inferior Frontal Sulcus
IBL	Instruction Based Learning
IQR	InterQuartile Range
LFC	Lateral Frontal Cortex
LI	Limbic network
MDD	Major Depressive Disorder
MD	Multiple-Demand
MPRAGE	Magnetization Prepared Rapid Gradient Echo
MRI	Magnetic Resonance Imaging
OCD	Obsessive Compulsive Disorder
PFC	PreFrontal Cortex
PPI	PsychoPhysiological Interaction
QIDS	Quick Inventory of Depressive Symptomatology
RF	Radio Frequency
ROI	Region Of Interest
RMSE	Root Mean Square Error
RNT	Repetitive Negative Thinking
SEM	Standard Error of the Mean
SD	Standard Deviation
SEM	Standard Error of the Mean
SM	SomatoMotor
SMA	Supplementary Motor Area

SN	Salience Network
SNR	Signal to Noise Ratio
SOS	Self-ordered switching
SPM	Statistical parametric mapping
TBI	Traumatic Brain Injury
TE	Echo Time
TR	Repetition Time
VA	Ventral Attention network
VS	ViSual network
VTs	Voluntary Task-Switching
WCST	Wisconsin Card Sorting Task

Abstract

I investigate the human brain systems involved in the cognitive control of behaviour. Using novel cognitive paradigms and brain imaging, I identify brain systems that support the flexible structuring of behaviour. I then observe how these systems are implicated in patients with depression as they respond to psilocybin therapy. In the first of three experiments, I observe the changes in healthy adult brain activation that are associated with task-switching. This demonstrated that remapping rules introduces a switching-cost to response speed and activates the multiple-demand (MD) network. Critically, switching-costs and MD activation were greater when the rules being remapped were of an abstract and higher-order nature. Going deeper, in the second experiment, I investigate how healthy adult brains mitigate switching-costs by structuring behaviour into efficient routines. I observe that learning to optimise and structure behaviour covaries with changes in MD and default mode network (DMN) activation alongside increases in between-network connectivity. These concurrent behavioural and neural adaptations imply that cognitive demand is minimised when behavioural routines are structured. Indeed, these mechanisms are known to have broad roles in flexibly adapting behaviour and, subsequently, they have been implicated in disorders such as depression. Using these insights, in the third experiment, I examine the neural basis of the treatment response to psilocybin in patients with depression. In two clinical trials, I find that treatment response covaried with global increases in between-network connectivity. Converging functional cartography measures indicated that this global shift in network organisation related to increased dynamic flexibility and integration of the MD and DMN. Together, the findings in this thesis indicate that a 'flexibly ordered brain', the adaptive sequencing of neurocognitive states, is a necessary feature of well-being and for successfully navigating the demands of daily life.

Contents page

Acknowledgements	5
Copyright Declaration	6
Declaration of originality	7
Statement of publications	9
Abbreviations	10
Chapter 1: Introduction	21
1.1 Cognitive control in the changing world	22
1.2 Task-switching experiments	27
1.3 Why is switching costly?	31
1.4 Task-switching in the brain	34
1.5 When cognition is not controlled	40
1.6 Aims and outline of the thesis	44
Chapter 2: Methods	47
2.1 MRI fundamentals	48
2.1.1 Magnets & radio waves	48
2.1.2 Image encoding	49
2.1.3 MRI modalities	50
2.1.4 The BOLD signal	51
2.2 fMRI data handling	52
2.2.1 Noise sources: Head motion & physiological fluctuation	52
2.2.2 fMRI preprocessing	53
2.2.3 Task activation modelling	55
2.2.4 Watershed by immersion clustering of activation maps	57
2.2.5 Connectivity modelling	60
2.2.5.1 Psychophysiological interactions	60
2.2.5.2 Pearson correlation	61
2.2.5.3 Modularity	62
2.2.5.4 Functional cartography	63
2.2.5.5 Dynamic flexibility	64
2.3 Modelling behaviour	65
2.3.1 Reaction time	65
2.3.2 Behavioural complexity	66
2.3.3 Measuring depression severity	67

Chapter 3: Contrasting hierarchical and multiple-demand accounts of frontal lobe functional organisation during task-switching

3.1 Abstract	72
3.2 Introduction	73
3.3 Methods	76
3.3.1 Participants	76
3.3.2 Multilevel switching task design	76
3.3.3 Data acquisition & preprocessing	78
3.3.4 Activation modelling	79
3.3.5 Region of interest definition	80
3.4 Results	81
3.4.1 Behavioural performance relates to switch abstraction	81
3.4.2 Switching activates a broad network of brain regions	84
3.4.3 PFC activation during switching does not reveal a hierarchy	84
3.4.4 Supplemental voxelwise analysis of cognitive demands does not reveal an abstraction hierarchy or selectivity in PFC	86
3.4.5 Cascade and rule abstraction ROIs show no evidence of an abstraction hierarchy in PFC	87
3.5 Discussion	89
3.6 Limitations	94
3.7 Conclusions	95
3.8 Supplemental figures	97

Chapter 4: Optimisation of functional network resources when learning behavioural strategies for performing complex tasks

4.1 Abstract	102
4.2 Introduction	103
4.3 Methods	106
4.3.1 Participants.	106
4.3.2 Self-ordered switch (SOS) paradigms.	106
4.3.3 Behavioural analysis.	111
4.3.5 MRI acquisition and preprocessing.	112
4.3.6 Image data analysis.	113
4.3.7 fMRI activation modelling.	113
4.3.8 Modelling functional connectivity learning effects.	114
4.4 Results	116
4.4.1 SOS costs to reaction time decrease with practise.	116
4.4.2 Task performance relates to structured behavioural strategies	117
4.4.3 SOS activates multiple-demand cortex and reduces default mode network activation.	120

4.4.4 Multiple-demand and default mode network activation changes with learning.	120
4.4.5 Functional connectivity between frontoparietal networks and subcortical regions increases with learning.	124
4.5 Discussion	126
4.6 Limitations	131
4.7 Conclusion	133
4.8 Supplemental materials	134
4.8.1 Repetition bias	134
4.8.2 Number of switches does not correlate with task performance	134
4.8.3 Trial activation modelling	135
4.8.4 Switching activation between study differences	135
4.8.5 Switching greater than trial activation changes over learning stages	135

Chapter 5: Decreased brain modularity after psilocybin therapy for depression

141

5.1 Abstract	142
5.2 Introduction	144
5.3 Materials and Methods	146
5.3.1 Trial overviews	146
5.3.2 Participants	146
5.3.3 Interventions	146
5.3.4 Measuring depression severity	147
5.3.5 Measuring brain network modularity	147
5.3.6 Brain network characteristics	148
5.4 Results	151
Open-label trial	151
5.4.1 Depression severity	151
5.4.2 Decreased brain modularity following psilocybin therapy	151
5.4.3 Decreased modularity predicts long-term clinical outcomes	151
5.4.4 Decreased DMN & increased DMN-frontoparietal FC post-treatment	152
Double-blind randomised controlled trial	155
5.4.5 Psilocybin therapy has greater efficacy than escitalopram for treating depression	155
5.4.6 Increased brain network integration is specific to psilocybin therapy	155
5.4.7 Depression recovery correlates with increased cognitive network flexibility.	156
5.5 Discussion	159
5.6 Limitations	161
5.7 Conclusion	163
5.8 Supplemental materials	164

5.8.1 MRI acquisition	164
5.8.2 MRI preprocessing	164
5.8.3 Functional connectivity (FC)	165
5.8.4 Modularity	165
5.8.5 Functional cartography	167
5.8.6 Dynamic flexibility	167
Chapter 6: General Discussion	171
The flexibly ordered brain: The adaptive sequencing of neurocognitive states.	171
6.1 Synthesis and inference	171
6.2 Future directions	181
6.3 Conclusion	187
References	189

List of figures

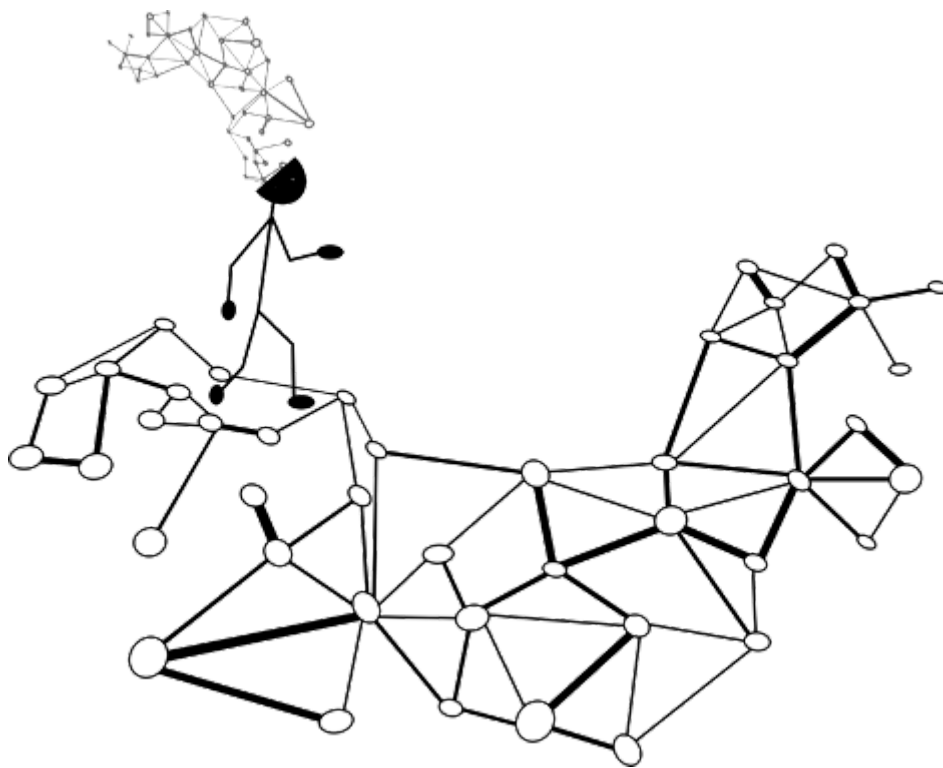
Chapter 1 graphic.....	21
Chapter 2 graphic.....	47
Chapter 3 graphic.....	71
Figure 3.1.....	75
Figure 3.2.....	83
Figure 3.3.....	88
Supplemental Figure 3.1.....	97
Supplemental Figure 3.2.....	98
Chapter 4 graphic.....	101
Figure 4.1.....	110
Figure 4.2.....	119
Figure 4.3.....	123
Figure 4.4.....	125
Supplemental Figure 4.1.....	137
Supplemental Figure 4.2.....	138
Supplemental Figure 4.3.....	139
Chapter 5 graphic.....	141
Figure 5.1.....	149
Figure 5.2.....	150
Figure 5.3.....	153
Figure 5.4.....	157
Figure 5.5.....	158
Chapter 6 graphic.....	171

“Everybody does have a book in them, but in most cases that's where it should stay.”

— **Christopher Hitchens**

Chapter 1: Introduction

This chapter overviews the major themes of this thesis. Here, I overview and interpret existing research before concluding by setting out the aims of the thesis.



The brain navigates between states of mind as it interacts with the world.

1.1 Cognitive control in the changing world

As we interact with the world, the human brain performs a staggering feat of information processing that involves the planning, monitoring and rapid adjustment of behaviour. These unrelenting actions can serve to work out the steps required to go from the brain's current state to a future desired state. For example, if the current state represents "thirst", then achieving the desired future state of "satiation" requires the planning and execution of a set of steps to make this possible.

These mental processes are known as *cognition* and are fundamental to human behaviour, intelligence and well-being (Diamond, 2013). Cognition facilitates the acquisition of knowledge, understanding of the world and involves the processing of information that is delivered through the body's senses (Miller & Cohen, 2001).

Within the cognitive sciences, *cognitive control* is a core area of active research. Cognitive control is a broad term that relates to the process of organising thoughts and behaviour to achieve certain goals (D. E. Meyer & Kieras, 1997; Miller, 2000; Miller & Cohen, 2001; Norman & Shallice, 1986). Naturally, cognitive control and the flexible adaptation of behaviour depend on the marshalling of the brain's neurophysiological resources (Cañas et al., 2003; Ionescu, 2012; Koch et al., 2018).

Cognitive control is necessary for using new information to update a strategy (Chamberlain et al., 2008; D. A. Grant & Berg, 1948; Hampshire & Owen, 2006) or during *task-switching* where attention shifts from one task to another (Jersild, 1927; Monsell, 2003). The research presented in this thesis uses novel task-switching paradigms and the century of task-switching research is overviewed in section 1.2 Task-switching experiments.

It is rare for humans to behave in a truly random or solely reflexive manner (Gauvrit et al., 2017; Kennedy, 2012). Even during creative expression, improvisation or

casual play, there will be overarching aims that require some level of cognitive control and the structuring of behaviour (Borman & Kurdek, 1987).

However, as will be discussed in section 1.3 “Why is switching costly?”, adapting behaviour and task-switching comes with a cost. Switching takes time and taxes the brain’s metabolic resources. This is empirically demonstrated in Chapter 3; however, in Chapter 4, I present novel evidence that the brain may overcome this cost, and make returns on the investment, by rapidly optimising behavioural routines toward efficient structures. Structuring behaviour may facilitate automation (performing without the need for cognitive control / conscious oversight) and learning for future use.

In recent decades, there has been a proliferation of brain imaging technologies that have provided the opportunity to gain insight into how the human brain supports cognition. Strikingly, it is now possible to capture the sequences of neurocognitive states that are indicative of healthy cognitive control and brain functioning. Subsequently, these techniques can offer real insight into thought disorders that may underlie psychiatric symptoms that negatively impact health and well-being.

Being able to observe these actions of the brain, *in vivo*, is remarkable. In this thesis, I use functional magnetic resonance imaging (fMRI) to observe changes in brain function during the adaptation of behaviour in healthy adults (Chapters 3 and 4) and changes in brain function as patients with depression respond to a psilocybin therapy clinical trial intervention (Chapter 5). fMRI is sensitive to the brain’s haemodynamics and is, therefore, an indirect measure of neuronal activity. An overview of how fMRI data can be acquired and modelled is overviewed in Chapter 2.

fMRI research is still in its infancy, or perhaps *just* entering its adolescence. fMRI has the potential to generate real insight, however, interpreting the data is challenging. The early days of fMRI research (1990s-2000s) were largely dominated by a zeitgeist of modular localisation where researchers sought to “map” specific

cognitive functions to isolated activations of specific brain regions. For example, some have proposed that task-switching can be specifically mapped to the frontal poles (Koechlin et al., 1999) or that response inhibition can be mapped to the right inferior frontal gyrus (IFG) (Aron et al., 2004; Rubia et al., 2003).

Modular frameworks of the brain are fruitful, and some amount of regional specialisation is well-established (e.g., the occipital cortex tends to process visual information in sighted people, see Gizewski et al. (2003)). However, the ambition to assign every cognitive or psychological function to unique portions of brain tissue is fundamentally flawed.

In contrast, most contemporary views consider the brain as a spatially distributed functional network (S. M. Smith et al., 2009) that dynamically adapts to the current context, cognitive demand and with learning (Betzel & Bassett, 2017). Dynamic network perspectives are supported by a mass of evidence that also undermines the modular accounts of brain function. This includes repeated observations that individual brain regions respond to diverse cognitive contexts (Erika-Florence et al., 2014; Fedorenko et al., 2013; Hampshire et al., 2010) and the observations that brain regions are embedded within spatially distributed dynamic networks (Braga et al., 2013; Leech et al., 2012; Margulies et al., 2016; S. M. Smith et al., 2009; Vidaurre et al., 2017). Together, these findings suggest that it is unlikely that brain regions do, or can, operate in isolation.

The brain's networks of information processors appear to dynamically adapt to the current context, cognitive demand and with learning (Betzel, 2020; Betzel & Bassett, 2017; Hampshire et al., 2016; Matthews & Hampshire, 2016; Parkin et al., 2015). Performing tasks have been associated with distinct and changing expressions of network connectivity (Finc et al., 2020; Khambhati et al., 2018) and, subsequently, broad sets of cognitive task states can be reliably predicted via differences in the whole-brain connectivity network features that they evoke (Soreq et al., 2019, 2021).

This richer and more complex view of brain function has arrived, in part, due to advances in MRI hardware, experimental design and the incorporation of multivariate statistics and network science. Subsequently, there now is the potential, and ambition, to characterise the neurophysiological basis of the sequence of functional states (Vidaurre et al., 2017) that correspond to mental phenomena such as mind-wandering (Smallwood et al., 2021) or the encoding and manipulation of information (Finc et al., 2020; Khambhati et al., 2018; Soreq et al., 2019).

During resting-state fMRI, attention is thought to be “unconstrained”. No explicit task is assigned to the subject; however, attention will still be externally constrained by the context of lying down inside of a noisy machine and, in some cases, by the instructions to stare at a fixation cross. Typically, functional connectivity (FC) analyses of resting-state fMRI data are static. They define time-averaged statistical dependencies between brain region timecourses. fMRI FC does require many transformations and abstractions of the recorded imaging data, which are themselves an indirect measurement of neuronal activity. However, static FC measures during resting-state can offer a high level of precision, predictive power and mechanistic insight. This includes the reliably precise mapping of interdigitated networks in individuals (Braga & Buckner, 2017), prediction of sociodemographic information (S. M. Smith et al., 2015), or individual brain’s via their connectomic ‘fingerprints’ (Finn et al., 2015). Similarly, task-based FC can also be used to accurately predict the particular cognitive task that a brain is currently performing (Cole, Reynolds, et al., 2013; Soreq et al., 2019, 2021) and even how well the task is being performed (Rosenberg et al., 2016; Soreq et al., 2021)

The absence of an explicit task in resting-state fMRI opens up many opportunities to investigate clinical populations who may struggle to perform a task due to cognitive impairment. Moreover, multiple resting-state measurements can be completed throughout a clinical trial without introducing task-learning confounds (Zhang et al., 2021). Many studies have used static FC measures to characterise univariate differences between patients and healthy controls (Matthews &

Hampshire, 2016; Parkes et al., 2021). Static FC also has the potential for generating multivariate feature spaces that can be used to predict clinical status (Lanka et al., 2020), symptom severity (Madhavan et al., 2019) and even the onset of neurodegeneration in Huntington's disease (Mason et al., 2018) or biomarkers of dementia (Teipel et al., 2015).

Pertinently, there are open questions regarding psychiatric disorders, such as depression or schizophrenia, where patients can exhibit complex symptomatic timecourses (Beard et al., 2016). In these populations, symptom severity can fluctuate and be significantly impacted by adverse events (Keller et al., 2007) or clinical intervention (Carhart-Harris et al., 2016, 2021). It is important to therefore understand the sequences of neurocognitive states that reflect particular symptoms and predict how they will change over time. This could include the tendency to ruminate and engage in repetitive negative thinking (RNT), for example (Lydon-Staley et al., 2019).

Subsequently, dynamic analyses of fMRI FC may represent an important avenue for understanding psychiatric disorders. Dynamic FC analyses can characterise, for example, the tendency for the brain to dwell or transition between states (Parkes et al., 2021). Notably, the spontaneous fluctuations in connectivity patterns (states) during rest have been demonstrated to show consistent temporal structure (Chang & Glover, 2010; Roberts et al., 2019; Vidaurre et al., 2017). Moreover, by considering the range of states expressed, and the transition rates between them, links between reduced dynamic flexibility and the severity of autism spectrum disorder (ASD) (Watanabe & Rees, 2017) and depressive symptoms (Goodman et al., 2021) have been discovered. These examples may indicate that the tendency to engage in repetitive stereotyped behaviours or engage in maladaptive rumination are linked to 'rigid' brain dynamics that are characterised by reduced state transitions.

As will be discussed in section 1.5 When cognition is not controlled, healthy brain functioning may rest on a curious mixture of stability and flexibility (Barbey, 2018; Bassett et al., 2018). Brain state stability may be important for completing tasks and

being resistant to destabilizing distractors (Finc et al., 2020) but flexibility may be important for adapting to changing environmental demands (Ezaki et al., 2020; Ferguson et al., 2017). Quantifying and characterising the balance between these dynamic properties may provide an important avenue for understanding clinical populations, symptom severity and response to treatment. In Chapter 5, I examine this notion further by capturing FC changes in patients with depression as they respond to psilocybin therapy treatment.

The remaining sections of this introduction chapter go further into the insights into the brain and behaviour that have been gained from task-switching paradigms. The relationship between these mechanisms for guiding and optimising behaviour are then discussed in relation to the symptomatic profiles associated with neuropsychiatric disorders. Finally, the aims and outlines of the thesis are presented.

1.2 Task-switching experiments

Task-switching refers to the switching of attention between distinct cognitive processes (Monsell, 2003). This ability is necessary for engaging with multiple competing demands or to react to changes in the environment. A century has passed since the early task-switching paradigms were developed (Culler, 1912; Jersild, 1927). Since then, a rich literature has emerged which has motivated a proliferation of experimental paradigms (Kiesel et al., 2010) and conceptual interpretations of the experimental findings (Vandierendonck et al., 2010).

Task-switching research may have generated the most replicated finding in the cognitive sciences: *Switching attention between tasks comes with a cost* (Monsell, 2003; G. Wylie & Allport, 2000). Researchers will typically compare the average difference in RT between consecutive trials of the same task (repeat-trials) and RT on the trial(s) that immediately follows a switch (switch-trials). Switching-costs capture a transient drop in behavioural performance following a switch between tasks. Typically, a drop in response speed is observed but response errors can also

occur. Switching-cost magnitudes will vary depending on the experimental manipulation but tend to be in the order of 100's of milliseconds (Kiesel et al., 2010; Koch et al., 2018).

Task-switching paradigms have evolved and diversified over 100 years. The early paradigms defined switching-costs by contrasting the performance between single-task and multiple-task blocks (Jersild, 1927; Spector & Biederman, 1976). For example, a single-task block could involve multiplying two numbers on each trial. A multiple-task block would then involve alternating between adding or multiplying numbers on consecutive trials. Jersild (1927) found that response times during the mixed-task blocks were slower than the single-task blocks and he attributed this effect to time-consuming top-down reconfiguration processes.

While this was a reasonable interpretation, Jersild's experimental descendants criticised the single/mixed-task paradigms due to the imbalance in working memory demand between the block types (Logan, 2007; R. D. Rogers & Monsell, 1995; Spector & Biederman, 1976; Sudevan & Taylor, 1987). The mixed-task block requires instructions from multiple tasks to be simultaneously held in mind and is, therefore, more demanding than the single-task blocks. Subsequently, the differences in performance become challenging to interpret as they likely represent mixtures of switching-costs and cognitive load.

In response to this, paradigms were developed to compare repeat-trials and switch-trials within the same block in an attempt to balance working memory demands. An early example of this was attempted using predictable sequences of tasks (R. D. Rogers & Monsell, 1995; Spector & Biederman, 1976) or by cuing the task before each trial (Meiran, 1996; Sudevan & Taylor, 1987). These paradigms are similar to the mixed-task blocks but combined sequences of repeat-trials with switch trials within blocks (e.g., ABABABBAA).

Despite this, the predictability of the switches inherent in these designs may limit the ability to measure the 'true' switch-cost due to potential preparatory

mechanisms. This is particularly relevant as switch-costs reliably decrease as the amount of time between trials increases (Arrington & Logan, 2004, 2005; Meiran, 1996; R. D. Rogers & Monsell, 1995). To mitigate preparatory effects, the predictability of upcoming switches has been manipulated using random task-cueing paradigms (Meiran, 1996; Sudevan & Taylor, 1987). These paradigms sequence tasks in a random order and cue each trial (Meiran, 1996; Miyake et al., 2004) or block of trials (Altmann & Gray, 2008; Barceló et al., 2007). In practice, these sequences are likely pseudo-randomised to balance the number of repeat and switch trials.

The block-cueing, or intermittent instruction, paradigms have the additional benefit of measuring variations in performance related to restarting the same task, where no switch has occurred between blocks. Tellingly, “restart-costs”, where no switching has occurred, present a similar magnitude to the switch-cost (Gopher et al., 2000; Waszak et al., 2003). More broadly, these block-cueing paradigms share many qualities with those paradigms used to assess novel instruction based learning (IBL) (Cole, Laurent, et al., 2013; Hampshire et al., 2016, 2019). NB:- On a conceptual level, it could be argued that starting or restarting the same task is very similar to switching between tasks. In both cases, a shift in the attentional state is required. There may be, therefore, opportunities to study switching using single-task fMRI block paradigms by examining the rest-to-task transitions (Scott et al., 2015).

A general criticism of most task-switching paradigms is that they lack endogenous acts of cognitive control. That is to say, subjects aren’t actively coordinating their behaviour but “simply” reacting to the exogenous events that are imposed by the experiment. These tasks do tap an important facet of cognitive control because humans are required to react to events in the environment. However, greater ecological validity may be achieved if paradigms were designed to assess acts of endogenous cognitive control where participants are deciding on how best to coordinate their behaviour (this possibility is explored in Chapter 4).

A step toward this front came from the more recently developed voluntary task-switching (VTS) paradigms from Arrington and Logan (2004). Here, participants are instructed to randomly pick one of two tasks to perform on each trial but to aim for an overall equal number of responses to each task. Interestingly, VTS switch-costs were smaller than some previous task-cueing experiments but costs were still in the 100's of ms range and were also mitigated by increased interval duration (Arrington & Logan, 2004, 2005).

Others have interpreted smaller VTS costs as evidence in favour of switching-costs representing a context-dependent mixture of top-down and bottom-up components (Koch et al., 2018; Vandierendonck et al., 2010). The decreased VTS cost is likely to be evidence of preparatory behaviours; however, despite the instruction to randomly pick tasks, individuals are still likely to plan their switches in advance and would develop non-random strategies to minimize cognitive load. Indicative of this, it is particularly revealing that subjects performing VTS did exhibit a repetition-bias, and opted to repeat the same task more often than task-switching (Arrington & Logan, 2004).

Despite the broad range of experimental designs, switching-costs have been consistently reported between the 100 - 900 ms range. There have been careful efforts to design paradigms that provide 'pure' measures of switching-costs by controlling working memory demands, predictability and controllability. Despite the decades of research that has produced a rich literature, many open questions remain. As will be discussed further in section 1.3 "Why is switching costly?", there are ongoing debates and competing theories regarding the basis of switching-costs.

1.3 Why is switching costly?

The impressive replication of switching-costs is noteworthy. However, despite a large literature exploiting a range of experimental paradigms, the neurocognitive mechanisms responsible for this phenomenon continue to be keenly debated.

As a slight aside, there is some evidence that Macaque monkeys may not exhibit switching-costs. Stoet and Snyder (2003) showed evidence from 2 monkeys that task-switching did not come with a measurable cost to performance. This begs the question, why are humans unable to eliminate switching-costs after extensive training (Berryhill & Hughes, 2009) and even after performing tens of thousands of trials (Stoet & Snyder, 2007)?

As is common in the cognitive sciences, the switching-cost debate revolves around the opposing views of *top-down* reconfiguration (Meiran, 1996; R. D. Rogers & Monsell, 1995) vs. *bottom-up* interference (Allport et al., 1994). Top-down proponents argue that switching-costs occur due to the time required for the active reconfiguration and preparation of the brain's resources (Arrington & Logan, 2004, 2005; Meiran, 1996; R. D. Rogers & Monsell, 1995; Rubinstein et al., 2001). The cognitive control mechanisms proposed to be associated with task-switching includes the active inhibition of the outgoing task state (Mayr, 2002; Mayr et al., 2006), memory retrieval (Mayr & Kliegl, 2000) and updates to working memory (E. E. Smith et al., 2001) during the activation of the incoming task state.

Evidence in support of the top-down interpretation comes from the observations that switching-costs occur during VTS as this indicates an endogenous component (Arrington & Logan, 2004, 2005). Furthermore, switching-cost magnitudes can diminish with learning (Yeung & Monsell, 2003b) and with increasing preparation time during the switch (Meiran, 1996; R. D. Rogers & Monsell, 1995). Moreover, subjects are less likely to VTS under high cognitive load conditions (Demanet et al., 2010).

While these findings indicate top-down components of switching-costs, they do not rule out the possibility of bottom-up contributions. Long preparation times between tasks (R. D. Rogers & Monsell, 1995) and even extensive training on task-switching paradigms (Berryhill & Hughes, 2009; Kramer et al., 1999; Liefvooghe et al., 2010; Stoet & Snyder, 2007) will mitigate switching-costs but not eliminate them, a residual cost will persist.

Proponents of bottom-up interpretations of switch-costs, propose that human cognitive architecture is limited by an 'inertia' where competing task representations can passively interfere with each other (Allport et al., 1994; Allport & Wylie, 1999; Waszak & Hommel, 2007). Essentially, they argue that switching-costs result from remapping failures. The bottom-up perspective is supported by the aforementioned residual switch-cost, indicative of passive task representation decay (Allport et al., 1994) and also by "asymmetrical" switching-costs where switching to the well-practised task generates a larger cost than switching to the novel task (Allport et al., 1994; Jersild, 1927; Yeung & Monsell, 2003b, 2003a).

Strict theoretical dichotomies can be misleading. In this case, there is evidence that supports either of these opposing top-down and bottom-up accounts of switching-costs. The contemporary view contextualises switching-costs as mixtures of components (Vandierendonck et al., 2010), e.g., active control mechanisms and passive interference (Arrington, 2008; Demanet et al., 2010; Meiran et al., 2000; Ruthruff et al., 2001). This view reconciles both accounts and suggests that switching-costs are complex and depend on the nature of paradigm (i.e., trial tempo, interval length, task familiarity, difficulty, feedback etc) and learning (Koch et al., 2018; Vandierendonck et al., 2010).

The evidence of many contributors to switching-costs suggests that it may be infeasible to consider switching-costs as one particular phenomenon. Rather, it may be pertinent to consider the contextually dependent mixtures of components that drive switching-costs and how these might vary with learning and feedback. There

is, therefore, scope to delineate kinds or modes of attentional switches as they may depend on distinct mechanisms (Hampshire & Owen, 2006). Moreover, stratifying attentional switching in this manner, and identifying component processes, may provide precise clinical markers of cognitive impairment (Gruszka et al., 2017).

1.4 Task-switching in the brain

This section considers the evidence for the neural mechanisms associated with cognitive control and task-switching. The large body of cognitive and behavioural research has motivated many brain imaging studies designed to understand the organisation of the brain and its role in coordinating thought and behaviour.

There is an overwhelming amount of evidence that associates the human frontal lobes with higher-order cognitive abilities, such as task-switching. Before conventional neuroimaging was developed, decades of neuropsychological research associated damage to the frontal lobe's tissues with impairments to various cognitive abilities, a.k.a. dysexecutive syndromes (Baddeley et al., 1997; Burgess et al., 2000; Duncan, 1986; Duncan et al., 1995; Milner, 1963; Nelson, 1976; Owen et al., 1990).

Neuropsychological research can deliver some specificity. For example, damage to dorsolateral prefrontal cortex (the top-and-side surfaces of the brain's frontal cortex) is associated with deficits in planning ability (Burgess et al., 2000), cognitive control (Glascher et al., 2012; Stuss et al., 2001) and cognitive flexibility (Grattan et al., 1994; Shallice & Burgess, 1991). Whereas damage to the orbitofrontal and medial prefrontal cortices (front of the brain, behind and above the eyes) is associated with poor emotional regulation, impulsivity and impaired learning from feedback (Angrilli et al., 1999; Berlin, 2004).

The advent of brain imaging, and fMRI in particular, allowed for a proliferation of neurocognitive experiments that were designed to reveal how the brain supports the mental faculties (Poldrack, 2012). Converging with prior neuropsychology research, in vivo brain imaging experiments provided further support for the frontal lobe's critical role in higher-order cognitive abilities (Duncan, 2001; Koechlin et al., 2003; Miller, 2000; Miller & Cohen, 2001; Ramnani & Owen, 2004).

The early task-switching experiments suggest that the precuneus, anterior insula, supplementary motor area (SMA), dorsolateral prefrontal cortex (DLPFC), inferior frontal sulcus (IFS) and frontal poles increase in activation during task-switching (Braver et al., 2003; Dove et al., 2000; Koechlin et al., 1999). Even though this is a broad set of distributed brain regions, spanning multiple lobes, a large focus was placed on the anterior prefrontal cortex (aPFC) (Koechlin et al., 1999; Ramnani & Owen, 2004).

The success of previous neuropsychological lesion-symptom mapping carried over into the new paradigm of brain imaging and function-region mapping. The aim of which was to find specific mappings between psychological or cognitive constructs to the activities of specific brain regions. It was typical to carefully design experimental paradigms that allowed controlled contrasts between activation patterns. This endeavour was considered successful if the contrast isolated an activation difference to specific brain regions.

Under a modular zeitgeist, the early task-switching fMRI experiments did much to crown the aPFC as *the* functional locale for task-switching (Koechlin et al., 1999). These findings were partly responsible for the precipitation of models that sought to describe the functional organization of the frontal lobe's as hierarchical.

The information cascade model (Koechlin et al., 2003) and abstraction hierarchy model (Badre & D'Esposito, 2007) propose that the functional properties of the frontal lobes vary along an anterior-posterior axis with anterior regions subserving processes of an increasingly abstract order. Moreover, aPFC regions are proposed to exert control over the activities of their posterior neighbours in a hierarchical manner until the signal reaches the precentral gyrus (most posterior portion of the frontal lobes). At which point, a concrete motor action is enacted. Subsequently, these hierarchical models may predict that the aPFC would selectively activate during task-switching (as task-switching is an example of an abstract higher-order process). This proposed organisation of PFC is considered to reflect the hierarchical nature of cognition itself (Botvinick, 2008).

Several studies have used evidence from fMRI task contrasts to support a rostral-caudal organisation of the PFC (Badre, 2008; Badre & D'Esposito, 2007; Koechlin et al., 2003; Koechlin & Hyafil, 2007). These fMRI findings are also supported by recordings of cortical electrophysiology indicating that tasks with greater abstract control demands evoke activity in increasingly anterior frontal regions (Voytek et al., 2015).

Indeed, the complex white matter anatomy of the frontal lobes (Catani et al., 2012; Rojkova et al., 2016) facilitates the occurrence of transmodal cortical regions that may integrate multiple sensory inputs from the rest of the brain (Margulies et al., 2016; Margulies & Smallwood, 2017; Miller & Cohen, 2001). Furthermore, imaging and histological markers have demonstrated anatomical hierarchies that parallel the functional organisation of the frontal lobes with rostral regions exhibiting thicker cortices, and denser cell body counts, relative to caudal regions (Thiebaut de Schotten et al., 2016).

However, the contemporary views of brain function do not fully support the notion of functional-anatomical one-to-one mappings (Duncan, 2013; Fedorenko et al., 2013; Hampshire & Sharp, 2015). In particular, dynamic network accounts present a notable challenge to the hierarchical models of PFC organisation. It could be argued that hierarchical models depend on findings generated from overly narrow experimental hypotheses and overly specific interpretations that map specific brain regions to specific functions (Lorenz et al., 2017).

This one-to-one mapping issue is not specific to the brain imaging of task-switching. It is also clear in the motor inhibition literature, for example. A consistent finding during response inhibition is increased activation of the right inferior frontal gyrus (rIFG). Subsequently, the rIFG has been denoted the module, or functional locale of 'inhibition' (Aron et al., 2004; Aron & Poldrack, 2006; Rubia et al., 2003). However, the rIFG activates during a broad array of contexts that are unlikely to involve inhibitory action (Erika-Florence et al., 2014; Hampshire et al., 2007,

2012). Moreover, frontal regions, including rIFG, tend to form distributed functional networks with other brain regions that co-activated in response to various cognitive demands (Erika-Florence et al., 2014; Fedorenko et al., 2013; Hampshire et al., 2016; Parkin et al., 2015). The dynamic network perspective of inhibition undermines previously held beliefs regarding one-to-one mappings and recasts the activation of the rIFG within a broader context of cognitive control.

More broadly, the contemporary network views are at odds with the modular perspective of brain function. The network view depicts the brain's functioning as a set of spatially distributed networks (S. M. Smith et al., 2009) that are hierarchically organised (Margulies et al., 2016; Vidaurre et al., 2017) and dynamically adapt to support present cognitive demand (Bassett et al., 2011; Betzel & Bassett, 2017; Cole, Reynolds, et al., 2013; Hampshire et al., 2016; Khambhati et al., 2018; Parkin et al., 2015; Soreq et al., 2021; G. R. Wylie et al., 2004).

Indeed, Jon Duncan's multiple-demand (MD) account of cognition and the frontal lobes not only reconciles the contemporary network view of brain function it also provides a plausible alternative to the hierarchical models of the frontal lobes (Duncan, 2013). The MD account posits a distributed neurocognitive resource that spans the lateral frontoparietal cortices, the anterior cingulate cortex (ACC), insula and aPFC (Fedorenko et al., 2013). In response to cognitive demand, MD regions rapidly form network coalitions that can flexibly recruit more resources as cognitive demand increases. Critically, these networks have broad spatial profiles and their connectivity patterns offer impressive predictive power on which is being performed and how well it is being performed (Rosenberg et al., 2016; Soreq et al., 2021)

The demand imposed by cognitive tasks is often associated with the fMRI activation of the MD regions. As has been suggested, additional resources will be flexibly recruited in response to increasing cognitive demand (Duncan, 2013). This has been observed as increased MD activation and an increase in the spatial extent of the activation pattern, with an emphasis of increasingly anterior PFC regions as demand increases (Crittenden & Duncan, 2014; Hampshire et al., 2016; Jiang & Kanwisher,

2003; Parkin et al., 2015). Note, this is in contrast to the hierarchical models that may predict *selective* activation of increasingly anterior modules as demand increases.

There is evidence that the MD system can be reliably sub-divided into two functional networks based on clusters of intrinsic FC (Dosenbach et al., 2007). These dual-networks are hypothesised to support distinct aspects of task implementation, maintenance and error handling (Dosenbach et al., 2006, 2008). Specifically, a network of dorsolateral frontoparietal (FP) regions rapidly adapt to subserve “moment-to-moment” adaptive control and contain fine-grained task representations (Crittenden et al., 2016; Woolgar, Hampshire, et al., 2011; Woolgar, Thompson, et al., 2011). The cingulo-opercular (CO) network of the dorsomedial frontal cortex, insula and frontal operculum regions exhibit sustained activation during tasks and has been suggested to maintain coarse-grained task representations of the task over longer timescales (Dosenbach et al., 2006).

Consistent with this model, the FP and CO show distinct dynamics as novel rules are learnt and implemented during instruction based learning (IBL). Here, both CO and FP activation increased during rule encoding (Mohr et al., 2016). However, IBL rules tend to be simple and are subject to rapid automation. As this occurs, there is a concurrent increase in response speed as FP activation decays and CO activation is sustained over time (Hampshire et al., 2016, 2019; Mohr et al., 2016).

The evidence for the MD dual-networks and the IBL findings are particularly revealing. Cognitive control allows for behavioural routines to be rapidly compiled and this is associated with a widespread reconfiguration of the brain’s global functional state (Hampshire et al., 2019). However, once the behavioural routine is established, cognitive demand diminishes and lateral FP activation decays. Substantiating this, analyses of working memory change detection paradigms indicate widespread activation increases during encoding are followed by maintenance states that are defined by patterns of low-activation and specific

patterns of increased functional connectivity which depend on items being held in mind (Soreq et al., 2019).

Similarly, task-switching involves switching between states of attention and this involves widespread shifts in the brain's functional organisation. For example, transitions between low and high cognitive demands involve shifts in the integration and segregation between many functional systems (Khambhati et al., 2018) and transitions toward increased demands require more 'energy', in terms of network control (Braun et al., 2019). These changes in global connectivity (Khambhati et al., 2018) map onto individual differences in switching-costs and specific connectivity increases between the FP and CO networks can relate to decreases in switching-costs (Yin et al., 2018).

To summarise, the behavioural costs associated with task-switching are accompanied by distributed activation increases, and particularly in the lateral frontoparietal regions associated with the MD system. This suggests that task-switching does present an increase in cognitive demand. However, drawing on the wider cognitive control research, similar findings are observed during other contexts that involve adaptations to the attentional state. The dynamics associated with neurocognitive flexibility are associated with widespread shifts in the brain's functional organisation. Efficiently adapting to changing environmental demands and optimising behaviour are likely to be useful in minimising cognitive demand.

The structuring of behaviour is an important human characteristic. There are open questions regarding the relationship between neurocognitive flexibility and the structuring of behaviour. fMRI paradigms, in particular, typically control when events occur and this limits the extent to which individuals can naturally behave. How efficient behaviours are structured presents an opportunity to gain neurocognitive insights that may bear decent ecological validity and clinical insight.

1.5 When cognition is not controlled

The complexity of the human brain's structure and functional organisation is its defining feature. The natural heterogeneity in the development of the brain (Deoni et al., 2011; Finn et al., 2015) is likely to predispose individuals to vary in cognitive ability (Rosenberg et al., 2016; S. M. Smith et al., 2015; Soreq et al., 2021) and to clinically disorders (Crossley et al., 2014; Parkes et al., 2021).

Cognition and cognitive control are critical for health and well-being (Diamond, 2013) and cognitive deficits are a pervasive feature of psychiatric and neurological disorders (Millan et al., 2012). Brain disorders can be examined at many levels, from the genetic, cellular and pharmacological (Duman, 1997; Mullins & Lewis, 2017), through to the levels of fMRI regional activation, connectomics and behaviour (Grahek et al., 2018; Grattan et al., 1994; Hampshire et al., 2013; Harlalka et al., 2019). However, an emerging view is that the connectomic and systems-level measures of brain functioning can provide precise mechanistic explanations of disorders and even individual differences in clinical severity (Bassett et al., 2018; Matthews & Hampshire, 2016; Parkes et al., 2020, 2021).

In healthy adults, natural fluctuations in mood and well-being (Pierson et al., 2021; Simbula, 2010) could be used to gain insight into brain disorders. For example, analyses of the 'myConnectome' project, a dense-sampling fMRI study of an individual (Poldrack et al., 2015), indicates that increases in positive mood were accompanied by periods of increased integration and flexibility of the brain's networks (Betz et al., 2017; Mirchi et al., 2019). A similar relationship of between-network integration and well-being has been observed at scale whereby increased integration between higher-order networks and subcortical regions has been associated with positive markers of well-being (S. M. Smith et al., 2015).

Patients with major depressive disorders (MDD) exhibit detrimentally rigid thought patterns and negative biases regarding oneself and future prospects (Beck & Clark,

1988; Lyons & Carhart-Harris, 2018). Akin to the family of ‘internalising’ disorders such as anxiety, obsessive compulsive disorders (OCD) and addiction, patients with depression commonly present with poor cognitive flexibility (Rock et al., 2014). Indeed, one of the aims of cognitive behavioural therapy (CBT) is to train patients to recognize, interrupt and challenge negative thought patterns (Watkins et al., 2011).

Relative to healthy adults, MDD patients exhibit deficits during contexts that require task-switching (Meiran et al., 2011) and rule updating during the Wisconsin Card Sorting Task (WCST) or the Intra-Dimensional Extra-Dimensional Set Shifting (IDED) task (Taylor Tavares et al., 2007). These examples of cognitive rigidity have serious implications for the perseverative thinking and maladaptive rumination aspects of depression (Watkins, 2008). Perniciously, cognitive rigidity will also undermine the ability to update beliefs with positive events that may challenge a patient's negative view of themselves (Beck & Clark, 1988).

Resonant with depressive phenomenology, brain imaging studies suggest that the default mode network (DMN) is hyperactive in patients (Berman et al., 2011; Hamilton et al., 2011; Zhou et al., 2020). This system is strongly associated with introspection and inwardly focused self-referential processing (Andrews-Hanna et al., 2014; Margulies et al., 2016) and these psychological functions are overactive in depression (Koster et al., 2011; Lyons & Carhart-Harris, 2018).

Furthermore, the dynamic interactions between the brain's functional networks appear to be abnormally modular (segregated) and rigid over time in depression patients (Zheng et al., 2018). Lateral frontoparietal within-network connectivity is weakened in MDD patients (Ellard et al., 2018; Kaiser et al., 2015). Moreover, the dynamic interactions between the DMN and lateral frontoparietal networks are abnormal, with both greater DMN dominance (Hamilton et al., 2011) and weaker interactions between higher-order networks (Goodman et al., 2021) correlating with worse depressive symptoms. The abnormal network features of depression (Manoliu et al., 2014) are important because system-level dynamics may moderate the tendency to ruminate (Lydon-Staley et al., 2019) and support attention switching

between internal thought and the external world (Hasenkamp et al., 2012; Turnbull et al., 2020).

The consistency between depression symptomatology and abnormal brain function has been combined in models such as the non-reward attractor theory (Rolls, 2016) and disengagement theory (Koster et al., 2011) of depression. These models paint depression as resulting from a tendency to dwell in certain neurocognitive states and with difficulty in switching between attentional states (Holtzheimer & Mayberg, 2011). Indeed, dynamic analyses have indicated a lower probability of transitions between functional states in MDD patients, compared to healthy controls (Yao et al., 2019; Zhao et al., 2021), and these dwell states tend to feature the DMN system (Goodman et al., 2021).

In stark contrast to the salient features of depression, the acute action of psychedelics on brain function is characterised by widespread increases in FC between regions and a ‘disintegration’ of typical intrinsic functional network configurations (Atasoy et al., 2018; L.-D. Lord et al., 2019; Luppi et al., 2021; Roseman et al., 2014). It is thought that psychedelics, such as psilocybin, are agonists of the serotonin 5HT-2A receptor (Vollenweider et al., 1998) and this dysregulates brain function in a way that broadens the functional ‘state-space’; that is to say the set of functional states or modes of operation that the brain traverses.

Akin to a topographical landscape, with peaks and troughs, the energy landscape of the brain’s functional states (Ezaki et al., 2020; Watanabe & Rees, 2017) has certain paths or transitions that are more likely, or require less ‘energy’ to achieve. It is thought that the acute action of psychedelics effectively flattens the energy landscape and therefore reduces the energy threshold required to make transitions between brain states (Carhart-Harris & Friston, 2019). Subsequently, atypical thought patterns may become more likely, or, more precisely, all possible state trajectories become closer to equiprobable.

The acute state of lability may be the neurobiological basis of the creative and novel insights classically associated with psychedelics (Carhart-Harris & Friston, 2019; Roseman et al., 2018; Uthaug et al., 2018). Tellingly, the 5HT-2A receptor is most densely expressed in those higher-order cortical networks associated with cognitive control, flexibility and, as discussed, these same systems can function abnormally in patients with depression (Beliveau et al., 2020).

Subsequently, psychedelics present an intriguing therapeutic potential, particularly for patients with ‘internalising’ disorders that are characterised by cognitive rigidity. Acutely broadening the brain’s state-space and increasing network integration (L.-D. Lord et al., 2019; Luppi et al., 2021), within a therapeutic context, may facilitate beneficial updates, or perspective shifts, regarding oneself and one’s prospects (Andersen et al., 2021; Carhart-Harris & Friston, 2019). Moreover, successful interventions may benefit patients’ long-term well-being if these potentially positive shifts in brain functioning can carry over, sub-acutely (Roseman et al., 2018).

In recent years, a small but increasing number of clinical trials have demonstrated an impressive efficacy of psychedelic therapy in improving the well-being of patients with depression (Carhart-Harris et al., 2016, 2021; Palhano-Fontes et al., 2019), anxiety (Griffiths et al., 2016; Grob et al., 2011), obsessive compulsive disorder (OCD) (Moreno et al., 2006), eating disorders (Spriggs et al., 2020), and substance abuse issues (Garcia-Romeu et al., 2014; Krebs & Johansen, 2012; Winkelman, 2015). However, the neurobiological basis of this potential efficacy is incompletely understood. Therefore, generating markers of the brain’s dynamics that are sensitive to changes in patient well-being presents an important opportunity to develop our understanding of patient treatment response.

1.6 Aims and outline of the thesis

Here, I state the three research aims of this thesis. Each experimental chapter addresses these aims in turn and together they provide insight into how brain systems flexibly adapt to cognitive demand and how these systems are important for well-being. The first aim of the thesis is:

1 - To identify the brain systems associated with exogenous attention switching.

Specifically, I perform a task-based fMRI activation analysis of a novel task-switching paradigm (Chapter 3). This task systematically alters the cognitive demand imposed during attention switching in healthy adults. Given the increased cost associated with increasing cognitive demand, the second aim of the thesis is:

2 - To test how the brain optimises switching-cost demand by structuring behaviour.

Specifically, in two fMRI studies, I examine how healthy adults develop efficient behavioural routines during a novel self-ordered switching (SOS) paradigm (Chapter 4). These findings indicate that structuring behaviour is an important ability for minimising the cognitive demand of low-level elements of the routine to ‘free-up’ resources for the successful high-level executive control of behaviour. The flexible adaptation of the brain’s higher-order functional networks related to successful behavioral outcomes but this ability varied across healthy individuals. The third and final aim of the thesis, therefore, was:

3 - To test how cognitive control systems are implicated as patients with depression respond to a psilocybin therapy clinical intervention.

Specifically, in two fMRI clinical trials, I examine how brain function changes in depression patients who receive psilocybin therapy. The symptomatic and neural markers of cognitive flexibility significantly increased following psilocybin therapy.

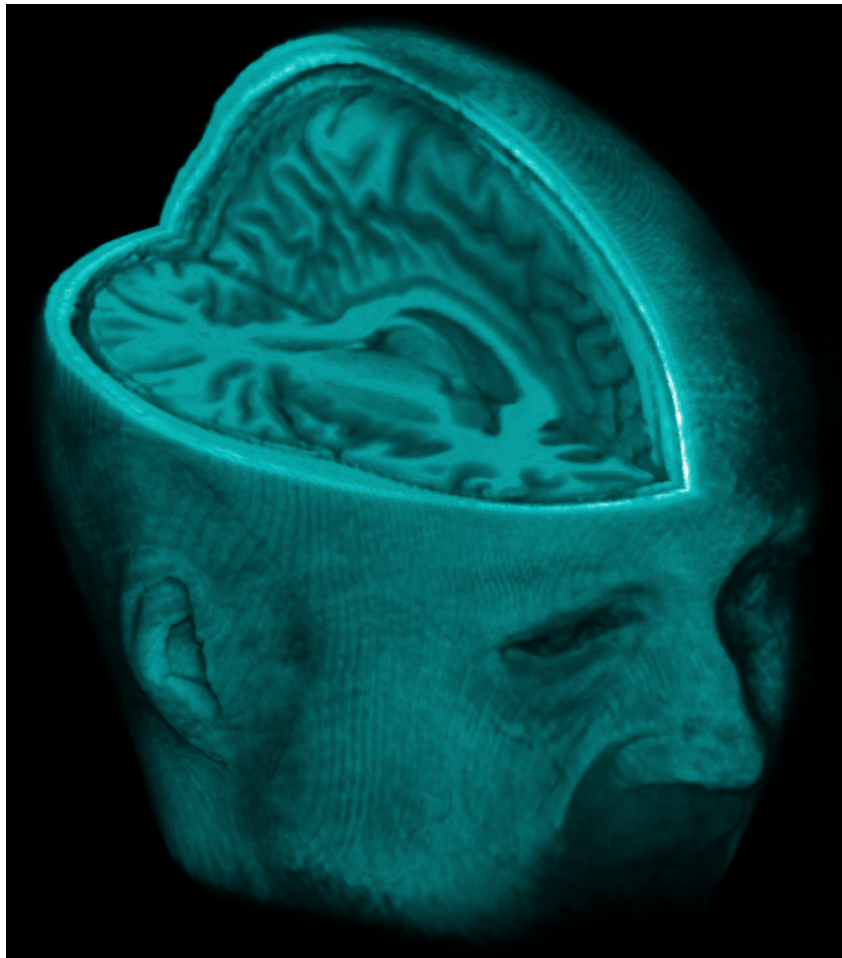
These neural markers predicted long-term increases in patient well-being and changes in functional dynamics were predominantly associated with the brain's higher-order networks including the DMN and the lateral frontoparietal networks.

“I would never die for my beliefs because I might be wrong.”

— **Bertrand Russel**

Chapter 2: Methods

I overview the experimental methods used in this thesis. This includes the fundamental principles of fMRI data acquisition; fMRI data handling; modelling brain activation and connectivity; the measurement of behavioural performance and assessing the clinical symptoms of depression.



You're using your brain to look at mine.

2.1 MRI fundamentals

2.1.1 Magnets & radio waves

The advent of MRI provided a valuable technique for measuring the structural and functional properties of the brain. MRI distinguishes itself from other medical imaging devices as it can produce high resolution images of the body without using potentially harmful ionising radiation (as is the case for x-ray and positron emission tomography) (Oldendorf & Oldendorf, 1988). These features make MRI an attractive research tool with high safety that can be used to acquire data from healthy and patient groups without strong clinical justification (Panych & Madore, 2018).

In its essence, MRI scanners create images using strong magnetic fields and radio wave frequency pulses to manipulate hydrogen protons in the body (Plewes & Kucharczyk, 2012). Protons have the physical property of angular momentum or 'spin'. The typically random direction of a proton's spin is susceptible to alignment by the strong magnetic field lines generated by an MRI scanner's main 'static' magnetic field lines (Bryant et al., 1991). Angular momentum is conserved and therefore protons precess around the field line (akin to a gyroscopic spinning top) (Caverly, 2015). The precessional frequency ('speed of gyration') depends on the nuclei of interests' gyromagnetic ratio (a constant value) and will increase with magnetic field strength, as formulated by the Larmor equation:

$$\omega = \gamma\beta_0, \tag{1}$$

where, the precessional frequency, ω , is proportional to the nuclei's gyromagnetic ratio, γ , and the strength of the magnetic field, denoted here as, β_0 , representing the scanner's static field (Hahn, 1950).

Knowing the precession frequency of a given nucleus and the magnetic field strength allows the 'resonance' to be put into MRI. Hydrogen is the target for conventional MRI. A precise amount of energy can be delivered to hydrogen protons

by emitting radio frequency pulses at the same frequency as the proton's precession frequency. This excitation causes the proton's spin to 'pull away' from the static field lines and towards the perpendicular/transverse plane, or even continue to the antiparallel/inverted position (if sufficient pulse power and time is applied) (Caverly, 2015). Once the radio frequency pulse ceases, the excited proton 'relaxes' and emits energy as it 'relaxes' back to the main magnetic field. This emitted energy is the basis of the signal that is measured in MRI (Irnich & Schmitt, 1995).

2.1.2 Image encoding

By exploiting the principle of magnetic resonance, an MRI scanner can construct 3-dimensional (3D) images of its subject. Similar to a digital camera, where pixels are assigned values, the scanner creates an image by 'filling' each voxel (3D pixel) of the 3D volume.

To build the 3D image volume, conventional MRI involves two processes: Slice selection and spatial encoding. Slice selection involves applying an additional magnetic gradient along the z-axis (G_z), typically, to the main static (β_0) field (Pipe, 1995). Following the Larmor equation, this induces linear variability in the precession frequency, with slower speeds at the bottom of the gradient. Therefore, the matching radio frequency pulse will selectively excite protons in the corresponding section of the gradient (Bieri & Scheffler, 2013).

The excited slice represents 1 2D grid of the 3D volume. Spatial encoding is then used to isolate the signal from each voxel along the x/y axes of the slice. Frequency encoding gradients (readout gradients) are also applied to alter the precession frequency along one axis (in a similar manner to the slice select gradient). To encode position along the second axis, phase-encoding gradients are applied to dephase (become out of sync) spin precession along the gradient. In contrast to the slice-select and frequency-encoding gradients, the phase-encoding gradients are not accompanied by the symmetric and reverse rephasing gradient during the pulse

sequence (Erasmus et al., 2004). Together, these gradients are systematically combined to isolate and measure the signal from each coordinate within the 3D grid.

2.1.3 MRI modalities

MRI can generate a broad range of image modalities, each of which are sensitive to different tissue properties (Basser et al., 1994; Haase, 1990; Weissman, 1986). The MR physicist can manipulate parameters within the imaging pulse sequence, such as the echo time (TE) and repetition (TR), which determine when and how protons are excited and when the signal is measured during its relaxation. Most of the contrast in an MR image of tissue comes from differences in fundamental parameters relating to the proton density and the T_1 , T_2 , T_2^* relaxation times (Gold et al., 2004). T_1 relaxation refers to the time required for longitudinal relaxation, or return to the steady precession around the B_0 field lines, after excitation, and T_1 -weighting of an image is controllable via the TR. T_2 and T_2^* both refer to the transverse relaxation and are controlled by the TE (Henoumont et al., 2009; Houmard et al., 1995).

In neuroimaging research, T_1 -weighted images (short TE, short TR) are most commonly used to generate anatomical images of the brain (Jack et al., 2008). T_1 -weighted images offer a high spatial resolution, isotropic voxels and moderate signal contrast between tissue types (e.g., grey matter, white matter).

In this thesis, I use T_1 -weighted images alongside a special kind of MRI contrast, functional MRI (fMRI). fMRI is sensitive to brain function and is acquired at speeds rapid enough to ‘freeze’ physiological motion (Ogawa et al., 1993). Gradient echo (GE) echo-planar imaging (EPI) acquisition schemes are typically used for fMRI. A ‘train’ of very short phase-encoding gradients are applied to generate a rapid sequence of ‘echoes’ within the TR (Feinberg & Yacoub, 2012). Each echo is the

point where transverse magnetisation has been refocused for a specific portion of the subject with a pair of alternating magnetic gradients.

The moderate temporal resolution of fMRI (1-3 seconds per measurement) is impressive but is substantially slower than the brain's electrophysiology. However, a tradeoff in spatial resolution is made by reducing the number of measurements required at each timepoint. This can be achieved in several ways including larger voxels and/or adding gaps between slices. The other tradeoff is in the sacrifice of the signal tissue contrast as periods in the sequence where tissue relaxation can occur are short (Poser et al., 2006). Subsequently, brain imaging experiments will typically acquire a high-resolution T1-weighted image in addition to the fMRI recordings (this is discussed further in section 2.2.2 fMRI preprocessing).

2.1.4 The BOLD signal

fMRI provides an indirect measure of brain function and relies on subtle differences in magnetisation susceptibility between oxygenated and deoxygenated haemoglobin in the blood (Logothetis & Wandell, 2004).

fMRI typically measures changes in the Blood Oxygen Level Dependent (BOLD) signal. In a curious twist of fate, haemoglobin that is carrying oxygen has a weaker interaction with a magnetic field (diamagnetic) than deoxygenated haemoglobin (paramagnetic) (Pauling & Coryell, 1936; Thulborn et al., 1982). Deoxygenated haemoglobin, therefore, introduces local inhomogeneities in the magnetic field where protons (excited and 'flipped down' onto the transverse plane) will exhibit a more rapid dephasing that is not recoverable by the refocusing gradients (or RF pulses). This results in lower signal magnitudes being measured in the presence of deoxyhemoglobin (Ogawa et al., 1993).

The brain requires a constant supply of oxygen and glucose metabolites. To manage fluctuations in demand, a dense and dynamic vasculature (Reichenbach et al., 1997; Tatu et al., 1998) reacts to the metabolic needs of local tissues (Chow et al., 2020). This accommodates the increased metabolic demand from neurons that are

‘firing’ action potentials at a greater rate than normal (Hasenstaub et al., 2010). The increase in blood flow increases the perfusion of oxygenated blood through the local tissue. This results in less transverse dephasing and an increase in MRI signal (Ogawa et al., 1993).

This neurovascular coupling is an important consideration for the modelling of fMRI because there is a temporal lag between the onset of increased neuronal activity, bolus perfusion and the subsequent measurable change in the BOLD signal. This lag is one of the limiting factors when it comes to the possible temporal resolution of fMRI (Friston et al., 2000; Woolrich et al., 2004).

2.2 fMRI data handling

2.2.1 Noise sources: Head motion & physiological fluctuation

In this thesis, I use fMRI to indirectly measure brain function. However, the fMRI signal has many sources to it, and they often can share statistical dependencies with our signal of interest, fluctuations in neural activity. There are too many ‘noise’ sources to be fully described here; however, the primary concerns relate to noise introduced by head motion and fluctuations in physiology.

On a basic level, head motion is problematic because we assume that the signal from a given voxel corresponds to the same portion of brain tissue over time. Therefore, motion mixes signals between adjacent regions. Beyond this, head motion introduces a panoply of corrupting influences that can persist over many measurements after the motion has occurred (10s of seconds) (Power et al., 2012).

The scanner’s homogeneous field is spoiled when a subject is introduced; however, additional shimming gradients are applied to counteract this (Stockmann & Wald, 2018) with the assumption that the subject is stationary during scanning. The interactions between the moving head and the magnetic field are very difficult to predict and result in issues relating to off-resonance, signal drop out and spin history artefacts (Muresan et al., 2002). All of these influences will bias our estimates

of brain function and can be particularly problematic when the study compares clinical groups that may present confounding differences in head motion (Makowski et al., 2019).

Physiological noise refers to changes in the MR signal that result from changes in breathing, heart rate and arousal (Orban et al., 2020). The simple act of breathing affects the homogeneity of the main field and changes in heart rate or breathing can become confounds if they occur in patterns that correlate with the experimental design (Brosch et al., 2002; Hu et al., 1995; Power et al., 2019). Relatedly, the systolic blood pressure waves can also induce physical motion of brain tissue and changes in heart rate and breathing rate will affect the BOLD signal (Chang et al., 2009).

2.2.2 fMRI preprocessing

Several preprocessing steps are applied to the raw imaging data before using the data to ask questions regarding brain function. The aims of preprocessing are to mitigate the noise introduced by head motion and physiology and to maximise the spatial similarity between individuals so that the equivalent brain regions can be compared across individuals.

A typical first step in fMRI preprocessing is slice-timing correction. fMRI data are typically acquired one slice at a time. This can mean that there can be large temporal offsets between the first and last slice collected within a single 3D volume. Slice-timing corrections use an interpolation across slices, within a volume, to weigh values against a reference slice (Sladky et al., 2011). The step is particularly important for sequences with longer TRs (Glasser et al., 2013).

To correct for in-scanner head-motion, a rigid-body transform is typically performed between each volume and a reference volume (e.g., the first, middle or last volume) (Friston et al., 1996). Rigid-body transforms allow 6 degrees of freedom (3 translations and 3 rotations along the X, Y, Z axes) to be optimised when minimising the difference between each pair of images. The root mean square error (RMSE) is a

typical objective function in this case. The resulting transforms estimate the parameters required to realign the series of 3D volumes collected over time (Power et al., 2015). These realignment parameters can be used as model predictors to account for time course variance associated with head motion (see section 2.2.3 Task activation modelling). Alternatively, this variance can be regressed out from each voxel and the analysis is then conducted on the residuals. This approach is common in resting-state analyses (Parkes et al., 2018).

It is important to overcome the natural variation in brain size and morphology when seeking to detect regional activation changes that are consistent across individuals. It is common to transform, or ‘register’, the MR images to a standard group space and coordinate system to facilitate statistical comparisons (Ashburner & Friston, 2000). However, BOLD EPI images exhibit poor tissue contrast and are typically acquired with a moderate spatial resolution (3mm^3) (Poser et al., 2006). This makes between-subject registrations particularly challenging. Instead, a T_1 -weighted image is also acquired during the experiment to generate a high-resolution anatomical image with moderate signal contrast for use during this registration stage. Specifically, the individual's motion-corrected EPI images are rigid-body transformed to the individual's high-resolution anatomical image (typically T_1 -weighted, 1mm^3) during a ‘co-registration’ step. This aligns the individual's T_1 -weighted image with their BOLD EPI volume timeseries.

To align the individual's data with the standard group space, an affine transform and non-linear warp is estimated using the anatomical image and this is then applied to the co-registered EPI's (Ashburner & Friston, 2000). Alternatively, a study-specific group space, or template, can first be generated as an intermediary step, where images are registered to the template followed by the transform to the standard group space (Ashburner, 2007; Ashburner & Friston, 2011).

A final spatial smoothing step is applied to improve the signal to noise ratio and aid the data to meet the assumptions of random field theory. Typically, voxel values are

replaced with the weighted mean of the surrounding voxel values calculated using a 3D gaussian kernel. An important consideration is the spatial extent of the kernel and this is controlled by the full-width at half-maximum (FWHM) value. Typically smoothing kernels in fMRI analyses of volumetric data range between 4-8 mm (Mikl et al., 2008).

2.2.3 Task activation modelling

The preprocessed fMRI data can be modelled in a myriad of ways. For task-fMRI, where individuals are scanned while performing tasks, model-based approaches are common. In particular, general linear modelling (GLM) can be used to arrange predictors for capturing aspects of the experiment, as they occurred in time (e.g., when a button is pressed) (Friston et al., 1994). However, the limited temporal resolution of fMRI and the neurovascular lag inherent in the BOLD signal, present an interesting modelling problem. To align the temporal resolution of the experimental events with the fMRI data, the model predictors (representing the onset and duration of events) are resampled (to equal the number of EPI volumes acquired) and convolved with a hemodynamic response function (HRF). This process improves the predictive power of the model (Mechelli et al., 2003).

The GLM allows for multiple predictors to be simultaneously modelled and for their independent effects on each voxel timecourse to be estimated (e.g., correct vs incorrect responses). It is also possible to include confound, or nuisance, predictors in the model to control for noise sources such as head motion or physiological predictors (Friston et al., 1996). This mass-univariate approach can be formulated as:

$$Y^T = \beta_0 + [H(X), E]\beta + e, \quad (2)$$

where, X , is the matrix of psychological timecourses and, $H(X)$, its HRF convolution. Y^T is the voxel timecourse and E is the matrix of nuisance regressors. β , the activation beta weights of interest, β_0 , is the model intercept and, e , the model residual.

Fitting the GLM to the timeseries from each voxel generates a beta coefficient, or contrast estimate, for each predictor in the model. For each predictor (condition), these contrast estimates can be arranged into a 3D map of the brain where each voxel now represents the relative change in the timeseries intensity given the multiple sources (predictors) or influences on that voxel timeseries. If a contrast estimate is positive, it is assumed that the condition of interest (predictor) is associated with an increase in functional activation of the brain tissue captured by that voxel (Monti, 2011).

Typically, classical inferential statistics are then used to generate group-level maps that represent consistent activation changes across individuals. This can be done with analysis of variance (ANOVA) models and t-tests, amongst others. Again, this is a mass-univariate approach meaning that an independent statistical test is applied to each voxel. The statistical maps can then be thresholded to isolate those voxels using statistical significance (the probability/confidence that repeating the experiment with an independent sample would yield the same result).

However, this statistical approach presents a non-trivial multiple-comparisons issue (Genovese et al., 2002). Say that we tolerate a typical 1/20 probability that our activation change is a false positive, we would expect to see a very large number of voxels, within the 10's of thousands being tested, to show a significant difference by random chance.

Subsequently, the significance threshold can be made more stringent (e.g., a 1/1000 chance); however, this will be at the cost of an increased false-negative rate. Instead, a correction for multiple comparisons can be applied at the level of individual voxels or on clusters of spatially adjacent voxels that are above threshold (Woo et al., 2014). Both are attempts to control the number of false-positives while preserving the true positives. One advantage of cluster-level correction is that the spatial dependency between voxels is considered to some extent (Woo et al., 2014).

Voxels are arbitrary units and therefore clusters of activation are more likely to reflect biologically plausible signals.

Corrections at the voxel or cluster level are typically performed with Bonferroni or a false discovery rate (FDR) correction. Bonferroni corrections are conservative as significance level (α) is redefined as the original α divided by the number of tests being carried out (e.g., number of voxels). This can result in a significant proportion of true positives being falsely rejected. The FDR correction takes a different perspective that adjusts probability values generated from a series of independent tests. Here, it is accepted that a proportion of the significant results are false-positives, instead of the proportion of all tests being conducted being false positives (Benjamini & Hochberg, 1995).

In chapters 3 and 4, I use task-activation analyses to define group-level activation changes predicted by the different experimental events. Unless otherwise stated, the activation maps were initially thresholded at an uncorrected voxelwise $p < 0.01$ threshold. The resulting maps were then cluster-corrected using an FDR $p < 0.05$.

Once thresholded, the spatial patterns of activation change can be compared across conditions or independent groups. Region of interest (ROI) and individual difference analyses can also be performed where activation magnitudes are sampled from specific brain regions and each individual and then correlated with measures relating to behavioural performance or clinical severity, for example (Poldrack, 2007).

2.2.4 Watershed by immersion clustering of activation maps

The large statistical spaces generated by fMRI data present a noteworthy challenge. What are the principled ways to summarise these data and derive valid insight? For task fMRI activation experiments, ROI analyses are commonly used. ROI analyses allow researchers to reduce the statistical space into manageable subsets of the data and ask focused questions about specific brain regions. However, while fruitful, the typical ROI analysis has been limited by an arduous manual process,

experimenter bias, anatomical expertise and susceptibility to post-hoc redefinitions for statistical significance (aka., “p-hacking”).

Consequently, the reproducibility across fMRI studies can inadvertently suffer (Lorenz et al., 2017). Despite the existence of many open-access tools for modelling MRI data, there is a surprising lack of tools that yield ROIs from functional activation maps in a data-driven and principled manner.

To address this, Eyal Soreq and I developed a toolbox designed to offer an unbiased data-driven alternative to ROI definition for fMRI task activation analysis. The “FUnctional and Structural Integration of Neuroimages” (FUSION) WaterShed (WS) toolbox is currently in preparation for public release. At the core of Fusion-WS is a variant of the classic watershed by immersion transform (F. Meyer & Beucher, 1990). The watershed transform is a naive but elegant approach that is well-suited to clustering fMRI task activation because it can efficiently segment continuous statistical spaces into discrete labels that can be non-gaussian in shape. Moreover, the number of clusters derived is parameter free, in the sense that there is no k to be defined. Three parameters are defined in this implementation: Filter threshold, flooding radius and merge threshold. Fusion-WS has default values for these parameters; however, appropriate values will depend on image resolution, focality of the activation clusters and the desired parsimony of the ROI set. These parameters are described below.

Watershed transforms are flexible as they can be applied to n -dimensional spaces; however, the Fusion-WS implementation is tailored to 3D spaces as these are typical in task fMRI analyses. The algorithm treats the activation map as a multi-dimensional surface where high and low intensities represent elevations, akin to a geographical landscape with peaks and troughs.

It is advisable to run Fusion-WS using an activation map that has been “preprocessed”. Random noise in the volume will result in a large ROI set containing many small ROIs with a high proportion that are unlikely to be relevant to the

experimental hypothesis. Subsequently, it is advisable to apply a cluster-correction or another threshold to remove redundant voxels from the space. The first stage in the Fusion-WS pipeline inspects the image and applies a basic internal threshold using the top 10% of the interquartile range (IQR) if the image has not been thresholded. NB:- This procedure is independently repeated for two-tailed maps on the positive and negative voxel intensities.

The second stage in Fusion-WS transforms the continuous space into a set of spatially independent connected components (ICC). For example, if there are two activation clusters that don't share any contiguous voxels, say one in the frontal lobe and another in the parietal lobe, then this would generate two independent components. The filter parameter is applied at this stage to exclude ICC below a given voxel volume. Watershed is then applied to each ICC's underlying voxel intensities, which are inverted so that activation peaks become troughs, or catchment basins (CB). An index is generated which ranks all voxels within the ICC. Iterating through this index, starting with the deepest CB, voxels are handled in 1 of 3 ways. It is first asked whether the voxel is proximal to another voxel with an existing label. This search is conducted within the flooding radius. If a label is found, then that voxel is assigned to the same label. If no label exists within the flooding radius, then a unique label is created for that voxel. If multiple labels are found then the label with the mode voxels in the neighbourhood is assigned in a "winner-takes-all" strategy. This procedure continues until all voxels within the ICC have been labelled. Finally neighbouring ROIs, within the ICC, are merged if they are below the merge threshold voxel volume.

The Fusion-WS functionality has been applied to task fMRI data in cognitive contexts, including working memory processes (Soreq et al., 2019), intelligence testing (Soreq et al., 2021) and IBL (Hampshire et al., 2019), and also for detecting activation differences in clinical contexts including Trichotillomania (J. E. Grant et al., 2018) and OCD (Hampshire et al., 2020).

In Chapter 3, a watershed transform is applied to a group-level fMRI activation map using the Fusion-WS implementation outlined above. Specifically, a large activation cluster in PFC was segmented into multiple ROIs. Critically, these data-driven ROIs closely correspond with previously reported activation coordinates (Crittenden & Duncan, 2014).

2.2.5 Connectivity modelling

The human brain exhibits complex network architectures, both in terms of the white-matter pathways (Catani & Thiebaut de Schotten, 2008) that carry action potentials between regions (S. S.-H. Wang et al., 2008) and the co-fluctuating activity patterns that occur between distributed regions (Bassett et al., 2013; Sporns & Kötter, 2004).

Activation studies provide information regarding which brain regions are involved in processes. Functional connectivity (FC) studies can allow insight into how information flows between regions as they interact (Friston, 2009). There are many ways to define FC between the timeseries within neuroimaging data. Undirected connectivity measures, based on timeseries similarity, are regularly used to indicate if activation between pairs of regions co-fluctuate over time. Typically, linear statistical dependencies are assessed using measures including Pearson correlation, cosine similarity or phase synchrony (H. E. Wang et al., 2014).

Multivariate approaches such as independent component analysis (ICA) can be applied to decompose the entire dataset into a set of spatially independent networks of voxels that covary over time (McKeown & Sejnowski, 1998). ICA has been applied to resting-state fMRI to identify the brain's intrinsic functional networks (S. M. Smith et al., 2009).

2.2.5.1 Psychophysiological interactions

ICA and bivariate correlations can be applied to task-fMRI data (Finc et al., 2020; Leech et al., 2011); however, psychophysiological interaction (PPI) modelling has the distinct advantage of measuring task-dependent *changes* in FC between pairs

of regions (Friston et al., 1997; Gitelman et al., 2003). The PPI uses GLM to estimate the change to intrinsic connectivity between a seed and target region (physiological) timecourses that is predicted by the task events (psychological).

A recent extension of the PPI framework simultaneously models multiple conditions of interest together. The generalised PPI (gPPI) has the advantage of greater sensitivity to measuring task-dependent FC (McLaren et al., 2012). I use the gPPI measure in Chapter 4 to examine how brain connectivity changes as healthy adults learn a complex task. The gPPI is modelled using the following:

$$Y^T = \beta_0 + [Y^S * H(X)]\beta_G + [Y^S, H(X), E]\beta + e, \quad (3)$$

where, X , is the matrix of psychological timecourses and, $H(X)$, its HRF convolution. Y^T is the target ROI timecourse and, Y^S , the seed ROI timecourse. E is the matrix of motion & tissue signal timecourses. β_G are the gPPI betas weights, β , the beta weights of no-interest, β_0 , is the model intercept and, e , the residual (McLaren et al., 2012). This procedure can be repeated for each pair of regions and the resulting gPPI's (beta weights) averaged to generate an undirected FC measure.

2.2.5.2 Pearson correlation

Brain function expressed in fMRI data can be summarised into network graphs. Nodes in the graph represent brain regions and the edges between nodes represent the strength of FC between regions (He & Evans, 2010; Lee et al., 2013). Graphs can be built using one of many FC measures but Pearson correlations are common (B. P. Rogers et al., 2007). Pearson correlations are defined by:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}, \quad (4)$$

where the correlation coefficient r is the covariance of variables x and y , as the sum of product deviations of each datapoint from their respective variable mean, divided

by their standard deviation (Pearson, 1896).

Pearson correlations are useful for describing simple bivariate linear relationships. The standardisation in its formulation provides a metric, ranging from -1 to 1, that is invariant and comparable across independent pairs of variables (given equal degrees of freedom). For example, a coefficient equal to 1 denotes a perfect linear relationship where both variables increase with 0 variance or deviation around the linear line of best fit. A coefficient close to 0 denotes a large variance from the line of best fit.

In brain imaging analyses, positive correlation coefficients between two timecourses (x , y) imply that the signals from two regions increase and decrease together over time. A negative correlation implies that when the signal in either timecourse increases the signal decreases in the other timecourse. Functional networks are typically considered as spatially distal regions whose timecourses positively correlate.

2.2.5.3 Modularity

Understanding how the brain's functional networks are organised remains an area of fervent research (Glasser et al., 2016; Schaefer et al., 2018; Yeo et al., 2011). However, it is clear that variations in functional connectome expression contain predictive information that is specific to individuals (Finn et al., 2015), individuals' task performance (Rosenberg et al., 2016; Soreq et al., 2021) and clinical status (Mason et al., 2018; Matthews & Hampshire, 2016).

Functional connectomes often represent large statistical spaces that can be burdensome to evaluate. However, composite measures that describe features of the connectome are widely used (Bassett et al., 2013; Bassett & Bullmore, 2006; Mattar et al., 2015).

Network modularity provides an elegant approach to describing functional connectomes. I use this metric in Chapter 5 to examine how brain network

organisation changes as patients with depression respond to psilocybin therapy. Modularity or, Q , indexes the degree of clustering and segregation within the network. Typically, a community detection algorithm (Blondel et al., 2008) is applied to a network graph whose objective is to maximise the extent to which the graph can be split into separate non-overlapping communities or modules (Newman, 2004, 2006; Newman & Girvan, 2004).

For the brain's functional connectome, clusters of connectivity (networks) are observable throughout. For example, regions within the visual cortex show preferential connectivity to each other than with the rest of the brain. The extent, and number, of clusters within the connectome determines the modularity score (Newman, 2006). Larger modularity scores indicate a greater degree of segregation between the brain's network. Here, I use modularity with the standard definition of:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j), \quad (5)$$

where A_{ij} represents the weight of FC (correlation) between ROI i and j , γ is the structural resolution free parameter (set to 1) and $\frac{k_i k_j}{2m}$ is the expected null FC defined with $k_i = \sum_j A_{ij}$ as the total FC across all connections with ROI i , c_i is the community to which ROI i is assigned, $\delta(c_i, c_j)$ is the Kronecker delta function and equals 1 if ROI i and j belong to the same community and 0 otherwise (Bassett et al., 2013) and $m = \frac{1}{2} \sum_{ij} A_{ij}$, is the total FC of the network.

Modularity scores can suffer from non-deterministic and near-degenerate graph partitions (meaning that there can be many solutions with similar optimality) and have a dependency on the sum of edge weights within the graph (e.g., FC within the connectome) (Rubinov & Sporns, 2011). To facilitate valid comparisons between scanning sessions and individuals, the community detection can be repeated (e.g., 100 times) and the partition with the largest modularity score normalised by the

mean modularity score generated from a distribution of randomly rewired (shuffled) networks (Maslov & Sneppen, 2002).

2.2.5.4 Functional cartography

The community detection procedure generates a community assignment for each node in the network graph. These labels can be used to determine the extent to which nodes are recruited to the functional networks that they typically ‘belong to’, as defined by healthy adults. For example, brain regions known to constitute the DMN should reliably form communities with each other.

In practical terms, an allegiance matrix, P , is calculated to represent the probability that two nodes i and j are assigned to the same community (Bassett et al., 2015; Mattar et al., 2015), defined here as:

$$P_{ij} = \frac{1}{O} \sum_{O=1}^O a_{ij}^{k,o}, \quad (6)$$

where O is the number of times the modularity algorithm was run. For each partition, $a_{ij}^{k,o}$ equals 1 if regions i and j belong to the same community.

Using the allegiance matrix, node scores can be summarised to generate functional network-level metrics such as recruitment (the extent to which nodes from the same network form communities) and integration (the extent to which nodes from two different systems form communities (Bassett et al., 2015; Mattar et al., 2015). Metrics such as these can be normalised against the mean values from randomly shuffled node-network assignments to account for differences in the number of regions in each network, or pair of networks (Finc et al., 2020).

2.2.5.5 Dynamic flexibility

The connectivity and graph analyses described thus far are ‘static’, i.e., they reflect a time-averaged connectivity pattern defined using all timepoints available (e.g., throughout a 10-minute recording). Dynamic functional connectivity can define how connectivity changes over time. In Chapter 5, I use a sliding-window approach

where connectivity matrices are estimated for each window in time (Hutchison et al., 2013).

Sliding window approaches can be used to estimate dynamic modularity (Lucas et al., 2016; Mucha et al., 2010) using a 3D $N \times N \times T$ FC matrix with FC estimated between each node N at each point in time or time-window T . Dynamic modularity can be defined as:

$$Q_{ML} = \frac{1}{2\mu} \sum_{ijlr} \left[\left(A_{ijl} - \gamma_l \frac{k_{il}k_{jl}}{2m_l} \right) \delta_{lr} + \delta_{ij} \omega_{jlr} \right] \delta(c_{il}, c_{jr}), \quad (7)$$

where $\mu = \frac{1}{2} \sum_{ijl} A_{ijl}$ is the total FC of the multilayer network, and $m_l = \frac{1}{2} \sum_{ij} A_{ijl}$ is the total FC of layer l , A_{ijl} is the FC between ROI i and j at layer l , $\frac{k_{il}k_{jl}}{2m_l}$ is the expected null FC at layer l . The two free γ and ω structural and temporal resolution parameters are used to scale the number of communities and strength of inter-layer edges, respectively. As is typical for dynamic modularity fMRI analyses, these parameters were both set to 1 (Finc et al., 2020; Harlalka et al., 2019; Mattar et al., 2015) in the dynamic modularity analysis reported in chapter 5.

The dynamic modularity estimates a community assignment for each node at each time-window in an $N \times T$ matrix. From this, metrics such as dynamic flexibility can be defined by simply the number of times a node changes its community allegiance over time, give the number of opportunities to observe change (Mattar et al., 2015):

$$f_i = 1 - \frac{1}{T-1} \sum_l^{T-1} \delta(c_{il}, c_{j,l+1}), \quad (8)$$

Flexibility scores close to 0 represent rigid nodes whose community allegiance is stable across time, scores close to 1 represent flexible nodes whose community allegiance regularly changes over time.

2.3 Modelling behaviour

2.3.1 Reaction time

Reaction time (RT) or response speed has been extensively used to model behaviour, attention and intelligence (Carlson et al., 1983; Prinzmetal et al., 2005). RT is a continuous measure representing the time between the presentation of a stimulus and determining and enacting the response (typically a button press) (Robbins, 2002).

RT is a proxy of cognitive load and, subsequently, can be experimentally manipulated and compared across individuals and groups (Barrouillet et al., 2007). RT can be modelled in many ways; however, a common approach is to summarise individual RT's within conditions of interest (e.g., for each difficulty level) and compared across conditions, between groups of subjects (Camicioli et al., 2008) or over time (Hampshire et al., 2016).

In my experimental chapters of this thesis, I use RT to compare changes in task performance as a function of switching demand in both chapters 3 and 4. I also use RT to inform the fMRI task activation models in chapters 3 and 4 and the fMRI gPPI FC models in chapter 4.

2.3.2 Behavioural complexity

Most cognitive experimental paradigms control when the order of events that the subject will respond to (Gratton et al., 2018; Rubinstein et al., 2001). However, in chapter 4, I utilise a novel experimental paradigm (designed by Dr Adam Hampshire) that requires subjects to structure their behaviour to manage competing demands. Essentially, this paradigm generates a subject-specific sequence of the four possible trial conditions. This presented the opportunity to quantify how behavioural structuring varied across individuals and how this relates to changes in brain function.

For this context, you could use an estimate of Shannon entropy (Shannon, 1948). It could be inferred that those behavioural sequences with high entropy values were generated by individuals who were performing randomly, as each condition may have occurred with a close to equal probability. However, instead, I estimate the algorithmic complexity of the subject's behaviour (Zenil et al., 2018). For random sequences, where no compression is possible, algorithmic complexity and Shannon entropy perform in a similar manner (Teixeira et al., 2011). However, for non-random sequences, algorithmic complexity has the advantage over Shannon entropy of being able to detect consistent structures within a sequence (Gauvrit et al., 2017).

Algorithmic complexity provides a computationally tractable estimation of Kolmogorov-Chaitin complexity (K) (Chaitin, 1966; Kolmogorov, 1968) by combining a block decomposition method (BDM) and coding theorem method (CTM) approach. Conceptually, algorithmic complexity is the size of the smallest algorithm capable of producing a given sequence of characters (for a full derivation see (Zenil, 2020; Zenil et al., 2018)).

To estimate the algorithmic complexity of an individual's behaviour in chapter 4, each behavioural timecourse was converted into a binary vector that coded the order that the subject performed the trial conditions. Sequence length naturally varied across individuals. To account for this, complexity scores were normalized by the mean algorithmic complexity calculated from 100 randomly shuffled versions of the data.

2.3.3 Measuring depression severity

In chapter 5, I examine the brain systems implicated in cognitive control (highlighted in chapters 3 and 4) in two resting-fMRI studies designed to assess changes in brain function as patients with depression respond to psilocybin therapy treatment.

An important component of these analyses relates to individual differences in depression severity scores, as was indexed by the Beck depression inventory (BDI) (Beck et al., 1996; Beck & Clark, 1988). In particular, I use scores generated using

the BDI-1A, a 21-item questionnaire. For each item, patients self-report agreement with one of four statements that carry 0-3 points. The total sum of scores represents the depression severity. The maximum score is 63, the clinical threshold is 20 points, >30 points is the severe depression category.

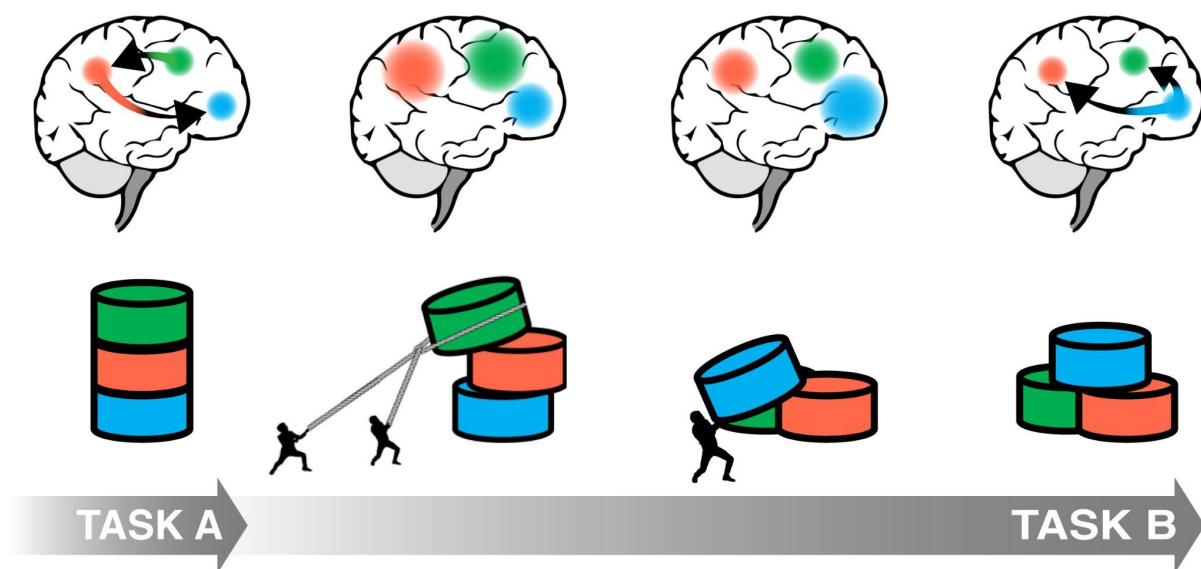
The BDI-1A is a well-validated measure (Beck et al., 1988), that taps cognitive, affective and somatic features of depression (Steer et al., 1987). Moreover, BDI has a broad agreement with other depression measures but has a stronger emphasis on cognitive features than other conventional measures such as the quick inventory of depressive symptomatology (QIDS) (Fried, 2017).

“If you want to improve, be content to be thought foolish and stupid”

— **Epictetus**

Chapter 3: Contrasting hierarchical and multiple-demand accounts of frontal lobe functional organisation during task-switching

In this chapter I analyse fMRI data from a novel task-switching experiment that was designed and collected by Dr Adam Hampshire. Addressing aim 1 of this thesis, I test which brain regions or networks support attention switching and how their activation varies with increasing cognitive demand. This dataset provided the opportunity to contrast two theoretical accounts of the frontal lobe's functional organisation. The findings from this experimental chapter are prepared for submission and, subsequently, the use of “we” refers to my predictions, analysis, data visualisation and interpretation.



The brain adapts during task-switching.

3.1 Abstract

There is an unresolved discrepancy between popular hierarchical and multiple-demand perspectives on the functional organisation of the human frontal lobes. Here, we tested alternative predictions of these perspectives with a novel fMRI switching paradigm. Each trial involved switching attention between stimuli, but at different levels of difficulty and abstraction. As expected, increasing response times were evident when comparing low-level perceptual switching to more abstract dimension, rule and task-switching. However, there was no evidence of an abstraction hierarchy within the prefrontal cortex (PFC). Nor was there recruitment of additional anterior PFC regions under increased switching demand. Instead, switching activated a widespread network of frontoparietal and cerebellar regions. Critically, the activity within PFC sub-regions uniformly increased with behavioural switch costs. We propose that both perspectives have some validity, but neither is complete. Too many studies have reported dissociations within MD for this volume to be functionally uniform, and the recruitment of more anterior regions with increased general difficulty cannot explain those results. Conversely, whilst reproducible evidence for a hierarchical functional organisation has been reported, this cannot be explained in terms of abstraction of representation or reconfiguration *per se*, because those interpretations generalise poorly to other task contexts.

Keywords: Cognitive control, task-switching, frontal lobes.

3.2 Introduction

The prefrontal cortex (PFC) is critical for human cognition and behaviour (Duncan, 2001; Fuster, 1999; Shallice & Evans, 1978), yet how regions within the PFC are functionally organised remains controversial. Broadly speaking, much of the debate is polarised around two competing perspectives.

One prominent class of models revolve around the notion that the PFC is organised hierarchically (Figure 1), with increasingly anterior PFC (aPFC) regions placing towards the top of the hierarchy and supporting cognitive processes of an increasingly higher-order (Badre, 2008, 2008; Badre et al., 2009; Koechlin et al., 2003; Ramnani & Owen, 2004). Evidence in support of this hypothesis has been provided from studies of executive dysfunction after lesions to aPFC (Badre et al., 2009; Duncan et al., 1995; Nelson, 1976; Reitan & Wolfson, 1994; Roca et al., 2010; Shallice et al., 2007), and by findings of increased functional MRI (fMRI) aPFC activation for tasks that require switching (Braver et al., 2003; Koechlin et al., 1999, 2003; Rushworth et al., 2002), sequential processes (Koechlin et al., 1999; Vendetti & Bunge, 2014), higher-order integrations (Parkin et al., 2015) and reversal learning (Hampshire & Owen, 2006).

However, there is increasing awareness that the generalisation of one-to-one mappings between cognitive processes and regional activations is problematic. More specifically, the same brain regions are often activated during tasks that operationally are distinct (Duncan, 2013; Duncan & Owen, 2000; Fedorenko et al., 2013) whilst brain regions attributed to specific functions are rarely found to activate in isolation (Cole et al., 2012; Daws, Scott, et al., 2020; Hampshire et al., 2010; Lorenz et al., 2017; Mouraux et al., 2011; Sharp et al., 2010; Yarkoni et al., 2011).

An alternative perspective that can account for this general activation profile states that a set of brain regions, referred to as “multiple-demand” cortex (MDC), forms a flexible cognitive resource (Duncan, 2013; Duncan & Owen, 2000; Fedorenko et al.,

2013). MDC is proposed to be recruited during diverse demanding tasks. In support of this view, aPFC, is often activated as part of a broader distributed network during diverse task contexts. Task-switching is one example (Cusack et al., 2010; Daws, Scott, et al., 2020; Dove et al., 2000), but others include target detection (Hampshire et al., 2007), motor-inhibition (Erika-Florence et al., 2014) and instruction-based learning (Hampshire et al., 2016, 2019).

aPFC activation has also been reported during tasks that lack abstraction or hierarchy by design, but that are challenging in other ways (Crittenden & Duncan, 2014; Jiang & Kanwisher, 2003). In order to account for studies that appeared to support PFC hierarchy, proponents of the globalist perspective (Dehaene et al., 1998) have proposed that progressively larger, and more anterior portions of the frontal lobes, including aPFC, may be recruited to the MDC network under increased task difficulty in general (Crittenden & Duncan, 2014; Jiang & Kanwisher, 2003).

In this study, we developed a ‘multilevel switching’ (ML) paradigm to test the alternative predictions of the hierarchical and MDC perspectives on PFC functional organisation. The ML paradigm allows a comparison of attentional switches that partially differentiate the abstraction of the information being updated and the amount of information being updated (general task difficulty). We gauged switch costs using reaction time (RT), instead of response errors, as RT provides a continuous behavioural index of difficulty for each trial that can be used to inform the fMRI activation modelling. Furthermore, switch costs to RT can occur without response errors (Daws, Scott, et al., 2020).

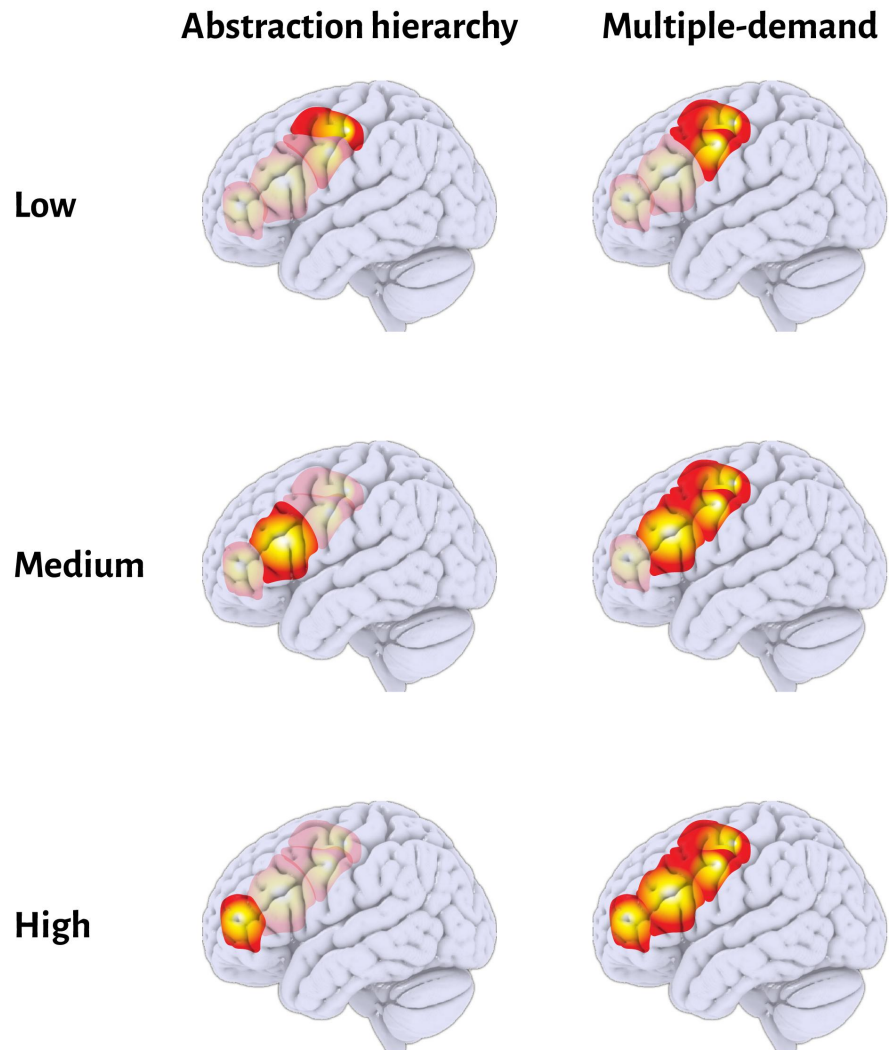


Figure 3.1. Schematic of alternate accounts of the frontal lobe's functional organisation. Left column depicts hierarchical models (Koechlin, Ody and Kouneiher, 2003; Badre, 2008) that denote increasingly anterior prefrontal cortex (aPFC) regions placing towards the top of the hierarchy and supporting cognitive processes of an increasingly higher-order (top-to-bottom rows). The right column depicts an alternate multiple-demand cortex (MDC) model (Duncan, 2013; Crittenden and Duncan, 2014). Here, MDC forms a distributed cognitive resource that flexibly recruits additional, and more anterior, frontal regions under increased task difficulty (top-to-bottom rows).

3.3 Methods

3.3.1 Participants

15 young healthy adults (8 female, mean age 25, ranging 20-40 years) participated in the study. This sample size is comparable to previous studies which have proposed differing accounts of the PFC's functional organisation: The information cascade model ($n=12$, Koechlin et al. (2003)), the abstraction hierarchy model ($n=19$, Badre and D'Esposito (2007)) and the MDC model ($n=13$, Crittenden and Duncan (2014)).

Effect sizes for aPFC activation are not reported in these studies; however, we have estimated them as ranging between 0.7-0.9 Cohen's D units, indicating a large effect. From a power perspective, there is an 80% probability of detecting the same effect with an effect size of $d=0.7$ (one-tailed) and $d=0.8$ (two-tailed) with the present sample.

All participants were right-handed English speakers with normal, or corrected to normal, eyesight. Volunteers were excluded if they had a history of neurological or psychiatric illness, were taking psychoactive medications or did not meet MRI safety criteria. Approval for this study was received by the Cambridge university research ethics committee and participants gave written informed consent prior to entering the fMRI scanner.

3.3.2 Multilevel switching task design

We designed the ML switching paradigm (Figure 3.2a) to compare attention switches that varied in difficulty and abstraction. The ML paradigm was designed for trials to be simple. Each trial presented one target and two probes. Participants responded with a button press and selected the appropriate probe based on the current mappings between the target and probes. These mappings were made explicit by text on the screen and by the stimuli themselves.

The individual trials were easy to perform; however, the important feature was that the mappings between the target and probes would switch on every trial. These switches would vary in difficulty, and level of abstraction, in terms of how much information needs to be updated. Behaviourally, it was expected that RT would be sensitive to this and would be shorter for low-level perceptual switches than for high-level switches that concerned tasks or rules.

Specifically, the target-probe mappings were manipulated across six domains: Tasks, Rules, Dimensions, Targets, Sides and Comparisons. Each of these domains had two levels and only one of these switches occurred between every consecutive trial. At the highest level, switches could occur between two tasks where the target and probes were either coloured shapes or words (Figure 3.2a). Within each task, switches could occur between two discrimination rules which determined whether participants should select the probe that matched the target or the probe that was the odd one out relative to the target and the other probe.

Switches could also occur between one of two dimensions of the stimuli within each task. For objects, the dimensions were shape and colour. For words, these were physical size and animate/inanimate. For example, the correct response for an object matching colour trial would be the probe object that matched the colour of the target object colour. The correct response for a word matching animate/inanimate trial with “horse” as the target and “pin” and “lion” as the probes, would be the “lion” probe.

At the lower perceptual levels, there were target-switches, where the target stimuli updated and the probes remained constant, side-switches, where the two probes simply switched L/R positions, and comparison-switches, where both probes switched between levels of the dimensions. For example, a comparison-switch across two object trials would involve switching the colour and shape of the probes. In figure 2a a comparison switch is described with a black and white schematic where the previous trial, t-1, presents a black circle target with a grey circle and a

black hexagon as the probes. During the current trial, the target remains and the probes are updated by a black circle and a grey hexagon as their colour and shape has been switched. Accordingly, a comparison-switch across two word trials could occur across the animate/inanimate dimension and one probe switches from “fly” to “pin” and the other probe switches from “car” to “lion” (Figure 3.2a).

Notably, the rule, dimension, and target-switches were designed to be similar insofar as they all differ by just one aspect to the comparison-switch, and in two aspects to each other. This design allowed the switch events to differ in terms of abstraction and the amount of information being updated on the screen.

One of the 6 switches (task, rule, dimension, side, target, comparison) occurred on every trial. However, comparison-switches were designed to require the minimal amount of switch reconfiguration demand and occurred on 35-40% of trials whereas each of the other 5 switch types occurred on 13-14% of trials. These frequencies were designed with fMRI model efficiency in mind to produce good separability between the more demanding switches and between those and the comparison-switches.

The paradigm was continuously performed for 16 minutes, without any resting periods, during fMRI recording. Responses required selecting the left or right probe with an MRI compatible button box held in the right hand using the 1st and middle fingers. The current trial continued until a response was made, at which point, the stimuli were removed from the screen. No response feedback was shown. The inter-trial, or response-to-trial, interval was 1400ms.

3.3.3 Data acquisition & preprocessing

The paradigm was programmed in Visual Basic and stimuli were projected on a screen, visible via a mirror, at the end of the scanner bore. fMRI was conducted using a 3 Tesla Siemens Scanner for 16 minutes while individuals performed ML switching. 480 T2*-weighted echo planar image blood oxygenation level dependent (BOLD) volumes with 32 3mm slices were acquired using the following parameters:

TR=2000ms, TE=30ms, FA=78°, 64 × 64 matrix, 192 × 192 mm field of view, 0.51 ms echo spacing, and 2232 Hz/Px bandwidth. A 1mm resolution MPAGE T1-weighted structural scan was also collected for each individual using the following parameters: TR=2250ms, TE=2.99ms, TI=900ms, 9° flip angle and a 256 × 240 × 192 matrix.

The first 10 functional images were discarded to account for equilibrium effects and the remaining images were pre-processed using a standard pipeline in SPM12 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience). Images were slice-timing and motion-corrected, spatially warped onto the standard Montreal Neurological Institute (MNI) template using the structural scan, and spatially smoothed with an 8 mm full width at half maximum Gaussian kernel.

3.3.4 Activation modelling

Participants's fMRI data were modelled at each voxel using general linear models (GLMs) in SPM12. Psychological predictors for each of the 6 switch types were defined using the onset of the trial stimuli being presented and the duration from this time and when a response was made, and these timings were convolved with the canonical haemodynamic response function. Head motion was modelled using 12 parameters derived from the rigid-body motion-correction stage concatenated with their 1st-order temporal derivatives.

ML switching was performed continuously throughout the run without resting periods. Subsequently, participant-level beta maps were estimated for each switching condition relative to the model constant (implicit baseline). Group-level voxelwise activation was modelled with a within-subject 1-way repeated-measures analysis of variance (rm-anova) with 6 levels (one for each switch domain).

To model activation increases that were general to attention switching, the switching events were collapsed via a positive effect of switching t-contrast was estimated. Effectively, this represents the mean activation increase across switches. Unless otherwise stated, the group-level activation maps were initially thresholded at

$p < 0.01$ uncorrected, followed by a cluster-level $p < 0.05$ false discovery rate (FDR) correction for multiple-comparisons.

Given the statistical concerns surrounding inflated false-positives and cluster-level inference raised by (Eklund et al., 2016) (although do see (Slotnick, 2017)), we conducted an additional non-parametric permutation test of voxelwise switching activation. Specifically, the mean beta map for switching events was calculated for each participant and subjected to an exhaustive 1-sample permutation procedure with sign-flipping and 5-mm variance smoothing via FSL's randomise command line tool (Winkler et al., 2014). Group-level maps were thresholded voxelwise using the default $p < 0.05$ family wise error (FWE) correction and threshold-free cluster enhancement (TFCE).

The parametric and non-parametric maps were compared for spatial similarity using dice coefficients (binarized using voxel significance) and Pearson correlation of the magnitudes from the significant voxel common to both maps.

3.3.5 Region of interest definition

Region of interest (ROI) analyses were conducted in the PFC to compare regional activation across the switching conditions. This was performed in two ways using a data-driven method and a priori coordinate of interest.

The data-driven method used our custom watershed transform to segment the group-level activation maps into discrete clusters (Grant et al., 2018; Soreq, Leech and Hampshire, 2019; Hampshire et al., 2020). This approach does not require k , or the number of clusters, to be defined. The watershed transform considers the 3D statistical volume of an activation map as a multi-dimensional surface where high and low intensities represent elevations. The algorithm iteratively “fills” independent catchment basins (local minima) with unique labels by flooding the various-independent local minima in the statistical volume until each voxel has been labelled. This results in the continuous statistical space being converted into a discrete ROI set that can be non-gaussian in shape, whilst accounting for

contiguous functional regions that have multiple maxima. These tools will be made available to the community in the near future as part of the Fusion watershed (FWS) toolbox.

The second ROI analysis used a priori MNI coordinates to define 5-mm spherical ROIs on the anatomical landmarks demarcated by the rule abstraction hierarchy (Badre & D'Esposito, 2007) and information cascade models (Koechlin et al., 2003). The rule abstraction hierarchy model proposes four ROIs: dorsal premotor (PMd, -30, -10, 68), pre-PMd (-37, 10, 34), inferior frontal sulcus (IFS, -51, 26, 24), aPFC (-36, 51, 9). The information cascade model proposes three ROIs: PMd (-32, -8, 53), pre-PMd (-43, 7, 20), IFS (-50, 35, 23).

3.4 Results

3.4.1 Behavioural performance relates to switch abstraction

Participant's median RTs for correct responses varied across the switching conditions (one-way rm-anova: $F(5, 70)=20.265$, $p<0.001$, $\eta_p^2=0.59$) and the pairwise comparisons (Figure 3.2b) partially differentiated switching abstraction from global difficulty as gauged by the behavioural switch cost. Specifically, the group-level RT was slowest for task-switches (median=2.375, interquartile range (IQR)=0.910), and was significantly slower than dimension-switches (paired t-test: $t_{14}=3.280$, $p=0.014$, confidence interval 95% (CI)=0.11 to 0.53, $d=0.85$), which was significantly slower than target-switches ($t_{14}=2.525$, $p=0.041$, CI=0.03 to 0.36, $d=0.65$), in turn. Target and rule-switch RT did not differ ($t_{14}=1.031$, $p=0.320$, CI=-0.18 to 0.06, $d=0.27$). RT for rule-switching was significantly slower than side-switching ($t_{14}=4.330$, $p=0.004$, CI=0.13 to 0.38, $d=1.12$), and side and comparison-switch RT did not differ ($t_{14}=0.625$, $p=0.542$, CI=-0.10 to 0.18, Cohen's $d=0.16$) (all p-values FDR-corrected).

Therefore, using RT as a behavioural index of cognitive demand controls general difficulty whilst probing switch abstraction across the target and rule conditions. It also provides a broader difficulty axis based on response speed, with low-level

perceptual comparisons and side-switches being the most simple, target, dimension and rule-switches having medium difficulty, and task-switches being the most demanding.

Response accuracy (Figure 3.2c) for all trials was near ceiling (median=95.802%, IQR=7.222), but did significantly vary across switch type (one-way rm-anova: $F(5,70)=6.712$, $p<0.001$, $\eta_p^2=0.32$). Group-median error rates were lowest for the comparison-switching (median=1.9%, IQR=5.2) and side-switching (median=3.0%, IQR=3.8), and these did not differ (paired t-test: $t_{14}=0.195$, $p=0.909$, $CI=-0.02$ to 0.03 , $d=0.06$). Relative to comparison-switching, significantly more errors were evoked by rule-switching ($t_{14}=6.116$, $p<0.001$, $CI=0.05$ to 0.10 , $d=1.43$), dimension-switching ($t_{14}=3.288$, $p=0.016$, $CI=0.02$ to 0.08 , $d=0.83$) and task-switching ($t_{14}=4.377$, $p=0.005$, $CI=0.03$ to 0.07 , $d=1.15$). Rule-switching errors were also greater than for target-switching ($t_{14}=3.288$, $p=0.016$, $CI=0.01$ to 0.07 , $d=0.65$), but rule-switching did not differ from dimension-switching ($t_{14}=1.358$, $p=0.267$, $CI=-0.01$ to 0.06 , $d=0.31$) or task-switching ($t_{14}=1.606$, $p=0.196$, $CI=-0.01$ to 0.05 , $d=0.39$) (all p-values FDR-corrected).

Individuals varied in the total number of trials performed and this ranged between 173-294 responses across participants (group-median=250, interquartile range (IQR)=68.75).

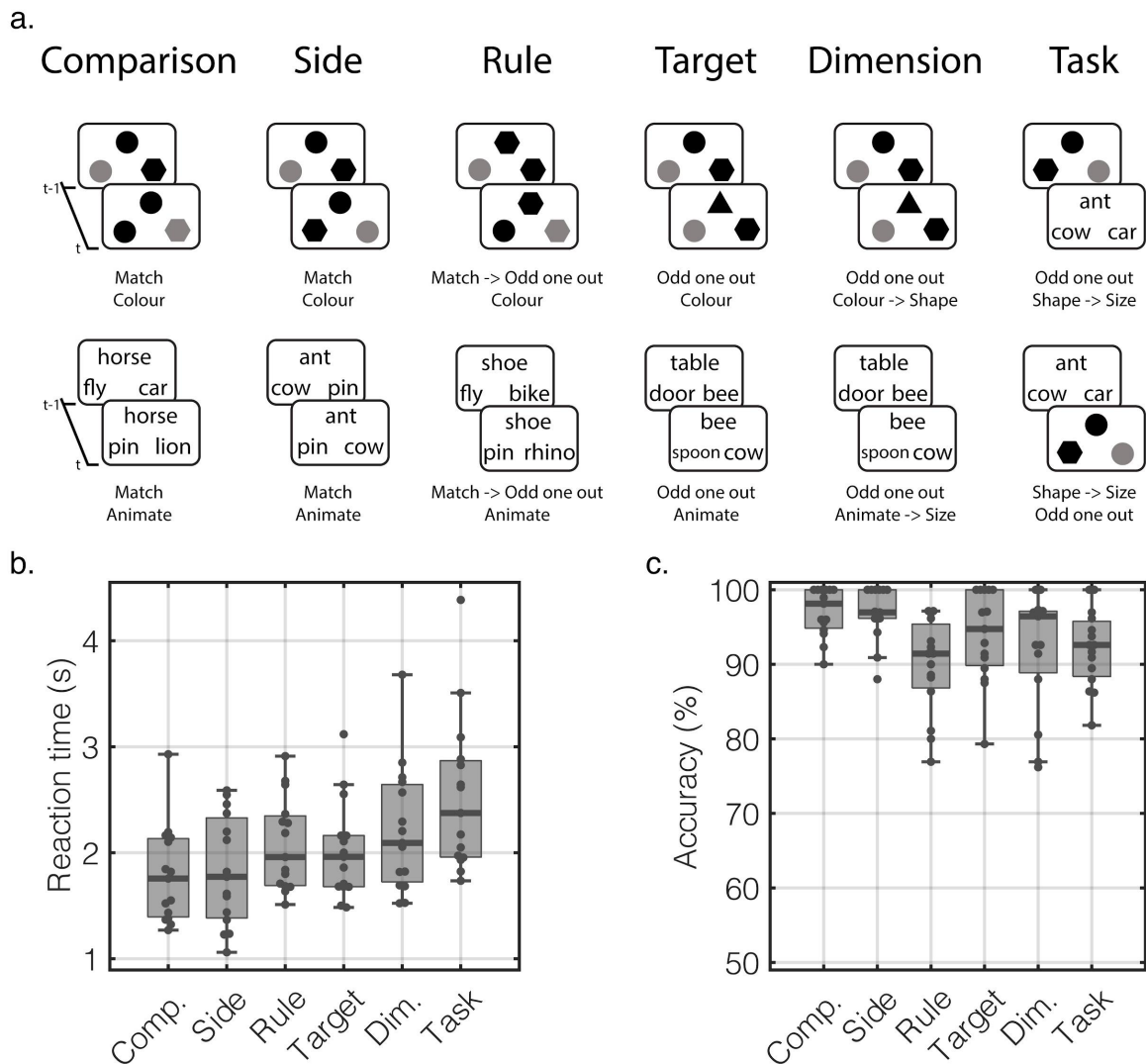


Figure 3.2 - Multilevel switching (ML) paradigm switch costs. a) Each trial (t) involved a switch relative to the previous trial ($t-1$). Switches could be low-level perceptual (comparison, side) or higher-level switches where the mappings between the target and the two probes changes, or task-switches between the words and object tasks. Examples of the 6 switch conditions are given here with pseudo-stimuli for the objects task (top row) and word task (bottom row). b) Reaction time boxplots in seconds (s) for each switch type. Each datapoint represents an individual's median reaction time during trials for each switch type where a correct response was made. c) The percentage of correct responses made for each condition. Each datapoint represents an individual. For b & c, boxplots are ordered by the group median reaction time. The central mark represents the median, the edges of the box are the 25th and 75th percentiles, dots outside of the whiskers are outliers.

3.4.2 Switching activates a broad network of brain regions

Activation during switching was modelled in a one-way rm-ANOVA (6-levels). Collapsing across switch types, switching increased activation across a broad set of lateral frontoparietal, anterior insula, occipital and cerebellar regions (Figure 3.3a). The inverse contrast rendered activation decreases in ventromedial PFC (vmPFC), posterior cingulate cortex (PCC) and in the temporal parietal junction (TPJ). These stimulus-driven switching activation changes replicate changes observed during self-ordered switching (Daws et al., 2020).

The equivalent permutation analysis on individuals mean switching activation generated a highly consistent group-level activation map (Supplemental Figure 3.1), both in terms of which voxels exhibited a significant effect ($DCE=0.716$) and the spatial variation in magnitude across the overlapping voxels (pearson: $r=0.867$, $p<0.001$, $CI=0.86$ to 0.87).

In a supplementary analysis, switching activation was collapsed within each task (objects, words) and then directly compared with t-contrasts (Supplemental Figure 3.2a). For the word greater than object contrast, significantly greater activation for switching between arrays of words was observed in left lateral PFC and medial parietal and occipital cortices. For the object greater than word contrast, significantly greater activation for switching between arrays of objects was observed in dorsolateral parietal regions, lateral occipital cortex and medial PFC.

3.4.3 PFC activation during switching does not reveal a hierarchy

The group-level activation result indicates that a broad network of brain regions show an increase in activation in response to general switching demands. To test for evidence of a hierarchy in the frontal lobes, we next compared the differences in activation magnitude across ROIs in the frontal lobes.

We segmented the switching activation contrast for the contrast of all switches (Figure 3.3a) using a custom watershed transform (see methods). This separated switching activation in the left frontal lobe into 5 ROIs (Figure 3.3b) that closely aligned with the proposed coordinates of the rule abstraction hierarchy (Badre and D'Esposito, 2007) and the information cascade model (Koechlin, Ody and Kouneiher, 2003): dorsal premotor (PMd), pre-PMd, IFS, aPFC and, additionally, the anterior insula (aINS).

For each ROI and participant, we used the group-level activation (Figure 3.3a) to extract a weighted-mean beta activation that was subsequently z-scored within-individual and compared in a two-way rm-anova with ROI (5 levels) and switch type (6 levels) as within-subject factors. There were significant main effects of ROI ($F(4, 56)=7.175$ $p<0.001$, $\eta_p^2=0.34$) and switch type ($F(5, 70)=7.190$, $p<0.001$, $\eta_p^2=0.34$). Critically though, there was no ROI * switch type interaction ($F(20, 280)=0.960$, $p=0.450$, $\eta_p^2=0.06$) (p-values Greenhouse-Geisser adjusted).

The main effect of ROI was examined by collapsing activation across switch type for each ROI (Figure 3.3c). Switching activation did increase along an anterior axis among the posterior frontal ROIs, in some accordance with the information cascade model (Koechlin, Ody and Kouneiher, 2003). Activation in the most-posterior PMd ROI was significantly less than in the pre-PMd (paired t-test: $t_{14}=-3.286$, $p=0.011$, $CI=0.27$ to 1.28 , $d=0.85$). While pre-PMd and IFS activation did not significantly differ ($t_{14}=-1.175$, $p=0.312$, $CI=-0.71$ to 0.21 , $d=0.30$), the IFS activation difference with PMd was larger in terms of effect size ($t_{14}=3.733$, $p=0.007$, $CI=0.44$ to 1.61 , $d=0.96$).

Despite this, activation in IFS was significantly greater than in aINS ($t_{14}=2.815$, $p=0.021$, $CI=0.16$ to 1.16 , $d=0.73$) and the aPFC ($t_{14}=6.494$, $p<0.001$, $CI=0.76$ to 1.52 , $d=1.68$) and, critically, activation in most-anterior aPFC ROI did not differ to the most-posterior PMd ROI ($t_{14}=0.359$, $p=0.725$, $CI=-0.58$ to 0.82 , $d=0.09$) (all ROI

comparisons FDR-corrected). This provides some evidence of a posterior-anterior hierarchy for switching predicted by the information cascade model (Koechlin, Ody and Kouneiher, 2003), but does not replicate the rule abstraction model (Badre and D'Esposito, 2007). However, these comparisons were for switching activation, in general.

The main effect of switch type was then explored by collapsing activation across ROI. Comparison and side-switching activation did not differ (paired t-test: $t_{14}=-0.393$, $p=0.700$, $CI=-0.55$ to 0.38 , $d=0.10$). Relative to side-switching, activation was significantly greater for rule-switching ($t_{14}=3.598$, $p=0.018$, $CI=0.35$ to 1.38 , $d=0.93$) and target-switching ($t_{14}=3.040$, $p=0.027$, $CI=0.13$ to 0.77 , $d=0.79$). rule-switching and target-switching activation did not differ ($t_{14}=1.840$, $p=0.173$, $CI=-0.07$ to 0.88 , $d=0.48$). Relative to target-switching, dimension ($t_{14}=-0.944$, $p=0.542$, $CI=-0.43$ to 0.17 , $d=0.24$) or task-switching ($t_{14}=-0.532$, $p=0.700$, $CI=-0.51$ to 0.30 , $d=0.14$) activation did not differ (all condition comparisons FDR-corrected).

3.4.4 Supplemental voxelwise analysis of cognitive demands does not reveal an abstraction hierarchy or selectivity in PFC

We ran an unconstrained voxelwise analysis that directly tested the effect of cognitive demand on activation (Figure 3.3e) by weighting each condition by the group-median RT order. This contrast did not render activation increases in frontopolar cortices for more demanding switches (e.g., dimension & task-switching). Nor did contrasting task-switching, the most difficult switch, selectively activate any region within the frontopolar cortex, when contrasted with the easiest condition, comparison-switching. Instead, areas of suprathreshold activation increase were distributed within a left lateralized set of regions comprising areas of multiple-demand cortex and the dorsal visual streams (Figure 3.3e). The inverse contrast did not render any regions of greater activation for less difficult switches at the cluster-corrected threshold.

Repeating the cognitive demand contrast, but for switching events separated by task (words, objects) generated the same result. Both tasks showed activation increases with switching demand across anterior and posterior PFC (Supplemental Figure 3.2b).

3.4.5 Cascade and rule abstraction ROIs show no evidence of an abstraction hierarchy in PFC

As a final step, we used the exact landmarks outlined by the cascade and rule theories to test for evidence of a functional hierarchy. 5mm spheres, centred on the ROI coordinates, were used to extract activation beta estimates from each switching condition and subject, and modelled in a ROI * switch type rm-anova. The information cascade ROIs (inferior frontal sulcus (IFS), dorsal premotor (PMd), and prePMd) showed a main effect of switching condition ($F(5, 70)=6.275$, $p=0.001$, $\eta_p^2=0.31$), and no main effect of ROI ($F(2, 28)=0.622$, $p=0.527$, $\eta_p^2=0.04$) or significant interaction ($F(10, 140)=1.500$, $p=0.201$, $\eta_p^2=0.10$) (p-values Greenhouse-Geisser adjusted). The hierarchy abstraction ROIs (aPFC, IFS, PMd, prePMd) showed a main effect of switching condition ($F(5, 70)=4.481$, $p=0.009$, $\eta_p^2=0.24$), ROI ($F(3, 42)=13.095$, $p<0.001$, $\eta_p^2=0.48$) and no significant interaction ($F(15, 210)=1.754$, $p=0.122$, $\eta_p^2=0.11$) (p-values Greenhouse-Geisser adjusted).

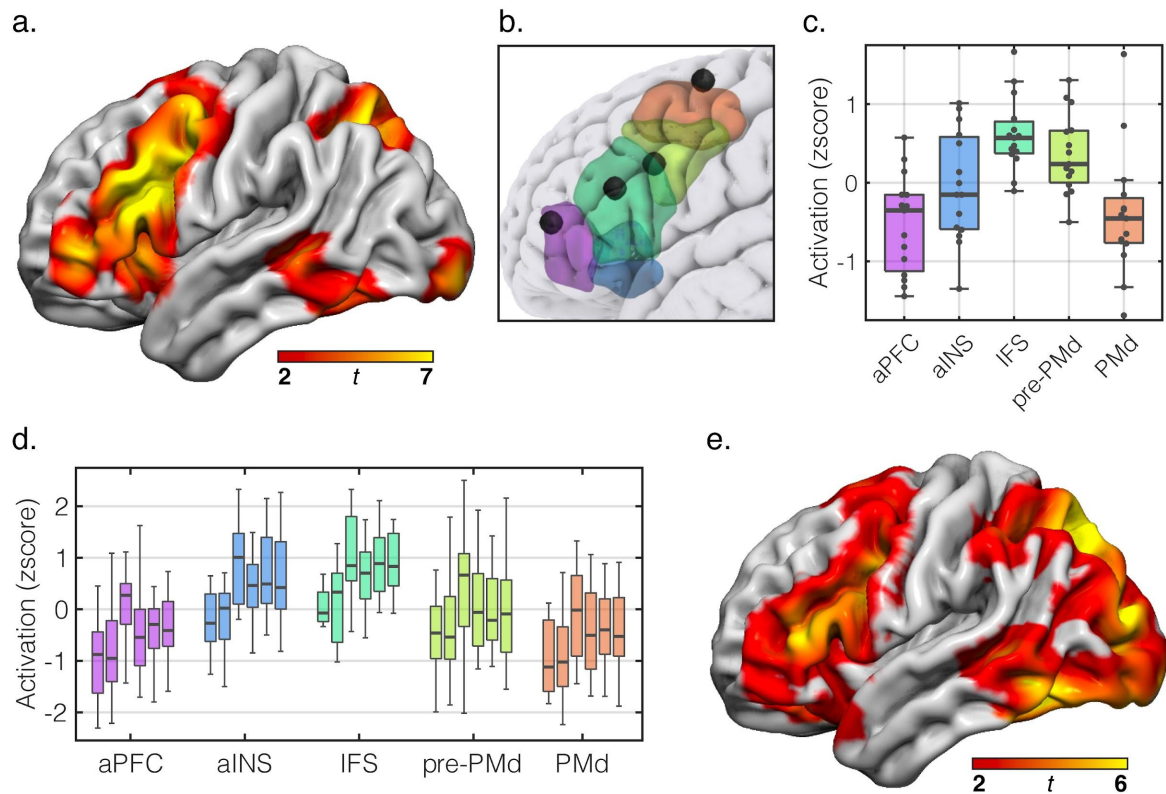


Figure 3.3 - Uniform activation increases in lateral frontal cortex (LFC) with increasing switching demands. a) Activation increases associated with switching, in general. b) LFC ROIs derived from a watershed transform of the switching activation map (a). Coordinates from the rule abstraction hierarchy (Badre and D'Esposito, 2007) are rendered as 5mm radius black sphere overlays. c) ROI mean activation across switching conditions (z-scored within subject). d) ROI mean activation for each ROI and condition (left-to-right: comparison, side, rule, target, dimension, task). e) Supplemental voxelwise t-contrast testing the effect of cognitive demand, indexed by RT, on activation. All maps are thresholded at $p < 0.01$ uncorrected, followed by a $p < 0.05$ cluster-correction. (aPFC=anterior prefrontal cortex, aINS=anterior insula, IFS=inferior frontal sulcus, PMd=dorsal premotor cortex).

3.5 Discussion

The aim of this study was to test predictions from alternative perspectives on human frontal lobe functional organisation. Neither perspective fully predicted the observed results.

A novel feature of the task design was that switches at different levels of abstraction were partially differentiated from general difficulty, as gauged by the response time costs. From the perspective of models that assert that the functions of the frontal lobes are organised hierarchically along a posterior-anterior abstraction axis (Koechlin, Ody and Kouneiher, 2003; Badre and D'Esposito, 2007), we predicted that switching costs associated with changes to rule mappings should activate more anterior frontal regions than those associated with stimulus dimension or target-switches. We also predicted that task-switches should specifically activate the most anterior PFC region (Koechlin et al., 1999; Dove et al., 2000; Rushworth, Passingham and Nobre, 2002; Braver, Reynolds and Donaldson, 2003). We could find no evidence to support a posterior-anterior functional axis. We also found no evidence that the aPFC plays a unique or specialised role in attentional switching, despite this being a prevailing view in the literature.

Instead, our findings demonstrate that cognitive control, responsible for the reconfiguration of attentional resources during switching, engages highly distributed neural systems. More specifically, switching-associated activation was evident throughout frontoparietal, subcortical and cerebellar regions. This included aPFC, even at the lowest levels of switching demand. This highly distributed pattern replicates several task-switching studies (Hampshire and Owen, 2006; Cusack, Mitchell and Duncan, 2010; Kim et al., 2012; Daws et al., 2020) and accords poorly with the notion that specific cognitive control processes are localisable to specific regions, including the tight mapping of switching to frontopolar cortex, as has been previously proposed (Koechlin et al., 1999; Rushworth, Passingham and Nobre, 2002; Braver, Reynolds and Donaldson, 2003).

Conversely, the multiple-demand perspective in its original formulation does predict that there should be patterns of distributed activation during switching conditions, in general (Duncan and Owen, 2000). However, it also has been proposed that the extent of this system is flexible and will recruit additional neural resources, typically anterior frontal regions, in response to general increases in cognitive demand (Duncan, 2013; Crittenden and Duncan, 2014).

This failure to validate either prediction presents an intriguing conundrum. A wealth of imaging studies have reported functional dissociations within the brain volume that comprises multiple demand cortex (Christoff et al., 2001; Ullsperger and von Cramon, 2001; Hampshire, Duncan and Owen, 2007; Parkin et al., 2015; Crittenden, Mitchell and Duncan, 2016; Mineroff et al., 2018). In many cases those dissociations have been robustly reproduced. The latter refinement of the MD perspective provides an explanation of why such dissociations might be evident. Therefore, it differs from hierarchical models as the increased anterior spread of task related activity within the PFC is predicted to be general for cognitive demands, irrespective of their abstraction (Jiang and Kanwisher, 2003; Crittenden and Duncan, 2014). The fact that no such anterior spread at higher switching demand was evident in the current study accords poorly with that explanation.

Indeed, for either of these accounts to be accurate, they must describe fundamental properties that generalise across contexts. When taken in combination, the results presented here and in previous studies accord with the notion that functional dissociations do exist within the lateral PFC, and that there may even be some level of hierarchy in network dynamics or regional function. However, the notion of an axis with concrete-abstract processing may be overspecified as it is based on results that do not generalise to tasks that probe related processes using somewhat different designs.

Arguably, this study did provide some evidence of a functional dissociation across lateral PFC regions. For example, the watershed transform parcelated the

task-evoked PFC activation into regions centered on distinct peak foci, and those ROIs aligned well with areas previously reported to functionally dissociate in hierarchical models (Koechlin, Ody and Kouneiher, 2003; Badre and D'Esposito, 2007). Furthermore, those regions showed differential sensitivity to switching. However, the difference in activation presented as an inverted U-shape along the posterior-anterior axis, with activation in the most anterior ROI (aPFC) similar to the most posterior ROI (PMd). Moreover, activation uniformly increased in all PFC regions with cognitive demand and regardless of abstraction.

A more accurate account of the data from this study comes from the multiple-demand perspective as originally couched, or relatedly, the global workspace model (Dehaene, Kerszberg and Changeux, 1998). However, when taking a broader perspective, the challenge is to determine the functional basis of previously observed dissociations, and possible hierarchy, within MDC.

We propose several possibilities for future investigation. One possible explanation is that the apparent spread of activity with increased difficulty is a variant on the “imager’s fallacy” (de Hollander et al., 2014). Specifically, application of a binary threshold to statistical maps can produce the impression of functional dissociations. Here, anterior PFC voxels were somewhat less active for switching in general, but activity increased with general difficulty. Although the latter effect was uniform across ROIs, if more anterior regions were just below and more posterior regions just above the statistical threshold at the lower switching level, this could lead to the illusion of a selective sensitivity of anterior region under higher switching demand. This same issue holds true for the notion of an increased spatial spread of MD with difficulty. This makes particular sense if MD is itself not considered to have a hard boundary, but instead, to capture a gradient of domain generality, where each brain region or voxel has some varying level of involvement across task contexts.

This latter notion of domain generality accords well with the observation of functional dissociations within MD, given the capacity of multivariate machine learning algorithms to identify tasks based on both activation and connectivity

patterns therein (Crittenden, Mitchell and Duncan, 2016; Soreq, Leech and Hampshire, 2019). Relatedly, we have argued that it may be inappropriate to impose strong functional dissociations on what may essentially be a multivariate system that supports tasks through a many-to-many mapping system (Hampshire and Sharp, 2015; Lorenz, Hampshire and Leech, 2017). Such a relationship would explain discrepancies in mappings from studies using similar but non-identical task designs, that is due to the imposition of overly narrow experimental designs and hypotheses based on narrow sampling of task designs (Lorenz, Hampshire and Leech, 2017).

A further potential explanation relates to a series of studies where we have reported reproducible evidence of functional hierarchies across the frontal cortices, in terms of connectivity and regional activation analyses (Erika-Florence, Leech and Hampshire, 2014; Parkin et al., 2015; Hampshire et al., 2016, 2019). These studies highlight that brain regions show changes in activation patterns and dynamics even when performing the same task but that are dependent on different stages of learning. Specifically, as individuals practise novel tasks their response speed increases and trial-evoked activation within PFC steadily decreases along an anterior-posterior axis. Taken together, these “learning curves” demonstrate aPFC recruitment during demanding contexts that can be transient as individual’s behaviours become optimised and, consequently, global cognitive demand falls.

From this, we speculate that within-task learning occurs at different rates for more and less complex task manipulations, and if not accounted for, and an average of all trials of a type is taken, that could produce the illusion of greater aPFC activity under more complex conditions, when differential placement on the learning curve may underlie such differences.

It is important to address the present methodological approach and determine if this precluded us from observing evidence of a hierarchy or multiple-demand dissociation. First, it could be argued that our ML switching paradigm provided switching-magnitudes and variability across switch types that were insufficient to

detect switching-specific aPFC activation or differences between PFC regions that would suggest a hierarchy. We consider false-negatives to be unlikely in this context as behavioural performance showed a significant variability across switch types. The comparison of the switches with low (side-switch) vs high (task-switching) abstraction showed a substantial ~700ms difference in RT. Furthermore, there was sufficient power to detect significant differences in aPFC activation for multiple contrasts; however, these same contrasts were also evident elsewhere.

Second, our comparison of activation across the PFC was conducted using ROIs that were constructed in a data-driven manner. Importantly, these ROIs showed a high correspondence with, and contained, coordinates previously defined by the cascade (Koechlin, Ody and Kouneiher, 2003) and abstraction hierarchy theories (Badre and D'Esposito, 2007). However, to abate potential concerns that the focused ROI analysis was biased, we performed an unconstrained whole-brain analysis which tested the effect of switching-abstraction on voxelwise activation. Both contrasts supplemented and extended the ROI analysis by demonstrating that switching-abstraction, and direct comparisons between the least and most challenging switches (task>comparison), increased activation across PFC and beyond in bilateral occipital cortices dorsal visual streams. Strikingly, this absence of evidence in support of a functional hierarchy was also the case when directly measuring activation from the exact MNI coordinates previously defined by both hierarchical models.

Finally, it could be argued that activation increases for more abstract switches are explainable by these switches being more challenging and therefore requiring individuals to attend to stimuli for longer periods. This is unlikely to be the case here, as we used individual trial RT to model activation per unit time, which essentially factors out activation differences due to visual-attentional processing times. Therefore, it is the case that even taking into account those differences, uniform increases in activation per unit time were evident for more difficult switching conditions.

3.6 Limitations

A central challenge to this investigation is in the attempt to dissociate abstraction from difficulty. It is challenging to separate these constructs on a conceptual level. Operationally, abstraction can refer to the hierarchical arrangement of multiple rules or policies that are applicable, override or conjunct with rules lower down in the hierarchy. Difficulty can be considered as the extent to which a task can be performed accurately and efficiently. Concurrently, difficulty relates to the extent to which competing tasks interfere or a task places a load on cognitive processes. Difficulty is often subjectively experienced as greater 'effort' and interacts with motivation.

It is reasonable to consider these constructs, abstraction and difficulty, as interrelated. Tasks that require greater abstraction tend to be more difficult as establishing the hierarchy and integration of multiple rules increases load. Matters are further complicated by task learning; whereby, the load is minimised as stimuli-response mappings, and mappings between rules, are reinforced with practice.

Despite this, the multilevel switching paradigm used here in Chapter 3 does attempt to *partially* dissociate abstraction and difficulty. Essentially, the amount of information being updated was modulated on every trial but the nature of the update varied. Abstraction was modulated by the nature of the change in the stimulus-response mapping while difficulty was modulated by the amount of information that changed on the screen. For example, rule or dimension switches (Figure 3.2) required an update of the current task mappings, such as switching from match to odd one out or from shape to colour. These switches involved minimal changes to the objects on screen (although features may have switched between the probes). In contrast, target switches did not change the task mappings but presented a new target on the screen and required an update of the mappings between the new target and the current probes.

In both cases, these “abstract” or “difficult” updates evoked similar increases to RT, response errors and frontal lobe activation. This raises an interesting possibility relating to whether these constructs are underpinned by the same processes or distinct processes evoke similar changes in behaviour and BOLD activation. The current investigation is limited as it may not be able to equivocate between these two possibilities.

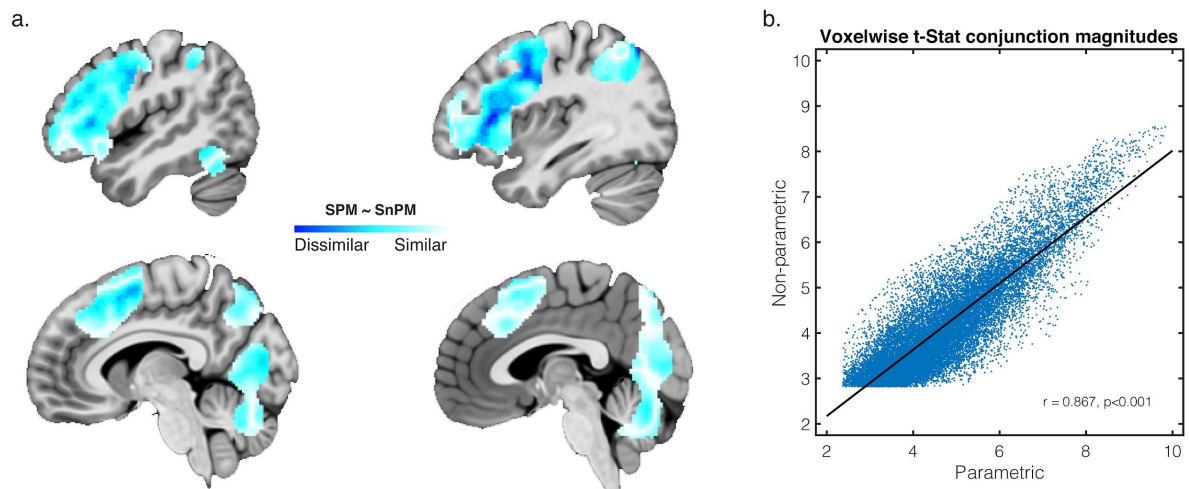
Relatedly, interpretations of cognitive processes from the observed fMRI BOLD signal should be *cautious*. Whatever the resolution that may be available for dissociating cognitive processes at the conceptual level, it is unlikely that this can be matched by fMRI. A central issue for fMRI is identifying the correct domain in which mental constructs can be appropriately mapped (Poldrack & Yarkoni, 2016). The spatial domain is complicated by the same brain regions activating to distinct processes and regions coactivating as part of distributed functional networks, a.k.a “one-to-many” and “many-to-many” mapping issues (Lorenz et al., 2017). The current resolution available to fMRI can be particularly problematic if distinct processes are supported by separate but densely overlapping networks. Hypothetically, the present experiment would not be able to dissociate an abstract vs a difficult update if these are separate processes and if they are supported by spatially overlapping networks (Soreq et al., 2021). The blurry lens of fMRI activation would generate similar observations, in this case. However, future work could be conducted within the domains of temporal dynamics and multivariate connectivity patterns which may provide the means to tease these processes apart (Soreq et al., 2019, 2021).

3.7 Conclusions

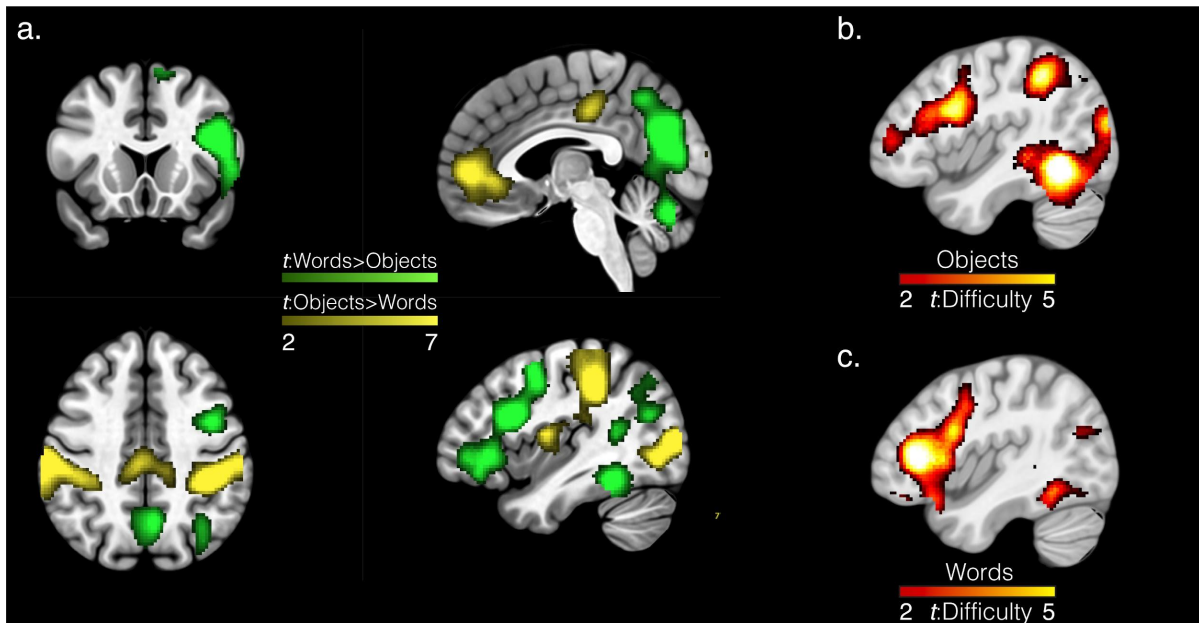
In summary, our results challenge predictions of prominent theories of abstraction and multiple-demand regarding the functional organisation of the frontal lobes. We demonstrate that neither of these models fully accounts for the distributed and uniform activation increases in the frontal lobes during switches with increasing

abstraction and cognitive demands. However, the MD perspective in its original form is closer to the mark when explaining the results reported here. Further work is required to better specify a fundamental organisation of the lateral PFC that can be demonstrated to generalise across task contexts.

3.8 Supplemental figures



Supplemental Figure 3.1 - Comparison of parametric and non-parametric activation maps. Group-level activations maps derived from two different approaches: A parametric approach using the SPM software and a non-parametric approach using the FSL randomise permutation tool. a) Activation similarity map of the intersection of significant voxels. Voxels are weighted by absolute differences between maps (normalised by the largest voxelwise difference). Light colours indicate greater similarity between maps. b) Scatterplot of voxel t-statistical values from the parametric and non-parametric maps. Each point represents a single voxel.



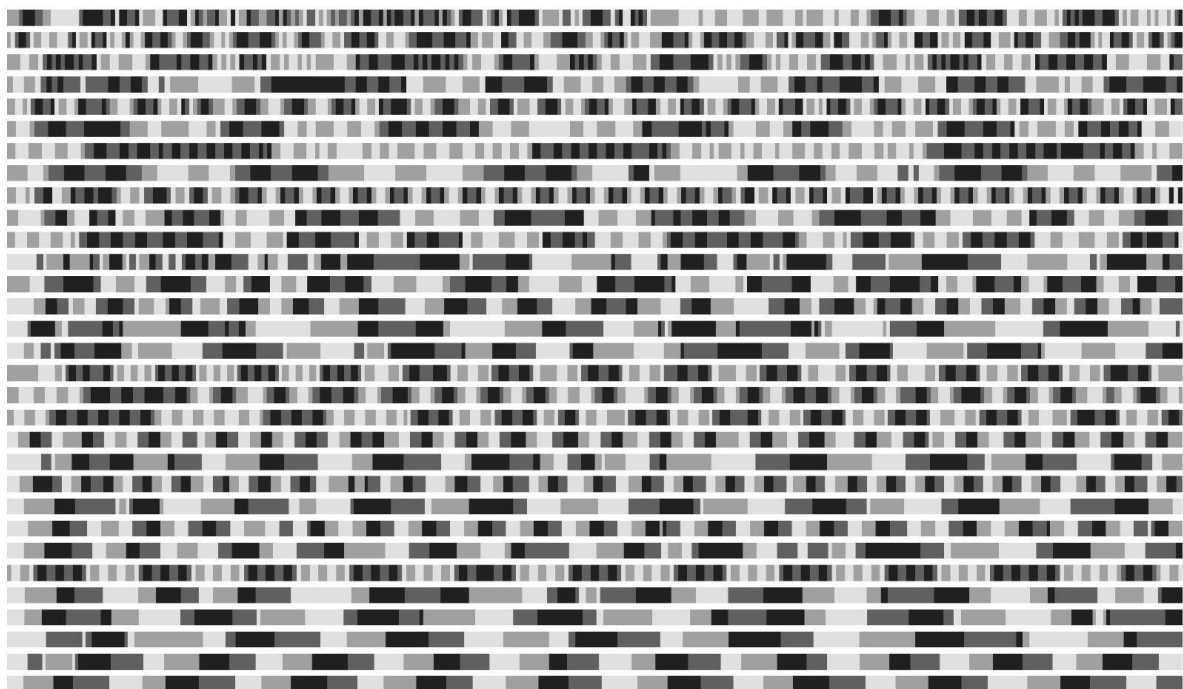
Supplemental Figure 3.2 - Switching activation contrasts broken down by task (words, objects). a) The average activation for switches that involved arrays of words or objects directly compared using t-contrasts. Green: Words > Objects; Yellow: Objects > Words. b/c) Cognitive demand t-contrasts representing activation increases for more demanding switch events. These contrasts are equivalent to the contrast rendered in Figure 3e but the switch events were separately modelled for objects (b) and words (c). All maps are thresholded using an initial $p < 0.01$ voxel threshold followed by an FDR cluster-level correction for multiple comparisons.

“Fortune favors the prepared mind.”

— **Louis Pasteur**

Chapter 4: Optimisation of functional network resources when learning behavioural strategies for performing complex tasks

Building on chapter 3, I analyse fMRI data from a novel task-switching experiment that was designed and collected by Dr Adam Hampshire. Addressing aim 2 of this thesis, in two studies, I examine how healthy adults learn to mitigate switching-costs by implementing structured behavioural routines. The findings from this experimental chapter are prepared for submission and, subsequently, the use of “we” refers to my predictions, analysis, data visualisation and interpretation.



The ability to create an efficient behavioural routine varies across individuals.

4.1 Abstract

We developed two novel self-ordered switching (SOS) fMRI paradigms to investigate how human behaviour and underlying network resources are optimised when learning to perform complex tasks with multiple goals. SOS was performed with detailed feedback and minimal pretraining (study 1) or with minimal feedback and substantial pretraining (study 2). In study 1, multiple-demand (MD) system activation became less responsive to routine trial demands but more responsive to the executive switching events with practice. Default Mode Network (DMN) activation showed the opposite relationship. Concomitantly, reaction time learning curves correlated with increased connectivity between functional brain networks and subcortical regions. This ‘fine-tuning’ of network resources correlated with progressively more routine and lower complexity behavioural structure. Furthermore, overall task performance was superior for people who applied structured behavioural routines with low algorithmic complexity. These behavioural and network signatures of learning were less evident in study 2, where task structure was established prior to entering the scanner. Together, these studies demonstrate how detailed feedback monitoring enables network resources to be progressively redeployed in order to efficiently manage concurrent demands.

Keywords: Cognitive control, learning, behavioural structure, task-switching, fMRI.

4.2 Introduction

The ability to rapidly learn complex tasks is a definitive feature of human cognition. Studies of cognitive control suggest that the human brain achieves this by rapidly adapting to encode behavioural routines (Finc et al., 2020; Khambhati et al., 2018), also known as a ‘task set’, comprising the combinations of rules, strategies and stimulus-response mappings that enable optimal performance (Sakai, 2008). This encoding is typically associated with distributed activation of the multiple-demand (MD) system (Duncan, 2013).

During instruction based learning (IBL), generating and establishing a simple routine imposes a cognitive demand that rapidly decreases with practise. This learning can be accompanied by a decay in MD activation, particularly in anterior prefrontal cortex (aPFC), as reaction time (RT) decreases (Hampshire et al., 2016, 2019). The established behavioural routine can be robustly identified by task-specific multivariate patterns of the brain’s functional connectivity (FC) (Soreq et al., 2019, 2021) and these patterns are predictive of behavioural performance (Rosenberg et al., 2016; Soreq et al., 2021). As individuals learn a task, FC is dynamic and becomes more stable with increased FC within cortical networks (Hampshire et al., 2016) and in their connectivity with subcortical regions (Antzoulatos & Miller, 2014; Bassett et al., 2011).

In everyday life, the optimal behavioural routine is often not evident when first performing tasks. This can be due to a lack of clear instruction, the inherent complexity of the task, or shifting demands. Consequently, it is necessary to monitor outcomes and progressively update behavioural routines in order to optimise behaviour (Sharp et al., 2010; Swick & Turken, 2002). This updating of routines is critical for flexible behaviour (Ionescu, 2012); however, it comes with a cost and is associated with increases in reaction time (RT) and heightened MD activation (Daws, Soreq, et al., 2020; Hampshire & Owen, 2006). Furthermore, it often is necessary to juggle multiple concurrent tasks, or to manage sub-routines of complex tasks, and in the process, to monitor multiple outcomes. Increased task

complexity poses a challenge for flexible cognitive systems due to their intrinsic limitations in capacity (Duncan, 2013). Under such conditions, humans have a tendency to impose hierarchical structure to their behavioural routines (Fallon et al., 2013).

This tendency towards self-ordered structure is underpinned by decisions regarding when and in what order to switch attention between competing tasks. Consequently, switching paradigms provide a natural choice for studying the cognitive control processes that enable flexible human behaviour. However, task-switching paradigms typically control when switches occur (Kiesel et al., 2010), which does not capture the self-ordered nature of real-world human behaviour (Hampshire & Owen, 2006). Despite this limitation, there is evidence that cognitive demand is minimised with learning and by structuring behaviour, where possible. Task-switching imposes a considerable cognitive demand but switching-costs to RT will decrease with practise (Yeung & Monsell, 2003b), although a residual cost persists in humans even after tens of thousands of trials (Stoet & Snyder, 2007). Moreover, switching-costs decrease with preparation time and also when switching events are predictable (R. D. Rogers & Monsell, 1995). During voluntary task-switching (VTS), participants are less likely to switch under high cognitive load (Demanet et al., 2010) and can exhibit a repetition bias despite being instructed to evenly and randomly switch between tasks (Arrington & Logan, 2004, 2005).

These findings indicate that both learning and the structuring of behaviour are important mechanisms for minimising cognitive demand when managing competing task demands. Here, we investigate this further using functional MRI (fMRI) and two novel “self-ordered switching” (SOS) paradigms. The SOS paradigm requires participants to learn through a process of feedback-driven trial and error how to manage their time between two tasks, each of which has two rules that the participants selects. To score maximum points, the participant must spend time on all four sub-tasks. In two studies, the concurrent optimisation of brain function and behaviour was examined as participants either actively learnt SOS demands based

on multiple feedback elements (study 1) or, as a control condition, performed SOS with minimal feedback only and after receiving extensive pre-training (study 2).

We tested the hypothesis that in study 1, where participants were performing the task for the first time and with detailed feedback, self-ordered behaviour would optimise with practise. It was predicted that RT would decrease and that behavioural routines would become increasingly structured. Using fMRI measures of network activation and FC, we hypothesised that brain function would vary as behaviour optimised. Specifically, we predicted that MD activation would decrease with practise for the simple trial events, but that these learning curves would be less rapid for the more demanding switching events. We predicted that FC increases between functional networks associated with cognition would accompany the learning of optimal behavioural routines, reflecting more efficient network processing. Finally, we expected that in study 2, where the task had been pre-trained and feedback was minimal, that participants would show little or no behavioural learning effects.

4.3 Methods

4.3.1 Participants.

31 healthy right-handed participants were recruited in total: **Study 1** - n=16, mean age 24.75 years, range 19-40, 10 female; **Study 2** - n=15 mean age 26.47 years, range 19-42, 11 female. Approval for this study was received by the Cambridge university research ethics committee and participants gave written informed consent. Participants were reimbursed £20 for their time.

4.3.2 Self-ordered switch (SOS) paradigms.

We designed two SOS paradigms that required participants to manage their time between two concurrent tasks that each had two subordinate rules in order to gain maximum points within 20 minutes. There was a matching task (MT) and an odd one out (OOO) task, both involving simple visual discriminations of objects (Figure 1a). The subordinate rules of each task were differentiated by the requirement to apply the rules to either colour or shape. During each trial, arrays of simple colour-shape objects were presented and participants made a visual discrimination response based on the current task and rule.

Critically, in order to keep gaining points, participants had to organise their time across all four sub-tasks. This meant that they chose whether to repeat the same condition (repeat-trial) or to switch between the tasks or the rules (task-switch, rule-switch). Responses were made with a button positioned under the right hand. Each trial continued until a response was made using either a left or right button press. The task and rule task did not change between trials unless the participant opted to task-switch or rule-switch by pressing the up or down button respectively.

When performing MT trials, a target and two probe objects (each with a shape and colour) were presented and the participant indicated via a left or right button press which of the two probes matched the target given the currently selected shape or colour rule. Similarly, during OOO trials, a 2x2 object grid was presented and the

participant indicated with a left or right button press which side of the display contained the object that differed from the others given the currently selected colour or shape rule.

The task and rule discrimination criteria were designed to be simple enough that individual trials could be performed with high accuracy and minimal effort. This ensured that individuals could focus on the executive requirement to learn how to divide their time across the tasks and rules in order to gain maximum points (see detailed criteria below). In both studies, SOS was continuously performed inside the scanner for 20 minutes with no resting periods. The studies differed with respect to the training administered and the response feedback received during the scanning session.

In study 1, participants played a short demo, during which they were required to undertake each of the four task-rule combinations and each of the switch types once (rule or task switch). On screen instructions were sequenced to orient the participant towards the feedback elements (described below) and controls for responding and switching. They then entered the scanner and undertook the SOS paradigm in a single continuous 20 minute acquisition. Both tasks were displayed on the screen with the MT at the top and the OOO task at the bottom (Figure 1c). The currently active task was indicated with a red border, the inactive task had a grey border (these border colours flipped when the task-switch button was pressed). The active rule was displayed with a text cue in the centre of the task panel (“COLOUR” or “SHAPE”, in yellow) and flipped when the rule-switch button was pressed. On each trial, new stimuli were displayed and participants were able to repeat the condition or could task-switch and/or rule-switch once prior to performing the trial within the selected panel. A feedback array was then immediately displayed for 1.5 seconds followed by the stimuli for the next trial (Figure 1c).

Participants learnt to optimally sequence their responses based on a hierarchically arranged set of performance meters (supplemental Figure 4.1a-b). These meters

were as follows. (1) An overall score meter was located at the centre of the screen alongside (2) two meters indicating the number of points currently accumulated for each of the two tasks. These were presented continuously on the screen. Additionally, the feedback array after each trial comprised (3) text indicating positive or negative feedback for the present trial response (“CORRECT” in green or “INCORRECT” in red) and (4) four meters, showing points currently accumulated on each of the sub-task (i.e., colour and shape for MT vs OOO).

These score elements interacted hierarchically. Specifically, sub-task meters would increase after a correct response to the corresponding condition (e.g., a correct response to a MT colour trial) and would decrease if an incorrect response was made. All meters ranged between one to ten and were initialised with five points at the start of the experiment.

Within each MT and OOO task, the aim was to respond to, and switch between, the task’s colour and shape rules in a sequence that would build up both rule meters to the maximum level of ten (supplemental Figure 4.1c-e), at which point, the two subtask meters would reset to five, and the corresponding high-order task meter and overall score incremented by 1. To encourage switching, if one subtask meter got further than two points ahead of the other, then the other meter decreased by a point (supplemental Figure 4.1d).

Switching between tasks was encouraged in a similar way to rule switching as the inactive task meter would decrease if the active task meter was two or more points ahead. Successfully dividing time between each task would build up both task meters, thereby increasing the overall score. When both task meters reached the maximum level of ten, both meters would reset to the initial position and a +10 point bonus was awarded to the overall score (supplemental Figure 4.1f).

Given that switching would be costly in terms of both effort and time, and the aim was to score as many points as possible per unit time, optimal behaviour could be defined as having a routine that used the minimal number of switches required to

ensure sufficient dividing of time between each task's sub-rules and between each task. Those who learnt the most efficient way of doing this, via trial and error, would be expected to score the most points during the experiment. An example sequence of events that a participant could perform is described in the supplementary materials.

In study 2, participants performed SOS with the same instructions, stimuli, rules and tasks; however, there were modifications in the presentation of the tasks, training and feedback. Most critically, immediately prior to the scanning session, participants received a 20 minute training session as opposed to a minimal demo. This extensive training session was designed to enable participants to arrive at a structured routine for performing the task prior to the scanner session. Additionally, the task and sub-task feedback scales from study 1 were only present during the training. Therefore, the only feedback available in the scanner was represented by the overall score number. There also were no bonus points awarded on increasing both task scales to their limit whilst in the scanner. These modifications were designed to minimise the information that the participant could use to further optimise their behavioural routine inside of the scanner whilst ensuring that they had sufficient information to remain engaged. Finally, as opposed to presenting both tasks simultaneously (study 1), only the currently selected MT or OOO task was presented on the screen (Figure 4.1e). This modification was intended to further control for activity related to switching between spatial locations as opposed to tasks.

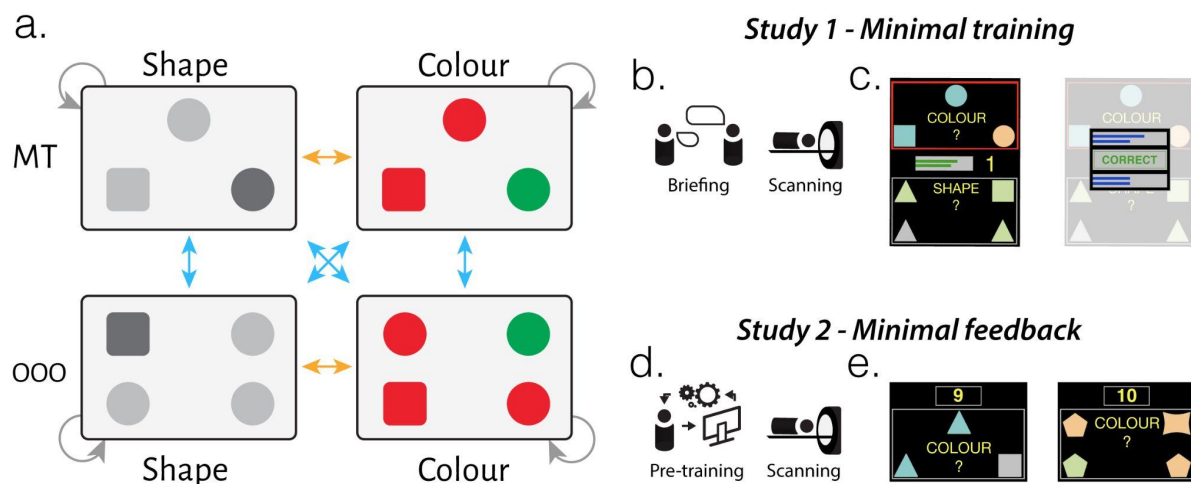


Figure 4.1. **Self-ordered switching (SOS) experiments.** a) Schematic of the two tasks (MT - Matching task, OOO - Odd one out) and the two rules (colour, shape) of the SOS paradigm. Grey arrows represent repeat-trials where the same condition was repeated. Task-switching and rule-switching are represented by the blue and orange arrows, respectively. b) In study 1, SOS was performed after minimal training with a short demo. c) Both tasks were simultaneously displayed on the screen and hierarchical response feedback was presented with progress meters on each task overall as well as each of the task's subordinate rules. d) In study 2, SOS was performed after extensive training with full hierarchical feedback and then performed in the scanner with minimal feedback involving the overall score only (e).

4.3.3 Behavioural analysis.

The RTs from correct trials were separated into three categories: Repeat-trials, rule-switch and task-switch. Repeat-trial RT was the time between the stimuli presentation and the L/R response being made. Rule-switch or task-switch RT was the time between the corresponding switch button being pressed and the subsequent L/R response. For each participant, RT's for each category were summarised using the median from correct trials. The number of trials completed would naturally vary across participants and response accuracy was summarised for each RT category as the percentage of correct responses made. Separate one-way repeated-measures analysis of variance (RM-ANOVA) models were independently conducted for the RT and accuracy measures.

Trial conditions could be repeated or switched between as each participant saw fit (Figure 4.1a). To quantify the extent to which participants structured their behaviour, we estimated the algorithmic complexity (Zenil et al., 2018) of each behavioural timecourse. Algorithmic complexity estimates the size of the smallest algorithm that could generate a given sequence and tends to be smaller for sequences with consistent structures vs. random sequences.

Here, we use the framework, outlined by Zenil et al. (2018), which combines a block decomposition method (BDM) and the coding theorem method (CTM) (Lempel & Ziv, 1976) to provide a reasonable estimate of Kolmogorov-Chaitin complexity (K) (Chaitin, 1966; Kolmogorov, 1968) for large strings in a computationally tractable way. The advantage of BDM is that it provides an estimate of K that not only considers statistical regularities, (e.g., as is captured by Shannon entropy), but it is also sensitive to segments of an algorithmic nature.

Timecourse length naturally varied across individuals, and to account for this, we normalised an individual's algorithmic complexity by the mean algorithmic complexity calculated from 100 randomly shuffled versions of the individual's data

(qualitatively similar results were obtained without normalisation). NB:- As expected, the shuffled data exhibited significantly greater algorithmic complexity (paired t-test: **Study 1** - $t_{15}=-9.775$, $p<0.001$, $d=3.54$; **Study 2** - $t_{14}=-35.329$, $p<0.001$, $d=14.29$).

A more focused behavioural analysis was conducted to examine how RT and algorithmic complexity varied over the course of the 20 minute experiment. Here, the behavioural timecourses were split into three non-overlapping time-windows (6.67 minutes of data), or “Learning stages”. The median RT for each category (correct trials) as well as the normalised algorithmic complexity (using the same procedure outlined above) were extracted from each learning stage.

4.3.5 MRI acquisition and preprocessing.

Whole-brain blood oxygenation level dependent (BOLD) fMRI recordings were acquired using the same Siemens 3T Trim Trio scanner and following parameters in both studies: TR=2s, TE=30ms, FA=78°, 3mm³ voxels, 3.75mm slice-gap, 32 axial slices. 630 volumes were acquired during the 20-minute SOS task run. A T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence was also acquired using the following parameters: TR=2250ms, TE=2.99ms, IT=900ms, FA=9°, 1mm³ voxels, 256 axial slices.

fMRI pre-processing was conducted using standard parameters with the Statistical Parametric Mapping 12 (SPM)¹ toolbox in Matlab 2018b². The first 10 volumes were discarded to account for T1-equilibrium effects. The remaining images were slice-time and motion-corrected, and the mean echo planar image (EPI) co-registered to the tissue segmented T1 image, normalised to 2mm³ MNI space and spatially smoothed (full half width maximum = 8 mm kernel).

¹ "SPM12 - Wellcome Trust Centre for Neuroimaging - UCL." Accessed July 30, 2019. <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>.

² "MATLAB - MathWorks - MATLAB & Simulink." Accessed July 30, 2019. <https://www.mathworks.com/products/matlab.html>.

4.3.6 Image data analysis.

Image Pre-processing was conducted using standard parameters with the Statistical Parametric Mapping 12 (SPM)³ toolbox in Matlab 2018b⁴. The first 10 volumes were discarded to account for T1-equilibrium effects. The remaining images were slice-time and motion-corrected, and the mean echo planar image (EPI) co-registered to the tissue segmented T1 image, normalised to 2mm³ MNI space and spatially smoothed (full half width maximum = 8 mm kernel).

4.3.7 fMRI activation modelling.

Participant's voxelwise fMRI data was modelled using General Linear Models (GLM) with the psychological events defined by responses to one of the four trial conditions (MT colour, MT shape, OOO colour, OOO shape) and when one of the two task-switches (MT to OOO, OOO to MT) or four rule-switches (MT colour to MT shape, MT shape to colour, OOO colour to shape, OOO shape to colour) occurred. Errors were infrequent but were captured in an additional predictor. Models in study 1 included an additional set of feedback events for when the rule or task bars de/incremented and when bonus points were awarded. These psychological predictors were convolved with the canonical haemodynamic response function (HRF) with the durations set to 0s.

Head motion and nuisance variables were modelled using the rigid-body realignment parameters, the mean timecourses from the white matter and cerebrospinal fluid tissue segmentations, alongwith their 1st order temporal derivatives. This resulted in 16 parameters. Additional spike regressors were included where framewise displacement (Power et al., 2012) exceeded the voxel acquisition size (3mm).

The 1st-level beta maps were estimated for each psychological predictor reactive to the model intercept (as there were no resting periods). Modelling against this

³ "SPM12 - Wellcome Trust Centre for Neuroimaging - UCL." Accessed July 30, 2019. <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>.

⁴ "MATLAB - MathWorks - MATLAB & Simulink." Accessed July 30, 2019. <https://www.mathworks.com/products/matlab.html>.

“implicit baseline” results in significant activation differences being relative to the average activation across time.

Additional 1st-level models were created to examine how activation varied over time. Specifically, the psychological predictors were segmented by the three learning stages defined in the behavioural analysis. The trial response events were collapsed into a single predictor, and switching events were collapsed into a single task-switch and a single rule-switch predictor. The same nuisance predictors were used in these models.

The group-level modelling of activation associated with switching involved a voxelwise one-way RM-ANOVA of the beta maps representing the two task-switch and four rule-switch conditions. To examine how switching activation varied over time, switch>repeat contrasts were generated from each learning stage and modelled with a group-level two-way RM-ANOVA with learning stage (3 levels) as the within-subject factor and study (2 levels) as a covariate. All group-level activation maps were initially thresholded at an uncorrected $p < 0.01$, followed by a $p < 0.05$ false discovery rate (FDR) cluster correction.

4.3.8 Modelling functional connectivity learning effects.

Generalised psychophysiological interaction (gPPI) models (Friston et al., 1997; McLaren et al., 2012) were used to examine how connectivity varied as SOS was performed. Specifically, gPPIs were fitted between pairs of mean timecourses from 274 ROIs defined by external atlases that parcellated the cortex, subcortex and cerebellum (Buckner et al., 2011; Fan et al., 2016; Schaefer et al., 2018). gPPI's were simultaneously estimated for the psychological events in each learning stage and pair of ROIs using the following:

$$Y^T = \beta_0 + [Y^S * H(X)]\beta_G + [Y^S, H(X), E]\beta + e,$$

where, X , is the matrix of psychological timecourses and, $H(X)$, its HRF convolution. Y^T is the target ROI timecourse and, Y^S , the seed ROI timecourse. E is the matrix of

motion & tissue signal timecourses. β_G are the gPPI betas weights, β , the beta weights of no-interest, β_0 , is the model intercept and, e , the residual. The resulting betas from each pair of ROIs were averaged ($[\text{model A} \rightarrow \text{B} + \text{model B} \rightarrow \text{A}] / 2$) and z-scored across learning stage. Subsequently, each dFC estimate represented the change in FC relative to that connection's baseline FC.

To estimate how gPPIs varied over the learning stages, generalised linear mixed effect (glme) models were fitted to each connection. This mass-univariate approach modelled each connection with the three learning stages as a fixed effect and participants as a random effect. The resulting model set was corrected for multiple-comparisons at a $p < 0.05$ FDR threshold.

4.4 Results

4.4.1 SOS costs to reaction time decrease with practise.

As expected, trials were performed with near-ceiling accuracy (**Study 1** - mean=94.87%, SD=4.18; **Study 2** - mean=94.69%, SD=4.04). Response accuracy did not vary across the repeat-trials, rule-switch or task-switch trials (one-way rm-anova: **Study 1** - $F_{2,30}=1.218$, $p=0.310$, $\eta_p^2=0.08$; **Study 2** - $F_{2,28}=1.867$, $p=0.173$, $\eta_p^2=0.12$).

In contrast, RT varied substantially between the repeat-trials, rule-switch and task-switch trials in both studies (one-way rm-anova: **Study 1** - $F_{2,30}=48.588$, $p<0.001$, $\eta_p^2=0.76$; **Study 2** - $F_{2,28}=52.886$, $p<0.001$, $\eta_p^2=0.79$). At a finer grain, in study 1 (Figure 4.2a), task-switching slowed response speed on average by 883ms (SD=538) (paired t-test: $t_{15}=6.573$, $p<0.001$, confidence interval 95% (CI)=0.60 to 1.17, Cohen's $d=1.64$), and rule-switches were 179ms (SD=273) faster ($t_{15}=-2.630$, $p=0.019$, CI=0.03 to 0.33, $d=0.66$), relative to repeat-trials (FDR corrected). In study 2 (Figure 4.2b), response speed slowed by 267ms (SD=143) after rule-switching (paired t-test: $t_{14}=7.214$, $p<0.001$, CI=0.19 to 0.35, $d=1.86$), and by a further 362ms (SD=268) after task-switching ($t_{14}=5.227$, $p<0.001$, CI=0.21 to 0.51, $d=1.35$) (FDR corrected).

In order to examine changes in switching-costs with learning, the trials were segmented by the three time-windows, or “learning stages”. Switching-costs for each learning stage were calculated by subtracting the mean repeat-trial RTs from the task-switch or rule-switch trial RTs. Task-switching costs varied over learning stages in both studies (one-way RM-ANOVA: **Study 1** - $F_{2,30}=3.579$, $p=0.040$, $\eta_p^2=0.19$; **Study 2** - $F_{2,28}=5.413$, $p=0.010$, $\eta_p^2=0.28$), reducing between the 1st and last stage by 369ms (SD=607) in study 1 (paired t-test: $t_{15}=-2.433$, $p=0.028$, CI=0.05 to

0.69, $d=0.61$), and by 278ms (SD=418) in study 2 ($t_{14}=-2.570$, $p=0.022$, CI=-0.51 to -0.05, $d=0.66$). Rule-switching costs did not vary significantly across learning stages in study 1 ($F_{2,30}=0.128$, $p=0.881$, $\eta_p^2=0.01$), but did in study 2 ($F_{2,28}=5.543$, $p=0.009$, $\eta_p^2=0.28$) and reduced by 217ms (SD=286) between the 1st and last learning stage ($t_{14}=-2.945$, $p=0.011$, CI=-0.38 to -0.06, $d=0.76$).

Individual differences in switching costs were examined further by correlating mean switch-trial RTs with the total number of switches performed across participants (supplemental Figure 4.2a-b). There was a significant negative correlation (Pearson: **Study 1** - $r_{15}=-0.557$, CI=-0.83 to -0.09, $p=0.025$; **Study 2** - $r_{14}=-0.658$, CI=-0.88 to -0.22, $p=0.008$). Therefore, even when self-ordered, there were substantial switching RT costs, the switching RT costs were optimised over time with learning, and individuals who made more switches had lower switch RT costs.

Repeat-trial RT varied between studies and with practise in a manner that indicated learning effects. A two-way rm-anova of repeat-trial RT showed main effects of study ($F_{1,29}=16.866$, $p<0.001$, $\eta_p^2=0.37$), learning stage ($F_{1,29}=309.770$, $p<0.001$, $\eta_p^2=0.91$) and a significant study * learning stage interaction ($F_{1,29}=14.516$, $p<0.001$, $\eta_p^2=0.33$). Repeat-trial RT was significantly slower in study 1 (two-sample t-test: $t_{29}=4.119$, CI = 0.29 to 0.85, $p<0.001$, $d=1.20$) but decreased at a greater rate between the first and last learning stages (two-sample t-test: $t_{29}=-2.878$, CI = -0.37 to -0.06, $p=0.007$, $d=0.93$).

4.4.2 Task performance relates to structured behavioural strategies

Notably, the number of points scored did not relate in a trivial manner to the total switches made. This was also the case when controlling for the number of trials completed (see supplemental materials for more details). However, the algorithmic complexity of individuals' behaviour (Figure 4.2c-d) had a significant negative correlation with points scored (partial Pearson: **Study 1** - $r_{15}=-0.646$, CI=-0.87 to -0.22, $p=0.017$; **Study 2** - $r_{14}=-0.805$, CI=-0.93 to -0.50, $p=0.002$) and the number of

response errors made (partial Pearson: **Study 1** - $r_{15}=-0.695$, $CI=-0.89$ to -0.31 , $p=0.004$; **Study 2** - $r_{14}=-0.785$, $CI=-0.93$ to -0.46 , $p<0.001$). NB:- These partial correlations control for the number of events completed. That is, people whose behavioural routines were more structured and whose trial-responses were more accurate achieved higher scores. Exemplar behavioural timecourses from each study are rendered in Figure 4.2e-f.

As predicted, algorithmic complexity was significantly lower in study 2, compared to study 1 (two-sample t-test: $t_{29}=-4.500$, $CI=-0.36$ to -0.14 , $p<0.001$, $d=1.26$). Moreover, study 2 exhibited a significantly greater repetition bias than study 1 (two-sample t-test: $t_{29}=6.643$, $CI=0.19$ to 0.36 , $p<0.001$, $d=1.53$).

Furthermore, within-subject behavior became increasingly structured as the SOS was practiced. Algorithmic complexity in the first behavioural learning stage negatively correlated with the sum of change across all three learning stages (partial Pearson **Study 1** - $r_{15}=-0.515$, $CI=-0.81$ to -0.03 , $p=0.0496$; **Study 2** - $r_{14}=-0.899$, $p<0.001$, $CI=-0.97$ to -0.72). This indicates that increases in behavioural structure were greater for those who were initially behaving in a less structured manner.

Therefore, individuals developed more structured behaviour with practice, being characterised by ‘chunks’ of repeat-trials that were consistently punctuated by switches between rules and tasks, and behaviour in study 2 began at a more structured level, due to structured routines developed during the extensive pre-training.

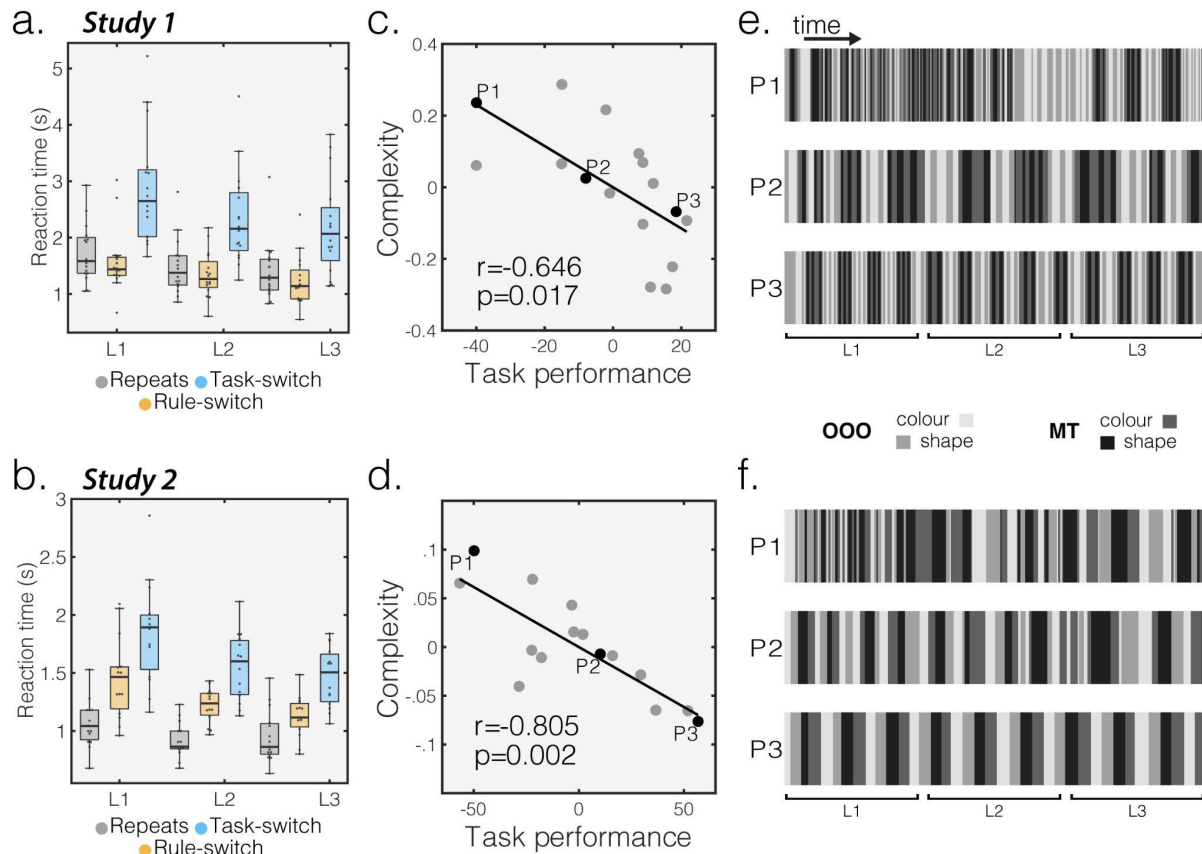


Figure 4.2. SOS switching costs and behavioural structuring. Switching-costs to reaction time (RT) decreased with practise in study 1 and 2 (a/b). Reaction times (RT) for each condition were summarised within each ~7 minute learning stage (L). Each point represents a participant's median RT for a condition. Boxplots represent the group median (central line), and the 25-75th percentiles (boxes). c/d) Normalised algorithmic complexity (y-axis) of participants' SOS behaviour against overall task performance, in terms of points scored (x-axis). The residuals are plotted from a partial correlation controlling for the number of trials completed. For each study, three participants are highlighted and their behavioural timecourses are displayed in e/f. Of those, participant 1 (P1) generated the most complex SOS behaviour and P3 the least.

4.4.3 SOS activates multiple-demand cortex and reduces default mode network activation.

We sought to identify brain regions where activation increased when individuals opted to switch between tasks or rules. Collapsing across switch-type, the cluster-corrected group-level t-contrast associated SOS with activation increases in the dorsolateral frontoparietal cortices, cerebellum, insula, putamen, and thalamus (Figure 4.3a). This pattern was consistent across studies (dice similarity = 0.58). The inverse t-contrast associated SOS with activation reductions in regions associated with the DMN (Figure 4.3b) and this pattern was consistent across studies (dice similarity = 0.56).

4.4.4 Multiple-demand and default mode network activation changes with learning.

Previous studies have examined activation changes when simple tasks are established through practise (Hampshire et al., 2016, 2019; Petersen et al., 1998; Ruge et al., 2019). To expand on this, voxelwise activation was compared across time during SOS. For each of the three learning stages, a 1st-level contrast was estimated for switching minus trial event activation. These contrasts were examined using group level random effects analyses with a 3 * 2 RM-ANOVA design where the learning stage was the within-subject factor and study was the between-subject factor.

The main effect of learning stage contrast rendered six bilateral regions where the difference between switching and trial activation significantly differed between the learning stages. For each cluster, a paired t-test determined the direction of the activation change using the mean bilateral activation from the 1st and last learning stages. Each cluster showed significant differences that survived bonferroni correction ($\alpha=0.0083=0.05/6$, see supplemental materials). We combined those regions whose activation increased (Figure 3c) into a composite MD region of interest (ROI) of the anterior cingulate, dorsolateral parietal cortex and occipital

cortex clusters. The MD ROI activation was significantly increased in the last learning stage, compared to the first (paired t-test: $t_{30}=7.230$, $CI=0.75$ to 1.34 , $p<0.001$, $d=0.68$). Those regions where activation decreased were combined into a composite DMN ROI and included the posterior cingulate, putamen and temporal lobe (Figure 3d). The DMN ROI activation was significantly decreased in the last learning stage, compared to the first (paired t-test: $t_{30}=7.432$, $CI=0.49$ to 0.85 , $p<0.001$, $d=1.15$).

Notably, MD activation increasing and DMN activation decreasing with learning is the exact opposite direction to what has been observed during simple stimulus-response learning tasks (Hampshire et al., 2019). To understand why this learning effect appeared to be inverted, we independently examined trial response activation and switching activation in the MD ROI across the learning stages (Figure 3e). The trial response events are more comparable to simple learning curve paradigms. Directly replicating these studies, activation in the MD ROI varied across the learning stages for trial responses (one-way RM-ANOVA: $F_{2,60}=16.043$, $p<0.001$, $\eta_p^2=0.35$) and was significantly decreased in the last learning stage, compared to the first (paired t-test: $t_{30}=-4.573$, $CI=-0.99$ to -0.38 , $p<0.001$, $d=0.82$). Conversely, MD ROI switching activation also varied across learning stage (one-way RM-ANOVA: **Task-switch** - $F_{2,60}=3.329$, $p=0.043$, $\eta_p^2=0.09$; **Rule-Switch** - $F_{2,60}=6.313$, $p=0.003$, $\eta_p^2=0.17$) but was significantly increased in the last learning stage, compared to the first (paired t-test: **Task-switch** - $t_{30}=2.224$, $CI=0.03$ to 0.67 , $p=0.034$, $d=0.40$; **Rule-switch** - $t_{30}=2.669$, $CI=0.09$ to 0.66 , $p=0.012$, $d=0.48$).

The opposite pattern was evident for the DMN ROI activation (Figure 3f). Specifically, activation varied across the learning stages for trial responses (one-way RM-ANOVA: $F_{2,60}=10.119$, $p<0.001$, $\eta_p^2=0.25$) but was significantly increased in the last learning stage compared to the first (paired t-test: $t_{30}=3.595$, $CI=0.11$ to 0.40 , $p=0.001$). Conversely, DMN ROI switching activation also varied across learning stages (one-way RM-ANOVA: **Task-switch** - $F_{2,60}=16.544$, $p<0.001$, $\eta_p^2=0.36$;

Rule-switch - $F_{2,60}=15.080$, $p<0.001$, $\eta_p^2=0.34$) but was significantly decreased in the last learning stage, compared to the first (paired t-test: **Task-switch** - $t_{30}=-4.398$, $CI=-0.73$ to -0.27 , $p<0.001$, $d=0.79$; **Rule-switch** - $t_{30}=-5.112$, $CI=-0.46$ to -0.20 , $p<0.001$, $d=0.92$).

The sum of differences in mean MD activation over learning stage negatively correlated with individuals' normalised algorithmic complexity (pearson, $r_{15}=-0.499$, $p=0.049$, $CI=-0.798$ to -0.004) in study 1 (Figure 3g) but not in study 2 ($r_{14}=-0.240$, $p=0.932$, $CI=-0.53$ to 0.49). The equivalent analysis of the DMN ROI activation (Figure 3h) did not correlate with algorithmic complexity (pearson, **Study 1** - $r_{15}=0.295$, $p=0.268$, $CI=-0.24$ to 0.69 ; **Study 2** - $r_{14}=0.290$, $p=0.294$, $CI=-0.26$ to 0.70).

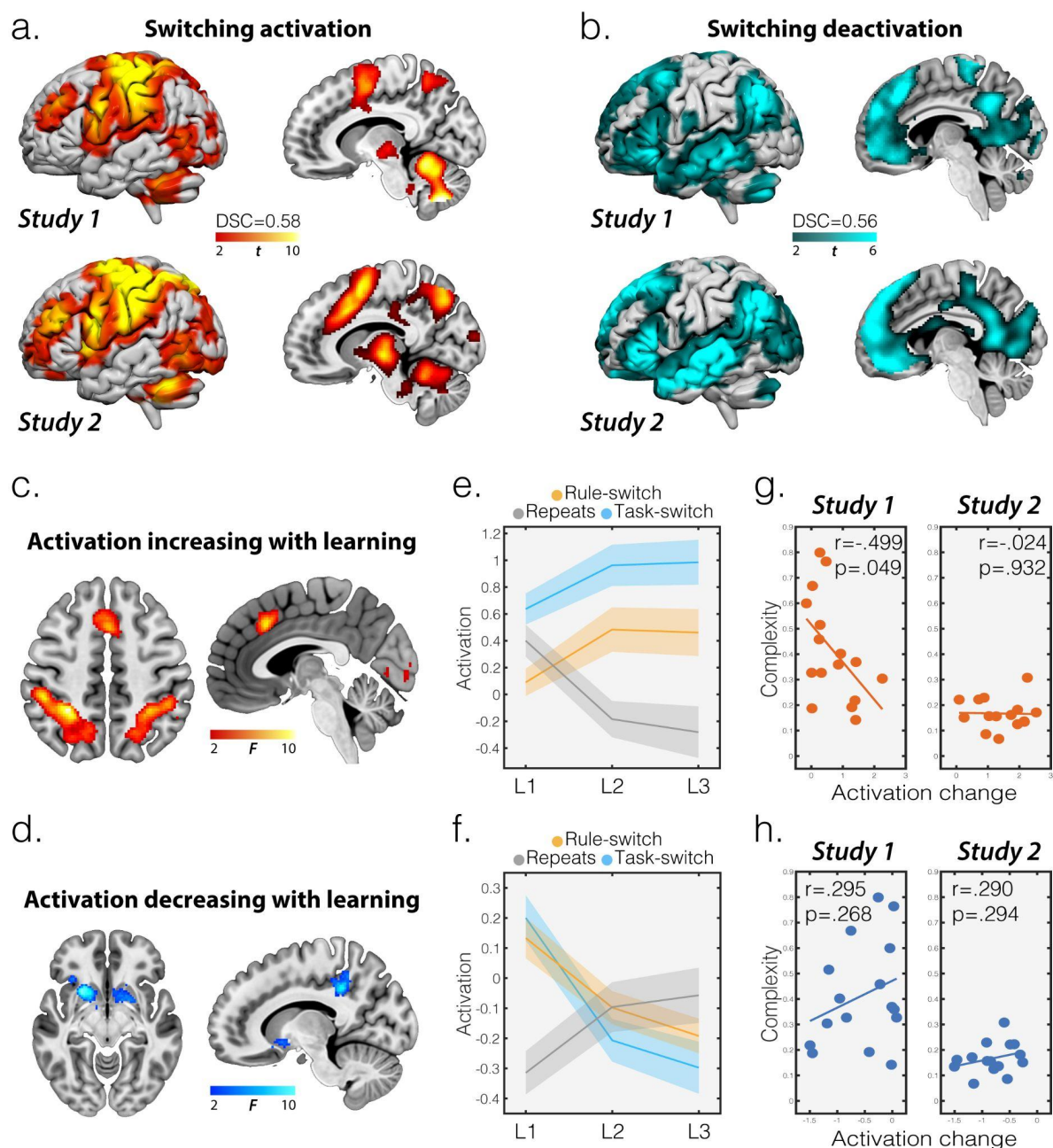


Figure 4.3. Multiple-demand and default mode activation varies with SOS learning. Group-level activation increases (a) and decreases (b) predicted by switching events. Dice similarity coefficients (DSC) represent the spatial similarity of significant voxels between the studies. d/e) Clusters of activation change over learning stages separated based on whether activation increased (c) or decreased (d). e/f) Mean activation independently examined for each learning stage (L) and condition for the MD (e) and DMN (f) composite ROIs (shaded area represents the standard error). g/h) Correlations between the sum of activation change over learning stage and normalised algorithmic complexity of SOS behaviour for the MD (g) and DMN (h) composite ROIs. All activation maps are thresholded at an uncorrected $p < 0.01$, followed by a $p < 0.05$ FDR cluster correction.

4.4.5 Functional connectivity between frontoparietal networks and subcortical regions increases with learning.

In accordance with the behavioural results showing that repeat-trial RT varied between studies and with practise (Figure 4.4a), the FC predicted by trial responses varied across learning stages in a manner that was largely specific to study 1. The mass-univariate glme modelling of the gPPI weighs across the learning stages identified connections whose FC significantly varied across learning stages. In study 1, the models that survived FDR correction for multiple comparisons were predominantly for connections between networks (Figure 4.4b) associated with cognition, such as the DMN, networks associated with MD (frontoparietal and attention networks), as well as with subcortical regions (Figure 4.4c). The glme models' t-statistic was used to determine whether the FC increased or decreased across the learning stages. This indicated that FC increased as the SOS task was practised in study 1. Notably, these effects were largely absent in study 2 (Figure 4.4d).

Finally, we estimated the relationship between the decreases in RT and FC over the learning stages. A mass univariate approach was taken where a partial correlation was calculated for each connection using the difference in gPPI and repeat-trial RT between the 1st and last learning stages. The partial correlation allowed this relationship to be estimated whilst controlling for the number of trials an individual completed. The resulting set of correlation coefficients were FDR-corrected for multiple comparisons. This unconstrained analysis showed significant correlations between the same systems identified in the previous analysis, demonstrating that a greater decrease in RT with learning correlated with a greater increase in FC. Again, these effects were specific to study 1 (Figure 4e) and largely absent in study 2 (Figure 4.4f).

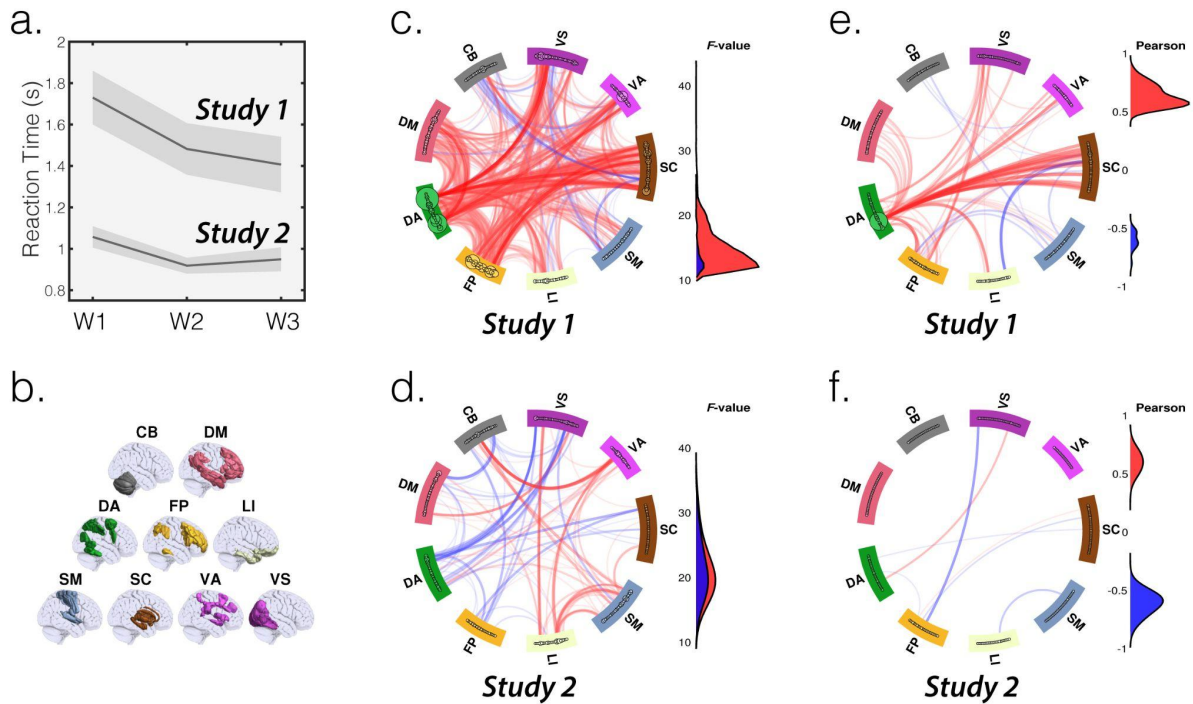


Figure 4.4. Increased between-network functional connectivity with learning. a) Reaction time (RT) decreased at a greater rate between the three windows (W) in study 1. The mean repeat-trial RT was plotted for each window with the standard error of the mean. b) Task-dependent functional connectivity (FC) changes for trial responses were estimated between brain regions and these were ordered by nine networks for visualisation purposes. c/d) Generalised linear mixed effects models identified connections whose strength significantly varied over the learning stages for study 1 and 2. The model t-statistic was used to render the connections in red or blue for positive or negative change, respectively. e/f) Partial correlations were used to estimate the relationship between changes in RT and FC over time for study 1 and 2 (controlling for the number of trials completed). Here, a positive correlation indicates an increase in response speed and FC. Each set of connectivities were FDR corrected and the surviving weights were rendered in the schemball and in the accompanying density distribution plot that was within-set normalised. CB=Cerebellum, DM=Default Mode, DA=Dorsal Attention, FP=Frontoparietal, LI=Limbic, SM=Somatomotor, SC=subcortical, VA=Ventral Attention, VS=Visual.

4.5 Discussion

The novel self-ordered switching paradigm developed here enabled new insights to be derived into the processes by which people learn complex multi-component tasks. In this executively challenging context, the learning process was characterised by the concurrent optimisation of behavioural and functional-network deployment; specifically, hierarchically structured behavioural routines that minimised switching costs were optimised based on the monitoring of multiple outcomes. Concurrently, functional-network resources were redeployed towards the executive switching points around which those routines were organised.

In everyday life we often are confronted with situations where multiple concurrent tasks must be performed. Furthermore, real-world tasks are often intrinsically hierarchical in their demands, requiring sub-goals to be completed in order to achieve overarching goals (Poljac et al., 2018). The majority of experimental paradigms that have been used to study cognitive control processes enforce careful control of the flow of events, for example, specifying when the participant must switch between sub-tasks (Kiesel et al., 2010; Koch et al., 2018). These types of design facilitate the careful balancing of conditions, but they lack ecological validity insofar as they do not provide the participant with the opportunity to apply their executive abilities in order to organise their own behaviour (Goldberg & Podell, 2000). This is important, because structured routines are a defining characteristic of human behaviour with disorganisation being symptomatic across prominent neurological and psychiatric populations (Averbeck et al., 2011; Collins & Koechlin, 2012; Fallon et al., 2013).

Moreover, the majority of cognitive control studies do not take into account the shifting role that brain systems play as a task is learnt. We argue that this is a critical oversight because the purpose of cognitive control systems is not just to support complex behaviour, but more to enable the learning of such behaviours in order that they can be performed with minimal effort.

Indeed, we and others have previously reported that the changes in brain activity and connectivity that are associated with learning during commonly applied cognitive control paradigms can dwarf the control processes that those paradigms were designed to examine (Badre et al., 2010; Erika-Florence et al., 2014). More specifically, in a recent sequence of articles, we and others reported that the learning of simple stimulus-response mappings from either instruction or feedback was accompanied by an ongoing reduction in event-related activity within frontoparietal networks associated with effortful cognition and a reduction in default mode network activity (Hampshire et al., 2016, 2019; Ruge et al., 2018, 2019; Ruge & Wolfensteller, 2010). This likely reflected a steady transition from deliberate to automated modes of behaviour.

The results of the current study extended that work by examining the shifting role of brain networks during the learning of tasks that are more ecologically valid insofar as they require multiple sub-tasks to be efficiently self-organised. When this learning process was supported by detailed ongoing feedback in study 1, the behavioural routines that participants tended to develop had several notable characteristics, these being the minimisation of number of switches performed, the hierarchical sequencing of task and rule switches, and the tendency towards algorithmically simpler sequences with practice. Interestingly, all three of these characteristics correlated with overall task performance, as quantified in number of points awarded across learning stages or in terms of individual differences in behavioural routines; however, from a simple computational perspective this would not necessarily have been the case.

For example, task switches could have been performed at the same frequency, or subordinate to rule switches whilst producing the same number of points per trial completed. However, participants in study 1 tended to organise the rule switches subordinate to the task switches, a structure that was implicitly encouraged by the structure of the feedback elements. Notably, task switches were more effortful, as evidenced by greater response time costs; therefore, the bias towards more rule

than task switches, and fewer switches overall, was adaptive insofar as it reduced overall executive costs. Relatedly, the application of algorithmically simpler routines was demonstrably less effortful, being accompanied by reduced switching costs and the achievement of higher overall scores. From this perspective, the development of hierarchically structured behaviour can be seen to reflect a tendency to optimise towards a mode of behaviour that requires minimal switching effort to produce maximum reward.

Indeed, evidence of a tendency to self-order behaviour in a structured manner even when such structure is not a task requirement has been seen previously in some previous studies. A prominent example is voluntary task-switching (VTS) paradigms, which instruct participants to randomly choose between two tasks on each trial. Interestingly, the original VTS reports imply that participants did not fully adhere to this instruction and exhibited a significant repetition bias (Arrington & Logan, 2004, 2005). This is evidence of the behavioural self-ordering that reduced the frequency of task-switching that may have resulted in an advantageous reduction in cognitive demand. Similarly, here, participants who were better practised at SOS (study 2) exhibited a significantly greater trial repetition bias, and also behaved with a more structured routine.

Relatedly, we have also observed that when identifying targets during a novel self-ordered intradimensional-extradimensional switching paradigm, participants tend to hierarchically order their search behaviour between stimuli of different dimensions, this in a context where there is no intrinsic computational advantage to sequencing the search by stimulus dimension but switching between dimensions nonetheless incurs a greater response time costs (Hampshire & Owen, 2006).

Perhaps the most striking result of the current study pertains to how the underlying network resources associated with repeat and switch trials changed as the behavioural routine was established. More specifically, the optimisation of a structured behavioural routine was concomitant with a progressive ‘fine-tuning’ of the brain’s functional activation that was more complex in its dynamics than that

which has previously been observed during simple stimulus-response learning (Hampshire et al., 2016, 2019; Ruge et al., 2019; Ruge & Wolfensteller, 2010). In accordance with those past findings, frontoparietal activation in response to the trials was observed to decrease as they were practised whilst the degree of default mode network deactivation was also seen to reduce. Unexpectedly though, we also observed the opposite pattern of change for switching events, with frontoparietal activation increasing and DMN activation decreasing as a function of learning. These executive switching events were effortful and critical for guiding behaviour and had to be conducted in a strategic manner. This distinguishes the SOS from most other task-switching and VTS paradigms.

Furthermore, the ACC and DLPFC regions identified here are part of the MD system and generally activate in response to increasing cognitive effort (Crittenden & Duncan, 2014). Therefore, this pattern of results likely reflects a redeployment of flexible and limited capacity network resources towards the executive switching events that enable structured behaviour as the simple stimulus-response mappings that constitute individual trials become automatic through practice. One intriguing possibility is that frontoparietal systems operate at the leading edge of the automatisisation process, enabling progressively more complex hierarchical behaviours to be developed. A testable prediction of this is that the switching activation would also decrease with enough practise. This would make an interesting prediction that could be tested across multiple SOS scanning sessions, perhaps with paradigms that have scope for deeper hierarchies that limited capacity resources can progressively develop.

In terms of individual differences, it is interesting to note that in study 1, the scale of change in trial vs switch activation over learning stages related to the complexity of the behavioural routine. Those participants who could establish an optimal balance between repeating a condition and making the costly switches behaved in a more structured manner, scored the most points and exhibited a greater ‘fine-tuning’ of frontoparietal activation away from simple trials to the executive switching events.

This indicates that optimal structured routines may also help accelerate the learning process.

The key difference between study 1 and study 2 was that the latter had more substantial training prior to entering the scanner and only minimal feedback within the scanner. These modifications were intended to provide a control condition where the potential for further optimisation of the behavioural routine was greatly reduced, with participants in study 2 being likely further along the 'learning curve'. The behavioural results supported this view, as in study 2, SOS was performed faster and with routines that had low algorithmic complexity. In this context, the changes in behaviour across time also were both less pronounced. RT dropped at a lower rate and the algorithmic complexity progressed towards a more stereotyped pattern that expressed more consistently repeating structures.

Intriguingly, the differences across studies in behavioural and network activity signatures of learning was also evident in the connectivity analyses. More specifically, in study 1 we observed widespread increases in FC between functional networks in the brain that have been associated with cognition and their connections with subcortical regions. A similar increase in network coherence was evident in the individual differences analysis, where greater reductions in RT were associated with a greater increase in network functional connectivity with learning. In contrast, FC changes with practise were largely absent in study 2. These findings accord with the notion that cognitive behaviours can through practice be supported by higher efficiency network states that are characterised by lower activity but increased connectivity (Hampshire et al., 2016; Soreq et al., 2019, 2021; Stokes, 2015).

More broadly, we believe that paradigms such as the SOS that enable participants the freedom to self-organise their behaviour have untapped potential for understanding normal and pathological cognitive control functions. The natural variation in the behavioural efficiency of healthy adults suggests that there is a potential to examine self-ordered behaviours in patient populations that suffer from

cognitive impairments. Indeed, high activity-low connectivity states have been reported in some dysexecutive populations when performing tasks that require cognitive control (Hampshire et al., 2013; Sheffield et al., 2021). Furthermore, it could be predicted that patients may take longer to generate a stable behavioural routine and that this routine would be more complex, inefficient and refined at a slower rate, compared to healthy controls. These features may dissociate across patient populations for example, Parkinson's patients show difficulty in updating strategies during set-shifting and can show a reduced tendency towards structured dimensional switching routines (Fallon et al., 2013), whereas patients with schizophrenia may prematurely implement a strategy before sufficiently exploring the options (Averbeck et al., 2011). Future work also should focus on the learning of structured routines over longer time frames and within paradigms that provide even greater potential for hierarchical self-organisation, and should examine how the responsiveness of functional networks evolves as deeper hierarchical behaviours are established.

4.6 Limitations

A limitation of the design in study 1 relates to the presentation of new stimuli following a task-switch or rule-switch. The visual information in supplemental figure 4.1 can be used as a reference to help clarify the order of the task events and timings. A trial began with new stimuli for the active task. If a response is made, the RT represents the time between the onset of the new stimuli and the response being made. When a response is made, the hierarchical feedback panel is presented and this is followed by the onset of the subsequent trial with new stimuli. If a rule-switch is made, the RT is the difference between the rule-switch button being pressed and the response being made. However, while the rule text would update ("shape" or "colour"), the stimuli did not update. This represents an error in the implementation of the task. Effectively, the user was already aware of the stimuli after they made the rule-switch and this was evident in the apparent "facilitation effect" observed in the rule-switch RT data.

The two tasks were both on display with one at the top and one at the bottom. Stimuli from both tasks were visible at all times. A similar error in the task implementation was present here because the stimuli did not update during a task-switch. Again, the stimuli only updated, within-task, after a trial response and feedback display event occurred. In contrast to rule switches, the separate task displays make it unlikely that participants had the stimuli of the task they were switching to in mind. This cannot be ruled out, however.

Despite this, a substantial task-switching cost to RT was evoked. Task-switch RT was the time between the task-switch button being pressed and the subsequent trial response. Task-switching is a reorienting of attention; however, paradigms wouldn't traditionally require a drastic shift in visual focus. Subsequently, some may argue that the 'pure cognitive' aspect of the task-switching process was confounded by the reorientation of visual attention to a greater extent than would be present in other paradigms.

Study 2 built on, and accounted for, the task implementation errors as stimuli updated after either a trial response or a switch was made. Moreover, tasks were shown in isolation and covered the entire display. The RT data followed the expected trend with switching-costs that followed rule-switches and significantly greater switching-costs following task-switching. Critically, there were multiple examples of evidence that suggested a robust replication between the studies, despite the differences in task implementation.

Fortunately, the task implementation issues do not undermine the inferences that can be made regarding the optimisation of behaviour and brain function. For example, the activation analyses modelled the points in time that buttons were pressed but the event durations were set to 0. Effectively, fitting a HRF to each event allowed for the prediction of activation changes associated with a user making a switch or trial response. Critically, both studies converged in terms of the voxelwise activation analyses.

More broadly, it is challenging to interpret changes in fMRI regional activation and relate them to cognitive constructs (Poldrack & Yarkoni, 2016). Here, switching events predicted activation increases in multiple-demand (MD) cortex but the activation magnitudes varied during the course of the experiment. The trend of decreasing MD activation predicted by trial events over time represents a robust replication of previous findings that are indicative of learning (Hampshire et al., 2016, 2019). However, the increase in MD activation over time for the switching events is novel. A dissociation between the trial response and switching activation timecourses for MD activation may represent evidence of a ‘fine-tuning’; whereby, the simpler parts of the routine are automated and greater attention is placed on the executive switch events that structure the behaviour. However, while this may be a reasonable inference to make from these data, it is important for future work to explore this further, given the novelty of this particular finding.

4.7 Conclusion

In summary, we developed SOS paradigms to examine how behaviour and brain states optimise to maximise efficiency. Behavioural performance and routines were sensitive to learning and as individuals learnt to manage SOS demands and this was reflected by an ongoing optimisation of the brain’s functional organisation. These findings characterise a fundamental property of human neurocognitive systems as the ongoing and concurrent self-optimisation of behavioural structure and network dynamics in order to maximise performance outcomes and minimise the use of neural resources.

4.8 Supplemental materials

4.8.1 Repetition bias

A repetition bias was calculated as the proportion of consecutive trials from the same condition relative to the total number of trials completed (supplemental Figure 4.2a). There was a significantly lower repetition bias in study 1 compared to study 2 (two-sample t-test: $t_{29}=-6.643$, $p<0.001$, $CI=-0.36$ to -0.19 , $d=1.53$). This indicates that participants in study 2 were more likely to repeat the conditions than switch between rules or tasks.

A two-way repeated-measures analysis of variance (RM-ANOVA) was used to explore the study differences between the repeat-trial and switch events (supplemental Figure 2b). Condition was the within-subject factor (repeat-trial, rule-switch, task-switch) and study was the between-subject factor (2 levels). There was a significant interaction of event number with study ($F_{2,58}=62.128$, $p<0.001$; $\eta_p^2=0.68$). Participants in study 1 performed fewer trials (two-sample t-test: $t_{29}=-5.262$, $p<0.001$, $CI=-81.87$ to -36.04 , $d=1.38$) and performed more rule-switches ($t_{29}=8.125$, $p<0.001$, $CI=49.85$ to 83.39 , $d=1.64$) than participants in study 2. The number of task-switches did not differ between studies ($t_{29}=0.024$, $p=0.981$, $CI=-11.426$ to 11.69 , $d=0.01$).

The ratio of task-to-rule-switching differed (two-sample t-test: $t_{29}=-8.118$, $p<0.001$, $CI=-0.71$ to -0.43 , $d=1.64$), with study 1 individuals (mean=0.294, $SD=0.139$) more likely to switch within-task than study 2 (mean=0.865, $SD=0.242$), who performed task-switching or rules-switching almost equally. Supplemental Figure 3e-f visualises these behavioural differences with graphs weighted by the mean probability of one event following another in each study.

4.8.2 Number of switches does not correlate with task performance

The number of switches performed did not significantly correlate with the overall points scored (pearson: Study 1 - rule-switch $r_{15}=0.473$, $CI=-0.03$ to 0.79 , $p=0.065$; **Study 1** - task-switch $r_{15}=0.213$, $CI=-0.32$ to 0.64 , $p=0.428$; ; **Study 2** - rule-switch $r_{14}=0.160$, $CI=-0.38$ to 0.62 , $p=0.569$; Study 2 - task-switch $r_{14}=0.034$, $CI=-0.49$ to 0.54 , $p=0.905$) (uncorrected). This was also the case when controlling for the number of trials completed (partial Pearson:

Study 1 - $r_{15}=0.407$, $CI=-0.11$ to 0.75 , $p=0.132$; **Study 2** - $r_{14}=0.151$, $CI=-0.39$ to 0.62 , $p=0.605$).

4.8.3 Trial activation modelling

The 1st level trial response contrasts maps were modelled at the group-level with a two-way repeated-measures (RM-ANOVA). The within-subject factors were task (match, odd one out) and rule (colour, shape). Group-level activation maps were thresholded at $p<0.01$ followed by a false discovery rate (FDR) cluster-level correction for multiple comparisons.

In both studies, the average effect of trial response contrast rendered activation in occipital cortex and in multiple-demand (MD) cortex, and included dorsolateral frontoparietal cortex, the supplementary motor area (SMA) anterior cingulate cortex (ACC) and anterior insula (supplemental Figure a/d). The main effect of task contrast showed an activation difference in the occipital cortex with greater activation during OOO trials (supplemental Figure 4b/e). The main effect of rule contrast showed an activation difference in occipital cortex with greater activation during shape trials (supplemental Figure 4c/f).

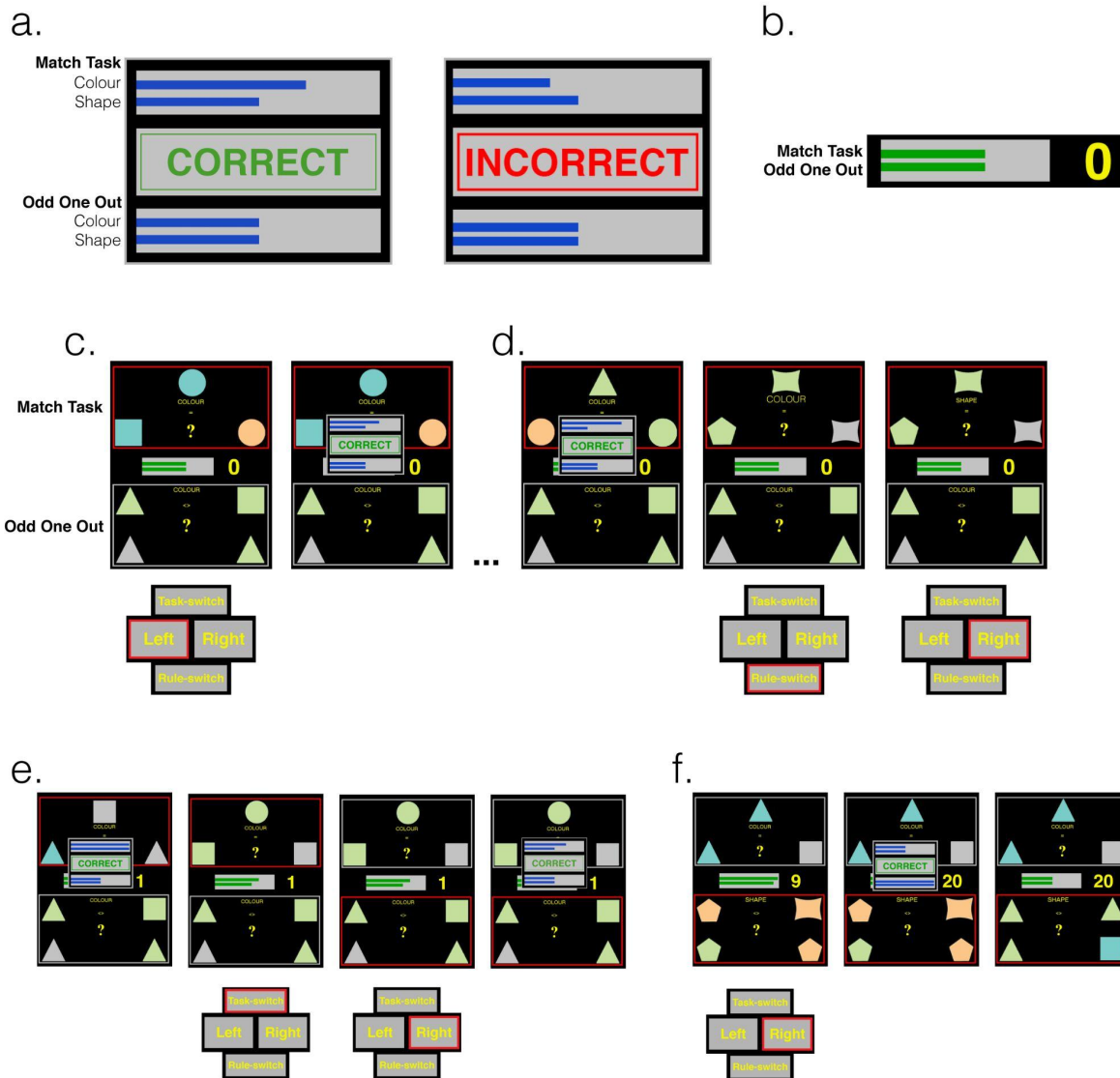
4.8.4 Switching activation between study differences

In the switching activation over learning stage analysis, the switch > trial contrast for each window was modelled in a two-way RM-ANOVA with learning stage as the within-subject factor (3 levels) and study (2 levels) as a covariate. The between-study differences in switching > trial activation was observed in the main effect of study contrast (supplemental Figure 4g). Study 1 showed greater activation in regions associated with the default mode network. Study 2 showed greater activation in regions associated with the multiple-demand (MD) network.

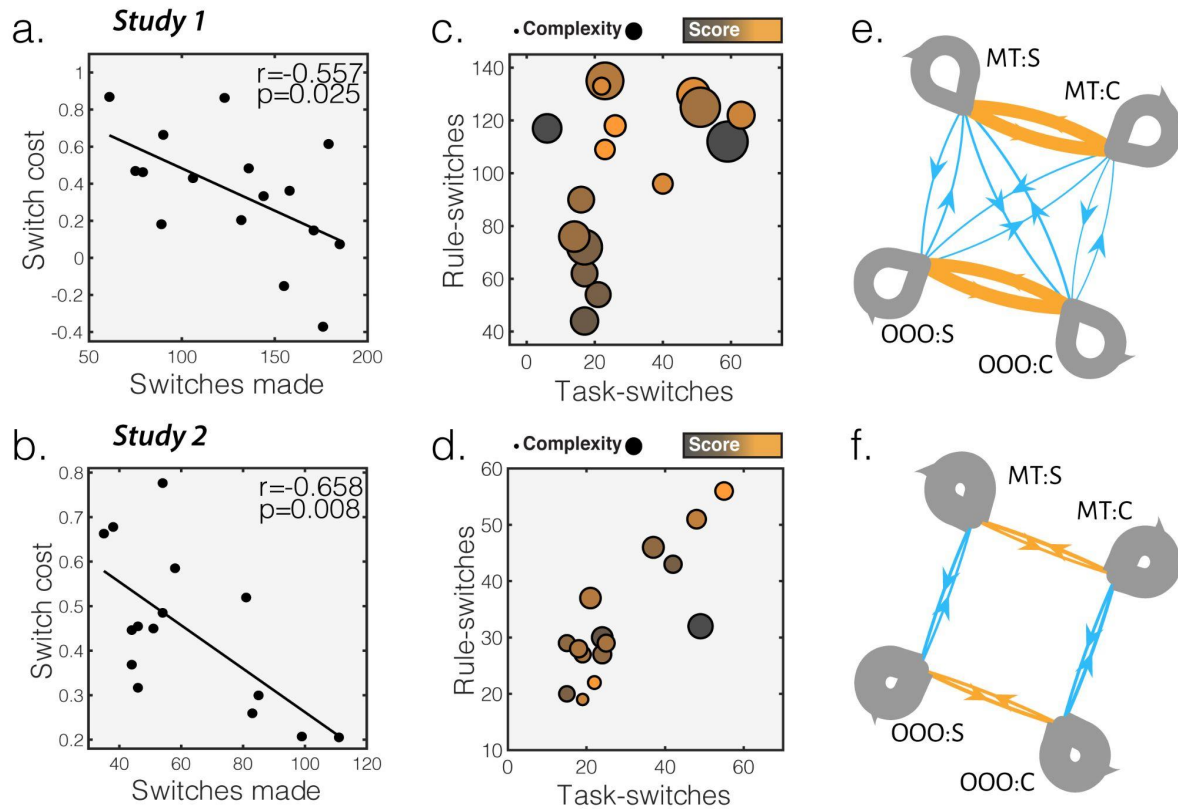
4.8.5 Switching greater than trial activation changes over learning stages

The main effect of learning stage contrast rendered 6 bilateral regions of interest (ROI). To determine the direction of the change, the activation was collapsed across hemispheres and compared with a paired t-test between the first and last learning stage. Activation in the

anterior cingulate (ACC), dorsolateral parietal cortex (DLPC) and occipital cortex (Occ.) was significantly increased between the first and last learning stages (paired t-test: DLPC - $t_{30}=4.657$, $p<0.001$, CI=1.65 to 0.64, $d=1.56$; ACC - $t_{30}=5.267$, $p<0.001$, CI=1.37 to 0.61, $d=1.59$; Occ. - $t_{30}=4.707$, $p<0.001$, CI=1.42 to 0.56, $d=1.57$). Activation decreased in the posterior cingulate cortex (PCC), putamen (Put.) and in the posterior temporal lobe (Temp.) (paired t-test: PCC - $t_{30}=4.097$, CI=0.31 to 0.94, $p<0.001$, $d=0.74$; Put. - $t_{30}=3.412$, CI=0.23 to 0.91, $p=0.0019$, $d=0.62$; Temp. - $t_{30}=5.401$, CI=-1.12 to 0.51, $p<0.001$, $d=0.97$). The comparisons made with these 6 ROIs survive bonferroni correction ($\alpha=0.0083$, $0.05/6$).



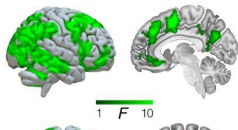
Supplemental Figure 4.1. SOS hierarchical response feedback (study 1). a) The feedback panel displayed response feedback along with progress meters for each condition. A correct response increased the bar for the given condition and an incorrect response decreased it. b) Progress meters were displayed for each task and increased when both meters for that task (a) maxxed out and reset. The task meters and overall score were displayed in between the two task panels. c) The SOS display and a correct response to a Match task (MT) colour trial. d) An example of the inactive rule being penalised, followed by a rule-switch. e) A participant has successfully switched between and filled both of the MT rule progress bars. The bars reset to the initial position and the task progress meter increases and overall score updates. Here, the participant opted to press the task-switch button and respond to the odd one out (OOO) task. Subsequently, the MT border turns grey and the OOO border turns red. f) Successfully switching between each task's rules and each task causes both task progress meters to reach the maximum. The overall score receives a +10 bonus and the task meters reset.



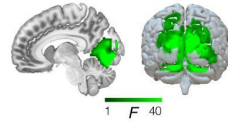
Supplemental Figure 4.2. a/b) Mean task-switch and rule-switch cost (y-axis, seconds) correlates with the total number of switches made in both studies. c/d) The number of rule-switches and task-switches completed. Each participant's datapoint size is weighted by the normalised algorithmic complexity of their behavioural timecourse. The colour represents the number of points scored (black/brown = low points, light-brown/orange = high points). e/f) Graphs representing the probability of repeat-trials for each condition and switches between them (grey = repeat-trials, orange = rule-switch, blue = task-switch). Edge widths represent the group average proportions of events. MT=Match task, OOO = odd one out, C = colour, S = shape.

Study 1

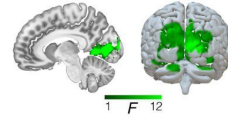
a - Average effect of trials



b - Main effect of task

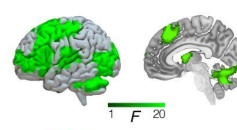


c - Main effect of rule

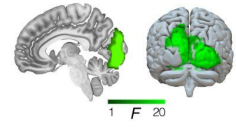


Study 2

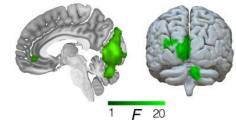
d - Average effect of trials



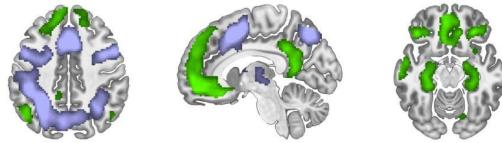
e - Main effect of task



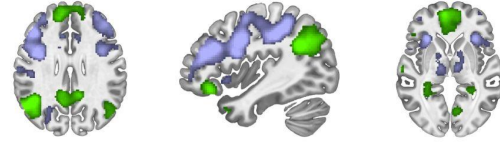
f - Main effect of rule



g.



Study 1 > Study 2
t: 1 5



Study 1 < Study 2
t: 1 5

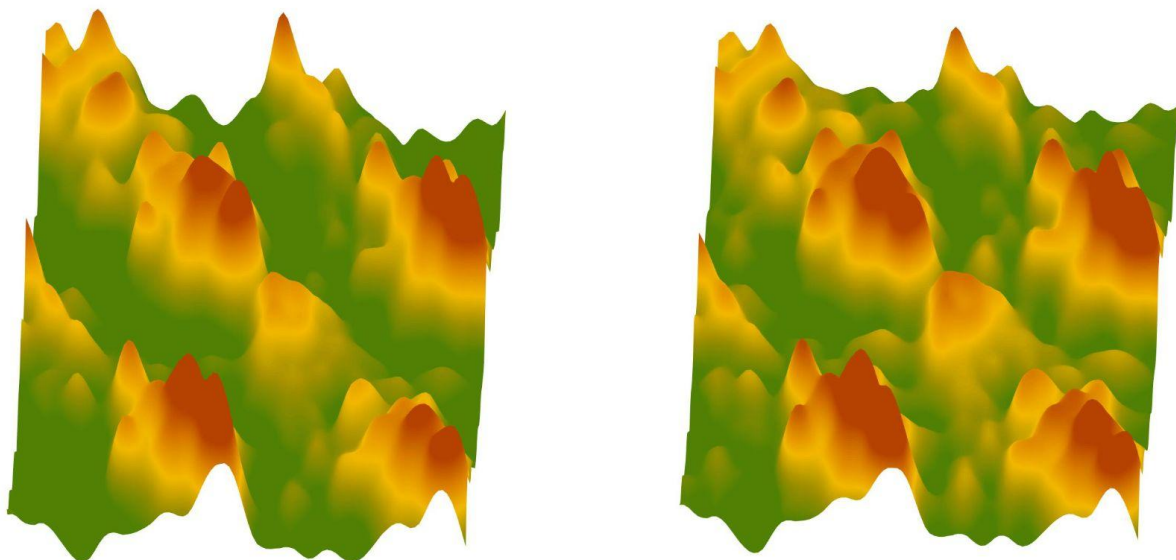
Supplemental Figure 4.3. a) Trial response activation modelling: A two-way repeated-measures voxelwise analysis of variance (RM-ANOVA) modelled activation change with task (match, odd one out) and rule (colour, shape) as the within-subject factors. This model was used to contrast the average effect of trials (a) and the main effects of task (b) or rule (c). This procedure was repeated for study 2 (d-f). g) The main effect of study contrast generated from the switch > trial across the three learning stage activation analysis. To determine the direction of the main effect, two t-contrasts were generated to display study 1 > 2 (green) and study 2 > 1 (lilac) differences. All maps were thresholded at $p < 0.01$ followed by a false discovery rate (FDR) cluster-level correction for multiple comparisons.

“Today I escaped from anxiety. Or no, I discarded it, because it was within me, in my own perceptions — not outside.”

- Marcus Aurelius

Chapter 5: Decreased brain modularity after psilocybin therapy for depression

The findings from previous chapters imply that neurocognitive flexibility is important for successfully adapting behaviour. In particular, it was the activation and connectivity of the MD, DMN and other higher-order brain systems that co-varied with task performance. In this chapter, I analyse resting-state fMRI data from two clinical trials of psilocybin therapy in patients with depression. These data were collected by Dr Robin Carhart-Harris. Addressing aim 3 of this thesis, I examine how recovery from depression relates to the brain systems associated with cognitive control. The findings from this experimental chapter are prepared for submission and, subsequently, the use of “we” refers to my predictions, analysis, data visualisation and interpretation.



Well-being may depend on brain systems that are integrated and flexible.

5.1 Abstract

Importance

Psilocybin therapy shows antidepressant potential. Our data suggests psilocybin may be effective by provoking less modular brain dynamics.

Objective

To assess the sub-acute impact of psilocybin on brain activity in patients with depression.

Design

Pre vs post-treatment resting-state functional MRI (fMRI) was recorded in two trials: 1) Open-label treatment-resistant depression (TRD) trial with baseline vs 1 day post-treatment fMRI (April-2015 to April-2016); 2) Two-arm double-blind RCT in major depressive disorder (MDD), fMRI baseline vs 3 week after psilocybin-therapy or 6 weeks of daily escitalopram (January-2019 to March-2020).

Setting

Study visits occurred at the NIHR Imperial Clinical Research Facility.

Participants

Adult male and female patients with TRD or MDD.

Intervention(s) (for clinical trials) or Exposure(s) (for observational studies)

Study 1: Two oral doses of psilocybin (10mg and 25mg, fixed order, 7 days apart). fMRI was recorded at baseline and one day after the 25mg dose. Study 2: either: 2 x 25mg oral psilocybin, 3 weeks apart, plus 6 weeks of daily placebo ('psilocybin-arm'), or 2 x 1mg oral psilocybin, 3 weeks apart, plus 6 weeks of daily escitalopram [10-20mg] ('escitalopram-arm'). fMRI was recorded at baseline and 3 weeks after the 2nd psilocybin dose which was the final day of the 6-week daily capsule ingestion.

Main Outcome(s) and Measure(s)

Beck Depression Inventory and fMRI network modularity.

Results

Study 1: In 16 adults (mean age [SD], 42.8 [10.1] years, 4 [25%] female), psilocybin therapy was associated with markedly decreased BDI scores at 1 week (mean difference, -21; 95% CI=[-27.3, -14.7], $P<.001$) and 6 months (mean difference, -14.19; 95% CI=[-21.3, -7.1], $P<.001$). Decreased network modularity at one day post-treatment correlated with treatment response at 6 months (pearson, 0.64; $P=.01$). Study 2: In 43 adults (42.7 [10.5] years, 14 [33%] female), antidepressant effects favoured the psilocybin-arm at 2 (mean difference, -8.76; 95% CI=[-13.6, -3.9], $P=.002$) and 6 weeks (mean difference, -8.78; 95% CI=[-15.6, -2.0], $P=.01$). Specific to the psilocybin-arm, improvements at the 6 week primary endpoint correlated with decreased network modularity (pearson, -0.42, $P=.025$).

Conclusions and Relevance

Robust and reliable antidepressant effects and consistent efficacy-related functional brain changes across two studies suggest that psilocybin therapy is a promising treatment option for depression with a plausible antidepressant action: decreased network modularity.

Trial registration

ClinicalTrials.gov identifier: NCT03429075

5.2 Introduction

Depression is a highly prevalent mental health condition (World Health Organisation, 2017), the incidence of which has increased during the Covid-19 pandemic (Tang et al., 2021), e.g., as reflected in increased prescriptions of antidepressant medications (Rabeea et al., 2021). However, even the best performing antidepressant drugs show modest efficacy (Hofmann et al., 2017), non-negligible side effects (Locher et al., 2017), discontinuation problems (Haddad, 1998), and high relapse rates (Steinert et al., 2014), highlighting the need for new improved treatments (Nutt & Carhart-Harris, 2020).

Patients with a diagnosis of depression often exhibit a negative cognitive bias, characterised by pessimism (Lyons & Carhart-Harris, 2018) poor ‘cognitive flexibility’ (Rock et al., 2014), rigid thought patterns (Beck & Clark, 1988) and negative fixations regarding oneself and future prospects (Lyons & Carhart-Harris, 2018). A number of authors have directly or indirectly taken inspiration from dynamical systems theory to describe depressive episodes as ‘attractor states’, i.e., stereotyped states with gravitational pull (Holtzheimer & Mayberg, 2011; Rolls, 2016).

Neuroimaging research has consistently found examples of abnormal brain functioning in depression, resonant with its phenomenology (Hamilton et al., 2015; Kendler, 2016; Nestler et al., 2002). A hierarchically supraordinate intrinsic brain network (Margulies et al., 2016), the default mode network (DMN), is associated with introspection and self-referential thinking (Andrews-Hanna et al., 2014) and these functions are often overactive in depression (Lyons & Carhart-Harris, 2018). Indeed, several studies have linked excessive engagement of DMN functioning with depressive symptomatology (Hamilton et al., 2011; Kaiser et al., 2015).

In addition to the DMN, other higher-order brain networks such as the executive (EN) and salience networks (SN) have been implicated in depression (Beevers, 2005; Ellard et al., 2018; Lydon-Staley et al., 2019). These networks are associated with the ‘cognitive control’ of thoughts and attention switching (Daws, Scott, et al., 2020; Kim et al., 2012) between introspective thought and an external focus (Turnbull et al., 2020). Such attentional switching is impaired in depression (Wilkinson & Goodyer, 2006). Tellingly, the serotonin 2A (5-HT_{2A}) receptor subtype, which is the key proteomic binding-site of ‘classic’ serotonergic psychedelic drugs, such as ‘psilocybin’ (Vollenweider et al., 1998), is most densely

expressed in a broad pattern of cortex that closely resembles a conjunction map of the DMN, EN and SN (Beliveau et al., 2020), corresponding to the upper half of the brain's principal hierarchical gradient (Margulies et al., 2016).

In the last 15 years, at least six separate clinical trials have reported impressive improvements in depressive symptoms with psilocybin therapy (Andersen et al., 2021). Included among these studies are: 1) an open-label trial in treatment-resistant depression (Carhart-Harris et al., 2017), and 2) a double-blind randomised controlled trial (DB-RCT) with an active comparator, the selective serotonin reuptake inhibitor (SSRI) and conventional antidepressant, escitalopram (Carhart-Harris et al., 2021). These two trials, which included pre and post-treatment functional magnetic resonance imaging (fMRI), are the focus of the present paper's analyses.

The therapeutic action of psilocybin and related psychedelics is incompletely understood. However, one model proposes that psychedelics cause a 5-HT_{2A} receptor induced dysregulation of spontaneous population-level neuronal activity, linked to a temporary 'disintegration' of intrinsic functional brain networks (Nutt et al., 2020) and a hypothesised decrease in the precision-weighting of internal predictive models instantiated by these functional modules (Carhart-Harris & Friston, 2019). One important corollary of modular 'disintegration' appears to be the broadening of the brain's functional repertoire of states or 'state-space' – commensurate with a broader or flatter global energy landscape (L.-D. Lord et al., 2019).

Here we hypothesize that the well-replicated finding of brain network disintegration and desegregation under psychedelics (Carhart-Harris et al., 2012; Luppi et al., 2021) will be apparent sub-acutely (Winkelman, 2015), in post-treatment resting-state fMRI data. We also hypothesise that this effect, consistent with a flatter energy landscape, will relate to improved depression outcomes, and also that it will not be observed after a course of the selective serotonin reuptake inhibitor (SSRI), escitalopram.

5.3 Materials and Methods

5.3.1 Trial overviews

The trial designs for and the main clinical outcomes of the open-label (Carhart-Harris et al., 2017) and DB-RCT (Carhart-Harris et al., 2021) have been previously published. Both were conducted at the NIHR Imperial Clinical Research Facility and received Imperial College London Sponsorship, NHS research ethics favourable opinion, Health Research Authority and MHRA approval. This work was done under a UK Home Office Schedule 1 Drug Licence.

5.3.2 Participants

For both trials, eligibility required a general practitioner (GP) confirmed diagnosis of unipolar MDD (16+ on the 21-item Hamilton Depression Rating scale [HAM-D]). The open-label trial had the additional criteria of treatment-resistant depression, as no improvement despite multiple courses of antidepressant medication (mean = 4.6 ± 2.6 past medications, range = 2-11) (Sackeim, 2001).

Exclusion criteria were: Immediate family or personal history of psychosis, risky physical health condition (physician-assessed), history of serious suicide attempts, positive pregnancy test and MRI contraindications. The DB-RCT had the additional exclusion criteria of selective serotonin reuptake inhibitor (SSRI) contraindications or previous escitalopram use. Eligible patients undertook telephone screening interviews, provided written informed consent and their mental and physical medical histories were thoroughly evaluated.

5.3.3 Interventions

Patients in the open-label trial attended a 1-day pre-treatment baseline session that included eyes-closed resting-state fMRI and clinical assessment (Figure 5.1a). This was followed by two psilocybin therapy dosing days (DD), separated by 1 week. A low-dose of psilocybin (10mg) was orally ingested on DD1 and followed by a high-dose dose (25mg) on DD2. The follow-up fMRI and clinical assessment occurred one day after DD2. Patients attended an on-site clinical assessment at 1-week post-DD2 and completed further clinical assessment electronically at 3 and 6 months

Of the 59 MDD patients in the DB-RCT, 30 were randomly allocated to the 'psilocybin-arm' and 29 to the 'escitalopram-arm' (Figure 5.1b). Patients attended a pre-treatment baseline eyes-closed resting-state fMRI. DD1, consisting of either 25mg psilocybin (psilocybin-arm) or 1mg psilocybin (escitalopram arm). All patients were informed that they would receive psilocybin, but were blind to the low-end dosage. DD2 occurred three weeks after DD1 and was a duplicate dosage. There was no dosage-crossover. Beginning one day post DD1, patients took daily capsules for 6 weeks and 1 day in total. For both conditions, one capsule per day was ingested for the first 3 weeks and two thereafter. Capsule content was either inert placebo (microcrystalline cellulose, psilocybin-arm) or escitalopram in the escitalopram-arm, 10mg for the first 3 weeks and 2 x 10mg = 20mg total thereafter.

5.3.4 Measuring depression severity

Beck Depression Inventory 1A (BDI-1A) scores were used to assess depression severity in both studies. This patient-rated measure captures a broader range of symptoms, with an additional focus on the cognitive features of depression, compared to other measures such as the QIDS-SR-16 (Fried, 2017). In the open-label trial, BDI was measured at baseline and 1 week, 3 months and 6 months post DD2. For the DB-RCT, BDI was measured at baseline and 2, 4 and 6 weeks post DD1.

5.3.5 Measuring brain network modularity

For each scanning session, resting-state fMRI was recorded using a 3T Siemens Tim Trio MRI scanner at Imanova, London, UK (see Supplemental Methods: MRI acquisition).

Our principal metric of interest was brain network modularity, a biological substrate that describes the degree of segregation between the brain's functional networks (also known as, the communities of brain regions) (Bassett et al., 2011). Preprocessed fMRI data were used to estimate functional connectivity (FC) matrices from 100 cortical regions as defined by a functional atlas (Schaefer et al., 2018) (see Supplemental Methods: MRI preprocessing; Functional connectivity) that were subjected to a commonly used community detection algorithm (Blondel et al., 2008). This step seeks to maximise the extent to which the brain regions can be segregated into non-overlapping communities or modules (Newman, 2004). See the Supplemental Methods: Modularity section for details on modularity estimation and normalisation.

The modularity metric used in the present work has been applied in many previous contexts, including depression studies (Finc et al., 2020; Lydon-Staley et al., 2019; Zheng et al., 2018), where high modularity scores indicate a greater degree of separation between brain networks. Scores were compared between imaging sessions and correlated with depression severity scores.

5.3.6 Brain network characteristics

The modularity metric assesses a particular property of global brain function. To gain a finer-grained perspective on changes to individual networks, we employed methods from functional cartography (Mattar et al., 2015). For study 1, we measured changes in *network recruitment* as the probability that brain regions of a network form communities with regions from the same network and *network integration* as the probability that regions form communities with regions from other networks (see eMethods: Functional cartography). For study 2, a dynamic community detection and flexibility analysis (Mucha et al., 2010) was applied using a sliding-window approach where FC was estimated for multiple windows in time, instead of for the entire scan. Network flexibility was then defined by the average number of times that brain regions within a given network changed their community affiliation across time (Mattar et al., 2015) (see Supplemental Methods: Dynamic flexibility).

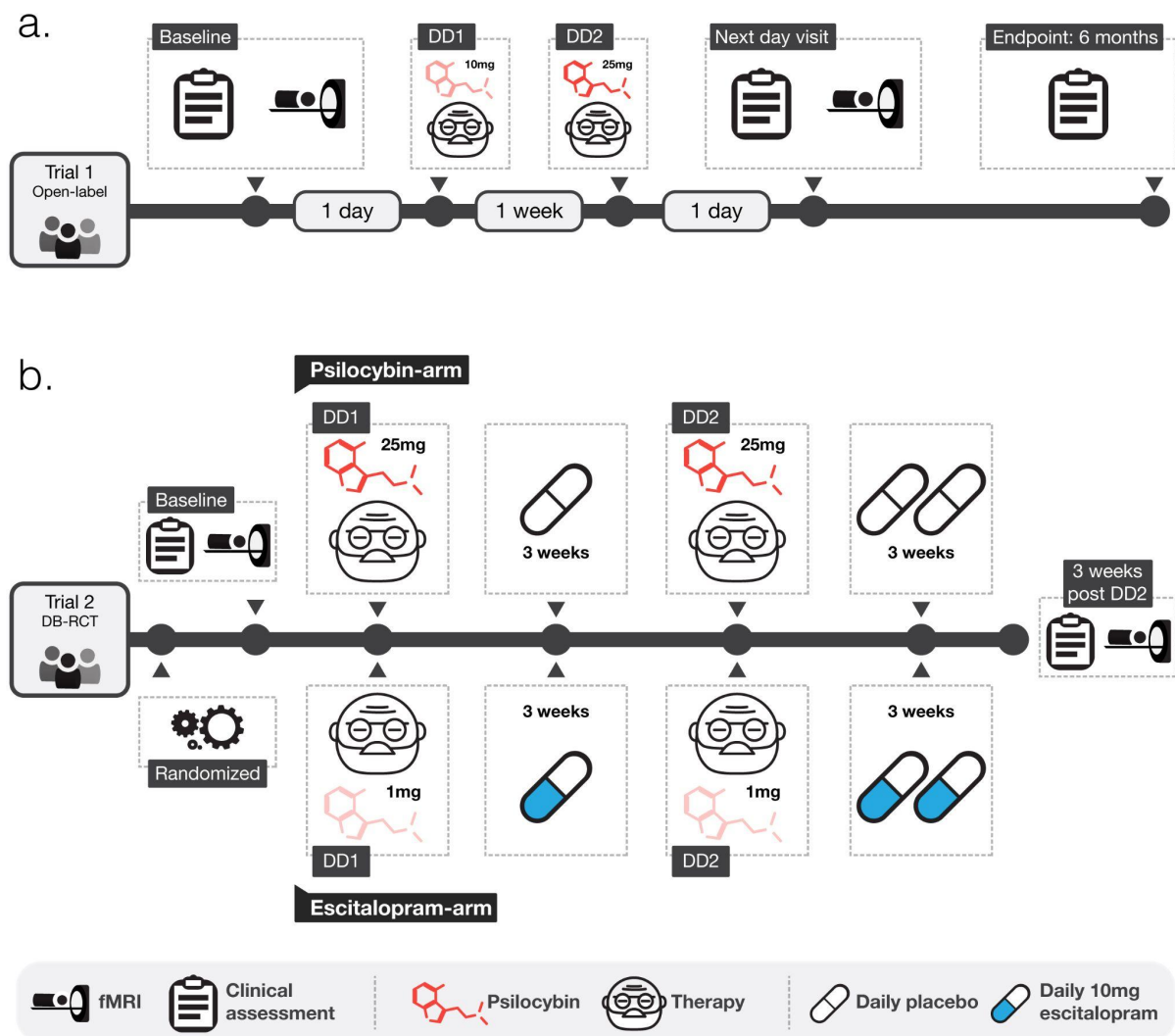


Figure 5.1. Trial timeline schematics. a) Open-label trial. Eligible patients attended a baseline resting-state fMRI and clinical assessment visit. This was followed by two psilocybin therapy dosing days (DD) separated by 1 week, which differed in dose strength (DD1=10mg, DD2=25mg, oral psilocybin). The post-treatment fMRI scan occurred one day after DD2. Remote clinical assessment continued for 6 months. b) DB-RCT. Patients attended a baseline resting-state fMRI and clinical assessment visit and were randomly assigned to the psilocybin-arm (top-branch) or escitalopram arm (bottom-branch). The psilocybin-arm involved 2x25mg psilocybin therapy DD's with 3 weeks of daily placebo capsules following each DD. The escitalopram-arm involved 2x1mg psilocybin therapy DD's with 3 weeks of 10mg daily escitalopram following DD1 and 20mg of escitalopram following DD2. Both groups attended a post-treatment fMRI and clinical assessment visit 6 weeks and one day after DD1.

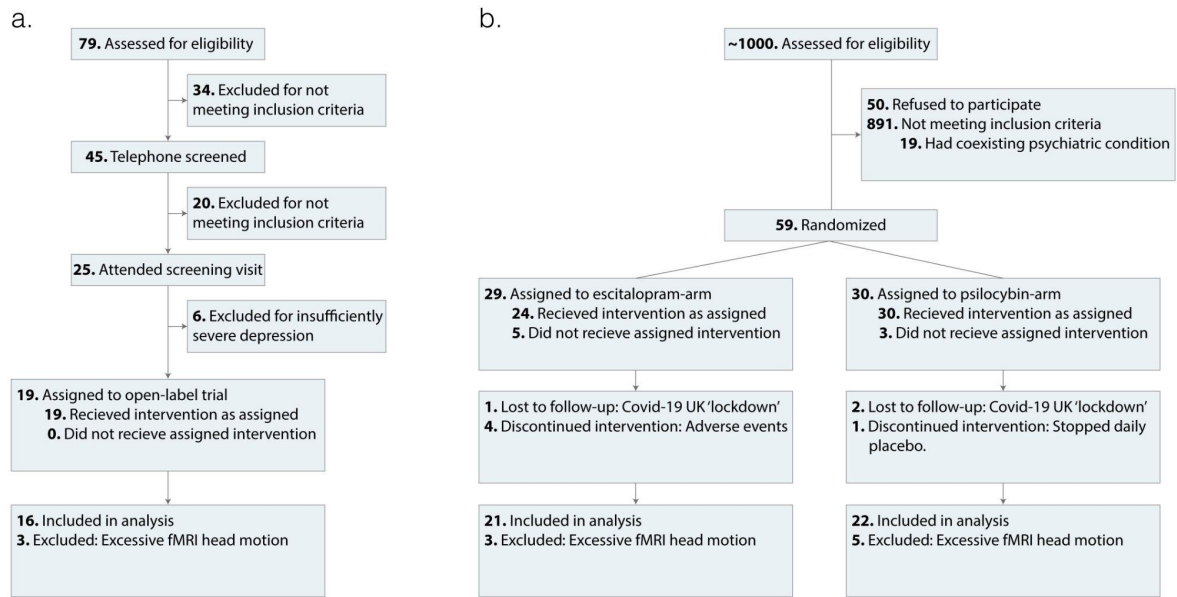


Figure 5.2. Recruitment flow diagram for the open-label (a) and DB-RCT (b).

5.4 Results

Open-label trial

5.4.1 Depression severity

Of the 19 patients recruited, 3 were excluded due to excessive fMRI head-motion. The remaining 16 patients (mean age=42.75, SD=10.15, 4 female) were the final analytical sample (Figure 2a). Baseline BDI scores indicated severe depression (mean BDI=34.81, SD=7.38). As previously reported, rapid and sustained reductions in depression severity were observed post-treatment (Carhart-Harris et al., 2017). Relative to baseline, significant BDI reductions were observed at 1 week (mean difference, -21.0 points; 95% CI, -27.30 to -14.71, $P<.001$) and this was still evident 6 months (mean difference, -14.19 points; 95% CI, -21.29 to -7.09, $P<.001$).

5.4.2 Decreased brain modularity following psilocybin therapy

Confirming our primary hypothesis prediction, brain network modularity was significantly reduced (Figure 5.3a) one day after psilocybin therapy (mean difference, -0.29; 95% CI 0.07 to 0.50, $P=.012$), indicative of an increased integration of brain networks.

5.4.3 Decreased modularity predicts long-term clinical outcomes

We hypothesised that decreased brain network modularity would relate to sustained improvements in depression severity following psilocybin therapy. To test this, we calculated Pearson correlations between the post-treatment brain modularity and the BDI scores from the 3 post-treatment timepoints (1 week, 3 months, 6 months). After false discovery rate (FDR) correction for multiple-comparisons, a strong significant correlation was observed at 6 months (Figure 5.3b - pearson, 0.64; $P=.023$). Although consistent with this, relationships at 3 months (0.46; $P=.114$) or 1 week (0.29; $P=.284$) did not survive correction. Furthermore, the pre vs post-treatment change in modularity significantly correlated with the change in BDI score at 6 months, relative to baseline (Figure 5.3c - pearson, 0.54; $P=.033$). These results indicate that decreased brain modularity relates to long-term improvements in depression symptom severity.

5.4.4 Decreased DMN & increased DMN-frontoparietal FC post-treatment

Consistent with previous work and our a priori assumptions, psilocybin therapy was related to significantly decreased DMN network recruitment (mean difference, -0.54; 95% CI, -0.92 to -0.15, $P=.009$), and increased integration between the DMN and multiple frontoparietal networks (DMN - EN, mean difference, 0.53; 95% CI, 0.15 to 0.90, $P=.01$; DMN - SN, mean difference, 0.55; 95% CI, 0.14 to 0.95, $P=.01$, FDR-corrected). A post-hoc exploratory analysis of network recruitment and integration was conducted between the cortical networks (Figure 5.3d/e). These results indicate a decrease in within-DMN connectivity alongside an increase in DMN connectivity to other higher-order networks, namely the EN and SN.

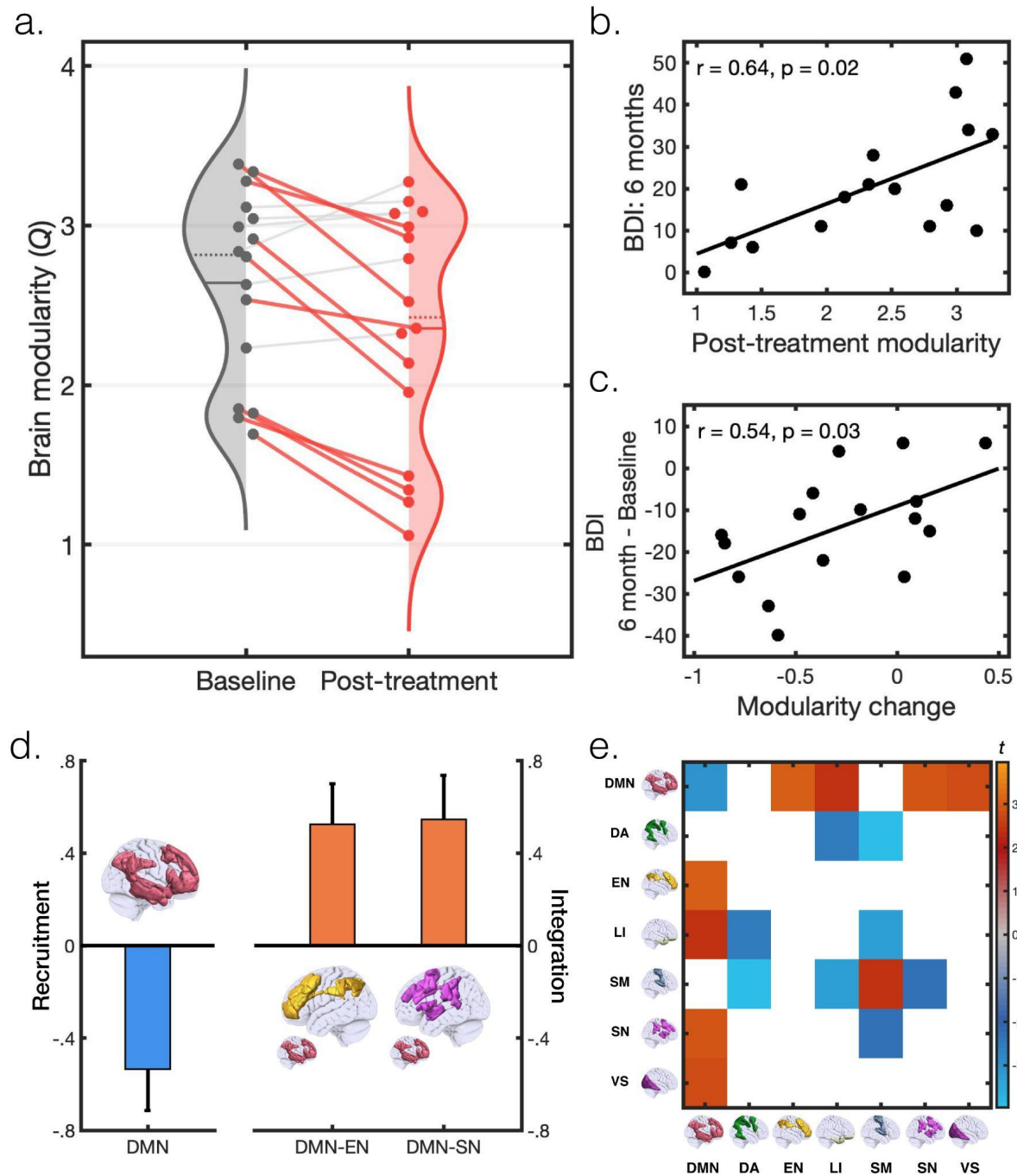


Figure 5.3. Open-label trial: Treatment-resistant depression patient responses to psilocybin therapy relate to decreases in brain network modularity one day post-treatment. a) Brain modularity (Q - normalised) significantly reduced, indicating a global increase in brain network integration. The solid and dotted lines represent the mean and median, respectively. Patient's data are connected by solid lines and rendered red if modularity decreased. b) Absolute post-treatment scan modularity correlated with absolute Beck depression inventory (BDI) scores at 6 months post-treatment. c) Change in brain modularity after treatment significantly correlated with the change in BDI scores between baseline and at 6 months. d) DMN recruitment decreased and its integration with the EN (gold) and SN (purple)

increased (bars represent the standard error of the mean) following psilocybin therapy. e) Paired t-statistics of the change in network recruitment (diagonal) and between network integration (off-diagonal). This exploratory analysis uses an uncorrected $P < .05$ threshold (Red=increase, Blue=Decrease). Comparisons do not survive multiple comparisons correction. DMN=Default Mode Network, DA=Dorsal attention, EN=Executive Network, LI=Limbic, SM=Somatomotor, SN=Saliency Network, VS=Visual.

Double-blind randomised controlled trial

Of the 59 MDD patients recruited, 29 were randomly allocated to the escitalopram-arm. Of those, 21 patients (mean age [SD], 40.9 [10.1], 6 [29%] female) were included in this imaging sample. 30 patients were randomly allocated to the psilocybin-arm. Of those, 22 patients (mean age [SD], 44.5 [11.0], 8 [36%] female) were included (Figure 2b).

5.4.5 Psilocybin therapy has greater efficacy than escitalopram for treating depression

Decreased depressive symptom severity was significantly greater under psilocybin than escitalopram, indicating superior efficacy of psilocybin therapy vs. escitalopram (Figure 5.4). This was confirmed within this sample in whom we had neuroimaging data by a significant arm x timepoint ANOVA interaction for the BDI scores (F , 4.47; $P=0.005$). FDR-corrected pairwise comparisons relative to baseline were significantly different at 2 weeks (mean difference, -8.73; 95% CI = -13.55 to -3.91, $P=0.002$), 4 weeks (mean difference, -7.79, 95% CI = -13.62 to -1.95, $P=0.013$) and at 6 weeks (mean difference, -8.78, 95% CI = -15.58 to -1.97, $P=0.013$), all favouring the psilocybin-arm.

5.4.6 Increased brain network integration is specific to psilocybin therapy

Confirming our primary hypothesis (Figure 5.5a-b) and replicating the findings of the open-label trial, brain network modularity significantly reduced following psilocybin therapy (mean difference, -0.39; 95% CI = -0.75 to -0.02, $P=0.039$). Individuals' changes in brain network modularity significantly correlated with changes in BDI scores at the 6 week primary endpoint (pearson, 0.42, $P=.025$, one-tailed).

Importantly, this replication was specific to the psilocybin-arm; in the escitalopram group (Figure 5.5d-e), modularity did not change from baseline to week 6 (mean difference, 0.01; 95% CI -0.35 to 0.33, $P=0.945$) and there was no significant relationship with changes in BDI scores (pearson, 0.08; $P=0.361$, one-tailed).

5.4.7 Depression recovery correlates with increased cognitive network flexibility.

Next, we examined the dynamic flexibility of the brain's canonical networks. This finer-grained metric summarises how often brain regions change their community allegiance during the course of an fMRI scan. Post-treatment changes in network flexibility were correlated with the changes in BDI score. Specifically, increased EN dynamic flexibility related to greater depression recovery at the 6 week primary endpoint for the psilocybin-arm (pearson, -0.76, $P=0.001$). Significant relationships predominantly involved the EN, SN and dorsal attention networks (Figure 5.5c). No significant correlations between BDI and dynamic flexibility were observed in the escitalopram-arm (Figure 5.5f).

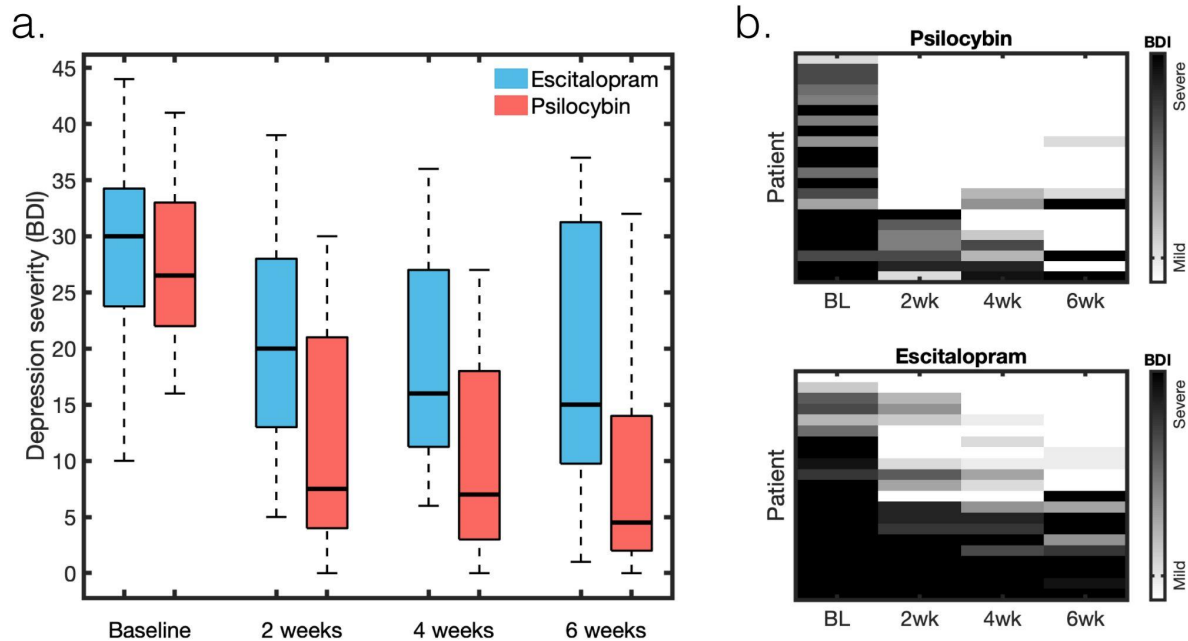


Figure 5.4. DB-RCT: Beck depression inventory (BDI) scores from each study arm. a) Boxplot for each timepoint at which BDI scores were measured. Decreases in depression severity from baseline were significantly greater in the psilocybin-arm (red) than in the escitalopram-arm (blue) at each timepoint. Median values are represented by the central marks, the 25th and 75th percentiles by the box edges and the whiskers extend to the data range. b) Individual patient BDI scores in a raster plot. Each row represents a single patient, each column represents a timepoint (BL=Baseline, wk=week). Rows were ordered by each patient's total BDI score sum across timepoints. Black/grey=Severe depression, white/light grey=Mild to no depression.

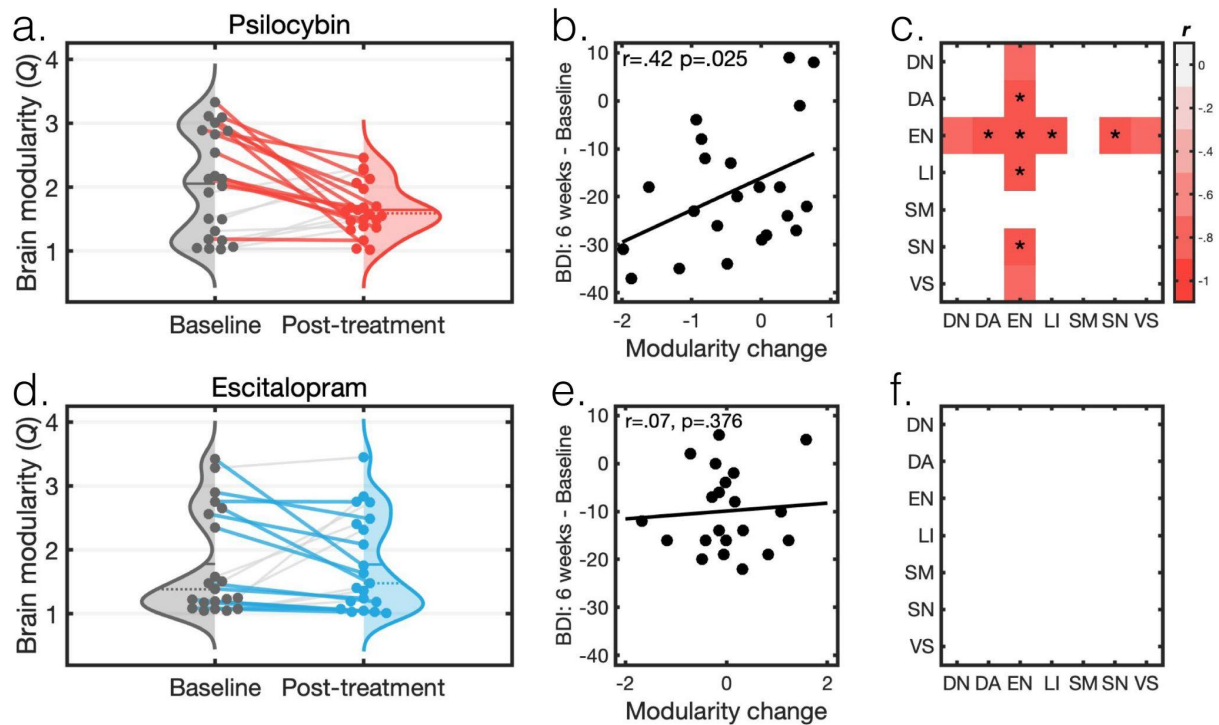


Figure 5.5. DB-RCT: Brain modularity reductions and individual differences are specific to psilocybin treatment. a) Significant decreases in brain network modularity (Q - normalised) in the psilocybin-arm post-treatment. The solid and dotted lines on the distributions represent the mean and median, respectively. Individual patient data are represented, connected with solid lines between sessions which were rendered red if modularity decreased between sessions. b) Patient individual differences (relative to baseline) correlate with improved treatment response at the 6-week primary endpoint. c) Significant correlations between changes in dynamic network flexibility and BDI (relative to baseline) at 6 weeks are coloured and those that survive FDR-correction are denoted with a * (white = $P>.05$). The equivalent analyses in the escitalopram-arm did not show significant session differences in brain modularity (e) and individual differences in BDI at 6 weeks (relative to baseline) did not correlate with modularity changes (e) or network flexibility (f). DN=Default Mode Network, DA=Dorsal attention, EN=Executive network, LI=Limbic, SM=Somatomotor, SN=Salience Network VS=Visual.

5.5 Discussion

In light of growing evidence for the antidepressant efficacy of psilocybin therapy (Carhart-Harris et al., 2021), these findings advance our understanding of possible underlying brain mechanisms. Across two separate trials decreased brain modularity was observed and correlated with improvements in depressive symptomatology. Moreover, this antidepressant action may be specific to psilocybin therapy, as no changes in modularity were observed post treatment with the conventional SSRI antidepressant, escitalopram.

Previous research into the acute brain action of psychedelics have revealed well-replicated changes in global brain function that are somewhat consistent with those observed here, i.e., an increased repertoire of inter-regional and between-network FC (L.-D. Lord et al., 2019; Luppi et al., 2021; Roseman et al., 2014). Our previous analysis had suggested some contrasting changes in the architecture of spontaneous brain function post psilocybin treatment for depression versus during the acute psychedelic state itself – i.e., spatially expanded DMN FC (post-treatment for TRD) versus acute DMN disintegration (Carhart-Harris et al., 2017). Other teams have, however, reported some evidence suggestive of increased inter-network FC 1 week and 1-month post-psilocybin (Barrett et al., 2020), as well as 1-day post-ayahuasca (Pasquini et al., 2020), including consistent increases in DMN-SN FC (Pasquini et al., 2020), albeit in healthy volunteers. The present findings greatly extend on previous work however, by showing robust, reliable and treatment-specific decreases in brain modularity post psilocybin therapy for depression that relate to antidepressant efficacy.

The present modularity metrics may be more sensitive indices of the antidepressant action of psilocybin than previously applied time-averaged within and between-network FC analyses (Carhart-Harris et al., 2017). Indeed, they may bear relevance to other FC metrics applied to acute-state psychedelic fMRI data (L.-D. Lord et al., 2019; Luppi et al., 2021; Roseman et al., 2014) where a general picture of increased global FC and a broadened state-space has emerged (Carhart-Harris & Friston, 2019). In this context, the results could be understood as a ‘carryover’ effect resembling brain dynamics associated with the acute psychedelic state, albeit at an attenuated level and in a specific population (i.e., depressed patients).

Previous research on resting-state activity in depression has found abnormal community structure (A. Lord et al., 2012) and heightened network modularity (Wei et al., 2017), correlating with symptom severity (Lydon-Staley et al., 2019; Ye et al., 2015). Additional work implies elevated FC between limbic regions such as the amygdala and high-level cortical regions in depression, correlating with ruminative symptoms (Feurer et al., 2021), as well as elevated within-DMN FC also correlating with rumination (Hamilton et al., 2011). Taken together, a model emerges of abnormally modular spontaneous brain function in depression that is effectively remediated by psilocybin therapy. According to various findings, the FC energy landscape or state-space in depression can be described as abnormally constricted, paralleling the narrow, internally focused, ruminative quality of mood and cognition in the disorder (Rolls, 2016). In contrast, psilocybin therapy appears to expand the brain's state-space, both acutely (Petri et al., 2014) and, as has been shown here, post-acutely in depressed patients, in a fashion that correlates with antidepressant outcomes. Moreover, this 'liberating' action of psilocybin is paralleled by subjective reports of emotional release via psychedelic therapy (Roseman et al., 2019; Watts et al., 2017) as well as sub-acute increases in behavioural optimism (Lyons & Carhart-Harris, 2018), cognitive flexibility (Murphy-Beiner & Soar, 2020), and psychological flexibility post psychedelic-use (Close et al., 2020; Davis et al., 2020; Zeifman et al., 2020).

We believe that this 'liberating' effect of psilocybin on cortical activity occurs via its direct agonist action on cortical 5-HT_{2A} receptors, dysregulating activity in regions rich in their expression. We believe chronic escitalopram does not have the same effect on brain modularity due to its more generalised action on the serotonin system and likely predominant effect on inhibitory postsynaptic 5-HT_{1A} receptors, which are richly expressed in limbic circuitry (Carhart-Harris & Nutt, 2017; Nutt et al., 2020).

Beyond the global decrease in network modularity post psilocybin, we observed functional changes in default mode, executive and salience network dynamics that are consistent with neurobiological models of depression (X. Wang et al., 2016). These higher-order frontoparietal networks house the highest density of the 5-HT_{2A} receptors, the principal action-site for serotonergic psychedelics (Beliveau et al., 2020; Vollenweider et al., 1998). High-level frontoparietal networks are implicated in the acute action of psychedelics, where they show reduced modularity and increased communication with regions ordinarily beyond their limits of allegiance (Carhart-Harris et al., 2012; L.-D. Lord et al., 2019; Luppi et al., 2021; Roseman et al., 2014).

The EN and SN have been associated with tasks requiring cognitive flexibility such as, planning (Hampshire et al., 2013), learning (Hampshire et al., 2019) and task-switching (Daws, Scott, et al., 2020; Kim et al., 2012); impaired functioning of these networks have been reported in depression (Lydon-Staley et al., 2019; Wei et al., 2017), and other disorders exhibiting cognitive inflexibility such as traumatic brain injury (Hellyer et al., 2015), autism spectrum disorder (Watanabe & Rees, 2017) and obsessive-compulsive disorder (Gu et al., 2007). Our results suggest that decreased modularity, or increased flexibility, of EN regions, following psilocybin therapy, is a key component of its therapeutic mechanism of action. We did not formally assess cognitive flexibility in the clinical trials reported here but we did observe improvements in general cognitive functioning post psilocybin treatment in the DB-RCT (Carhart-Harris et al., 2021).

Phase 3 clinical trial evidence will be required to achieve licensing for psilocybin therapy and pragmatic trials would help inform questions regarding treatment practicability and optimization (Andersen et al., 2021). For more mechanistic research, we would recommend using network modularity measures such as those employed here. Brain imaging research can generate large statistical spaces that can be burdensome to manage and susceptible to noise, contributing to the challenge of detecting reliable biomarkers. Composite measures of brain function, such as dynamic network modularity, combined with a symptoms-based approach, may constitute an improved approach in this regard, however (Andersen et al., 2021).

The dynamic flexibility analysis employed in trial 2 provided a useful perspective. However, it is limited by its requirement for fMRI scans with many timepoints. Timeseries need to be of sufficient length to be split into multiple time-windows that are themselves sufficiently long to compute reliable FC measures. It can be challenging to reliably collect high-quality data of sufficient length in patient cohorts. Advances to fMRI temporal resolution, however, may improve this issue in the near future.

5.6 Limitations

This investigation relies on several abstractions of the fMRI data. In particular, the data were transformed into Pearson correlation FC matrices which were then summarised by network modularity and other cartography measures, including

network integration and dynamic flexibility. Arguably, all measures of brain function may suffer from limited validity and specificity. fMRI activation analyses benefit from several decades of implementation and they still require careful interpretation given the inherent lags in neurovascular coupling and the array of non-neuronal confounding factors. Network-level and dynamic analyses of fMRI data require further development both in terms of the methodology and the conceptual interpretations. Nonetheless, the fervent research into network-level and dynamic descriptions of brain function has presented a significant potential for gaining understanding (Parkes et al., 2021).

Here, a hypothesis regarding a reduction in network modularity was confirmed in two independent studies. Evidence was then presented which suggested that the change in network organisation was meaningful as it correlated with patient outcomes. However, the modularity metric is multivariate and can therefore be influenced by many factors as it is derived from an underlying FC matrix. The limited specificity of network modularity was supplemented by the network-level cartography analyses which identified networks whose organisation change provided a plausible explanation of a modularity decrease.

The dynamic flexibility analysis in study 2 may suffer from several potential confounds. As mentioned, there is no consensus on how to perform these analyses as the approaches are still in development. It may be challenging to even achieve a consensus in the short-term given the variability in fMRI temporal resolution and SNR that protocols currently offer. However, it is reasonable to be guided by certain principles, such as those relating to statistical power. For example, sliding-window analyses can be used to assess dynamics, with respect to changes in the correlation weights over time. However, there are several considerations that will influence the degrees of freedom such as the number of ROIs, the window-size, window-overlap and number of windows. For dynamic flexibility measures, a sufficient number of TRs need to be available to reliably estimate a Pearson correlation within a window and for a sufficient number of windows to be defined.

Moreover, parameters for the multilayer modularity estimation such as the within and inter-layer coupling can vary.

Here, windows were set to 30 TRs with 50% overlap. For dynamic fMRI analyses, 30 TRs is a typical, but conservative, use of the data as it provides a reasonable amount of statistical power for estimating Pearson FC while maximising the number of windows available across time. Moreover, in line with other studies, the gamma and omega coupling parameters were set to 1 (Bassett et al., 2011; Finc et al., 2020). Effectively, this could be thought of as a default because the coupling strengths were not being up-weighted or down-weighted.

Despite this, these dynamic analyses are limited in that there may be many factors that may cause fluctuations in the cross-correlations of timeseries over time. Do these fluctuations represent something that is neuronal, stochastic noise, or both? While stochastic noise has not been ruled out by the current investigation, it is worth noting the strong correlations between the increased dynamic flexibility of specific networks and improved patient outcomes ($r \sim 0.7-0.8$). Critically, these orthogonal findings were specific to the psilocybin-arm, there was no evidence of this effect in the escitalopram-arm.

This investigation is limited as it does not rule out these potential confounds when it comes to dynamic analyses of fMRI data. There is scope for future research to investigate the nature of these 'dynamic' fluctuations via parameter sweeps of window size, window overlap and noise. Simulations of this kind may help to help understand the neural components of a fluctuating fMRI signal and to identify the optimal parameters for minimising noise and other confounding factors.

5.7 Conclusion

Depression presents considerable challenges to multiple stakeholders. Here, we identify a robust, reliable and potentially specific biomarker of response to psilocybin therapy for depression that may help to explain why it could become a valuable new treatment option.

5.8 Supplemental materials

5.8.1 MRI acquisition

Brain imaging was performed on a 3 T Siemens Tim Trio at Invicro, London, UK. Anatomical images were acquired using the ADNI-GO (Alzheimer's Disease Neuroimaging Initiative, Grand Opportunity) (Jack et al., 2008) recommended MPAGE parameters (1 mm isotropic voxels, TR = 2300 ms, TE = 2.98 ms, 160 sagittal slices, 256×256 in-plane FOV, flip angle = 9 degrees, bandwidth = 240 Hz/pixel, GRAPPA acceleration = 2).

In both studies, fMRI data were collected with T2*-weighted echo-planar images (EPI) with 3 mm isotropic voxels. In study 1, a 12 channel head coil was used to acquire 240 volumes in ~8 minutes: TR = 2000 ms, TE = 31 ms, 36 axial slices, flip angle = 80 degrees, bandwidth = 2298 Hz/pixel, GRAPPA acceleration = 2). In study 2, a 32-channel head coil was used to acquire 480 multiband-accelerated (2x) volumes were acquired in ~10 minutes: TR = 1250 ms, TE = 30 ms, 44 axial slices, flip angle = 70 degrees, bandwidth = 2232 Hz/pixel, GRAPPA acceleration = 2).

5.8.2 MRI preprocessing

Imaging data were preprocessed via a custom in-house pipeline composed of tools from the FMRIB Software Library (FSL) (S. M. Smith et al., 2004), Analysis of Functional NeuroImages (AFNI) (Cox, 1996), Freesurfer (Dale et al., 1999) and Advanced Normalisation Tools (ANTs) (Avants et al., 2009) packages. Patients were excluded if either fMRI scan contained >20% of volumes with a framewise displacement (FD) greater than 0.5 mm (Power et al., 2014).

Specifically, the following pre-processing stages were performed: 1) removal of the first three volumes; 2) de-spiking (3dDespike, AFNI); 3) slice time correction (3dTshift, AFNI); 4) motion correction (3dvolreg, AFNI) by registering each volume to

the volume most similar, in the least-squares sense, to all others; 5) brain extraction (BET, FSL); 6) rigid body registration to anatomical scans (BBR, FSL); 7) non-linear registration to 2 mm MNI brain (Symmetric Normalization (SyN), ANTS); 8) scrubbing (Power et al., 2012) - using a framewise displacement (FD) threshold of 0.5 mm, scrubbed volumes were replaced with the mean of the neighbouring volumes. 9) 6mm full width at half maximum (FWHM) gaussian spatial smoothing (3dBlurInMask, AFNI); 10) 0.01 to 0.08 Hz band-pass filtering (3dFourier, AFNI); 11) linear and quadratic de-trending (3dDetrend, AFNI); 12) Voxelwise nuisance regression with the 6 realignment motion regressors and 3 tissue signal regressors (ventricles (Freesurfer, eroded in 2 mm space), draining veins (DV) (FSL's CSF minus Freesurfer's Ventricles, eroded in 1 mm space) and local white matter (WM) (FSL's WM minus Freesurfer's subcortical grey matter (GM) structures, eroded in 2 mm space). Regarding local WM regression, AFNI's 3dLocalstat was used to calculate the mean local WM time-series for each voxel, using a 25 mm radius sphere centred on each voxel (Jo et al., 2010).

5.8.3 Functional connectivity (FC)

A functional atlas was used to separate the cerebral cortex into 100 regions of interest (ROI) (Schaefer et al., 2018). We then estimated the FC between each pair of ROIs by calculating a Pearson correlation coefficient between each pair of mean signal 'timecourses' (representing fluctuations in neural activity over time). This resulted in an FC matrix, with each element representing the connectivity strength between a pair of ROIs. Positive values were retained and Fisher transformed to z-scores. This procedure was repeated for each patient and scan (baseline, follow-up).

5.8.4 Modularity

The FC matrices were used to summarise the community structure and segregation of brain networks using a common Louvain-like community detection algorithm (Blondel et al., 2008) where the objective is to maximise the extent to which brain areas can be separated into non-overlapping communities or modules (Newman,

2004). The modularity quality function score, (Newman, 2004, 2006; Newman & Girvan, 2004), tends to be high when the brain exhibits a high segregation between its functional networks (e.g., strong clusters of FC within brain networks/communities with weak FC to the rest of the brain).

This approach has been commonly applied to fMRI data to characterise how brain function adapts in a number of contexts (Bassett et al., 2011; Finc et al., 2020), including in depression studies (Gong & He, 2015; Zheng et al., 2018). Here, modularity, Q , (Mucha et al., 2010; Newman, 2004) was defined in the standard way by:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j),$$

where A_{ij} represents the weight of FC (correlation) between ROI i and j , γ is the structural resolution free parameter (set to 1) and $\frac{k_i k_j}{2m}$ is the expected null FC defined with $k_i = \sum_j A_{ij}$ as the total FC across all connections with ROI i , c_i is the community to which ROI i is assigned, $\delta(c_i, c_j)$ is the Kronecker delta function and equals 1 if ROI i and j belong to the same community and 0 otherwise (Bassett et al., 2013) and $m = \frac{1}{2} \sum_{ij} A_{ij}$, is the total FC of the network.

To allow valid comparisons between patients and scans, the modularity scores were generated 100 times and the partition with the largest modularity score was normalised by the mean modularity generated from 100 randomly rewired (shuffled) FC matrices (Maslov & Sneppen, 2002). This common procedure was applied to account for the non-deterministic and near-degenerate partitions (solutions with differing but similar optimality) generated by Louvain algorithms and to account for modularity scores relating to the total sum of FC within the network (Rubinov & Sporns, 2011). This process was repeated independently for each patient and scan.

5.8.5 Functional cartography

The community detection procedure generates a community assignment to each ROI. We used these labels to determine the extent to which ROIs were recruited to the functional network that they typically ‘belong to’, as defined by healthy adults (e.g., DMN regions should reliably form communities with each other).

First, we created an allegiance matrix, P , (Bassett et al., 2015; Mattar et al., 2015) which represented the probability that two regions i and j were assigned to the same community across the 100 iterations of the modularity algorithm, defined here as:

$$P_{ij} = \frac{1}{O} \sum_{o=1}^O a_{ij}^{k,o},$$

where $O = 100$ as the number of partitions. For each partition, $a_{ij}^{k,o}$ equals 1 if regions i and j belong to the same community.

Using the allegiance matrix, we then calculated scores for each region that summarised the extent that regions from the same functional network form communities together (recruitment) or regions from two different systems form communities together (integration) (Bassett et al., 2015; Mattar et al., 2015). These functional cartography measures were then normalised against the mean values from 1000 randomly shuffled ROI-network assignments to account for differences in the number of regions in each network (Finc et al., 2020). Recruitment and integration scores were then averaged at the network level using 7 predefined cortical networks (Yeo et al., 2011).

5.8.6 Dynamic flexibility

A dynamic flexibility analysis was conducted in study 2 using a multilayer modularity estimation (Lucas et al., 2016; Mucha et al., 2010). FC matrices were generated from 30-volume sliding-windows (37.5 seconds of real-time) with 50% overlap

resulting in a $N \times N \times T$ matrix for each patient and scan. This window size is typical for estimating dynamic FC with fMRI (Leonardi & Van De Ville, 2015). Multilayer modularity, Q_{ML} , was estimated 100 times from each $N \times N \times T$ FC matrix by:

$$Q_{ML} = \frac{1}{2\mu} \sum_{ijlr} \left[\left(A_{ijl} - \gamma_l \frac{k_{il}k_{jl}}{2m_l} \right) \delta_{lr} + \delta_{ij} \omega_{jlr} \right] \delta(c_{il'} c_{jr}),$$

where $\mu = \frac{1}{2} \sum_{ijl} A_{ijl}$ is the total FC of the multilayer network, and $m_l = \frac{1}{2} \sum_{ij} A_{ijl}$ is the total FC of layer l , A_{ijl} is the FC between ROI i and j at layer l , $\frac{k_{il}k_{jl}}{2m_l}$ is the expected null FC at layer l . The two free γ and ω structural and temporal resolution parameters are used to scale the number of communities and strength of inter-layer edges, respectively. As is typical for dynamic modularity fMRI analyses, both were set to 1 (Finc et al., 2020; Harlalka et al., 2019; Mattar et al., 2015).

The multilayer modularity estimation generates an $N \times T$ matrix where each element represents the community assignment of each ROI at each layer (time-window). From this, the flexibility metric, f , can be simply calculated as the number of times an ROI changes its community allegiance, given the number of observations (Mattar et al., 2015):

$$f_i = 1 - \frac{1}{T-1} \sum_l^{T-1} \delta(c_{il'} c_{jl+1}),$$

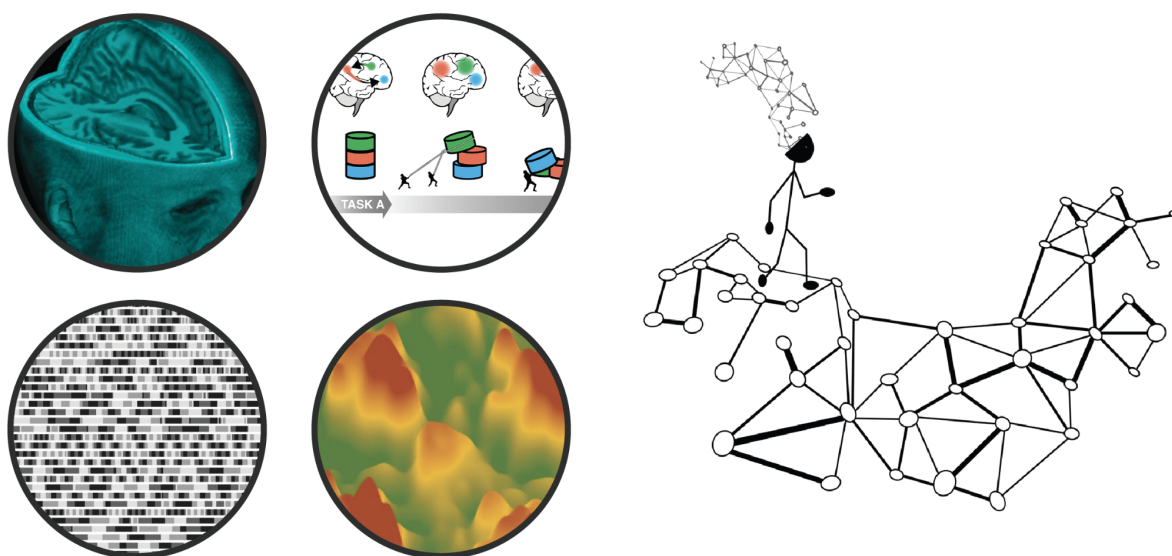
Flexibility scores close to 0 represent rigid ROIs whose community allegiance is stable across time, scores close to 1 represent flexible ROIs whose community allegiance regularly changes (highly flexible).

“If I have seen further it is by standing on the shoulders of Giants”

— Issac Newton

Chapter 6: General Discussion

This chapter provides the opportunity to synthesise the main findings from the research presented in this thesis. The first objective is to re-state the aims of the thesis and evaluate the extent to which they align with the experimental findings. From this, the main threads are drawn together to emphasise the novelty and importance of the potential inferences this is research. Finally, these insights are carried forward to a discussion of future directions that could expand on the present findings and address the new questions that they raise.



The flexibly ordered brain: The adaptive sequencing of neurocognitive states.

6.1 Synthesis and inference

This thesis investigated how human brain systems adapt to cognitive demand and how these systems are implicated in recovery from depression. The aims of this thesis were:

- 1 - To identify the brain systems associated with exogenous attention switching.*
- 2 - To test how the brain optimises switching-cost demand by structuring behaviour.*
- 3 - To test how cognitive control systems are implicated as patients with depression respond to a psilocybin therapy clinical intervention.*

The fMRI task activation analyses presented in Chapter 3 generated findings that imply that a distributed network of brain regions respond to attention switching. In contrast to what might be predicted by modular accounts of brain function, activation increased in many regions of the brain that are considered to be part of the multiple-demand (MD) network (Fedorenko et al., 2013). Activation increases were observed in large portions of the anterior and lateral prefrontal cortex, anterior insula and dorsal lateral parietal cortices. MD activation during cognitive demand isn't novel in itself. This finding is a further example that can be predicted from the distributed functional network accounts of cognitive control (Duncan, 2013).

However, the finding that MD activation increased in a somewhat uniform manner may be novel. This was not predicted either by the hierarchical accounts of frontal lobe organisation (Badre & D'Esposito, 2007; Koechlin et al., 2003) nor was it fully predicted by the MD account (Crittenden & Duncan, 2014; Duncan, 2013). There was no evidence of frontal lobe activation fractionating as cognitive demand increased. It is important to consider what this could represent given the reports of activation fractionating as a function of cognitive demand during instruction based

learning (IBL) (Hampshire et al., 2016; Mohr et al., 2016), working memory (Soreq et al., 2019) and planning paradigms (Hampshire et al., 2013).

Despite the modest sample size, the paradigm did deliver robust differences in terms of behavioural performance and functional activation. Moreover, these differences were explainable by the cognitive demand that the switching events presented. For example, the low-level perceptual switches were associated with smaller switching-costs and weaker MD activation relative to the higher-order updates to, targets, rules and stimuli sets. Subsequently, observing a uniform activation increase across MD regions is unlikely to represent a false negative.

It could be argued that this novel paradigm places a relatively high cognitive demand on the subjects. Switching events occurred on every trial and there was little opportunity to predict the kinds of switches that would occur in the sequence. This limits the extent to which cognitive demand can be mitigated through learning as seen during IBL (Hampshire et al., 2016) or by structuring behaviour as seen in Chapter 4 and during voluntary task switching (VTS) (Arrington & Logan, 2004, 2005). The task-set, or procedure, held in mind was being constantly updated. This is a context of high cognitive demand and attention. Subsequently, these findings could be considered explainable by the MD account of frontal lobe organisation because the ‘entire’ MD system is predicted to be recruited under higher cognitive demand (Crittenden & Duncan, 2014). Despite the experimental manipulations that varied demand, it could be that all of the task events were in the ‘upper end of cognitive demand and therefore all switches recruited most of the MD system.

Yet, this would not explain the lack of regional fractionation across the least or most demanding switching events, even if all of the switches generated high demand. If this is a reasonable assertion, then this underlines the potentially limited generalisability of both of the hierarchical or MD models of frontal lobe organisation. It is important to recognise this because these accounts seek to describe *fundamental* properties of the frontal lobe’s functional organisation. Therefore, it is

reasonable to expect to observe evidence of their claims in all contexts that involve frontal lobe function.

Beyond presenting a potential limitation of current frontal lobe models, there are broader messages from Chapter 3. These findings reinforce the distributed network properties of cognitive control. They imply that updating the representation of the present context is an important property of brain function that requires the synchronisation of many spatially distal regions, even when low-level perceptual updates are required. These findings, therefore, challenge the notion that task-switching is subserved by a specific region of aPFC, e.g., the frontal poles (Koechlin et al., 1999).

A novel feature of Chapter 3 was the finding that switching-costs vary with cognitive demand. The increased switching-cost and MD activation distinguished the low-level perceptual switches from the higher-order abstract switches. From this, it could be inferred that the abstract updates to rules and task-switching are more costly because a greater reconfiguration of the attentional state is required. This paints cognition as a hierarchical phenomenon whereby task representations are built from fine-grained representations that are grouped by increasing levels of abstraction. If so, a higher-order update to the representation may be costly due to the drastic overhaul of the entire hierarchy; whereas, a ‘tweak’ to the low-order mappings may be less costly as it may not necessitate a propagation throughout the entire task representation.

Despite BOLD’s indirect measure of neural activity, it is reasonable to draw - *cautious* - inferences which suggest that these high-level reconfigurations also present a greater regional metabolic demand (Logothetis, 2008; Logothetis & Wandell, 2004). Indeed, switching was accompanied by concurrent increases in MD activation and decreases in DMN activation. These potentially opposing mixtures of activation change would, therefore, make it difficult, and probably unreasonable, to draw strong conclusions about changes to global metabolism. However, it can be

argued that the metabolic demands of those MD regions were significantly greater during high-level reconfiguration, vs low perceptual remapping.

Given that switching is so costly, it follows that it would be advantageous to define, implement and refine behavioural routines that would minimise the cognitive demand of switching when managing competing task demands. Daily life requires the management of and switching between distinct competing goals. One could randomly engage with these goals but this would likely result in a greater than necessary amount of switching-cost and encourage inefficient modes of operation. Alternatively, an extreme minimisation of switching could allow a single task to be worked on to completion; however, this would be at the detriment to other tasks and may prohibit the adaptation of goals and priorities to the changing environmental demands.

Naturally, the efficiency of human behaviour will likely depend on the ability to flexibly order task sequences: Ordering behavioural routines to minimise cognitive demand but, at the same time, being flexible enough to adapt the routine by evaluating one's progress to the changing environmental demands. Understanding impairments to either of these aspects has clear implications for various clinical populations. For example, patients with obsessive compulsive disorder (OCD) can fail to flexibly remap task rules during reversal learning (Chamberlain et al., 2008) and patients with depression can exhibit hyperactive inwardly focused rumination (Rolls, 2016).

Chapter 4 examined how the healthy adult brain adapts to novel demands as behavioural routines are developed. This rare opportunity was afforded by the novel self-ordered switching (SOS) paradigms that were designed by Dr Adam Hampshire. The two SOS experiments allowed individuals to structure their behaviour whilst managing and switching between, competing demands. Building on, and replicating, the findings from Chapter 3, MD activation was not only predicted by switching events but it varied over time in a manner that was consistent with task-learning.

It is fascinating to consider the concurrent optimisation of the brain and behaviour as novel demands are practised. In most task fMRI experiments, the events are strictly controlled meaning that individuals are reacting rather than deciding. While this approach is valuable, the SOS paradigm is particularly novel as individuals could behave as they saw fit.

As would be predicted from IBL experiments (Hampshire et al., 2016; Mohr et al., 2016), individuals rapidly optimised their SOS behavioural performance in terms of faster response speed. Trial response speed increased and switching-costs decreased. Beyond this, there was evidence that behavioural routines became increasingly structured as SOS was practised. The structuring of behaviour is likely advantageous for multiple reasons. However, prior research has indicated that switching-costs are diminished when switching-events are predictable (R. D. Rogers & Monsell, 1995). In such cases, predictability follows from individuals observing the patterns of events generated by the paradigm. In contrast, predictability in Chapter 4 likely increased as individuals solidified their behavioural routines as they found their solutions.

Critically, the changes in MD activation that accompanied this learning process dissociated the simple trials and demanding switching events. Practising the simple visual discriminations on each trial was associated with ‘learning curves’ that have been observed in IBL experiments (Hampshire et al., 2016, 2019; Mohr et al., 2016) with MD activation decreasing with practice. Intriguingly, the opposite trend was observed for switching events as MD activation increased with practice.

Together, this may indicate a novel insight into learning and the brain’s adaptation of behaviour. As recognised in Chapter 3, task-switching is costly and particularly when the switch involves updates that are higher-order and abstract. The results from Chapter 4 may reflect an ability to rapidly learn and automate the low-level elements of a relatively complex behavioural routine (stimulus-response mappings of simple visual discriminations, in this case). This process may reflect a mechanism

of adaptive resource reallocation whereby the automation of the behavioural routine's simple elements effectively 'frees-up' resources toward the higher-order aspects relating to the overall strategy and performance of the current routine. In the SOS case, this would relate to deciding the optimal combinations of when and which switches to perform. Indeed, those who developed behavioural routines that were more structured exhibited greater adaptation of MD activation and were more successful at the task overall.

Evidence of the brain's flexible adaptation was also observed in the functional connectivity (FC) analyses. Here, a striking distinction between the two studies was revealed. The generalised psychophysiological interaction (gPPI) FC measure used here has the advantage of capturing *task-dependent* changes in FC that are over and above the intrinsic FC. When SOS was practised after minimal training, making a trial response was associated with increased FC between many regions with higher-order network associations. The increase over time was linked to practise effects as changes in FC strength co-varied with increases in response speed.

Taken together, this shift toward a state of increased connectivity and decreased activation reveals an interesting relationship between attentional states and learning. A similar observation has been made with working memory change detection paradigms (Soreq et al., 2019). Here, the encoding and maintenance of information are associated with widespread activation that attenuates as FC patterns emerge. Critically, these patterns are specific enough to reliably decode the type and amount of information being maintained (Soreq et al., 2019).

In Chapter 4, the adaptation in FC and concurrent optimisation of behaviour was specific to study 1. It is reasonable to consider this as an interesting neurocognitive feature because individuals in study 1 were performing SOS in the scanner after minimal training and were exposed to feedback. This was in contrast to study 2 where individuals performed SOS in the scanner after more robust pretraining and without feedback. Subsequently, there was a greater opportunity for in-task learning in study 1. Moreover, individuals in study 2 likely went in the scanner with a known

behavioural routine that remained stable over time (given the lack of in-task feedback). In support of this, response speed was faster and behavioural routines were significantly more structured in study 2 but response speed increased at a greater rate in study 1.

At first glance, the widespread FC effects observed in study 1 could be interpreted to reflect signal drift. It is common in fMRI data to observe linear trends that result from interactions between the temperature dependency of the magnetic field strength and heat being generated from hardware rapidly switching gradients during EPI acquisition (Foerster et al., 2005). However, this is an unlikely explanation of the between-study difference for multiple reasons. Detrending steps were taken during the data preprocessing and both studies were conducted on the same scanner and for the same period of time. Moreover, bidirectional effects were observed in the activation analyses and these were consistent with the system those regions belong to and also specific to the event type (e.g., trial vs switch).

Taken together, Chapter 4 demonstrates a concurrent optimisation of the brain and behaviour. The findings imply that the brain can adapt via learning and behavioural structuring to minimise cognitive demand with the outcome of improved behavioural performance. The lateral frontoparietal MD systems activated in response to cognitive demand but their markers of activation and FC varied with practise as simple processes, or components, within a complex task were learnt. This process may represent a general neurocognitive feature that has the benefit of minimizing switching-costs required for adapting to and simultaneously managing multiple tasks.

The self-ordered behaviour observed in Chapter 4 presented an interesting opportunity to consider variability across healthy individuals. Successfully performing the SOS could be achieved via a myriad of routines. Hypothetically, an individual could switch on every trial and score many points. If consistently performed, this would be represented by a simple repeating loop that may result in a highly structured routine. However, this procedure would likely impose a high and

sustained cognitive demand given the high frequency of switching events. Instead, the top-performing individuals opted to minimise switch frequency by grouping consecutive trials within each task and rule. Such a procedure may be more complex than simply switching on every trial; however, this approach appears to be effective in minimising cognitive demand as these novel routines were amenable to learning.

Furthermore, those top performers showed a greater adaptation to the task demands, in terms of activation of connectivity of higher-order functional networks. These individual differences in neurocognitive flexibility may present an interesting opportunity to gain insight into clinical populations; however, it can be challenging for patients to perform cognitive tasks in the scanner. Moreover, if the aim were to assess how brain function adapts to clinical intervention, it can be challenging to avoid task-learning confounds (Zhang et al., 2021).

Chapter 5 took the opportunity of a clinical context to build upon the inferences made thus far regarding neurocognitive flexibility. The findings presented in chapters 3 and 4 shed some light on how higher-order systems adapt to optimise behaviour. Arguably, cognitive rigidity is a salient feature of depression (Beck & Clark, 1988; Grahek et al., 2018) and patients can exhibit impaired attention switching abilities (Meiran et al., 2011; Wilkinson & Goodyer, 2006). Mirroring this, functional abnormalities of the higher-order networks have been associated with the cognitive flexibility that is observed in patients with depression (Goodman et al., 2021; Hamilton et al., 2011).

The recent observations made from Dr Robin Carhart-Harris' psilocybin for depression trials (Carhart-Harris et al., 2016, 2021) provided a rare opportunity to examine the changes in brain function that accompany the alleviation of patients' depression symptoms. Admittedly, there were no healthy controls to establish whether patients' brain function was abnormal at baseline. However, the changes in the brain's functional organisation that followed psilocybin therapy followed a

trajectory that was indicative of normal brain functioning. Indeed, these changes also co-varied with the long-term treatment responses.

From a global perspective, psilocybin therapy was followed by a decrease in network modularity which implies an increase in between-network integration. In these trials, the fMRI recordings weren't captured during the acute action of psilocybin. Instead, they were captured before and the day after (study 1) or three weeks after (study 2) the psilocybin therapy intervention.

The global decrease in modularity observed here appears to represent a 'carry-over' effect, an echo of the acute effect, perhaps. The acute action of psychedelics, such as psilocybin, pushes brain function toward a labile state defined by heightened integration and flexibility (Carhart-Harris & Friston, 2019). This state is markedly different from typical healthy brain function (Carhart-Harris et al., 2012). Moreover, the psychedelic state could be argued as the polar opposite of the functional regime of depression as is implied by the observations of heightened between-network segregation (Wei et al., 2017; Ye et al., 2015) and reduced dynamic flexibility (Goodman et al., 2021), relative to healthy controls.

The two studies in Chapter 5 associated the rapid, substantial and sustained treatment response of psilocybin to greater increases between-network integration. The shared variance in patient outcomes with changes in network modularity suggests that this composite measure of the brain's functional organisation is clinically relevant. However, it was advantageous to delve deeper and address whether this global shift was underpinned by a general shift in brain networks or was specific to certain functional systems.

There are multiple converging lines of enquiry that made it logical to predict that the brain's higher-order systems would feature in this finer-grained enquiry. The set of lateral frontoparietal networks (including the MD network) and the DMN express the greatest density of psilocybin's principal site of serotonergic action, the 5HT-2A receptor (Beliveau et al., 2020). Furthermore, these systems are tightly associated

with cognitive control abilities (Duncan, 2013; Soreq et al., 2021) and can function abnormally in depression (Beevers, 2005; Ellard et al., 2018; Lydon-Staley et al., 2019). Moreover, evidence from Chapters 3 and 4 indicate that these systems flexibly adapt to changing environmental demands and that their flexibility is associated with the ability to efficiently self-organise behaviour.

Subsequently, the functional cartography findings were predictable and interpretable. In particular, hyperactive DMN functioning has been associated in rumination in depression patients (Hamilton et al., 2011). The day after psilocybin therapy within-DMN, FC had significantly weakened and between-network connectivity between the DMN and lateral frontoparietal higher-order systems increased. Admittedly, these more specific markers were not replicated in study 2; however, this may result from differences in the study protocols as the post-psilocybin scan was at three weeks in study 2. Study 2 also used a different fMRI protocol that allowed for twice as many volumes to be acquired. This afforded the dynamic sliding-window analysis in study 2 that also implicated the MD (executive) network via an increase in dynamic flexibility which correlated with improved patient outcomes.

Taken together, replicating the global reduction in network modularity and the converging findings from the cartography analyses indicate that an increase in the integration and flexibility of the brain's functional networks may underlie the psilocybin treatment response in patients with depression. It could be argued that these markers imply a healthy shift in the sequencing of neurocognitive states. For example, if a healthy control group was available, it would be reasonable to predict that the significant differences at baseline would be attenuated following the intervention.

6.2 Future directions

This investigation has presented many further avenues to explore. I am at wonder with our ability to generate in vivo neuroimaging data. The scale at which these data are being acquired is unprecedented. As alluded to in the introduction chapter, neuroimaging research is rapidly advancing and our understanding of how best to use this technology to yield insights on brain function are maturing. Concurrently, improvements in MR hardware to boost sampling frequency and SNR are being met with advances in computational power and data modelling.

While these advances are promising, it is the interpretation of these data that may present the biggest challenge. Continued interdisciplinary efforts that bridge philosophy, psychology, cognitive science and the neurosciences are required to develop our conceptual frameworks of the brain and mind. This should aid the field to design experiments that ask questions that generate clearer insights. For example, the move away from the early modular zeitgeist to the current epoch of dynamic network perspectives on brain function seems to be entirely the correct direction.

Looking forward, I anticipate that the short-to-medium term objectives will relate to refining the data modelling approaches. Ideally, it will be possible to capture precise mechanistic descriptions that govern the sequences of neurocognitive states involved in the brain's self-organization principles and how these adapt whilst interacting with the environment. To this end, the dynamic approaches to modelling neuroimaging data are particularly exciting. However, these may require further development and will become more viable as MR temporal resolution and SNR increase.

In the long term, the central challenge to neuroimaging will be whether individuals can be characterised in a specific manner. It is reasonable to assume that it will be possible, one day, to infer the appropriate course of action for a psychiatric patient

after a conventional brain scan. The increasing interest in the dense-sampling of individuals may present an avenue toward this end.

Concerning the research presented in this thesis, I believe that there is a clear need to develop paradigms that capture naturalistic behaviours. A limitation of many paradigms in use today is that they are designed to repeatedly measure a process in isolation over a 5-20 minute period. It is difficult to conjure up a real-world scenario that this would closely relate to. Even during repetitive monotonous work, people strategize, strike up conversations or daydream. It is the case that repetitive, or contrived, aspects of in-scanner paradigms are a deliberate consequence of the poor SNR of fMRI meaning that many measurements are required to achieve a requisite amount of statistical certainty. It is also the case that these paradigms will continue to generate useful insight into brain function.

However, future experimenters should consider incorporating elements of self-ordered behaviours into their fMRI paradigms. Ecological validity will always be an issue when it comes to the inferences that can be made from psychological and cognitive research involving humans in controlled environments. It's like performing chemistry with dirty test tubes. However, the ability to navigate the varying demands of daily life likely represents an important handle on mental health and well-being. Dr Adam Hampshire's SOS paradigm that was featured in Chapter 4 demonstrates that it is possible to capture meaningful insight when individuals perform tasks in a relatively free manner. To my knowledge, the SOS paradigm is unique in this respect. Further research should be conducted to capture how people self-order their behaviour particularly in contexts like the SOS where the problem space exhibits a near-degeneracy with many possible solutions that result in similar outcomes.

It will be illuminating to explore self-ordered behaviours within clinical populations and cluster these data to identify sub-types of maladaptive behaviour. It would be useful to capture how individuals strategize their behaviour and make use of their cognitive resources within naturalistic contexts. This could, potentially, be achieved

by setting errands for a subject to complete at the home or within a virtual context. Identifying these maladaptive features could be useful for artificial intelligence (AI) assisted healthcare. For example, daily routines for at-home patients could be tailor-made to avoid the impact of the cognitive control impairments associated with their clinical status.

A limitation of the cross-sectional studies presented here is that variation between individuals could result from fluctuations in well-being or motivation. Even in healthy adults, factors such as sleep, caffeination, stress and blood sugar influence cognitive ability and the brain's haemodynamics. Capturing individual fluctuations in initiatives such as the myConnectome project and the Midnight Scan Club has generated an exciting new research avenue toward this end. A holistic dense-sampling approach to research may become an important alternative to the group average perspectives on data modelling. There may even be scope for developing subject-specific models of changes to the brain's functional organisation over various timescales.

Perhaps there are “rhythms” to the brain and cognitive ability, in the sense that there are natural variations that occur over weeks and months. At the peaks and troughs of these fluctuations, you may find differences in network organisation, cognitive ability and well-being. Cross-sectional comparisons aren't well-equipped to deliver insights in this case. However, dense-sampling approaches could be used to track these macro-fluctuations and observe whether the frequency of these changes varies between healthy adults and across patient cohorts. In Chapter 5, there was evidence to suggest the fMRI data acquired after the intervention could be used to predict the long-term treatment response. Similarly, it would be interesting to consider if a dense-sampling approach could be used to predict the likelihood of a future decline in well-being and clinical status. A proactive warning of upcoming declines in well-being would be a game-changer for the treatment of disorders including depression and schizophrenia.

The psilocybin for depression trials suggests that there is a great potential for psychedelic therapies. It will be important for these clinical trials to progress to phase 3 trials where larger numbers of patients can be evaluated. One of the challenges of brain imaging research is the large statistical spaces they can generate. Subsequently, it can be advantageous to use composite scores that can elegantly capture a useful and reliable index of the brain's functional organisation. Graph theoretical measures, such as network modularity and others, have the potential utility for making neuroimaging analyses of larger-scale clinical trials more tractable.

Many open questions relate to the action of psychedelics on brain function. Arguably, there is a need for the field to make up for the time lost to misinformed political climates. As mentioned, large-scale clinical trials are the next frontier. However, the findings from this thesis suggest that it may be informative to incorporate direct assessments of cognitive control into future trials. Concerning depression, it would be reasonable to predict that some of the cognitive deficits observed at baseline would attenuate following psilocybin therapy. For example, it would be reasonable to predict greater switching-costs at baseline, in comparison to controls. Moreover, it could be predicted that individual differences in treatment responses would correspond to a greater reduction in switching-costs. This potential evidence of increased cognitive flexibility would neatly build upon the increases in network flexibility that was identified in Chapter 5. The challenge would be developing operationally simple paradigms that can be performed multiple times by patient cohorts without the potentially confounding effects of difficulty or learning.

There are other open questions relating to the specificity of psilocybin and other psychedelic therapies. Arguably, the recent proliferation of psychedelic research suggests that the greatest potential lies in those “internalising” disorders such as addiction, anxiety, depression, eating disorders and OCD. These disorders exhibit symptomatic features relating to rigid thought patterns and repetitive behaviours. Further work is required to determine the validity of this transdiagnostic potential.

While the serotonin hypothesis of depression is not baseless, it may portray an overly simplistic picture of brain function and well-being. The mediocre efficacy of conventional antidepressants does not do much to support the notion that increasing serotonin availability is sufficient for treating depression. There are likely many other contextual factors that modulate well-being including diet, sleep quality, exercise and being embedded within a nurturing social context.

One future direction of research is to explore if non-pharmacological approaches could also encourage flexible brain function in a manner that improves patient well-being. It is interesting to consider whether exposing patients to new and stimulating experiences, that may come with travel, skill learning or with acquiring more fulfilling employment, could positively influence patient well-being. The logistics of this would be prohibitive due to the individualised need and would require an “empirical life coach”, perhaps. However, a dense-sampling clinical approach could be taken. Here, focused questions could be addressed. For example, it could be asked how brain dynamics and mood fluctuate over weeks with and without periods of intensive aerobic exercise, meditation or in relation to sleep quality.

Could “healthy” brain dynamics be trained? Closed-loop neurofeedback paradigms could be used to train individuals to express certain functional dynamics, such as functional state-spaces that are broad and flexible. It could be useful to incorporate elements of attention switching into the paradigm. Given the symptomatic profile of depression, learning to recognise and override maladaptive rumination is a common aim of cognitive behavioural therapy (CBT). Could CBT be enhanced with real-time neurofeedback? In theory, markers of the individual's current focus of attention could be developed to crudely represent periods of “external” or “internal” focus. From this, the proportion of time in each state could be monitored and it would be interesting to ask whether transitions could be encouraged. Could this kind of ‘flexible dynamics training’ decrease the probability of maladaptive rumination in depression?

On a theoretical level, there is the potential to develop models that characterize clinical disorders by the brain's functional dynamics. Given the current understanding, it may be reasonable to consider a continuum of clinical disorders with 'hyper-stability' on the one end and 'hyper-flexibility' on the other. Healthy adults, for example, would be placed in the middle with optimal metastability, e.g., a mixture of stability and flexibility. Patient cohorts suffering from cognitive rigidity (e.g., Anxiety, ASD, Depression, OCD) would place on in a hyper-stable region. In contrast, cohorts with attention deficits and/or impaired belief updating, such as in ADHD or Schizophrenia, where maintaining stable focused attentional states is challenging, would place on the hyper-flexible region. The brain under the acute action of psychedelics would be placed on the extreme edge of hyper-flexibility.

Clearly, this continuum is conjecture and would require extensive empirical investigation. However, this theoretical model may be useful as it would imply that psychedelics have utility for patient cohorts in the hyper-stable region because they nudge brain function toward the hyper-flexible region. It would also predict that psychedelic therapies would *not* be effective for patient cohorts that are situated in the hyper-flexible region.

6.3 Conclusion

In conclusion, this thesis provides novel insight into the nature of cognitive control. The extraordinary ability to rapidly structure thoughts and behaviour is critical for successfully navigating the varied and unpredictable demands of daily life. Yet, our understanding of these mechanisms is incomplete. Cognitive control and task-switching are supported by a broad network of brain regions that rapidly form coactivating coalitions to process information and generate context-dependent routines to meet novel overarching goals. These neurocognitive mechanisms are complex but this research suggests that it is beneficial for higher-order systems to behave dynamically and flexibly. The flexible ordering of neurocognitive states appears to be an important feature of intelligent behaviour and well-being.

“Emancipate yourself from the idea of a celestial dictatorship and you have taken the first step to becoming free.”

— **Christopher Hitchens**

References

- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). MIT Press.
- Allport, A., & Wylie, G. (1999). *Task-switching: Positive and negative priming of task-set*. Oxford University Press.
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, 115(3), 602–639.
<https://doi.org/10.1037/0033-295X.115.3.602>
- Andersen, K. A. A., Carhart-Harris, R., Nutt, D. J., & Erritzoe, D. (2021). Therapeutic effects of classic serotonergic psychedelics: A systematic review of modern-era clinical studies. *Acta Psychiatrica Scandinavica*, 143(2), 101–118.
<https://doi.org/10.1111/acps.13249>
- Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance: The brain's default network. *Annals of the New York Academy of Sciences*, 1316(1), 29–52. <https://doi.org/10.1111/nyas.12360>
- Angrilli, A., Palomba, D., Cantagallo, A., Maietti, A., & Stegagno, L. (1999). Emotional impairment after right orbitofrontal lesion in a patient without cognitive deficits. *NeuroReport*, 10(8), 1741–1746.
<https://doi.org/10.1097/00001756-199906030-00021>

- Antzoulatos, E. G., & Miller, E. K. (2014). Increases in Functional Connectivity between Prefrontal Cortex and Striatum during Category Learning. *Neuron*, 83(1), 216–225. <https://doi.org/10.1016/j.neuron.2014.05.005>
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8(4), 170–177. <https://doi.org/10.1016/j.tics.2004.02.010>
- Aron, Adam. R., & Poldrack, R. A. (2006). Cortical and Subcortical Contributions to Stop Signal Response Inhibition: Role of the Subthalamic Nucleus. *Journal of Neuroscience*, 26(9), 2424–2433. <https://doi.org/10.1523/JNEUROSCI.4682-05.2006>
- Arrington, C. M. (2008). The effect of stimulus availability on task choice in voluntary task switching. *Memory & Cognition*, 36(5), 991–997. <https://doi.org/10.3758/MC.36.5.991>
- Arrington, C. M., & Logan, G. D. (2004). The Cost of a Voluntary Task Switch. *Psychological Science*, 15(9), 610–615. <https://doi.org/10.1111/j.0956-7976.2004.00728.x>
- Arrington, C. M., & Logan, G. D. (2005). Voluntary Task Switching: Chasing the Elusive Homunculus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 683–702. <https://doi.org/10.1037/0278-7393.31.4.683>
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1), 95–113. <https://doi.org/10.1016/j.neuroimage.2007.07.007>
- Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry—The Methods. *NeuroImage*, 11(6), 805–821. <https://doi.org/10.1006/nimg.2000.0582>
- Ashburner, J., & Friston, K. J. (2011). Diffeomorphic registration using geodesic

- shooting and Gauss–Newton optimisation. *NeuroImage*, 55(3), 954–967.
<https://doi.org/10.1016/j.neuroimage.2010.12.049>
- Atasoy, S., Vohryzek, J., Deco, G., Carhart-Harris, R. L., & Kringelbach, M. L. (2018). Common neural signatures of psychedelics: Frequency-specific energy changes and repertoire expansion revealed using connectome-harmonic decomposition. In *Progress in Brain Research* (Vol. 242, pp. 97–120). Elsevier.
<https://doi.org/10.1016/bs.pbr.2018.08.009>
- Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTs). *Insight J*, 2, 1–35.
- Averbeck, B. B., Evans, S., Chouhan, V., Bristow, E., & Shergill, S. S. (2011). Probabilistic learning and inference in schizophrenia. *Schizophrenia Research*, 127(1–3), 115–122. <https://doi.org/10.1016/j.schres.2010.08.009>
- Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology*, 11(2), 187–194.
<https://doi.org/10.1037/0894-4105.11.2.187>
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5), 193–200.
<https://doi.org/10.1016/j.tics.2008.02.004>
- Badre, D., & D’Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, 19(12), 2082–2099.
<https://doi.org/10.1162/jocn.2007.19.12.2082>
- Badre, D., Hoffman, J., Cooney, J. W., & D’Esposito, M. (2009). Hierarchical

- cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, 12(4), 515–522. <https://doi.org/10.1038/nn.2277>
- Badre, D., Kayser, A. S., & D’Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron*, 66(2), 315–326.
<https://doi.org/10.1016/j.neuron.2010.03.025>
- Barbey, A. K. (2018). Network Neuroscience Theory of Human Intelligence. *Trends in Cognitive Sciences*, 22(1), 8–20. <https://doi.org/10.1016/j.tics.2017.10.001>
- Barceló, F., Periáñez, J. A., & Nyhus, E. (2007). An information theoretical approach to task-switching: Evidence from cognitive brain potentials in humans. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/neuro.09.013.2007>
- Barrett, F. S., Doss, M. K., Sepeda, N. D., Pekar, J. J., & Griffiths, R. R. (2020). Emotions and brain function are altered up to one month after a single high dose of psilocybin. *Scientific Reports*, 10(1), 2214.
<https://doi.org/10.1038/s41598-020-59282-y>
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 570–585.
<https://doi.org/10.1037/0278-7393.33.3.570>
- Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1), 259–267.
[https://doi.org/10.1016/S0006-3495\(94\)80775-1](https://doi.org/10.1016/S0006-3495(94)80775-1)
- Bassett, D. S., & Bullmore, E. (2006). Small-World Brain Networks. *The Neuroscientist*, 12(6), 512–523. <https://doi.org/10.1177/1073858406293182>
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton,

- S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108(18), 7641–7646. <https://doi.org/10.1073/pnas.1018985108>
- Bassett, D. S., Wymbs, N. F., Rombach, M. P., Porter, M. A., Mucha, P. J., & Grafton, S. T. (2013). Task-Based Core-Periphery Organization of Human Brain Dynamics. *PLoS Computational Biology*, 9(9), e1003171. <https://doi.org/10.1371/journal.pcbi.1003171>
- Bassett, D. S., Xia, C. H., & Satterthwaite, T. D. (2018). Understanding the Emergence of Neuropsychiatric Disorders With Network Neuroscience. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(9), 742–753. <https://doi.org/10.1016/j.bpsc.2018.03.015>
- Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2015). Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*, 18(5), 744–751. <https://doi.org/10.1038/nn.3993>
- Beard, C., Millner, A. J., Forgeard, M. J. C., Fried, E. I., Hsu, K. J., Treadway, M. T., Leonard, C. V., Kertz, S. J., & Björgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, 46(16), 3359–3369. <https://doi.org/10.1017/S0033291716002300>
- Beck, A. T., & Clark, D. A. (1988). Anxiety and depression: An information processing perspective. *Anxiety Research*, 1(1), 23–36. <https://doi.org/10.1080/10615808808248218>
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients. *Journal of*

Personality Assessment, 67(3), 588–597.

https://doi.org/10.1207/s15327752jpa6703_13

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100.

[https://doi.org/10.1016/0272-7358\(88\)90050-5](https://doi.org/10.1016/0272-7358(88)90050-5)

Beevers, C. (2005). Cognitive vulnerability to depression: A dual process model. *Clinical Psychology Review*, 25(7), 975–1002.

<https://doi.org/10.1016/j.cpr.2005.03.003>

Beliveau, V., Ozenne, B., Strother, S., Greve, D. N., Svarer, C., Knudsen, G. M., & Ganz, M. (2020). The structure of the serotonin system: A PET imaging study. *NeuroImage*, 205, 116240.

<https://doi.org/10.1016/j.neuroimage.2019.116240>

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300.

<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>

Berlin, H. A. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain*, 127(5), 1108–1126. <https://doi.org/10.1093/brain/awh135>

Berman, M. G., Peltier, S., Nee, D. E., Kross, E., Deldin, P. J., & Jonides, J. (2011). Depression, rumination and the default network. *Social Cognitive and Affective Neuroscience*, 6(5), 548–555. <https://doi.org/10.1093/scan/nsq080>

Berryhill, M. E., & Hughes, H. C. (2009). On the minimization of task switch costs

- following long-term training. *Attention, Perception & Psychophysics*, 71(3), 503–514. <https://doi.org/10.3758/APP.71.3.503>
- Betzel, R. F. (2020). Network neuroscience and the connectomics revolution. *ArXiv:2010.01591 [q-Bio]*. <http://arxiv.org/abs/2010.01591>
- Betzel, R. F., & Bassett, D. S. (2017). Multi-scale brain networks. *NeuroImage*, 160, 73–83. <https://doi.org/10.1016/j.neuroimage.2016.11.006>
- Betzel, R. F., Satterthwaite, T. D., Gold, J. I., & Bassett, D. S. (2017). Positive affect, surprise, and fatigue are correlates of network flexibility. *Scientific Reports*, 7(1), 520. <https://doi.org/10.1038/s41598-017-00425-z>
- Bieri, O., & Scheffler, K. (2013). Fundamentals of balanced steady state free precession MRI: Fundamentals of Balanced SSFP MRI. *Journal of Magnetic Resonance Imaging*, 38(1), 2–11. <https://doi.org/10.1002/jmri.24163>
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Borman, K. M., & Kurdek, L. A. (1987). Grade and Gender Differences In and the Stability and Correlates Of the Structural Complexity of Children's Playground Games. *International Journal of Behavioral Development*, 10(2), 241–251. <https://doi.org/10.1177/016502548701000207>
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12(5), 201–208. <https://doi.org/10.1016/j.tics.2008.02.009>
- Braga, R. M., & Buckner, R. L. (2017). Parallel Interdigitated Distributed Networks

- within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron*, 95(2), 457-471.e5. <https://doi.org/10.1016/j.neuron.2017.06.038>
- Braga, R. M., Sharp, D. J., Leeson, C., Wise, R. J. S., & Leech, R. (2013). Echoes of the Brain within Default Mode, Association, and Heteromodal Cortices. *Journal of Neuroscience*, 33(35), 14031–14039. <https://doi.org/10.1523/JNEUROSCI.0570-13.2013>
- Braun, U., Harneit, A., Pergola, G., Menara, T., Schaefer, A., Betzel, R. F., Zang, Z., Schweiger, J. I., Schwarz, K., Chen, J., Blasi, G., Bertolino, A., Durstewitz, D., Pasqualetti, F., Schwarz, E., Meyer-Lindenberg, A., Bassett, D. S., & Tost, H. (2019). *Brain state stability during working memory is explained by network control theory, modulated by dopamine D1/D2 receptor function, and diminished in schizophrenia* [Preprint]. Neuroscience. <https://doi.org/10.1101/679670>
- Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural Mechanisms of Transient and Sustained Cognitive Control during Task Switching. *Neuron*, 39(4), 713–726. [https://doi.org/10.1016/S0896-6273\(03\)00466-5](https://doi.org/10.1016/S0896-6273(03)00466-5)
- Brosch, J. R., Talavage, T. M., Ulmer, J. L., & Nyenhuis, J. A. (2002). Simulation of human respiration in fMRI with a mechanical model. *IEEE Transactions on Biomedical Engineering*, 49(7), 700–707. <https://doi.org/10.1109/TBME.2002.1010854>
- Bryant, R. G., Mendelson, D. A., & Lester, C. C. (1991). The magnetic field dependence of proton spin relaxation in tissues. *Magnetic Resonance in Medicine*, 21(1), 117–126. <https://doi.org/10.1002/mrm.1910210114>
- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. T. (2011). The

organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5), 2322–2345.

<https://doi.org/10.1152/jn.00339.2011>

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38(6), 848–863. [https://doi.org/10.1016/S0028-3932\(99\)00134-7](https://doi.org/10.1016/S0028-3932(99)00134-7)

Camicioli, R. M., Wieler, M., de Frias, C. M., & Martin, W. R. W. (2008). Early, untreated Parkinson's disease patients show reaction time variability. *Neuroscience Letters*, 441(1), 77–80.

<https://doi.org/10.1016/j.neulet.2008.06.004>

Cañas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics*, 46(5), 482–501.

<https://doi.org/10.1080/0014013031000061640>

Carhart-Harris, R. L., Bolstridge, M., Rucker, J., Day, C. M. J., Erritzoe, D., Kaelen, M., Bloomfield, M., Rickard, J. A., Forbes, B., Feilding, A., Taylor, D., Pilling, S., Curran, V. H., & Nutt, D. J. (2016). Psilocybin with psychological support for treatment-resistant depression: An open-label feasibility study. *The Lancet Psychiatry*, 3(7), 619–627.

[https://doi.org/10.1016/S2215-0366\(16\)30065-7](https://doi.org/10.1016/S2215-0366(16)30065-7)

Carhart-Harris, R. L., Erritzoe, D., Williams, T., Stone, J. M., Reed, L. J., Colasanti, A., Tyacke, R. J., Leech, R., Malizia, A. L., Murphy, K., Hobden, P., Evans, J., Feilding, A., Wise, R. G., & Nutt, D. J. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin.

Proceedings of the National Academy of Sciences, 109(6), 2138–2143.

<https://doi.org/10.1073/pnas.1119598109>

Carhart-Harris, R. L., & Friston, K. J. (2019). REBUS and the Anarchic Brain: Toward a Unified Model of the Brain Action of Psychedelics. *Pharmacological Reviews*, 71(3), 316–344. <https://doi.org/10.1124/pr.118.017160>

Carhart-Harris, R. L., Giribaldi, B., Watts, R., Baker-Jones, M., Murphy-Beiner, A., Murphy, R., Martell, J., Blemings, A., Erritzoe, D., & Nutt, D. J. (2021). Trial of Psilocybin versus Escitalopram for Depression. *New England Journal of Medicine*, 384(15), 1402–1411. <https://doi.org/10.1056/NEJMoa2032994>

Carhart-Harris, R. L., & Nutt, D. J. (2017). Serotonin and brain function: A tale of two receptors. *Journal of Psychopharmacology (Oxford, England)*, 31(9), 1091–1120. <https://doi.org/10.1177/0269881117725915>

Carhart-Harris, R. L., Roseman, L., Bolstridge, M., Demetriou, L., Pannekoek, J. N., Wall, M. B., Tanner, M., Kaelen, M., McGonigle, J., Murphy, K., Leech, R., Curran, H. V., & Nutt, D. J. (2017). Psilocybin for treatment-resistant depression: FMRI-measured brain mechanisms. *Scientific Reports*, 7(1), 13187. <https://doi.org/10.1038/s41598-017-13282-7>

Carlson, J. S., Jensen, C. M., & Widaman, K. F. (1983). Reaction time, intelligence, and attention. *Intelligence*, 7(4), 329–344. [https://doi.org/10.1016/0160-2896\(83\)90008-9](https://doi.org/10.1016/0160-2896(83)90008-9)

Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., & Thiebaut de Schotten, M. (2012). Short frontal lobe connections of the human brain. *Cortex*, 48(2), 273–291. <https://doi.org/10.1016/j.cortex.2011.12.001>

- Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 1105–1132.
<https://doi.org/10.1016/j.cortex.2008.05.004>
- Caverly, R. H. (2015). MRI Fundamentals: RF Aspects of Magnetic Resonance Imaging (MRI). *IEEE Microwave Magazine*, 16(6), 20–33.
<https://doi.org/10.1109/MMM.2015.2419753>
- Chaitin, G. J. (1966). On the Length of Programs for Computing Finite Binary Sequences. *Journal of the ACM*, 13(4), 547–569.
<https://doi.org/10.1145/321356.321363>
- Chamberlain, S. R., Menzies, L., Hampshire, A., Suckling, J., Fineberg, N. A., del Campo, N., Aitken, M., Craig, K., Owen, A. M., Bullmore, E. T., Robbins, T. W., & Sahakian, B. J. (2008). Orbitofrontal Dysfunction in Patients with Obsessive-Compulsive Disorder and Their Unaffected Relatives. *Science*, 321(5887), 421–422. <https://doi.org/10.1126/science.1154433>
- Chang, C., Cunningham, J. P., & Glover, G. H. (2009). Influence of heart rate on the BOLD signal: The cardiac response function. *NeuroImage*, 44(3), 857–869.
<https://doi.org/10.1016/j.neuroimage.2008.09.029>
- Chang, C., & Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, 50(1), 81–98.
<https://doi.org/10.1016/j.neuroimage.2009.12.011>
- Chow, B. W., Nuñez, V., Kaplan, L., Granger, A. J., Bistrong, K., Zucker, H. L., Kumar, P., Sabatini, B. L., & Gu, C. (2020). Caveolae in CNS arterioles mediate neurovascular coupling. *Nature*, 579(7797), 106–110.

<https://doi.org/10.1038/s41586-020-2026-1>

Close, J. B., Hajien, E. C., Watts, R., Roseman, L., & Carhart-Harris, R. L. (2020).

Psychedelics and psychological flexibility – Results of a prospective web-survey using the Acceptance and Action Questionnaire II. *Journal of Contextual Behavioral Science*, 16, 37–44.

<https://doi.org/10.1016/j.jcbs.2020.01.005>

Cole, M. W., Laurent, P., & Stocco, A. (2013). Rapid instructed task learning: A new

window into the human brain's unique capacity for flexible cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 13(1), 1–22.

<https://doi.org/10.3758/s13415-012-0125-7>

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S.

(2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, 16(9), 1348–1355. <https://doi.org/10.1038/nn.3470>

Cole, M. W., Yarkoni, T., Repovs, G., Anticevic, A., & Braver, T. S. (2012). Global

Connectivity of Prefrontal Cortex Predicts Cognitive Control and Intelligence. *Journal of Neuroscience*, 32(26), 8988–8999.

<https://doi.org/10.1523/JNEUROSCI.0536-12.2012>

Collins, A., & Koechlin, E. (2012). Reasoning, Learning, and Creativity: Frontal Lobe

Function and Human Decision-Making. *PLoS Biology*, 10(3), e1001293.

<https://doi.org/10.1371/journal.pbio.1001293>

Cox, R. W. (1996). AFNI: Software for Analysis and Visualization of Functional

Magnetic Resonance Neuroimages. *Computers and Biomedical Research*, 29(3), 162–173. <https://doi.org/10.1006/cbmr.1996.0014>

Crittenden, B. M., & Duncan, J. (2014). Task Difficulty Manipulation Reveals Multiple

- Demand Activity but no Frontal Lobe Hierarchy. *Cerebral Cortex*, 24(2), 532–540. <https://doi.org/10.1093/cercor/bhs333>
- Crittenden, B. M., Mitchell, D. J., & Duncan, J. (2016). Task Encoding across the Multiple Demand Cortex Is Consistent with a Frontoparietal and Cingulo-Opercular Dual Networks Distinction. *The Journal of Neuroscience*, 36(23), 6147–6155. <https://doi.org/10.1523/JNEUROSCI.4590-15.2016>
- Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., & Bullmore, E. T. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain*, 137(8), 2382–2395. <https://doi.org/10.1093/brain/awu132>
- Culler, A. J. (1912). *Interference and Adaptability: An Experimental Study of Their Relation with Special Reference to Individual Differences*. Science Press.
- Cusack, R., Mitchell, D. J., & Duncan, J. (2010). Discrete Object Representation, Attention Switching, and Task Difficulty in the Parietal Lobe. *Journal of Cognitive Neuroscience*, 22(1), 32–47. <https://doi.org/10.1162/jocn.2009.21194>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, 9(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- Davis, A. K., Barrett, F. S., & Griffiths, R. R. (2020). Psychological flexibility mediates the relations between acute psychedelic effects and subjective decreases in depression and anxiety. *Journal of Contextual Behavioral Science*, 15, 39–45. <https://doi.org/10.1016/j.jcbs.2019.11.004>
- Daws, R. E., Scott, G., Soreq, E., Leech, R., Hellyer, P. J., & Hampshire, A. (2020). *Optimisation of brain states and behavioural strategies when learning*

- complex tasks* [Preprint]. Neuroscience.
<https://doi.org/10.1101/2020.06.17.156570>
- Daws, R. E., Soreq, E., Li, Y., Sandrone, S., & Hampshire, A. (2020). *Contrasting hierarchical and multiple-demand accounts of frontal lobe functional organisation during task-switching* [Preprint]. Neuroscience.
<https://doi.org/10.1101/2020.06.27.175133>
- Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, 95(24), 14529–14534. <https://doi.org/10.1073/pnas.95.24.14529>
- Demanet, J., Verbruggen, F., Liefoghe, B., & Vandierendonck, A. (2010). Voluntary task switching under load: Contribution of top-down and bottom-up factors in goal-directed behavior. *Psychonomic Bulletin & Review*, 17(3), 387–393.
<https://doi.org/10.3758/PBR.17.3.387>
- Deoni, S. C. L., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., Williams, S. C. R., & Murphy, D. G. M. (2011). Mapping Infant Brain Myelination with Magnetic Resonance Imaging. *Journal of Neuroscience*, 31(2), 784–791. <https://doi.org/10.1523/JNEUROSCI.2106-10.2011>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3), 99–105. <https://doi.org/10.1016/j.tics.2008.01.001>
- Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., Fox, M. D., Snyder, A. Z., Vincent, J. L., Raichle, M. E.,

- Schlaggar, B. L., & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073–11078. <https://doi.org/10.1073/pnas.0704320104>
- Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006). A Core System for the Implementation of Task Sets. *Neuron*, 50(5), 799–812. <https://doi.org/10.1016/j.neuron.2006.04.031>
- Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & Yves von Cramon, D. (2000). Prefrontal cortex activation in task switching: An event-related fMRI study. *Cognitive Brain Research*, 9(1), 103–109. [https://doi.org/10.1016/S0926-6410\(99\)00029-4](https://doi.org/10.1016/S0926-6410(99)00029-4)
- Duman, R. S. (1997). A Molecular and Cellular Theory of Depression. *Archives of General Psychiatry*, 54(7), 597. <https://doi.org/10.1001/archpsyc.1997.01830190015002>
- Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3(3), 271–290. <https://doi.org/10.1080/02643298608253360>
- Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2(11), 820–829. <https://doi.org/10.1038/35097575>
- Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, 80(1), 35–50. <https://doi.org/10.1016/j.neuron.2013.09.015>
- Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, 33(3), 261–268.

[https://doi.org/10.1016/0028-3932\(94\)00124-8](https://doi.org/10.1016/0028-3932(94)00124-8)

- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483. [https://doi.org/10.1016/S0166-2236\(00\)01633-7](https://doi.org/10.1016/S0166-2236(00)01633-7)
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28), 7900–7905. <https://doi.org/10.1073/pnas.1602413113>
- Ellard, K. K., Zimmerman, J. P., Kaur, N., Van Dijk, K. R. A., Roffman, J. L., Nierenberg, A. A., Dougherty, D. D., Deckersbach, T., & Camprodon, J. A. (2018). Functional Connectivity Between Anterior Insula and Key Nodes of Frontoparietal Executive Control and Salience Networks Distinguish Bipolar Depression From Unipolar Depression and Healthy Control Subjects. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(5), 473–484. <https://doi.org/10.1016/j.bpsc.2018.01.013>
- Erasmus, L. J., Hurter, D., Naude, M., Kritzinger, H. G., & Acho, S. (2004). A short overview of MRI artefacts. *South African Journal of Radiology*, 8(2), 13. <https://doi.org/10.4102/sajr.v8i2.127>
- Erika-Florence, M., Leech, R., & Hampshire, A. (2014). A functional network perspective on response inhibition and attentional control. *Nature Communications*, 5(1), 4073. <https://doi.org/10.1038/ncomms5073>
- Ezaki, T., Fonseca dos Reis, E., Watanabe, T., Sakaki, M., & Masuda, N. (2020). Closer to critical resting-state neural dynamics in individuals with higher fluid intelligence. *Communications Biology*, 3(1), 52.

<https://doi.org/10.1038/s42003-020-0774-y>

Fallon, S. J., Williams-Gray, C. H., Barker, R. A., Owen, A. M., & Hampshire, A.

(2013). Prefrontal Dopamine Levels Determine the Balance between Cognitive Stability and Flexibility. *Cerebral Cortex*, 23(2), 361–369.

<https://doi.org/10.1093/cercor/bhs025>

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S.,

Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex*, 26(8), 3508–3526. <https://doi.org/10.1093/cercor/bhw157>

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41), 16616–16621.

<https://doi.org/10.1073/pnas.1315235110>

Feinberg, D. A., & Yacoub, E. (2012). The rapid development of high speed, resolution and precision in fMRI. *NeuroImage*, 62(2), 720–725.

<https://doi.org/10.1016/j.neuroimage.2012.01.049>

Ferguson, M. A., Anderson, J. S., & Spreng, R. N. (2017). Fluid and flexible minds: Intelligence reflects synchrony in the brain's intrinsic network architecture.

Network Neuroscience, 1(2), 192–207.

https://doi.org/10.1162/NETN_a_00010

Feurer, C., Jimmy, J., Chang, F., Langenecker, S. A., Phan, K. L., Ajilore, O., &

Klumpp, H. (2021). Resting state functional connectivity correlates of rumination and worry in internalizing psychopathologies. *Depression and Anxiety*, da.23142. <https://doi.org/10.1002/da.23142>

- Finc, K., Bonna, K., He, X., Lydon-Staley, D. M., Kühn, S., Duch, W., & Bassett, D. S. (2020). Dynamic reconfiguration of functional brain networks during working memory training. *Nature Communications*, 11(1), 2435. <https://doi.org/10.1038/s41467-020-15631-z>
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., & Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, 18(11), 1664–1671. <https://doi.org/10.1038/nn.4135>
- Foerster, B. U., Tomasi, D., & Caparelli, E. C. (2005). Magnetic field shift due to mechanical vibration in functional magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(5), 1261–1267. <https://doi.org/10.1002/mrm.20695>
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197. <https://doi.org/10.1016/j.jad.2016.10.019>
- Friston, K. J. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2), e33. <https://doi.org/10.1371/journal.pbio.1000033>
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage*, 6(3), 218–229. <https://doi.org/10.1006/nimg.1997.0291>
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210.

<https://doi.org/10.1002/hbm.460020402>

Friston, K. J., Mechelli, A., Turner, R., & Price, C. J. (2000). Nonlinear responses in fMRI: The Balloon model, Volterra kernels, and other hemodynamics.

NeuroImage, 12(4), 466–477. <https://doi.org/10.1006/nimg.2000.0630>

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J., & Turner, R. (1996).

Movement-Related effects in fMRI time-series: Movement Artifacts in fMRI.

Magnetic Resonance in Medicine, 35(3), 346–355.

<https://doi.org/10.1002/mrm.1910350312>

Fuster, J. M. (1999). Synopsis of function and dysfunction of the frontal lobe. *Acta*

Psychiatrica Scandinavica, 99(s395), 51–57.

<https://doi.org/10.1111/j.1600-0447.1999.tb05983.x>

Garcia-Romeu, A., Griffiths, R. R., & Johnson, M. W. (2014). Psilocybin-occasioned

mystical experiences in the treatment of tobacco addiction. *Current Drug*

Abuse Reviews, 7(3), 157–164.

<https://doi.org/10.2174/1874473708666150107121331>

Gauvrit, N., Zenil, H., Soler-Toscano, F., Delahaye, J.-P., & Brugger, P. (2017).

Human behavioral complexity peaks at age 25. *PLOS Computational Biology*,

13(4), e1005408. <https://doi.org/10.1371/journal.pcbi.1005408>

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps

in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4),

870–878. <https://doi.org/10.1006/nimg.2001.1037>

Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling

regional and psychophysiologic interactions in fMRI: The importance of

hemodynamic deconvolution. *NeuroImage*, 19(1), 200–207.

[https://doi.org/10.1016/s1053-8119\(03\)00058-2](https://doi.org/10.1016/s1053-8119(03)00058-2)

Gizewski, E. R., Gasser, T., de Greiff, A., Boehm, A., & Forsting, M. (2003).

Cross-modal plasticity for sensory and motor activation patterns in blind subjects. *NeuroImage*, 19(3), 968–975.

[https://doi.org/10.1016/S1053-8119\(03\)00114-9](https://doi.org/10.1016/S1053-8119(03)00114-9)

Glascher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Paul,

L. K., & Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, 109(36), 14681–14686.

<https://doi.org/10.1073/pnas.1206608109>

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub,

E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171–178. <https://doi.org/10.1038/nature18933>

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B.,

Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 80, 105–124.

<https://doi.org/10.1016/j.neuroimage.2013.04.127>

Gold, G. E., Han, E., Stainsby, J., Wright, G., Brittain, J., & Beaulieu, C. (2004).

Musculoskeletal MRI at 3.0 T: Relaxation times and image contrast. *AJR. American Journal of Roentgenology*, 183(2), 343–351.

<https://doi.org/10.2214/ajr.183.2.1830343>

Goldberg, E., & Podell, K. (2000). Adaptive Decision Making, Ecological Validity, and

- the Frontal Lobes. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 56–68. [https://doi.org/10.1076/1380-3395\(200002\)22:1;1-8;FT056](https://doi.org/10.1076/1380-3395(200002)22:1;1-8;FT056)
- Gong, Q., & He, Y. (2015). Depression, Neuroimaging and Connectomics: A Selective Overview. *Biological Psychiatry*, 77(3), 223–235. <https://doi.org/10.1016/j.biopsych.2014.08.009>
- Goodman, Z., A Bainter, S., Kornfeld, S., Chang, C., S Nomi, J., & Q Uddin, L. (2021). Whole-Brain Functional Dynamics Track Depressive Symptom Severity. *Cerebral Cortex*, bhab047. <https://doi.org/10.1093/cercor/bhab047>
- Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General*, 129(3), 308–339. <https://doi.org/10.1037/0096-3445.129.3.308>
- Grahek, I., Everaert, J., Krebs, R. M., & Koster, E. H. W. (2018). Cognitive Control in Depression: Toward Clinical Models Informed by Cognitive Neuroscience. *Clinical Psychological Science*, 6(4), 464–480. <https://doi.org/10.1177/2167702618758969>
- Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. <https://doi.org/10.1037/h0059831>
- Grant, J. E., Daws, R., Hampshire, A., & Chamberlain, S. R. (2018). An fMRI Pilot Study of Cognitive Flexibility in Trichotillomania. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 30(4), 318–324. <https://doi.org/10.1176/appi.neuropsych.18030038>
- Grattan, L. M., Bloomer, R. H., Archambault, F. X., & Eslinger, P. J. (1994). Cognitive

- flexibility and empathy after frontal lobe lesion. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, 7(4), 251–259.
- Gratton, G., Cooper, P., Fabiani, M., Carter, C. S., & Karayanidis, F. (2018). Dynamics of cognitive control: Theoretical bases, paradigms, and a view for the future. *Psychophysiology*, 55(3), e13016. <https://doi.org/10.1111/psyp.13016>
- Griffiths, R. R., Johnson, M. W., Carducci, M. A., Umbricht, A., Richards, W. A., Richards, B. D., Cosimano, M. P., & Klinedinst, M. A. (2016). Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of Psychopharmacology*, 30(12), 1181–1197. <https://doi.org/10.1177/0269881116675513>
- Grob, C. S., Danforth, A. L., Chopra, G. S., Hagerty, M., McKay, C. R., Halberstadt, A. L., & Greer, G. R. (2011). Pilot Study of Psilocybin Treatment for Anxiety in Patients With Advanced-Stage Cancer. *Archives of General Psychiatry*, 68(1), 71. <https://doi.org/10.1001/archgenpsychiatry.2010.116>
- Gruszka, A., Hampshire, A., Barker, R. A., & Owen, A. M. (2017). Normal aging and Parkinson's disease are associated with the functional decline of distinct frontal-striatal circuits. *Cortex*, 93, 178–192. <https://doi.org/10.1016/j.cortex.2017.05.020>
- Gu, B.-M., Park, J.-Y., Kang, D.-H., Lee, S. J., Yoo, S. Y., Jo, H. J., Choi, C.-H., Lee, J.-M., & Kwon, J. S. (2007). Neural correlates of cognitive inflexibility during task-switching in obsessive-compulsive disorder. *Brain*, 131(1), 155–164. <https://doi.org/10.1093/brain/awm277>
- Haase, A. (1990). Snapshot flash mri. Applications to t1, t2, and chemical-shift

- imaging. *Magnetic Resonance in Medicine*, 13(1), 77–89.
<https://doi.org/10.1002/mrm.1910130109>
- Haddad, P. (1998). The SSRI discontinuation syndrome. *Journal of Psychopharmacology*, 12(3), 305–313.
<https://doi.org/10.1177/026988119801200311>
- Hahn, E. L. (1950). Nuclear Induction Due to Free Larmor Precession. *Physical Review*, 77(2), 297–298. <https://doi.org/10.1103/PhysRev.77.297.2>
- Hamilton, J. P., Farmer, M., Fogelman, P., & Gotlib, I. H. (2015). Depressive Rumination, the Default-Mode Network, and the Dark Matter of Clinical Neuroscience. *Biological Psychiatry*, 78(4), 224–230.
<https://doi.org/10.1016/j.biopsych.2015.02.020>
- Hamilton, J. P., Furman, D. J., Chang, C., Thomason, M. E., Dennis, E., & Gotlib, I. H. (2011). Default-Mode and Task-Positive Network Activity in Major Depressive Disorder: Implications for Adaptive and Maladaptive Rumination. *Biological Psychiatry*, 70(4), 327–333.
<https://doi.org/10.1016/j.biopsych.2011.02.003>
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: Inhibition and attentional control. *NeuroImage*, 50(3), 1313–1319.
<https://doi.org/10.1016/j.neuroimage.2009.12.109>
- Hampshire, A., Daws, R. E., Neves, I. D., Soreq, E., Sandrone, S., & Violante, I. R. (2019). Probing cortical and sub-cortical contributions to instruction-based learning: Regional specialisation and global network dynamics. *NeuroImage*, 192, 88–100. <https://doi.org/10.1016/j.neuroimage.2019.03.002>

- Hampshire, A., Duncan, J., & Owen, A. M. (2007). Selective Tuning of the Blood Oxygenation Level-Dependent Response during Simple Target Detection Dissociates Human Frontoparietal Subregions. *Journal of Neuroscience*, 27(23), 6219–6223. <https://doi.org/10.1523/JNEUROSCI.0851-07.2007>
- Hampshire, A., Hellyer, P. J., Parkin, B., Hiebert, N., MacDonald, P., Owen, A. M., Leech, R., & Rowe, J. (2016). Network mechanisms of intentional learning. *NeuroImage*, 127, 123–134. <https://doi.org/10.1016/j.neuroimage.2015.11.060>
- Hampshire, A., Highfield, R. R., Parkin, B. L., & Owen, A. M. (2012). Fractionating Human Intelligence. *Neuron*, 76(6), 1225–1237. <https://doi.org/10.1016/j.neuron.2012.06.022>
- Hampshire, A., MacDonald, A., & Owen, A. M. (2013). Hypoconnectivity and Hyperfrontality in Retired American Football Players. *Scientific Reports*, 3(1), 2972. <https://doi.org/10.1038/srep02972>
- Hampshire, A., & Owen, A. M. (2006). Fractionating Attentional Control Using Event-Related fMRI. *Cerebral Cortex*, 16(12), 1679–1689. <https://doi.org/10.1093/cercor/bhj116>
- Hampshire, A., & Sharp, D. J. (2015). Contrasting network and modular perspectives on inhibitory control. *Trends in Cognitive Sciences*, 19(8), 445–452. <https://doi.org/10.1016/j.tics.2015.06.006>
- Hampshire, A., Zadel, A., Sandrone, S., Soreq, E., Fineberg, N., Bullmore, E. T., Robbins, T. W., Sahakian, B. J., & Chamberlain, S. R. (2020). Inhibition-Related Cortical Hypoconnectivity as a Candidate Vulnerability Marker for Obsessive-Compulsive Disorder. *Biological Psychiatry: Cognitive*

Neuroscience and Neuroimaging, 5(2), 222–230.

<https://doi.org/10.1016/j.bpsc.2019.09.010>

Harlalka, V., Bapi, R. S., Vinod, P. K., & Roy, D. (2019). Atypical Flexibility in Dynamic Functional Connectivity Quantifies the Severity in Autism Spectrum Disorder. *Frontiers in Human Neuroscience*, 13, 6.

<https://doi.org/10.3389/fnhum.2019.00006>

Hasenkamp, W., Wilson-Mendenhall, C. D., Duncan, E., & Barsalou, L. W. (2012).

Mind wandering and attention during focused meditation: A fine-grained temporal analysis of fluctuating cognitive states. *NeuroImage*, 59(1),

750–760. <https://doi.org/10.1016/j.neuroimage.2011.07.008>

Hasenstaub, A., Otte, S., Callaway, E., & Sejnowski, T. J. (2010). Metabolic cost as a unifying principle governing neuronal biophysics. *Proceedings of the National Academy of Sciences*, 107(27), 12329–12334.

<https://doi.org/10.1073/pnas.0914886107>

He, Y., & Evans, A. (2010). Graph theoretical modeling of brain connectivity. *Current Opinion in Neurology*, 23(4), 341–350.

<https://doi.org/10.1097/WCO.0b013e32833aa567>

Hellyer, P. J., Scott, G., Shanahan, M., Sharp, D. J., & Leech, R. (2015). Cognitive Flexibility through Metastable Neural Dynamics Is Disrupted by Damage to the Structural Connectome. *Journal of Neuroscience*, 35(24), 9050–9063.

<https://doi.org/10.1523/JNEUROSCI.4648-14.2015>

Henoumont, C., Laurent, S., & Vander Elst, L. (2009). How to perform accurate and reliable measurements of longitudinal and transverse relaxation times of MRI contrast media in aqueous solutions. *Contrast Media & Molecular Imaging*,

- 4(6), 312–321. <https://doi.org/10.1002/cmmi.294>
- Hofmann, S. G., Curtiss, J., Carpenter, J. K., & Kind, S. (2017). Effect of treatments for depression on quality of life: A meta-analysis. *Cognitive Behaviour Therapy*, 46(4), 265–286. <https://doi.org/10.1080/16506073.2017.1304445>
- Holtzheimer, P. E., & Mayberg, H. S. (2011). Stuck in a rut: Rethinking depression and its treatment. *Trends in Neurosciences*, 34(1), 1–9. <https://doi.org/10.1016/j.tins.2010.10.004>
- Houmard, J. A., Smith, R., & Jendrasiak, G. L. (1995). Relationship between MRI relaxation time and muscle fiber composition. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, 78(3), 807–809. <https://doi.org/10.1152/jappl.1995.78.3.807>
- Hu, X., Le, T. H., Parrish, T., & Erhard, P. (1995). Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magnetic Resonance in Medicine*, 34(2), 201–212. <https://doi.org/10.1002/mrm.1910340211>
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A., Keilholz, S., Kiviniemi, V., Leopold, D. A., de Pasquale, F., Sporns, O., Walter, M., & Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, 80, 360–378. <https://doi.org/10.1016/j.neuroimage.2013.05.079>
- Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology*, 30(2), 190–200. <https://doi.org/10.1016/j.newideapsych.2011.11.001>
- Irnich, W., & Schmitt, F. (1995). Magnetostimulation in MRI. *Magnetic Resonance in*

- Medicine*, 33(5), 619–623. <https://doi.org/10.1002/mrm.1910330506>
- Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., L. Whitwell, J., Ward, C., Dale, A. M., Felmlee, J. P., Gunter, J. L., Hill, D. L. G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., ... ADNI Study. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*, 27(4), 685–691. <https://doi.org/10.1002/jmri.21049>
- Jersild, A. T. (1927). *Mental set and shift*. *Archives of psychology*. Nabu Press.
- Jiang, Y., & Kanwisher, N. (2003). Common Neural Mechanisms for Response Selection and Perceptual Processing. *Journal of Cognitive Neuroscience*, 15(8), 1095–1110. <https://doi.org/10.1162/089892903322598076>
- Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A., & Cox, R. W. (2010). Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *NeuroImage*, 52(2), 571–582. <https://doi.org/10.1016/j.neuroimage.2010.04.246>
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. *JAMA Psychiatry*, 72(6), 603. <https://doi.org/10.1001/jamapsychiatry.2015.0071>
- Keller, M. C., Neale, M. C., & Kendler, K. S. (2007). Association of Different Adverse Life Events With Distinct Patterns of Depressive Symptoms. *American Journal of Psychiatry*, 164(10), 1521–1529. <https://doi.org/10.1176/appi.ajp.2007.06091564>
- Kendler, K. S. (2016). The Phenomenology of Major Depression and the

- Representativeness and Nature of DSM Criteria. *American Journal of Psychiatry*, 173(8), 771–780. <https://doi.org/10.1176/appi.ajp.2016.15121509>
- Kennedy, W. G. (2012). Modelling Human Behaviour in Agent-Based Models. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-Based Models of Geographical Systems* (pp. 167–179). Springer Netherlands. https://doi.org/10.1007/978-90-481-8927-4_9
- Khambhati, A. N., Medaglia, J. D., Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. (2018). Subgraphs of functional brain networks identify dynamical constraints of cognitive control. *PLOS Computational Biology*, 14(7), e1006234. <https://doi.org/10.1371/journal.pcbi.1006234>
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, 136(5), 849–874. <https://doi.org/10.1037/a0019842>
- Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A Meta-Analysis. *Human Brain Mapping*, 33(1), 130–142. <https://doi.org/10.1002/hbm.21199>
- Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin*, 144(6), 557–583. <https://doi.org/10.1037/bul0000144>
- Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, 399(6732), 148–151. <https://doi.org/10.1038/20178>

- Koechlin, E., & Hyafil, A. (2007). Anterior Prefrontal Function and the Limits of Human Decision-Making. *Science*, 318(5850), 594–598.
<https://doi.org/10.1126/science.1142995>
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, 302(5648), 1181–1185.
<https://doi.org/10.1126/science.1088545>
- Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information *. *International Journal of Computer Mathematics*, 2(1–4), 157–168. <https://doi.org/10.1080/00207166808803030>
- Koster, E. H. W., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011). Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review*, 31(1), 138–145. <https://doi.org/10.1016/j.cpr.2010.08.005>
- Kramer, A. F., Hahn, S., & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica*, 101(2–3), 339–378.
[https://doi.org/10.1016/S0001-6918\(99\)00011-6](https://doi.org/10.1016/S0001-6918(99)00011-6)
- Krebs, T. S., & Johansen, P.-Ø. (2012). Lysergic acid diethylamide (LSD) for alcoholism: Meta-analysis of randomized controlled trials. *Journal of Psychopharmacology*, 26(7), 994–1002.
<https://doi.org/10.1177/0269881112439253>
- Lanka, P., Rangaprakash, D., Gotoor, S. S. R., Dretsche, M. N., Katz, J. S., Denney, T. S., & Deshpande, G. (2020). MALINI (Machine Learning in NeuroImaging): A MATLAB toolbox for aiding clinical diagnostics using resting-state fMRI data.

- Data in Brief*, 29, 105213. <https://doi.org/10.1016/j.dib.2020.105213>
- Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting-state fMRI: A review of methods and clinical applications. *AJNR. American Journal of Neuroradiology*, 34(10), 1866–1872. <https://doi.org/10.3174/ajnr.A3263>
- Leech, R., Braga, R., & Sharp, D. J. (2012). Echoes of the Brain within the Posterior Cingulate Cortex. *Journal of Neuroscience*, 32(1), 215–222. <https://doi.org/10.1523/JNEUROSCI.3689-11.2012>
- Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: Distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(9), 3217–3224. <https://doi.org/10.1523/JNEUROSCI.5626-10.2011>
- Lempel, A., & Ziv, J. (1976). On the Complexity of Finite Sequences. *IEEE Transactions on Information Theory*, 22(1), 75–81. <https://doi.org/10.1109/TIT.1976.1055501>
- Leonardi, N., & Van De Ville, D. (2015). On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage*, 104, 430–436. <https://doi.org/10.1016/j.neuroimage.2014.09.007>
- Liefooghe, B., Demanet, J., & Vandierendonck, A. (2010). Persisting activation in voluntary task switching: It all depends on the instructions. *Psychonomic Bulletin & Review*, 17(3), 381–386. <https://doi.org/10.3758/PBR.17.3.381>
- Locher, C., Koechlin, H., Zion, S. R., Werner, C., Pine, D. S., Kirsch, I., Kessler, R. C., & Kossowsky, J. (2017). Efficacy and Safety of Selective Serotonin Reuptake Inhibitors, Serotonin-Norepinephrine Reuptake Inhibitors, and

- Placebo for Common Psychiatric Disorders Among Children and Adolescents: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, 74(10), 1011. <https://doi.org/10.1001/jamapsychiatry.2017.2432>
- Logan, G. D. (2007). What it costs to implement a plan: Plan-level and task-level contributions to switch costs. *Memory & Cognition*, 35(4), 591–602. <https://doi.org/10.3758/BF03193297>
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453(7197), 869–878. <https://doi.org/10.1038/nature06976>
- Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD Signal. *Annual Review of Physiology*, 66(1), 735–769. <https://doi.org/10.1146/annurev.physiol.66.082602.092845>
- Lord, A., Horn, D., Breakspear, M., & Walter, M. (2012). Changes in Community Structure of Resting State Functional Connectivity in Unipolar Depression. *PLoS ONE*, 7(8), e41282. <https://doi.org/10.1371/journal.pone.0041282>
- Lord, L.-D., Expert, P., Atasoy, S., Roseman, L., Rapuano, K., Lambiotte, R., Nutt, D. J., Deco, G., Carhart-Harris, R. L., Kringelbach, M. L., & Cabral, J. (2019). Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. *NeuroImage*, 199, 127–142. <https://doi.org/10.1016/j.neuroimage.2019.05.060>
- Lorenz, R., Hampshire, A., & Leech, R. (2017). Neuroadaptive Bayesian Optimization and Hypothesis Testing. *Trends in Cognitive Sciences*, 21(3), 155–167. <https://doi.org/10.1016/j.tics.2017.01.006>
- Lucas, G. S., Bazzi, M., Jutla, I. S., & Mucha, P. J. (2016). *A generalized Louvain method for community detection implemented in MATLAB*.

<http://netwiki.amath.unc.edu/GenLouvain>

- Luppi, A. I., Carhart-Harris, R. L., Roseman, L., Pappas, I., Menon, D. K., & Stamatakis, E. A. (2021). LSD alters dynamic integration and segregation in the human brain. *NeuroImage*, 227, 117653.
<https://doi.org/10.1016/j.neuroimage.2020.117653>
- Lydon-Staley, D. M., Kuehner, C., Zamoscik, V., Huffziger, S., Kirsch, P., & Bassett, D. S. (2019). Repetitive negative thinking in daily life and functional connectivity among default mode, fronto-parietal, and salience networks. *Translational Psychiatry*, 9(1), 234.
<https://doi.org/10.1038/s41398-019-0560-0>
- Lyons, T., & Carhart-Harris, R. L. (2018). More Realistic Forecasting of Future Life Events After Psilocybin for Treatment-Resistant Depression. *Frontiers in Psychology*, 9, 1721. <https://doi.org/10.3389/fpsyg.2018.01721>
- Madhavan, R., Joel, S. E., Mullick, R., Cogsil, T., Niogi, S. N., Tsiouris, A. J., Mukherjee, P., Masdeu, J. C., Marinelli, L., & Shetty, T. (2019). Longitudinal Resting State Functional Connectivity Predicts Clinical Outcome in Mild Traumatic Brain Injury. *Journal of Neurotrauma*, 36(5), 650–660.
<https://doi.org/10.1089/neu.2018.5739>
- Makowski, C., Lepage, M., & Evans, A. C. (2019). Head motion: The dirty little secret of neuroimaging in psychiatry. *Journal of Psychiatry & Neuroscience: JPN*, 44(1), 62–68. <https://doi.org/10.1503/jpn.180022>
- 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., & 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.
<https://doi.org/10.3389/fnhum.2013.00930>
- Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., Bezgin, G., Eickhoff, S. B., Castellanos, F. X., Petrides, M., Jefferies, E., & Smallwood, J. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, 113(44), 12574–12579.
<https://doi.org/10.1073/pnas.1608282113>
- Margulies, D. S., & Smallwood, J. (2017). Converging evidence for the role of transmodal cortex in cognition. *Proceedings of the National Academy of Sciences*, 114(48), 12641–12643. <https://doi.org/10.1073/pnas.1717374114>
- Maslov, S., & Sneppen, K. (2002). Specificity and Stability in Topology of Protein Networks. *Science*, 296(5569), 910–913.
<https://doi.org/10.1126/science.1065103>
- Mason, S. L., Daws, R. E., Soreq, E., Johnson, E. B., Scahill, R. I., Tabrizi, S. J., Barker, R. A., & Hampshire, A. (2018). Predicting clinical diagnosis in Huntington’s disease: An imaging polymarker: Predicting Clinical Diagnosis in HD. *Annals of Neurology*, 83(3), 532–543. <https://doi.org/10.1002/ana.25171>
- Mattar, M. G., Cole, M. W., Thompson-Schill, S. L., & Bassett, D. S. (2015). A Functional Cartography of Cognitive Systems. *PLOS Computational Biology*, 11(12), e1004533. <https://doi.org/10.1371/journal.pcbi.1004533>
- Matthews, P. M., & Hampshire, A. (2016). Clinical Concepts Emerging from fMRI Functional Connectomics. *Neuron*, 91(3), 511–528.

- <https://doi.org/10.1016/j.neuron.2016.07.031>
- Mayr, U. (2002). Inhibition of action rules. *Psychonomic Bulletin & Review*, 9(1), 93–99. <https://doi.org/10.3758/BF03196261>
- Mayr, U., Diedrichsen, J., Ivry, R., & Keele, S. W. (2006). Dissociating Task-set Selection from Task-set Inhibition in the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, 18(1), 14–21. <https://doi.org/10.1162/089892906775250085>
- Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1124–1140. <https://doi.org/10.1037/0278-7393.26.5.1124>
- McKeown, M. J., & Sejnowski, T. J. (1998). Independent component analysis of fMRI data: Examining the assumptions. *Human Brain Mapping*, 6(5–6), 368–372.
- McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, 61(4), 1277–1286. <https://doi.org/10.1016/j.neuroimage.2012.03.068>
- Mechelli, A., Henson, R. N. A., Price, C. J., & Friston, K. J. (2003). Comparing event-related and epoch analysis in blocked design fMRI. *NeuroImage*, 18(3), 806–810. [https://doi.org/10.1016/s1053-8119\(02\)00027-7](https://doi.org/10.1016/s1053-8119(02)00027-7)
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1423–1442. <https://doi.org/10.1037/0278-7393.22.6.1423>
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component Processes in Task Switching.

Cognitive Psychology, 41(3), 211–253.

<https://doi.org/10.1006/cogp.2000.0736>

Meiran, N., Diamond, G. M., Toder, D., & Nemets, B. (2011). Cognitive rigidity in unipolar depression and obsessive compulsive disorder: Examination of task switching, Stroop, working memory updating and post-conflict adaptation. *Psychiatry Research*, 185(1–2), 149–156.

<https://doi.org/10.1016/j.psychres.2010.04.044>

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review*, 104(1), 3–65.

<https://doi.org/10.1037/0033-295X.104.1.3>

Meyer, F., & Beucher, S. (1990). Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1), 21–46.

[https://doi.org/10.1016/1047-3203\(90\)90014-M](https://doi.org/10.1016/1047-3203(90)90014-M)

Mikl, M., Marecek, R., Hlustík, P., Pavlicová, M., Drastich, A., Chlebus, P., Brázdil, M., & Krupa, P. (2008). Effects of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging*, 26(4), 490–503.

<https://doi.org/10.1016/j.mri.2007.08.006>

Millan, M. J., Agid, Y., Brüne, M., Bullmore, E. T., Carter, C. S., Clayton, N. S., Connor, R., Davis, S., Deakin, B., DeRubeis, R. J., Dubois, B., Geyer, M. A., Goodwin, G. M., Gorwood, P., Jay, T. M., Joëls, M., Mansuy, I. M., Meyer-Lindenberg, A., Murphy, D., ... Young, L. J. (2012). Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. *Nature Reviews Drug Discovery*, 11(2), 141–168.

<https://doi.org/10.1038/nrd3628>

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1(1), 59–65. <https://doi.org/10.1038/35036228>

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, 24(1), 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>

Milner, B. (1963). Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes. *Archives of Neurology*, 9(1), 90. <https://doi.org/10.1001/archneur.1963.00460070100010>

Mirchi, N., Betzel, R. F., Bernhardt, B. C., Dagher, A., & Mišić, B. (2019). Tracking mood fluctuations with functional network patterns. *Social Cognitive and Affective Neuroscience*, 14(1), 47–57. <https://doi.org/10.1093/scan/nsy107>

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, 115(2–3), 123–142. <https://doi.org/10.1016/j.actpsy.2003.12.004>

Mohr, H., Wolfensteller, U., Betzel, R. F., Mišić, B., Sporns, O., Richiardi, J., & Ruge, H. (2016). Integration and segregation of large-scale brain networks during short-term task automatization. *Nature Communications*, 7(1), 13217. <https://doi.org/10.1038/ncomms13217>

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)

Monti, M. M. (2011). Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach. *Frontiers in Human Neuroscience*, 5, 28.

<https://doi.org/10.3389/fnhum.2011.00028>

Moreno, F. A., Wiegand, C. B., Taitano, E. K., & Delgado, P. L. (2006). Safety, Tolerability, and Efficacy of Psilocybin in 9 Patients With Obsessive-Compulsive Disorder. *The Journal of Clinical Psychiatry*, 67(11), 1735–1740. <https://doi.org/10.4088/JCP.v67n1110>

Mouraux, A., Diukova, A., Lee, M. C., Wise, R. G., & Iannetti, G. D. (2011). A multisensory investigation of the functional significance of the “pain matrix”. *NeuroImage*, 54(3), 2237–2249. <https://doi.org/10.1016/j.neuroimage.2010.09.084>

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science*, 328(5980), 876–878. <https://doi.org/10.1126/science.1184819>

Mullins, N., & Lewis, C. M. (2017). Genetics of Depression: Progress at Last. *Current Psychiatry Reports*, 19(8), 43. <https://doi.org/10.1007/s11920-017-0803-9>

Muresan, L., Renken, R., Roerdink, J. B., & Duifhuis, H. (2002). *Position-history and spin-history artifacts in fMRI time series* (A. V. Clough & C.-T. Chen, Eds.; p. 444). <https://doi.org/10.1117/12.463613>

Murphy-Beiner, A., & Soar, K. (2020). Ayahuasca’s ‘afterglow’: Improved mindfulness and cognitive flexibility in ayahuasca drinkers. *Psychopharmacology*, 237(4), 1161–1169. <https://doi.org/10.1007/s00213-019-05445-3>

Nelson, H. E. (1976). A Modified Card Sorting Test Sensitive to Frontal Lobe Defects. *Cortex*, 12(4), 313–324. [https://doi.org/10.1016/S0010-9452\(76\)80035-4](https://doi.org/10.1016/S0010-9452(76)80035-4)

- Nestler, E. J., Barrot, M., DiLeone, R. J., Eisch, A. J., Gold, S. J., & Monteggia, L. M. (2002). Neurobiology of Depression. *Neuron*, 34(1), 13–25.
[https://doi.org/10.1016/S0896-6273\(02\)00653-0](https://doi.org/10.1016/S0896-6273(02)00653-0)
- Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E*, 70(5), 056131. <https://doi.org/10.1103/PhysRevE.70.056131>
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577–8582.
<https://doi.org/10.1073/pnas.0601602103>
- Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 026113.
<https://doi.org/10.1103/PhysRevE.69.026113>
- Norman, D. A., & Shallice, T. (1986). Attention to Action: Willed and Automatic Control of Behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation* (pp. 1–18). Springer US.
https://doi.org/10.1007/978-1-4757-0629-1_1
- Nutt, D., & Carhart-Harris, R. (2020). The Current Status of Psychedelics in Psychiatry. *JAMA Psychiatry*.
<https://doi.org/10.1001/jamapsychiatry.2020.2171>
- Nutt, D., Erritzoe, D., & Carhart-Harris, R. (2020). Psychedelic Psychiatry's Brave New World. *Cell*, 181(1), 24–28. <https://doi.org/10.1016/j.cell.2020.03.020>
- Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical Journal*, 64(3),

- 803–812. [https://doi.org/10.1016/S0006-3495\(93\)81441-3](https://doi.org/10.1016/S0006-3495(93)81441-3)
- Oldendorf, W., & Oldendorf, W. (1988). Advantages and Disadvantages of MRI. In W. Oldendorf & W. Oldendorf, *Basics of Magnetic Resonance Imaging* (pp. 125–138). Springer US. https://doi.org/10.1007/978-1-4613-2081-4_9
- Orban, C., Kong, R., Li, J., Chee, M. W. L., & Yeo, B. T. T. (2020). Time of day is associated with paradoxical reductions in global signal fluctuation and functional connectivity. *PLOS Biology*, 18(2), e3000602. <https://doi.org/10.1371/journal.pbio.3000602>
- Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, 28(10), 1021–1034. [https://doi.org/10.1016/0028-3932\(90\)90137-D](https://doi.org/10.1016/0028-3932(90)90137-D)
- Palhano-Fontes, F., Barreto, D., Onias, H., Andrade, K. C., Novaes, M. M., Pessoa, J. A., Mota-Rolim, S. A., Osório, F. L., Sanches, R., dos Santos, R. G., Tófoli, L. F., de Oliveira Silveira, G., Yonamine, M., Riba, J., Santos, F. R., Silva-Junior, A. A., Alchieri, J. C., Galvão-Coelho, N. L., Lobão-Soares, B., ... Araújo, D. B. (2019). Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: A randomized placebo-controlled trial. *Psychological Medicine*, 49(4), 655–663. <https://doi.org/10.1017/S0033291718001356>
- Panych, L. P., & Madore, B. (2018). The physics of MRI safety: Physics of MRI Safety. *Journal of Magnetic Resonance Imaging*, 47(1), 28–43. <https://doi.org/10.1002/jmri.25761>
- Parkes, L., Fulcher, B., Yücel, M., & Fornito, A. (2018). An evaluation of the efficacy,

reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage*, 171, 415–436.

<https://doi.org/10.1016/j.neuroimage.2017.12.073>

Parkes, L., Moore, T. M., Calkins, M. E., Cook, P. A., Cieslak, M., Roalf, D. R., Wolf, D. H., Gur, R. C., Gur, R. E., Satterthwaite, T. D., & Bassett, D. S. (2021).

Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Translational Psychiatry*, 11(1), 232. <https://doi.org/10.1038/s41398-021-01342-6>

Parkes, L., Satterthwaite, T. D., & Bassett, D. S. (2020). Towards precise resting-state fMRI biomarkers in psychiatry: Synthesizing developments in transdiagnostic research, dimensional models of psychopathology, and normative neurodevelopment. *ArXiv:2006.04728 [q-Bio]*.

<http://arxiv.org/abs/2006.04728>

Parkin, B. L., Hellyer, P. J., Leech, R., & Hampshire, A. (2015). Dynamic Network Mechanisms of Relational Integration. *Journal of Neuroscience*, 35(20), 7660–7673. <https://doi.org/10.1523/JNEUROSCI.4956-14.2015>

Pasquini, L., Palhano-Fontes, F., & Araujo, D. B. (2020). Subacute effects of the psychedelic ayahuasca on the salience and default mode networks. *Journal of Psychopharmacology*, 34(6), 623–635.

<https://doi.org/10.1177/0269881120909409>

Pauling, L., & Coryell, C. D. (1936). The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. *Proceedings of the National Academy of Sciences*, 22(4), 210–216.

<https://doi.org/10.1073/pnas.22.4.210>

- Pearson, K. (1896). VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 187, 253–318. <https://doi.org/10.1098/rsta.1896.0007>
- Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., & Vaccarino, F. (2014). Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101), 20140873. <https://doi.org/10.1098/rsif.2014.0873>
- Pierson, E., Althoff, T., Thomas, D., Hillard, P., & Leskovec, J. (2021). Daily, weekly, seasonal and menstrual cycles in women’s mood, behaviour and vital signs. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-020-01046-9>
- Pipe, J. G. (1995). Spatial Encoding and Reconstruction in MRI with Quadratic Phase Profiles. *Magnetic Resonance in Medicine*, 33(1), 24–33. <https://doi.org/10.1002/mrm.1910330105>
- Plewes, D. B., & Kucharczyk, W. (2012). Physics of MRI: A primer. *Journal of Magnetic Resonance Imaging*, 35(5), 1038–1054. <https://doi.org/10.1002/jmri.23642>
- Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, 2(1), 67–70. <https://doi.org/10.1093/scan/nsm006>
- Poldrack, R. A. (2012). The future of fMRI in cognitive neuroscience. *NeuroImage*, 62(2), 1216–1220. <https://doi.org/10.1016/j.neuroimage.2011.08.007>
- Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., Gorgolewski, K. J., Luci, J., Joo, S. J., Boyd, R. L., Hunicke-Smith, S., Simpson, Z. B., Caven, T., Sochat, V., Shine, J. M., Gordon, E., Snyder, A. Z.,

- Adeyemo, B., Petersen, S. E., ... Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, 6(1), 8885. <https://doi.org/10.1038/ncomms9885>
- Poldrack, R. A., & Yarkoni, T. (2016). From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. *Annual Review of Psychology*, 67(1), 587–612.
<https://doi.org/10.1146/annurev-psych-122414-033729>
- Poljac, E., Kiesel, A., Koch, I., & Müller, H. (2018). New perspectives on human multitasking. *Psychological Research*, 82(1), 1–3.
<https://doi.org/10.1007/s00426-018-0970-2>
- Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, 55(6), 1227–1235. <https://doi.org/10.1002/mrm.20900>
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154.
<https://doi.org/10.1016/j.neuroimage.2011.10.018>
- Power, J. D., Lynch, C. J., Silver, B. M., Dubin, M. J., Martin, A., & Jones, R. M. (2019). Distinctions among real and apparent respiratory motions in human fMRI data. *NeuroImage*, 201, 116041.
<https://doi.org/10.1016/j.neuroimage.2019.116041>
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in

- resting state fMRI. *NeuroImage*, 84, 320–341.
<https://doi.org/10.1016/j.neuroimage.2013.08.048>
- Power, J. D., Schlaggar, B. L., & Petersen, S. E. (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage*, 105, 536–551. <https://doi.org/10.1016/j.neuroimage.2014.10.044>
- Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology. General*, 134(1), 73–92. <https://doi.org/10.1037/0096-3445.134.1.73>
- Rabeea, S. A., Merchant, H. A., Khan, M. U., Kow, C. S., & Hasan, S. S. (2021). Surging trends in prescriptions and costs of antidepressants in England amid COVID-19. *DARU Journal of Pharmaceutical Sciences*.
<https://doi.org/10.1007/s40199-021-00390-z>
- Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, 5(3), 184–194. <https://doi.org/10.1038/nrn1343>
- Reichenbach, J. R., Venkatesan, R., Schillinger, D. J., Kido, D. K., & Haacke, E. M. (1997). Small vessels in the human brain: MR venography with deoxyhemoglobin as an intrinsic contrast agent. *Radiology*, 204(1), 272–277.
<https://doi.org/10.1148/radiology.204.1.9205259>
- Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, 4(3), 161–198. <https://doi.org/10.1007/BF01874891>
- Robbins, T. (2002). The 5-choice serial reaction time task: Behavioural pharmacology and functional neurochemistry. *Psychopharmacology*,

- 163(3–4), 362–380. <https://doi.org/10.1007/s00213-002-1154-7>
- Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., & Breakspear, M. (2019). Metastable brain waves. *Nature Communications*, 10(1), 1056. <https://doi.org/10.1038/s41467-019-08999-0>
- Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain*, 133(1), 234–247. <https://doi.org/10.1093/brain/awp269>
- Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, 44(10), 2029–2040. <https://doi.org/10.1017/S0033291713002535>
- Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity in the human brain by fMRI. *Magnetic Resonance Imaging*, 25(10), 1347–1357. <https://doi.org/10.1016/j.mri.2007.03.007>
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231. <https://doi.org/10.1037/0096-3445.124.2.207>
- Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell’Acqua, F., & Thiebaut de Schotten, M. (2016). Atlasing the frontal lobe connections and their variability due to age and education: A spherical deconvolution tractography study. *Brain Structure and Function*, 221(3), 1751–1766. <https://doi.org/10.1007/s00429-015-1001-3>
- Rolls, E. T. (2016). A non-reward attractor theory of depression. *Neuroscience & Biobehavioral Reviews*, 68, 47–58.

<https://doi.org/10.1016/j.neubiorev.2016.05.007>

Roseman, L., Haijen, E., Idialu-Ikato, K., Kaelen, M., Watts, R., & Carhart-Harris, R. (2019). Emotional breakthrough and psychedelics: Validation of the Emotional Breakthrough Inventory. *Journal of Psychopharmacology*, 33(9), 1076–1087. <https://doi.org/10.1177/0269881119855974>

Roseman, L., Leech, R., Feilding, A., Nutt, D. J., & Carhart-Harris, R. L. (2014). The effects of psilocybin and MDMA on between-network resting state functional connectivity in healthy volunteers. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00204>

Roseman, L., Nutt, D. J., & Carhart-Harris, R. L. (2018). Quality of Acute Psychedelic Experience Predicts Therapeutic Efficacy of Psilocybin for Treatment-Resistant Depression. *Frontiers in Pharmacology*, 8, 974. <https://doi.org/10.3389/fphar.2017.00974>

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, 19(1), 165–171. <https://doi.org/10.1038/nn.4179>

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *NeuroImage*, 20(1), 351–358. [https://doi.org/10.1016/S1053-8119\(03\)00275-1](https://doi.org/10.1016/S1053-8119(03)00275-1)

Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*, 56(4), 2068–2079. <https://doi.org/10.1016/j.neuroimage.2011.03.069>

- Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763–797.
<https://doi.org/10.1037/0096-1523.27.4.763>
- Ruge, H., Karcz, T., Mark, T., Martin, V., Zwosta, K., & Wolfensteller, U. (2018). On the efficiency of instruction-based rule encoding. *Acta Psychologica*, 184, 4–19. <https://doi.org/10.1016/j.actpsy.2017.04.005>
- Ruge, H., Schäfer, T. A., Zwosta, K., Mohr, H., & Wolfensteller, U. (2019). Neural representation of newly instructed rule identities during early implementation trials. *ELife*, 8, e48293. <https://doi.org/10.7554/eLife.48293>
- Ruge, H., & Wolfensteller, U. (2010). Rapid Formation of Pragmatic Rule Representations in the Human Brain during Instruction-Based Learning. *Cerebral Cortex*, 20(7), 1656–1667. <https://doi.org/10.1093/cercor/bhp228>
- Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the Human Medial Frontal Cortex in Task Switching: A Combined fMRI and TMS Study. *Journal of Neurophysiology*, 87(5), 2577–2592.
<https://doi.org/10.1152/jn.2002.87.5.2577>
- Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance*, 27(6), 1404–1419. <https://doi.org/10.1037/0096-1523.27.6.1404>
- Sackeim, H. A. (2001). The definition and meaning of treatment-resistant depression. *The Journal of Clinical Psychiatry*, 62 Suppl 16, 10–17.
- Sakai, K. (2008). Task Set and Prefrontal Cortex. *Annual Review of Neuroscience*,

- 31(1), 219–245. <https://doi.org/10.1146/annurev.neuro.31.060407.125642>
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, 28(9), 3095–3114. <https://doi.org/10.1093/cercor/bhx179>
- Scott, G., Hellyer, P. J., Hampshire, A., & Leech, R. (2015). Exploring spatiotemporal network transitions in task functional MRI: Exploring Spatiotemporal Network Transition. *Human Brain Mapping*, 36(4), 1348–1364. <https://doi.org/10.1002/hbm.22706>
- Shallice, T., Stuss, D. T., Picton, T. W., Alexander, M. P., & Gillingham, S. (2007). Multiple effects of prefrontal lesions on task-switching. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/neuro.09.002.2007>
- Shallice, T., & Burgess, P. W. (1991). DEFICITS IN STRATEGY APPLICATION FOLLOWING FRONTAL LOBE DAMAGE IN MAN. *Brain*, 114(2), 727–741. <https://doi.org/10.1093/brain/114.2.727>
- Shallice, T., & Evans, M. E. (1978). The Involvement of the Frontal Lobes in Cognitive Estimation. *Cortex*, 14(2), 294–303. [https://doi.org/10.1016/S0010-9452\(78\)80055-0](https://doi.org/10.1016/S0010-9452(78)80055-0)
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., & Mehta, M. A. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proceedings of the National*

Academy of Sciences, 107(13), 6106–6111.

<https://doi.org/10.1073/pnas.1000175107>

Sheffield, J. M., Mohr, H., Ruge, H., & Barch, D. M. (2021). Disrupted Salience and Cingulo-Opercular Network Connectivity During Impaired Rapid Instructed Task Learning in Schizophrenia. *Clinical Psychological Science*, 9(2), 210–221. <https://doi.org/10.1177/2167702620959341>

Simbula, S. (2010). Daily fluctuations in teachers' well-being: A diary study using the Job Demands–Resources model. *Anxiety, Stress & Coping*, 23(5), 563–584. <https://doi.org/10.1080/10615801003728273>

Sladky, R., Friston, K. J., Tröstl, J., Cunnington, R., Moser, E., & Windischberger, C. (2011). Slice-timing effects and their correction in functional MRI. *NeuroImage*, 58(2), 588–594. <https://doi.org/10.1016/j.neuroimage.2011.06.078>

Slotnick, S. D. (2017). Cluster success: FMRI inferences for spatial extent have acceptable false-positive rates. *Cognitive Neuroscience*, 8(3), 150–155. <https://doi.org/10.1080/17588928.2017.1319350>

Smallwood, J., Turnbull, A., Wang, H., Ho, N. S. P., Poerio, G. L., Karapanagiotidis, T., Konu, D., Mckeown, B., Zhang, M., Murphy, C., Vatansever, D., Bzdok, D., Konishi, M., Leech, R., Seli, P., Schooler, J. W., Bernhardt, B., Margulies, D. S., & Jefferies, E. (2021). The neural correlates of ongoing conscious thought. *IScience*, 24(3), 102132. <https://doi.org/10.1016/j.isci.2021.102132>

Smith, E. E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P., & Koeppe, R. A. (2001). The neural basis of task-switching in working memory: Effects of performance and aging. *Proceedings of the National Academy of Sciences*,

- 98(4), 2095–2100. <https://doi.org/10.1073/pnas.98.4.2095>
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31), 13040–13045. <https://doi.org/10.1073/pnas.0905267106>
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E. J., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., & Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience*, 18(11), 1565–1567. <https://doi.org/10.1038/nn.4125>
- Soreq, E., Leech, R., & Hampshire, A. (2019). Dynamic network coding of working-memory domains and working-memory processes. *Nature Communications*, 10(1), 936. <https://doi.org/10.1038/s41467-019-08840-8>
- Soreq, E., Violante, I. R., Daws, R. E., & Hampshire, A. (2021). Neuroimaging evidence for a network sampling theory of individual differences in human intelligence test performance. *Nature Communications*, 12(1), 2072. <https://doi.org/10.1038/s41467-021-22199-9>

- Spector, A., & Biederman, I. (1976). Mental Set and Mental Shift Revisited. *The American Journal of Psychology*, 89(4), 669. <https://doi.org/10.2307/1421465>
- Sporns, O., & Kötter, R. (2004). Motifs in brain networks. *PLoS Biology*, 2(11), e369. <https://doi.org/10.1371/journal.pbio.0020369>
- Spriggs, M. J., Kettner, H., & Carhart-Harris, R. L. (2020). Positive effects of psychedelics on depression and wellbeing scores in individuals reporting an eating disorder. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*. <https://doi.org/10.1007/s40519-020-01000-8>
- Steer, R. A., Beck, A. T., Riskind, J. H., & Brown, G. (1987). Relationships between the Beck Depression Inventory and the Hamilton Psychiatric Rating Scale for Depression in depressed outpatients. *Journal of Psychopathology and Behavioral Assessment*, 9(3), 327–339. <https://doi.org/10.1007/BF00964561>
- Steinert, C., Hofmann, M., Kruse, J., & Leichsenring, F. (2014). Relapse rates after psychotherapy for depression – stable long-term effects? A meta-analysis. *Journal of Affective Disorders*, 168, 107–118. <https://doi.org/10.1016/j.jad.2014.06.043>
- Stockmann, J. P., & Wald, L. L. (2018). In vivo B 0 field shimming methods for MRI at 7 T. *NeuroImage*, 168, 71–87. <https://doi.org/10.1016/j.neuroimage.2017.06.013>
- Stoet, G., & Snyder, L. H. (2003). Executive control and task-switching in monkeys. *Neuropsychologia*, 41(10), 1357–1364. [https://doi.org/10.1016/S0028-3932\(03\)00048-4](https://doi.org/10.1016/S0028-3932(03)00048-4)
- Stoet, G., & Snyder, L. H. (2007). Extensive practice does not eliminate human switch costs. *Cognitive, Affective, & Behavioral Neuroscience*, 7(3), 192–197.

<https://doi.org/10.3758/CABN.7.3.192>

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405.

<https://doi.org/10.1016/j.tics.2015.05.004>

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, 39(8), 771–786.

[https://doi.org/10.1016/S0028-3932\(01\)00013-6](https://doi.org/10.1016/S0028-3932(01)00013-6)

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 89–103. <https://doi.org/10.1037/0096-1523.13.1.89>

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 99(25), 16354–16359.

<https://doi.org/10.1073/pnas.252521499>

Tang, F., Liang, J., Zhang, H., Kelifa, M. M., He, Q., & Wang, P. (2021). COVID-19 related depression and anxiety among quarantined respondents. *Psychology & Health*, 36(2), 164–178. <https://doi.org/10.1080/08870446.2020.1782410>

Tatu, L., Moulin, T., Bogousslavsky, J., & Duvernoy, H. (1998). Arterial territories of the human brain: Cerebral hemispheres. *Neurology*, 50(6), 1699–1708.

<https://doi.org/10.1212/WNL.50.6.1699>

Taylor Tavares, J. V., Clark, L., Cannon, D. M., Erickson, K., Drevets, W. C., & Sahakian, B. J. (2007). Distinct Profiles of Neurocognitive Function in Unmedicated Unipolar Depression and Bipolar II Depression. *Biological*

- Psychiatry*, 62(8), 917–924. <https://doi.org/10.1016/j.biopsych.2007.05.034>
- Teipel, S., Drzezga, A., Grothe, M. J., Barthel, H., Chételat, G., Schuff, N., Skudlarski, P., Cavedo, E., Frisoni, G. B., Hoffmann, W., Thyrian, J. R., Fox, C., Minoshima, S., Sabri, O., & Fellgiebel, A. (2015). Multimodal imaging in Alzheimer's disease: Validity and usefulness for early detection. *The Lancet Neurology*, 14(10), 1037–1053.
- [https://doi.org/10.1016/S1474-4422\(15\)00093-9](https://doi.org/10.1016/S1474-4422(15)00093-9)
- Teixeira, A., Matos, A., Souto, A., & Antunes, L. (2011). Entropy Measures vs. Kolmogorov Complexity. *Entropy*, 13(3), 595–611.
- <https://doi.org/10.3390/e13030595>
- Thiebaut de Schotten, M., Urbanski, M., Batrancourt, B., Levy, R., Dubois, B., Cerliani, L., & Volle, E. (2016). Rostro-caudal Architecture of the Frontal Lobes in Humans. *Cerebral Cortex*, cercor;bhw215v1.
- <https://doi.org/10.1093/cercor/bhw215>
- Thulborn, K. R., Waterton, J. C., Matthews, P. M., & Radda, G. K. (1982). Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochimica Et Biophysica Acta*, 714(2), 265–270.
- [https://doi.org/10.1016/0304-4165\(82\)90333-6](https://doi.org/10.1016/0304-4165(82)90333-6)
- Turnbull, A., Karapanagiotidis, T., Wang, H.-T., Bernhardt, B. C., Leech, R., Margulies, D., Schooler, J., Jefferies, E., & Smallwood, J. (2020). Reductions in task positive neural systems occur with the passage of time and are associated with changes in ongoing thought. *Scientific Reports*, 10(1), 9912.
- <https://doi.org/10.1038/s41598-020-66698-z>
- Uthaug, M. V., van Oorsouw, K., Kuypers, K. P. C., van Boxtel, M., Broers, N. J.,

- Mason, N. L., Toennes, S. W., Riba, J., & Ramaekers, J. G. (2018). Sub-acute and long-term effects of ayahuasca on affect and cognitive thinking style and their association with ego dissolution. *Psychopharmacology*, 235(10), 2979–2989. <https://doi.org/10.1007/s00213-018-4988-3>
- Vandierendonck, A., Liefoghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, 136(4), 601–626. <https://doi.org/10.1037/a0019791>
- Vendetti, M. S., & Bunge, S. A. (2014). Evolutionary and Developmental Changes in the Lateral Frontoparietal Network: A Little Goes a Long Way for Higher-Level Cognition. *Neuron*, 84(5), 906–917. <https://doi.org/10.1016/j.neuron.2014.09.035>
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48), 12827–12832. <https://doi.org/10.1073/pnas.1705120114>
- Vollenweider, F. X., Vollenweider-Scherpenhuyzen, M. F. I., Bäbler, A., Vogel, H., & Hell, D. (1998). Psilocybin induces schizophrenia-like psychosis in humans via a serotonin-2 agonist action: *NeuroReport*, 9(17), 3897–3902. <https://doi.org/10.1097/00001756-199812010-00024>
- Voytek, B., Kayser, A. S., Badre, D., Fegen, D., Chang, E. F., Crone, N. E., Parvizi, J., Knight, R. T., & D'Esposito, M. (2015). Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. *Nature Neuroscience*, 18(9), 1318–1324. <https://doi.org/10.1038/nn.4071>
- Wang, H. E., BÃ©nar, C. G., Quilichini, P. P., Friston, K. J., Jirsa, V. K., & Bernard, C. (2014). A systematic framework for functional connectivity measures.

- Frontiers in Neuroscience*, 8. <https://doi.org/10.3389/fnins.2014.00405>
- Wang, S. S.-H., Shultz, J. R., Burish, M. J., Harrison, K. H., Hof, P. R., Towns, L. C., Wagers, M. W., & Wyatt, K. D. (2008). Functional trade-offs in white matter axonal scaling. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(15), 4047–4056. <https://doi.org/10.1523/JNEUROSCI.5559-05.2008>
- Wang, X., Öngür, D., Auerbach, R. P., & Yao, S. (2016). Cognitive Vulnerability to Major Depression: View from the Intrinsic Network and Cross-network Interactions. *Harvard Review of Psychiatry*, 24(3), 188–201. <https://doi.org/10.1097/HRP.0000000000000081>
- Waszak, F., & Hommel, B. (2007). The costs and benefits of cross-task priming. *Memory & Cognition*, 35(5), 1175–1186. <https://doi.org/10.3758/BF03193487>
- Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, 46(4), 361–413. [https://doi.org/10.1016/S0010-0285\(02\)00520-0](https://doi.org/10.1016/S0010-0285(02)00520-0)
- Watanabe, T., & Rees, G. (2017). Brain network dynamics in high-functioning individuals with autism. *Nature Communications*, 8(1), 16048. <https://doi.org/10.1038/ncomms16048>
- Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, 134(2), 163–206. <https://doi.org/10.1037/0033-2909.134.2.163>
- Watkins, E. R., Mullan, E., Wingrove, J., Rimes, K., Steiner, H., Bathurst, N., Eastman, R., & Scott, J. (2011). Rumination-focused cognitive–behavioural therapy for residual depression: Phase II randomised controlled trial. *British*

Journal of Psychiatry, 199(4), 317–322.

<https://doi.org/10.1192/bjp.bp.110.090282>

- Watts, R., Day, C., Krzanowski, J., Nutt, D., & Carhart-Harris, R. (2017). Patients' Accounts of Increased "Connectedness" and "Acceptance" After Psilocybin for Treatment-Resistant Depression. *Journal of Humanistic Psychology*, 57(5), 520–564. <https://doi.org/10.1177/0022167817709585>
- Wei, M., Qin, J., Yan, R., Bi, K., Liu, C., Yao, Z., & Lu, Q. (2017). Abnormal dynamic community structure of the salience network in depression: Abnormal Salience Network in Depression. *Journal of Magnetic Resonance Imaging*, 45(4), 1135–1143. <https://doi.org/10.1002/jmri.25429>
- Weissman, M. M. (1986). Understanding the Clinical Heterogeneity of Major Depression Using Family Data. *Archives of General Psychiatry*, 43(5), 430. <https://doi.org/10.1001/archpsyc.1986.01800050028003>
- Wilkinson, P. O., & Goodyer, I. M. (2006). Attention difficulties and mood-related ruminative response style in adolescents with unipolar depression. *Journal of Child Psychology and Psychiatry*, 0(0), 060907041921003-???. <https://doi.org/10.1111/j.1469-7610.2006.01660.x>
- Winkelman, M. (2015). Psychedelics as Medicines for Substance Abuse Rehabilitation: Evaluating Treatments with LSD, Peyote, Ibogaine and Ayahuasca. *Current Drug Abuse Reviews*, 7(2), 101–116. <https://doi.org/10.2174/1874473708666150107120011>
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92, 381–397. <https://doi.org/10.1016/j.neuroimage.2014.01.060>

- Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, 91, 412–419.
<https://doi.org/10.1016/j.neuroimage.2013.12.058>
- Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive Coding of Task-Relevant Information in Human Frontoparietal Cortex. *Journal of Neuroscience*, 31(41), 14592–14599.
<https://doi.org/10.1523/JNEUROSCI.2616-11.2011>
- Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *NeuroImage*, 56(2), 744–752. <https://doi.org/10.1016/j.neuroimage.2010.04.035>
- Woolrich, M. W., Behrens, T. E. J., & Smith, S. M. (2004). Constrained linear basis sets for HRF modelling using Variational Bayes. *NeuroImage*, 21(4), 1748–1761. <https://doi.org/10.1016/j.neuroimage.2003.12.024>
- World Health Organisation. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. World Health Organisation.
- Wylie, G., & Allport, A. (2000). Task switching and the measurement of ‘switch costs’. *Psychological Research*, 63(3–4), 212–233.
<https://doi.org/10.1007/s004269900003>
- Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2004). Don’t think of a white bear: An fMRI investigation of the effects of sequential instructional sets on cortical activity in a task-switching paradigm. *Human Brain Mapping*, 21(4), 279–297.
<https://doi.org/10.1002/hbm.20003>
- Yao, Z., Shi, J., Zhang, Z., Zheng, W., Hu, T., Li, Y., Yu, Y., Zhang, Z., Fu, Y., Zou, Y., Zhang, W., Wu, X., & Hu, B. (2019). Altered dynamic functional connectivity in

- weakly-connected state in major depressive disorder. *Clinical Neurophysiology*, 130(11), 2096–2104.
<https://doi.org/10.1016/j.clinph.2019.08.009>
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–670. <https://doi.org/10.1038/nmeth.1635>
- Ye, M., Yang, T., Qing, P., Lei, X., Qiu, J., & Liu, G. (2015). Changes of Functional Brain Networks in Major Depressive Disorder: A Graph Theoretical Analysis of Resting-State fMRI. *PLOS ONE*, 10(9), e0133775.
<https://doi.org/10.1371/journal.pone.0133775>
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165. <https://doi.org/10.1152/jn.00338.2011>
- Yeung, N., & Monsell, S. (2003a). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 455–469. <https://doi.org/10.1037/0096-1523.29.2.455>
- Yeung, N., & Monsell, S. (2003b). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 29(5), 919–936. <https://doi.org/10.1037/0096-1523.29.5.919>
- Yin, S., Deák, G., & Chen, A. (2018). Coactivation of cognitive control networks during task switching. *Neuropsychology*, 32(1), 31–39.

<https://doi.org/10.1037/neu0000406>

Zeifman, R. J., Wagner, A. C., Watts, R., Kettner, H., Mertens, L. J., & Carhart-Harris, R. L. (2020). Post-Psychedelic Reductions in Experiential Avoidance Are Associated With Decreases in Depression Severity and Suicidal Ideation. *Frontiers in Psychiatry*, 11, 782.

<https://doi.org/10.3389/fpsy.2020.00782>

Zenil, H. (2020). A Review of Methods for Estimating Algorithmic Complexity: Options, Challenges, and New Directions. *Entropy*, 22(6), 612.

<https://doi.org/10.3390/e22060612>

Zenil, H., Hernández-Orozco, S., Kiani, N., Soler-Toscano, F., Rueda-Toicen, A., & Tegnér, J. (2018). A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity. *Entropy*, 20(8), 605. <https://doi.org/10.3390/e20080605>

Zhang, J., Kucyi, A., Raya, J., Nielsen, A. N., Nomi, J. S., Damoiseaux, J. S., Greene, D. J., Horovitz, S. G., Uddin, L. Q., & Whitfield-Gabrieli, S. (2021). What have we really learned from functional connectivity in clinical populations? *NeuroImage*, 118466.

<https://doi.org/10.1016/j.neuroimage.2021.118466>

Zhao, L., Wang, D., Xue, S.-W., Tan, Z., Wang, Y., & Lian, Z. (2021). Aberrant state-related dynamic amplitude of low-frequency fluctuations of the emotion network in major depressive disorder. *Journal of Psychiatric Research*, 133, 23–31. <https://doi.org/10.1016/j.jpsychires.2020.12.003>

Zheng, H., Li, F., Bo, Q., Li, X., Yao, L., Yao, Z., Wang, C., & Wu, X. (2018). The dynamic characteristics of the anterior cingulate cortex in resting-state fMRI

of patients with depression. *Journal of Affective Disorders*, 227, 391–397.

<https://doi.org/10.1016/j.jad.2017.11.026>

Zhou, H.-X., Chen, X., Shen, Y.-Q., Li, L., Chen, N.-X., Zhu, Z.-C., Castellanos, F. X.,

& Yan, C.-G. (2020). Rumination and the default mode network:

Meta-analysis of brain imaging studies and implications for depression.

NeuroImage, 206, 116287.

<https://doi.org/10.1016/j.neuroimage.2019.116287>