2021

Individual variability in value-based decision making: behavior, cognition, and functional brain topography

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

INDIVIDUAL VARIABILITY IN VALUE-BASED DECISION MAKING: BEHAVIOR, COGNITION, AND FUNCTIONAL BRAIN TOPOGRAPHY

by

CLAUDIO A. TORO SEREY

B.S., Northern Kentucky University, 2013
M.A., Boston University, 2019

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2021

Approved by

First Reader _____

Joseph T. McGuire, Ph.D.
Assistant Professor of Psychological and Brain Sciences

Second Reader _____

David Somers, Ph.D.
Professor of Psychological and Brain Sciences

Third Reader _____

Marc Howard, Ph.D.
Professor of Psychological and Brain Sciences

DEDICATION

To my wife Angelina, who has tirelessly encouraged my growth, and my daughter Zoe, who unknowingly guides it.

ACKNOWLEDGMENTS

I am indebted to a large number of individuals who have been fundamental to the completion of this work. First, I want to thank my doctoral advisor, Joseph T. McGuire, for his patient guidance and generosity over the past five years. From the smallest snippet of code to the most challenging aspects of neuroimaging research, his tutelage, trust, and mentorship helped me grow beyond what I believed myself capable of. I would also like to acknowledge David Somers, Marc Howard, and Sam Ling, who filled my graduate experience with invaluable advice, thoughtful academic feedback, and often-necessary cheerful comments.

I would like to acknowledge the great community of graduate students who supported me and helped me succeed in graduate school. Special thanks go to Ilona Bloem, Sean Tobyne, James Brissenden, Allen Chang, Ian Bright, Rachel Wehr, Louis Vinke, and Kathryn Devaney. I would also like to thank present and former members of the McGuire lab for the many discussions that benefited this work, and Leah Bakst in particular for providing unconditional encouragement and advice during our time as office mates.

Above all, I would like to thank friends and family for their perennial love, and for lifting me out of emotional depths that threatened my ability to achieve this goal. Their belief in me has been a fundamental ingredient in everything I have accomplished. I would like to particularly thank my Abuelo, Luis Hernán Serey, for fomenting my desire to understand the mind and the organ from which it emerges. Lastly, I would like to thank my wife Angelina and daughter Zoe, without whom I am nothing.

INDIVIDUAL VARIABILITY IN VALUE-BASED DECISION MAKING: BEHAVIOR, COGNITION, AND FUNCTIONAL BRAIN TOPOGRAPHY

CLAUDIO TORO SEREY

Boston University Graduate School of Arts and Science, 2021

Major Professor: Joseph T. McGuire, Assistant Professor of Psychology

ABSTRACT

Decisions often require weighing the costs and benefits of available prospects. Value-based decision making depends on the coordination of multiple cognitive faculties, making it potentially susceptible to at least two forms of variability. First, there is heterogeneity in brain organization across individuals in areas of association cortex that exhibit decision-related activity. Second, a person's preferences can fluctuate even for repetitive decision scenarios. Using functional magnetic resonance imaging (fMRI) and behavioral experiments in humans, this project explored how these distinct sources of variability impact choice evaluation, localization of valuation in the brain, and the links between valuation and other cognitive phenomena.

Group-level findings suggest that valuation processes share a neural representation with the "default network" (DN) in medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC). Study 1 examined brain network variability in an open dataset of resting-state fMRI (n=100) by quantitatively testing the hypothesis that the spatial layout of the DN is unique to each person. Functional network topography was well-aligned across individuals in PCC, but highly idiosyncratic in mPFC. These results

highlighted that the apparent overlap of cognitive functions in these areas should be evaluated within individuals.

Study 2 examined variability in the integration of rewards with subjective costs of time and effort. Two computerized behavioral experiments (total n=132) tested how accept-or-reject foraging decisions were influenced by demands for physical effort, cognitive effort, and unfilled delay. The results showed that people's willingness to incur the three types of costs differed when they experienced a single type of demand, but gradually converged when all three were interleaved. The results could be accounted for by a computational model in which contextual factors altered the perceived cost of temporal delay.

Finally, Study 3 asked whether the apparent cortical overlap between valuation effects and the DN persisted after accounting for individual variability in brain topography and behavior. Using fMRI scans designed to evoke valuation and DN-like effects (n=18), we reproduced the idiosyncratic network topography from Study 1, and observed valuation-related effects in individually identified DN regions. Collectively, these findings advance our taxonomic understanding of higher-order cognitive processes, suggesting that seemingly dissimilar valuation and DN-related functions engage overlapping cortical mechanisms.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA ……. Analysis of Variance

ARI ……. Adjusted Rand Index

BOLD ……Blood Oxygen Level Dependent

dlPFC ……Dorsolateral Prefrontal Cortex

dmPFC ……..Dorsal Medial Prefrontal Cortex

DN ……..Default Network

EV ……...Expected Value

fMRI ……Functional Magnetic Resonance Imaging

FWHM ……..Full-Width Half-Maximum

GLM ……..General Linear Model

HCP ……Human Connectome Project

MNI ……Montreal Neurological Institute

mPFC ……Medial Prefrontal Cortex

PCC ……Posterior Cingulate Cortex

RPE ……Reward Prediction Error

SP ……Spectral Partitioning

SV ……..Subjective Value

TR ……...Repetition Time

vmPFC ……Ventral Medial Prefrontal Cortex

VS ……..Ventral Striatum

CHAPTER ONE: INTRODUCTION

Preamble

Value-based decision making emerges from the integration of many cognitive faculties spanning attention, perception, and memory. This makes it potentially susceptible to at least two forms of variability both within and across individuals. First, at a neuroscientific level, there is a marked heterogeneity in brain organization across individuals in areas of association cortex that exhibit decision-related activity. Second, at a behavioral level, responses to the same decision-making situation can shift between or within individuals depending on what information is either explicitly highlighted or covertly attended.

Recent evidence suggests that brain networks related to higher order functions can be spatially misaligned across individuals (Braga & Buckner, 2017; Gordon, Laumann, Adeyemo, et al., 2017), which results in obscuring of intricate functional organization during traditional group averaging (Fedorenko, Duncan, & Kanwisher, 2012; Michalka, Kong, Rosen, Shinn-Cunningham, & Somers, 2015; Tobyne, Osher, Michalka, & Somers, 2017). This topographic variability is at least in part distinct from sulcal delineation and other structural features (Conroy, Singer, Guntupalli, Ramadge, & Haxby, 2013), and has been proposed as an important predictor of individual differences in behavior (Bijsterbosch et al., 2018; Kong et al., 2018).

Behavioral variability can stem from the unique perspective people adopt during decision making scenarios. Work in behavioral economics has shown that choices can be molded by manipulating the salience of features in a decision's context (Kahneman &

Tversky, 1984; Prelec & Loewenstein, 1991) or an individual's interpretation of the task (Feher da Silva & Hare, 2020). We can therefore enrich our understanding of subjective valuation by probing the interactions between contextual manipulations and how individuals cognitively construe a decision scenario.

The goal of the present doctoral thesis will be to understand how these distinct types of individual variability (and their interaction) impact our understanding of choice evaluation, localization of value in the brain, and the links between valuation and other cognitive phenomena.

Shortcomings of a Group-based Taxonomy of Neurocognitive Processes

Regions of human medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) have consistently been found to participate in diverse aspects of value-based decision making (Bartra, McGuire, & Kable, 2013; Clithero & Rangel, 2014). However, overlapping swaths of cortex have also been implicated in various cognitive functions such as memory (Schacter, Addis, & Buckner, 2007), self-referential thought (Gusnard, Akbudak, Shulman, & Raichle, 2001), and affective regulation (Reddan, Wager, & Schiller, 2018), speaking against a dedicated role of these areas in valuation processes.

Diverse cognitive functions are often grouped under a common theme in functional brain networks of regions distributed throughout cortex, with each network subserving general aspects of cognition, perception, and behavior (Power et al., 2011; Uddin, Yeo, & Spreng, 2019; Yeo et al., 2011). Among these, the network that covers most of value-related mPFC and PCC is called the default network (DN; Buckner, Andrews-Hanna, & Schacter, 2008; Buckner & DiNicola, 2019). This network was

originally identified as a set of regions that cohesively reduced their activity when attention had to be exteriorly directed in order to confront demanding tasks (Laird et al., 2009; Shulman et al., 1997), but was later observed when individuals were scanned during wakeful resting (Greicius, Krasnow, Reiss, & Menon, 2003). Based on these two observations, researchers have theorized that the DN is in charge of introspective processes.

However, even though many of the aforementioned effects are in concert with this type of introspective dynamics, the connection between these and the increases in activity evoked by subjectively-valuable exterior stimulation is less clear. Fig. 1.1 shows aggregated activity across studies from meta-analyses of subjective valuation (Bartra et al., 2013) and default-related deactivation (Laird et al., 2009), conveying a high degree of overlap between these two phenomena. Indeed, a recent meta-analysis that examined this qualitative juxtaposition failed to dissociate these effects in mPFC and portions of PCC (Acikalin, Gorgolewski, & Poldrack, 2017). The apparent correspondence of these phenomena at the neural level has prompted theories of a common cognitive mechanism that supports all these functions (Euston, Gruber, & McNaughton, 2012; Northoff & Hayes, 2011). This has fueled interest in the degree to which task-related effects truly overlap spatially with the DN (DiNicola, Braga, & Buckner, 2020).

However, conclusions from group-level functional colocalization are limited by our ability to align brain activity patterns across individuals (Guntupalli, Feilong, & Haxby, 2018). Individual-specific studies of brain organization have qualitatively described spatially variable organization of functional networks that encompass both

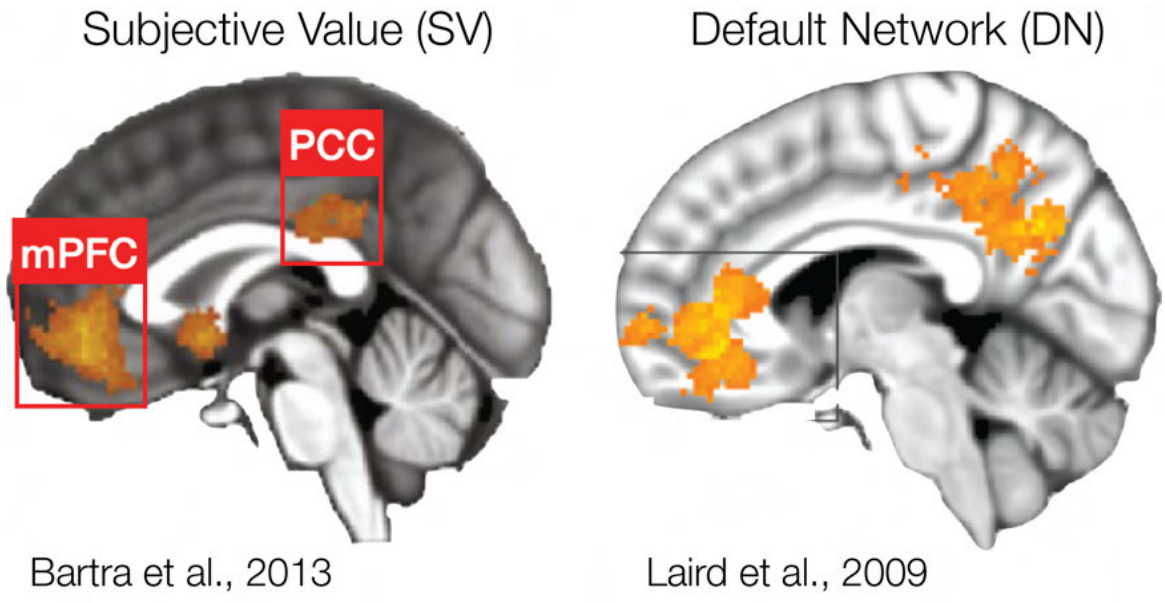

Figure 1.1: Meta-analytic results aggregating studies of subjective valuation and the default network. On the left are regions that increase their activity with obtained reward. The right shows areas that decreased their activity with increased demands across studies. These group-level analyses suggest an apparent overlap of these phenomena in medial prefrontal (mPFC) and posterior cingulate (PCC) cortices.

mPFC and PCC (Gordon, Laumann, Adeyemo, et al., 2017; Laumann et al., 2015). In fact, careful examination of canonical DN in individuals has shown that the core regions of the network—including mPFC and PCC--can be sub-divided into two interwoven networks whose layouts differ across people (Braga & Buckner, 2017; Braga, Van Dijk, Polimeni, Eldaief, & Buckner, 2019), and whose functional roles can be dissociated (DiNicola, Braga, & Buckner, 2020). The problem is accentuated in ventral mPFC, which displays considerable idiosyncratic cortical folding (Lopez-Persem, Verhagen, Amiez, Petrides, & Sallet, 2019; Mackey & Petrides, 2014). It is therefore possible that the apparent meta-analytic overlap between the DN and valuation is due to a lack of

consideration for this variability in mPFC and PCC, and might be better examined at the individual level.

Variability in Choice Behavior

The traditional approach to mapping function to brain structures using functional magnetic resonance imaging (fMRI) often assumes that contrasts among conditions remain consistent throughout an experimental session (Huettel, Song, & McCarthy, 2014). However, people often change their attitudes towards repetitive choices. This is often observed in people's willingness to exert effort in everyday situations. For example, we might choose to take the effortful route home that we normally avoid rather than waiting for a delayed train, even if both avenues take the same amount of time. In neuroeconomics, demands for both time and effort have traditionally been identified as costs that detract from the net subjective value of rewards (Westbrook & Braver, 2015). Indeed, decision makers tend to prefer low-effort and immediate rewards over larger rewards that demand longer delays and higher effort (Ainslie, 1975; Hull, 1943; Kool & Botvinick, 2018; Kool, McGuire, Rosen, & Botvinick, 2010; Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009; Westbrook, Kester, & Braver, 2013). However, understanding the sporadic attractiveness of effort exemplified above has been elusive, and presents an important research avenue (Inzlicht, Shenhav, & Olivola, 2017).

Recent theoretical and experimental accounts have suggested that the perceived subjective costs of effort and delay may stem from missing out on concurrent possibilities, meaning that decision makers track the opportunity cost of time when evaluating whether to pursue costly behavioral avenues (Kurzban, Duckworth, Kable, &

Myers, 2013; Otto & Daw, 2019). Flexible preferences for effort might therefore arise from the availability of other types of demands in the same choice environment (Kool & Botvinick, 2018). Traditional accounts in behavioral economics and neuroeconomics are limited in addressing this possibility, as measuring the cost of effort and delay has often been done by offering higher payouts for an increment of the respective demand (Westbrook & Braver, 2015). Single choice foraging paradigms provide an ecological solution to this problem, by asking individuals whether the current offer outweighs what could be obtained and experienced during that time instead (Mobbs, Trimmer, Blumstein, & Dayan, 2018). The sequential nature of these tasks can help us capture an important source of variability, namely that choices to similar or repetitive scenarios can fluctuate depending on what features of a decision are explicitly highlighted (Kahneman & Tversky, 1984) or covertly construed (Feher da Silva & Hare, 2020), by encouraging individuals to emphasize different elements of the decision context (such as the availability of alternative ways to achieve rewards in the environment, or the characteristics of the present offer). The rich modeling tools afforded to foraging paradigms can help us computationally identify such emphases (Stephens & Krebs, 1986).

Understanding the degree of choice variability in this type of ecological scenario, as well as the underlying reason for such dynamics (i.e. what individuals might covertly attend to), can improve our ability to control for such fluctuations when mapping valuation processes to the brain.

Interactions Between Topographic and Behavioral Variability

Individual differences in fine-grained functional brain organization can complicate the mapping of task effects. Careful experimentation in individuals has shown that canonical Broca's area is adjacent to a domain-general region that can confound group-level neural localization of language with task difficulty (Fedorenko, Duncan, & Kanwisher, 2013). Individualized investigations have also revealed that attention-related areas of lateral PFC have interleaved sections that are biased for visual and auditory information, a spatial pattern that is sensitive to the spatial or temporal emphasis of a short-term memory task (Michalka et al., 2015; Noyce, Cestero, Michalka, Shinn-Cunningham, & Somers, 2017; Tobyne et al., 2017). This susceptibility of functional topography to the nature of the task has been further demonstrated in recent work by Salehi et al. (2019), who suggested that functional regions can spatially reconfigure within individuals depending on the cognitive perspective induced by different tasks--even though the activity pattern associated with each task is reliable over time. In the domain of value-based decision making, different sub-regions of mPFC appear to evaluate outcomes that are truly desirable to an individual versus those that are only selected because of their congruity with the task's goals (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Shenhav & Karmarkar, 2019). Moreover, the variability of network layouts across people has been linked to differences in trait-like psychological characteristics (Bijsterbosch et al., 2018; Kong et al., 2018).

These findings imply that an individual's cognitive construal of a task can interact with idiosyncratic brain topography to produce variability in functional brain mapping

results. Even for people with similar functional topography, or within the same individual over time, activity patterns in valuation cortex could differ if individuals covertly engage with (or attend to) decision making tasks differently. For instance, some individuals might make decisions according to a decision rule rather than outcome value, which could potentially free up cognitive resources and allow them to introspect. Shifts between these perspectives could also occur within individuals as they gain experience with the task. These distinct cognitive approaches to the same task could potentially recruit different functional networks that are identifiable within individuals, but are otherwise obscured by group averaging. Therefore, it is important that we account for these behavioral and neural sources of variability before we conclude that valuation is among the many processes ascribed to the DN.

Organization of Dissertation

Subsequent sections of this dissertation are organized into four chapters. In Chapters 2-4, I use a combination of neuroimaging and behavioral studies in humans to examine both cortical topographic variability and behavioral variability in value-based decision making. Both types of variability have implications for assessing the overlap between default and valuation networks. The final chapter summarizes the research findings, and presents possible avenues for future study.

Chapter 2 takes a first quantitative step in addressing the present issue by testing the hypothesis that the topography of the DN in the vicinity of valuation cortex is heterogeneous across individuals. Using fMRI, this work first establishes that a gradient-based network clustering algorithm, spectral partitioning (SP), can efficiently capture an

individual's functional network topography. Using this method, it shows that that the functional topography of mPFC is composed of at least two sub-networks, one of which appears to correspond to the DN by means of a meta-analytic anchor. Further, even though the spatial organization of these sub-networks shares common organizational principles across individuals (i.e. interdigitated along a dorsal-ventral axis), their precise layout is heterogeneous across individuals (particularly in mPFC), but reliable across an individual's scanning sessions and days. The refined estimate of the DN can be consistently mapped to sulci within ventral mPFC (vmPFC), but not so in PCC and dorsal mPFC. Finally, the resulting layouts are validated against a recently proposed subdivision of DN that has fruitfully distinguished among cognitive phenomena (Braga & Buckner, 2017; DiNicola et al., 2020).

Chapter 3 focuses on possible choice variability either within or across individuals. This study tests the hypothesis that preferences should vary depending on whether individuals face a single type or multiple types of behavioral demands, and probes what computational factors drive these effects. Based on two foraging experiments and computational modeling, it first shows that preferences for mental effort are stable when individuals face a single form of demand, but fade over time when other types of costs are also experienced in the same environment. Second, based on a combination of traditional foraging and reinforcement learning models, it suggests that these dynamics are not due to people's inability to estimate the value of each option relative to its environment, but instead arise because each demand biases the estimate of the opportunity cost of time. I propose that this bias arises from the way each task

modulates the subjective passage of time, which could hypothetically be recalibrated when multiple tasks are interleaved.

Chapter 4 combines the knowledge gained from previous chapters, and tests the hypothesis that the valuation network can be spatially distinguished from the DN in mPFC, PCC, and striatum when accounting for sources of individual variability. Using fMRI, it first introduces evidence that the same idiosyncratic subdivision of the DN observed in Chapter 2 also subserves working memory and decision-making tasks. Second, and contrary to expectations, it finds that the spatially refined DN within these partitionings facilitates multiple types of valuation processes. Third, it suggests that striatum cannot be discretely paired to each DN sub-network, in line with anatomically-based theories that this subcortical region instead mirrors a dorsal-ventral axis of prefrontal cortex (Haber & Knutson, 2010). Finally, it shows that this alignment was not modulated by people's degree of engagement with the decision-making task over time.

CHAPTER TWO: SPECTRAL PARTITIONING IDENTIFIES INDIVIDUAL HETEROGENEITY IN THE FUNCTIONAL NETWORK TOPOGRAPHY OF VENTRAL AND ANTERIOR MEDIAL PREFRONTAL CORTEX [1]

Introduction

Human medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) are jointly associated with a diverse set of cognitive processes (Hiser & Koenigs, 2018; Kragel et al., 2018), and contain subregions that are part of the brain's default network (DN; Buckner, Andrews-Hanna, & Schacter, 2008; Buckner & DiNicola, 2019). DN regions are characterized by a decrease in BOLD activity during externally oriented tasks that require attention or cognitive control, in comparison with less-demanding task conditions or periods of rest (Buckner et al., 2008; Laird et al., 2009; McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003). The DN can also be identified on the basis of a distinctive pattern of inter-region correlations in resting-state fMRI data (Buckner et al., 2008; Fox et al., 2005; Greicius et al., 2003; Yeo et al., 2011).

Many different cognitive task manipulations evoke patterns of brain activity that overlap with DN regions in ventral and anterior mPFC and in PCC. Examples include manipulations of self-referential thinking (Gusnard et al., 2001; Mitchell, Banaji, & Macrae, 2005), memory (Euston et al., 2012; Schacter et al., 2007), affective regulation (Reddan et al., 2018; Schiller, Levy, Niv, LeDoux, & Phelps, 2008), and subjective valuation (Bartra et al., 2013; Clithero & Rangel, 2014; Kable & Glimcher, 2007; Levy,

[1] This work has been previously published as Toro-Serey, C., Tobyne, S.M., & McGuire, J.T. (2020) Spectral partitioning identifies individual heterogeneity in the functional network topography of ventral and anterior medial prefrontal cortex. *NeuroImage*, 205

Lazzaro, Rutledge, & Glimcher, 2011). Some of these task-related effects are thought to reflect processes integral to the functional role of the DN, such as internally oriented cognition, scene construction, and self-projection (Buckner & Carroll, 2007; Hassabis & Maguire, 2007). For other task-related effects, such as subjective valuation (i.e., greater BOLD activity in response to more highly valued choice prospects and outcomes, relative to prospects and outcomes that are less highly valued), the degree of overlap with DN regions is only partial and the reason for the overlap is less obvious (Acikalin et al., 2017). Insofar as the DN shares a subset of nodes in common with the distributed brain systems that support valuation and other functions, this has potential to inform our theoretical understanding of the cognitive operations involved in those functions (Northoff & Hayes, 2011). As a result, there is widespread interest in understanding the degree to which DN regions overlap topographically with task-related effects (Buckner & DiNicola, 2019; DiNicola, Braga, & Buckner, 2020; Spreng, 2012).

However, strong conclusions about functional colocalization require consideration of individual-level heterogeneity in topographic patterns of brain activity. A recognized limitation of group averaging and meta-analysis is that the functional topography of individual brains can be misaligned and blurred (Fedorenko et al., 2012; Guntupalli et al., 2018; Michalka et al., 2015; Tobyne, Sean M. et al., 2018; Wang et al., 2015; Woo et al., 2014), exaggerating the apparent overlap across domains. This concern is especially pronounced in ventral mPFC, which is subject to considerable idiosyncratic cortical folding (Lopez-Persem et al., 2019; Mackey & Petrides, 2014; Zilles, Palomero-Gallagher, & Amunts, 2013) and inter-subject functional variability (Mueller et al.,

2013). An alternative approach is to focus on analyses at the individual-participant level. Individual-level analyses of fMRI data have identified idiosyncratic, reliable, and valid patterns of functional organization that would be blurred in aggregative estimates (Gordon, Laumann, Gilmore, et al., 2017; Gratton et al., 2018; Laumann et al., 2015; Michalka et al., 2015; Tobyne, Sean M. et al., 2018), and subject-specific network arrangements have been found to predict behavioral characteristics (Kong et al., 2018). Recent work has uncovered fine-grained subdivisions within the DN using both data-driven clustering and individually customized seed-based connectivity analysis (Braga & Buckner, 2017; Braga et al., 2019). It is therefore possible that the apparent overlap of the DN with task-related effects might, in some cases, be attributable to low effective spatial resolution, and that the organization of mPFC and PCC might be better understood at the individual level. An important first step in investigating this possibility, and the goal of the present study, is to quantify the degree of variability in the topography of the DN within mPFC and PCC across a large sample of individuals.

A useful way to characterize individual-specific brain organization is to examine patterns of resting-state functional connectivity. Connectome-based analyses of resting-state functional connectivity have been fruitful in identifying individualized functional subregions that correspond well to task-induced activity patterns (Gordon, Laumann, Gilmore, et al., 2017; Laumann et al., 2015; Smith et al., 2009; Tobyne et al., 2018). A functional connectome can be represented in the form of a network, and graph theoretic methods can be applied to analyze the network's structure (Bassett & Sporns, 2017; Rubinov & Sporns, 2010). In the context of network analysis, community detection

algorithms subdivide brain networks into sets of nodes that share more connections with each other than with the rest of the network (Fortunato & Hric, 2016; Garcia, Ashourvan, Muldoon, Vettel, & Bassett, 2018). Here we use the technique of spectral partitioning (SP), an efficient community detection algorithm that deterministically subdivides a network into two communities (Belkin & Niyogi, 2003; Chung, 1997; Fiedler, 1975). SP has previously been used to characterize the posterior-anterior functional gradient of the insula using resting-state fMRI data (Tian & Zalesky, 2018), and was shown to robustly and reliably separate both simulated and actual primate ECoG networks (Toker & Sommer, 2019). We use SP here to identify subsets of nodes within mPFC and PCC that share spontaneously covarying temporal activation patterns during rest.

Here, we aimed to subdivide mPFC and PCC into individual- specific DN and non-DN communities, and to quantify the degree of topographic heterogeneity in the resulting community structure over time and across individuals. We did this by capitalizing on the respective strengths of meta-analysis and subject-specific analyses of brain networks. We used a data-driven network-analysis procedure to identify two communities that each spanned both mPFC and PCC in each individual participant. We found that the resulting communities had a stereotyped topographic layout within PCC (according to a label-agnostic similarity metric), whereas their layout in mPFC was variable across individuals but stable across test/re-test. We took advantage of the more consistent configuration within PCC to assign meta-analysis-derived labels to the two communities. Because our data-driven method established correspondence between PCC

subregions and mPFC subregions, the labels defined in PCC could then be indexed into the more heterogeneous community structure of mPFC in each individual.

The outline of this chapter is as follows. First, we defined a search space by selecting parcels from an established brain atlas (Glasser et al., 2016) that corresponded to previously defined DN and limbic networks on the medial cortical wall (Yeo et al., 2011). A cortical surface-based meta-analysis of the DN and valuation literatures identified a parcel in PCC that was DN-specific at the aggregate level. Valuation was selected as an example of a cognitive domain in which group-average activity patterns overlap extensively with DN regions on the medial surface, despite being segregable elsewhere (Acikalin et al., 2017). We then derived a functional connectivity network of all the surface vertices within the search space for each of 100 individual resting-state fMRI data sets from the Human Connectome Project (HCP; Van Essen et al., 2012), and used the SP algorithm to subdivide each individual's network into DN and non-DN communities (labeled according to which community included the meta-analytically identified DN-specific parcel in PCC). Focusing on individual vertices in the search space rather than the parcels (as is typical in brain network analyses) allowed us to finely delineate the topographic extent of each community. The resulting communities varied topographically across individuals, while also appearing to follow common organizational principles. Test-retest analyses showed that these partitionings were similar across scanning days within (but not between) individuals, and that individual-level idiosyncrasy was greater in mPFC. Partitionings obtained from the SP algorithm had higher test-retest reliability than did analogous results from seed-based functional

connectivity. We observed a trend for the DN community to be located within principal sulci in ventral mPFC and left PCC, but in gyri within superior mPFC and right PCC. Lastly, we describe how the structure of the resulting automatically defined DN and non-DN communities both aligns with and differs from a recently proposed scheme for identifying subdivisions within the DN (Braga & Buckner, 2017; Braga et al., 2019). Our work highlights the usefulness of estimating brain effects at the individual level in mPFC and PCC, and provides a new framework and tool set for future investigations of overlap across cognitive domains.

Methods

Search space

For all analyses, we defined our search space based on the 17-network parcellation proposed by (Yeo et al., 2011). First, we selected vertices on the medial cortical surface that were contained by the DN and limbic networks in HCP's 32,000 vertex surface space (fs_LR_32k). Next, we overlaid those networks on a parcellated atlas of the human cortical surface (360 regions; Glasser et al., 2016), and retained a set of parcels that covered approximately the same brain regions (visually inspected, retaining parcels that appeared to have at least 15% overlap). This resulted in a search space that consisted of 40 parcels across hemispheres (Table 2.1). The search space in each hemisphere was naturally divided into two spatially non-contiguous clusters in PCC and mPFC, facilitating the examination of each region separately.

Meta-analysis

We used a novel approach to cortical surface parcel-based meta-analysis to assess whether individual parcels within the search space were preferentially associated with subjective valuation or with decreased activity during externally oriented tasks, which served to operationalize the DN. For subjective valuation, we gathered peak activation coordinates from 200 studies that reported positive effects in contrasts of higher-value minus lower-value outcomes or prospects (Bartra et al., 2013). For the DN, we acquired coordinates from 80 studies that reported reductions in BOLD during externally directed tasks compared to a baseline (Laird et al., 2009). The coordinates represent areas that exceeded the statistical significance threshold in each original study. For each study, we created an indicator map in standard volumetric space (MNI152, 1 mm resolution) which contained values of 1 in a 10 mm radius sphere around each reported activation peak, and values of 0 elsewhere (Wager, Lindquist, Nichols, Kober, & Van Snellenberg, 2009). The indicator map for each study was then projected to a standard cortical mesh (fsaverage, 160,000 vertices, projfrac-max from 0 to 1 by 0.25, registered using mni152.register.dat) using FreeSurfer's mri_vol2surf (Dale, Fischl, & Sereno, 1999a; Fischl, Sereno, & Dale, 1999) (http://surfer.nmr.mgh.harvard.edu/). We then resampled the Glasser et al. (2016) parcellation to fsaverage, and tallied how many studies had positive indicator values intersecting with each cortical parcel (the details of the resampling procedure are described in https://wiki.humanconnectome.org/display/PublicData/HCPþUsersþFAQ #HCPUsersFAQ-9.HowdoImapdatabetweenFreeSurferandHCP, and were implemented using a custom script available at https://github.com/stobyne/Spherical-Surface-

Swapper). Two studies from the subjective valuation corpus were removed because they did not contain activation peaks that overlapped with cortex, leaving a final number of 198 studies.

To test for parcels that were significantly more strongly associated with one domain than the other, we performed per-parcel chi-squared tests comparing the proportion of studies with activation in that parcel between the two domains. We permuted the study domain labels (DN or valuation) 5000 times while preserving the total number of studies in each domain, and on each iteration stored the maximum resulting chi- squared statistic across all parcels. This gave us a null distribution of 5000 maximum chi-squared values. The 95th percentile of this distribution served as an FWE-corrected significance threshold to evaluate unpermuted chi-squared values.

Resting-state fMRI data

Our fMRI analyses used resting-state fMRI data from the Human Connectome Project (Van Essen et al., 2012) Q6 release (N =100, randomly sampled from the total pool of 469 available subjects). The Washington University Institutional Review Board approved all experimental procedures, and all subjects provided written informed consent in accordance with the guidelines set by the institution. Each subject′s data was acquired over two days at Washington University in St. Louis on a Siemens CONNECTOM Skyra MRI scanner (Siemens, Erlangen, Germany). Four resting state runs (repetition time = 0.720 s, echo time = 33.1 ms, flip angle = 52º, multiband factor =8, 72 slices, 2 mm isotropic voxels) each comprised 1200 time points (14 min 24 s) for a total of 4800 time points. Two runs were acquired on each day, with the phase encoding direction set to left-

right for one run and right-left for the other. Only subjects with both left-right and right-left phase encoding for each day were included (i.e. subjects with four resting-state fMRI sessions). In addition, only datasets with low motion levels (under 1.5 mm) and fewer than 5% of points over 0.5 mm framewise displacement (FD; Power et al., 2014) were used. See (Van Essen et al., 2012) for more details about the data acquisition protocol.

Data initially underwent the HCP minimal preprocessing pipeline (Glasser et al., 2013), which included gradient nonlinearity correction, motion correction, EPI distortion correction, high-pass filtering (0.0005 Hz threshold), MNI152-based normalization, surface reconstruction, and mapping of functional data to a standardized cortical surface model (details can be found in Glasser et al., 2013). In addition, data underwent temporal denoising based on independent components (FMRIB's ICA-based X-noiseifier, FIX; Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Data were further preprocessed using an in-house pipeline described previously (Tobyne et al., 2017). Steps (in order) included linear interpolation across high motion timepoints with over 0.5mm of FD, band-pass filtering (allowed frequencies ranged from 0.009 to 0.08 Hz), and temporal denoising via mean grayordinate signal regression (Burgess et al., 2016). Interpolation of high-motion time points was performed to avoid temporal smoothing of noisy signal from head motion into the filtered signal during the bandpass procedure. After filtering and denoising, the interpolated high-motion time points were censored by deletion and each run was temporally demeaned. The processed time series had a median of 4799 time points (minimum = 4661) across participants. Each subject's brain was comprised of 32k

standard grayordinates per hemisphere (combined in a CIFTI file). We retained only the cortical surfaces, which resulted in 59,412 total surface vertices per subject.

Network definition

All network analyses were performed using the igraph package (v. 1.1.2; https://igraph.org/r/; Csardi & Nepusz, 2006) in R (v. 3.4.1; https://www.r-project.org/; R Core Computing Team, 2017). To establish each subject's network, we selected all the vertices contained within the mPFC/PCC search space (n = 4801 per subject; mPFC = 2854, PCC = 1947) and computed the Pearson correlation of the time series for every pair of vertices. All correlation values were transformed using Fisher's r to z. This produced a weighted network for each subject, in which the nodes were surface vertices and the edge weights were the correlations among them. Edges mostly consisted of positive correlations (mean proportion positive = 0.65, SD = 0.03). We chose not to threshold the network, as the SP algorithm is well equipped to operate on complete (i.e. fully-connected) weighted graphs (Chung, 1997). However, our results were unchanged if we retained only significant correlations ($p < 0.05$, uncorrected) in the weight matrices. Next, we took the exponential of the z-transformed correlations so that all weights became positive while maintaining their ordinal ranks. Ensuring that all edges were positive facilitated the construction of the graph Laplacian (see below), which requires all off-diagonal elements to have the same sign by design. We generated and analyzed network weight matrices at four levels: (1) for each subject's full concatenated dataset (up to 4800 TRs); (2) on each step of a sliding window analysis (see Section 2.7 for more details); (3)

for the concatenated time series for the two runs on each day (up to 2400 TRs); and (4) for each run separately (up to 1200 TRs).

Community detection

Communities (i.e. clusters) were identified using the SP algorithm (Belkin & Niyogi, 2003; Chung, 1997; Fiedler, 1975; Higham, Kalna, & Kibble, 2007). First, each network was represented as an n x n network weight matrix *W* as described above (where n equals the number of vertices in the search space, 4,801). The matrix was then transformed into its symmetric normalized Laplacian form

$$L = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$$

where *I* is an identity matrix of size n, and *D* is a diagonal matrix containing the strength of each vertex (i.e. the sum of its edge weights with all other vertices). This resulted in a matrix wherein each entry was the negative normalized value of the connection (from 0 to 1) between any two vertices relative to their combined connectivity strength, and with ones along the diagonal. The transformation ensures that every row sums to approximately zero. We then computed the eigenvalues and eigenvectors of the symmetric normalized Laplacian matrix, and used the eigenvector associated with the second-to-lowest eigenvalue (traditionally called the 'Fiedler vector') to divide the network into two. The Fiedler vector consists of a set of positive and negative values and is binarized by sign to partition the network into two similar-sized communities (Fiedler, 1975). In this way, SP avoids producing communities that are too small to be physiologically meaningful (for example, small sets of vertices that are spuriously

correlated due to measurement noise). Given that this data-driven method does not label the two communities or establish correspondence across participants, we defined each individual's "DN" community as that which contained the majority of the vertices in the DN-specific PCC parcel identified in our meta-analysis (area 7 m). The completeness of the graphs ensured that SP did not face the issues associated with its use in sparse networks (Fortunato and Hric, 2016).

In order to evaluate the validity of the resulting partitionings across community-detection methods, we also estimated network communities using the more traditional approach of modularity maximization (Garcia et al., 2018), based on the algorithm from Clauset et al. (2004). The method heuristically iterates through many possible combinations of vertices, and selects the partitioning that maximizes the within-community edge weights, relative to a random network containing the same number of edges and communities. Unlike SP, modularity can fractionate a network into more than two communities. Agreement between the partitions provided by the bounded (SP) and unbounded (modularity) community detection methods would suggest the results are not distorted by the restriction of SP to binary partitionings.

Partition evaluation

We used the Adjusted Rand index (ARI) to evaluate the stability and topographical heterogeneity of the communities within and across individuals (Hubert & Arabie, 1985), which was calculated using the "mcclust" package in R (Fritsch, 2012). The ARI is a metric that quantifies the similarity between two alternative clusterings of the same data. The base of the ARI is computed by the formula

$$\frac{a+b}{a+b+c+d}$$

where a is the number of pairs of nodes that were grouped together in both partitionings, b is the number that were grouped separately, and c and d denote the number of pairs grouped together (separately) in one partitioning, but separately (together) in the other. Therefore, the ARI estimates the fraction of all possible node pairs that had the same status (connected or not) in both partitionings (with the denominator equal to n(n – 1)/2). The resulting ratio is adjusted against a baseline given by the expectation assuming independent partitionings to yield an index that ranges from 0 to 1, where 0 denotes the value expected by chance. This means that even though differences are heavily penalized, positive ARI values compare favorably against chance clustering (and the index can take negative values if the ratio given by the formula above falls below the chance level). In short, the ARI quantifies the chance-corrected agreement between any two partitions while being agnostic to the labeling scheme.

We performed a number of comparisons among partitions. First, we computed the degree of agreement between SP and modularity maximization per subject. SP and modularity maximization have been previously found to show a tendency toward underfitting and overfitting, respectively, in their community detection performance in a diverse set of network types (Ghasemian, Hosseinmardi, & Clauset, 2019), so alignment between the two algorithms would increase our confidence in the validity of the resulting partitionings. Next, we compared the subject-level SP partitionings across individuals, and calculated the mean pairwise ARI for the group. We then performed the same evaluation for PCC and mPFC separately, and examined whether there were differences

in overall agreement within these regions by performing a paired permutation analysis. For each individual and region, we took the mean ARI with all 99 other individuals, then took the difference between regions to get an ARI difference per subject. On each of 5000 permutations each subject's ARI difference was independently sign-flipped and the group mean difference was added to a null distribution. The unpermuted group mean difference was then evaluated against this permuted distribution.

To identify vertices whose community assignment was more stable or more variable, we performed a sliding window analysis (20 min windows, 1 min increments, median number of windows per subject = 37, range = 35 37), comparing each window's resulting partitioning against the partitioning derived from the subject's whole data set. A 20-min window has previously been found to yield relatively stable and unbiased estimates of individual-level brain network characteristics (Gordon et al., 2017). We assessed whether the magnitude of the Fiedler vector value for a given vertex (for the full subject-level data set) was associated with the stability of that vertex's sub-network assignment across time windows. To do this, we fit a mixed effects logistic regression model, in which the dependent variable was the proportion of times each vertex participated in the DN community across windows, and the explanatory variables included a random effect of subject and a fixed effect of the Fiedler vector value for that vertex (derived from their full time series). Based on this significant relationship, we identified a threshold Fiedler vector value for each subject, such that empirical above-threshold vertices were persistently associated with either DN or non-DN more than 99% of the time.

We then estimated the level of agreement between network partitions estimated using data across individual scan days (with 2 days per participant). If the functional organization estimated by SP is indeed individual-specific, we should see higher agreement within individual (test/re-test across days) than across individuals. We tested this idea by computing the ratio of the mean ARI within and between individuals. Ratios close to one would denote similar within-participant and across- participant alignment, whereas ratios considerably higher than one would suggest that partitions were more similar within-participant than across participants. We then extended this idea by computing the agreement across individual runs (4 per subject). Similar to the day-based analysis, we assessed whether run-level data showed higher agreement within-subject than between subjects.

Seed-based resting-state functional connectivity versus community detection

We evaluated the performance of the SP algorithm in comparison to a simpler partitioning approach based on seed-based functional connectivity. Independently for each day (2 per individual), we estimated each subject′s DN partition in mPFC based on its vertex-wise functional correlations (Pearson) with the spatially averaged activity across all vertices in the PCC search space. We used the whole PCC region because it is traditionally thought to be a prominent node of the DN (Buckner et al., 2008), and is a common area for researchers to place seeds for vertex- and volume-based connectivity analyses (e.g. Fox et al., 2005). We compared these seed-based maps with the unthresholded Fiedler vectors produced by SP, with the sign of the Fiedler vector oriented so the DN community was marked by positive values in every subject. We

calculated three sets of across-day similarity values for each individual: 1) between the two seed-based maps; 2) between the two SP-based maps; and 3) between seed- and SP-based maps. Because the values in the maps were continuous-valued (and not categorical labels, which would be amenable to ARI), we quantified the similarity between maps in terms of the spatial Spearman correlation across vertices. These spatial correlations were meant to determine the test/re-test reliability of each approach, as well as the overall level of agreement between them. For 8 subjects, the communities produced with one of the days' data sets had split coverage of area 7 m, and our community labeling scheme for the Fiedler vector produced a sign mismatch across days. ARI is robust to such labeling issues, but the inconsistency produced strong negative correlations of the Fiedler vector across days for these individuals. Visual inspection showed that the community layout was well aligned across days, and so we matched the labeling of their partitionings based on the day that sufficiently covered area 7 m.

The two methods were expected to produce somewhat similar results, but the one displaying greater within-subject agreement across days should be preferred (for a discussion on the stability of functional networks see Kong et al. (2018) and Gratton et al. (2018)). We therefore compared the within-subject spatial correlation coefficients produced by each method through a paired permutation analysis. For each of the 100 individuals, we computed the difference in inter-day correlations between methods, randomized the sign of these values 5000 times, and computed the mean of these differences on each iteration. The empirical difference in means was then evaluated against this permuted distribution.

Associations with sulcal morphology

Next we asked whether the location of the DN and non-DN communities was systematically related to sulcal morphology. Based on a previous report of individual alignment of DN within sulci in ventral mPFC (vmPFC; Lopez-Persem et al., 2019), we subdivided our search space into three regions: vmPFC, which matched the ROI used by Lopez-Persem and colleagues (2019; areas 25, s32, a24, 10v, 10r, p32, and OFC); superior mPFC (sup-mPFC), encompassing the remaining dorsal areas in our mPFC space; and all of the PCC search space. We used each subject's curvature maps from the HCP (transformed to fs_LR 32k space), in which cortical depth is quantified by negative numbers for sulci and positive numbers for gyri. For each individual, we computed the Spearman correlation between curvature and the unthresholded Fiedler vector values in each region and hemisphere separately (6 correlations per individual). Since the DN was indicated by positive FV values, negative correlations meant that the DN was more likely to be contained in sulci, with non-DN located in gyri. Finally, we collected all individual correlations for each combination of hemisphere and region, and performed a one-sample t-test on each set to determine whether correlations were significantly different from 0 in our group (6 tests total).

Results

Meta-analysis

We performed a coordinate-based meta-analysis to identify cortical surface parcels within mPFC and PCC that were preferentially associated with the DN or with subjective valuation. Volumetric coordinates from 80 studies with task deactivation

contrasts and 198 studies with valuation contrasts were projected onto a cortical surface, and mapped to discrete parcels from a multimodal cortical parcellation (Glasser et al., 2016) to produce a list of brain areas reported per study. The 40 parcels considered were limited to the medial portion of the default and limbic networks defined by the Yeo et al. (2011) 17-network parcellation. Domain-specificity was tested by first permuting the domain labels across studies (DN or valuation) to create a null distribution for the maximum chi-squared statistic in the search space (see Methods for details). The null distribution was used to identify regions that were reported significantly more often in one literature or the other.

Fig. 2.1 shows the proportion of times each parcel was reported for each domain, as well as the significant differences between domains. The 95th percentile of the permuted chi-squared distribution was 8.87. Based on this threshold, area 7 m in PCC/precuneus was the only parcel to show a preferential association with the DN bilaterally (Left: observed $\chi^2 = 10.07$, $p = 0.029$; Right: observed $\chi^2 = 18.89$, $p < 0.001$). The adjacent area v23 exhibited a similar effect, albeit only unilaterally (Right: observed $\chi^2 = 11.51$, $p = 0.011$; Left: observed $\chi^2 = 8.25$, $p = 0.067$). There appeared to be a bilateral preference toward valuation effects in mPFC area 25 (Left: observed $\chi^2 = 12.91$, $p = 0.005$; Right: observed $\chi^2 = 12.83$, $p = 0.005$); however, closer inspection suggested this effect was driven by subcortical foci centered in adjacent ventral striatum. No other parcels were preferentially implicated in valuation relative to DN. We therefore selected area 7 m as an interpretable, bilateral reference point for labeling DN and non-DN communities in the analyses that follow. We note that the area labeled 7 m in the parcellation used here

(Glasser et al., 2016) is different from (and located inferiorly on the medial surface to) the non-DN area 7 m discussed in previous work (Andrews-Hanna, Reidler, Huang, & Buckner, 2010).

Individual-level DN and non-DN communities

Within the mPFC/PCC search space, we estimated the topography of the DN for each individual. Using each individual's full time series (approximately 4800 total TRs from four 14-min scanning runs acquired over two days), we calculated the full vertex-to-vertex correlation matrix for the 4801 surface vertices in the search space. We represented each individual's correlation matrix in the form of a network, with cortical surface vertices as nodes and transformed correlation values as edge weights. We then applied the SP community detection algorithm to partition the network into two cohesive functional communities.

Fig. 2.2 shows a representative partitioning of the search space for a single participant (100307), with additional examples presented in the first two columns of Fig. 2.3. The SP algorithm subdivides a network according to the positive versus negative values in the Fiedler vector (the eigenvector related to the second-to-lowest eigenvalue of the network's normalized Laplacian matrix, see Methods). Since this is a data-driven approach, there is no a priori labeling for the two communities. We assigned the DN label to the community that contained the majority of the DN-specific PCC parcel from the meta-analysis (7 m). We oriented each individual's Fiedler vector so positive values corresponded to the DN community (K. H. Nenning et al., 2017), and were assigned a value of 1 in the binarized partitionings (with 0 denoting non-DN). In qualitative terms,

the resulting patterns contained substantial DN coverage in posterior PCC (as dictated by our labeling strategy), with non-DN vertices in anterior PCC. The mPFC region tended to include DN vertices in its ventral-anterior and dorsal-anterior areas, with a persistent non-DN pocket between them. This non-DN section extended posteriorly into pregenual cingulate cortex (area a24). We note that the addition of restrosplenial cortex (an area commonly regarded as part of canonical DN) to the search space did not change these results; as expected, that area tended to be largely assigned to the DN community (Fig. 2.3).

Before evaluating the degree of generalizibility of this topographic pattern across individuals, we examined the validity of the partitionings by comparing them to results from an alternative community detection algorithm, modularity maximization (Clauset et al., 2004). Modularity seeks to find the set of communities that maximizes within-community connection weights relative to a null model. Since modularity is not constrained to a predetermined number of communities, it was capable of finding more than two in our data set. We quantified the cross-method agreement in terms of the Adjusted Rand Index (ARI; see Methods), which measures the proportion of node pairs in a network that were either clustered together or separately in both partitionings, while being agnostic to labeling schemes and controlling for chance clustering. The ARI normally takes values ranging from 0 to 1, with 0 indicating chance agreement (but can take negative values if the similarity falls below chance). Fig. 2.4 contains examples of ARI values in real and simulated contexts.

The two clustering methods had high agreement (mean ARI = 0.87, SD = 0.13). Modularity showed a tendency to produce additional communities (median = 3, range = 2, 5). However, the additional communities encompassed a small number of vertices (median = 16.5, IQR = 6–41.5) compared to the principal two (median = 4783.5, IQR = 4759.5–4795), suggesting that a binary partitioning provided a reasonable approximation of the network's true community structure.

Next, we examined the similarity of SP-based partitionings across individuals by computing the ARI between every pair of subjects, and found modestly above-chance agreement overall (mean = 0.13, SD = 0.05). Qualitative inspection of the community organization showed good alignment for PCC, whereas the pattern in mPFC was consistent but shifted topographically across subjects. To quantify this heterogeneity in mPFC, we calculated the between-subject ARI for each region separately (Fig. 2.5). The functional topography of PCC was better aligned across individuals (mean = 0.19, SD = 0.09) than mPFC (mean = 0.1, SD =0.05; paired permutation, $p < 0.001$; Cohen's D = 1.26).

Pattern variability over time

We next sought to estimate whether individual vertices had a stable or unstable community affiliation over time. We did so by performing a sliding window analysis on each subject′s full time series (20 min windows shifting by 1 min). We compared the partitioning derived from each window with the partitioning computed using the entire time series (Fig. 2.6). Our focus here was not on the overall level of agreement (which is expected to be high given the use of overlapping data), but on differences in stability

across nodes. The sliding window analysis provided a means to identify nodes that were highly variable, and allowed us to determine whether these variable nodes followed a specific spatial structure.

The mean ARI along each subject's time series was significantly higher for PCC (mean = 0.59; SD = 0.14) than mPFC (mean = 0.5; SD = 0.13; paired permutation, $p < 0.001$; Cohen's D = 0.65). A subset of nodes showed exceptionally high stability, in that they were assigned to the same community in every time window. The percentage of stable nodes ranged from 0 to 73% across individuals (median = 49.5%, IQR = 29% - 60.25%).

We next tested whether the continuous-valued Fiedler vector (before binarization into discrete communities) carried information about the stability of individual nodes. There is precedent in the literature for the idea that the magnitude (and not just the sign) of the Fiedler vector values conveys important information about the role of each node in the network (Gkantsidis, Mihail, & Zegura, 2003; Tian & Zalesky, 2018). Therefore, we tested whether the magnitude of the eigenvector values was associated with the stability of nodes over time. Specifically, we estimated the proportion of DN affiliations per node as a function of Fiedler vector values, using a logistic mixed effects model (Fig. 2.6). The model identified a positive significant relationship between these features ($\beta = 217.02$, SE = 0.67, $p < 0.001$), signifying that vertices with higher absolute Fiedler vector values were more persistent in their relationship with their corresponding community over time. These analyses suggest that there is potential value in thresholding the Fiedler vector as a means to identify reliable DN and non-DN vertices on an individual subject basis. We

therefore thresholded each subject's Fiedler vector to produce these refined maps. For each individual, we estimated the threshold by selecting the empirical smallest absolute Fiedler vector value that yielded an average stability across supra- threshold nodes of 99%, for positive (mean = 0.0132, SD = 0.006) and negative (mean = -0.0139, SD = 0.0069) values separately. Individuals without such stable nodes (n = 19) were not thresholded, and were included in the subsequent analyses in unthresholded form. The median proportion of retained vertices per individual was 0.49 (IQR = 0.29 0.65). Sub-threshold vertices were set to zero in Fiedler vector maps and 0.5 in the binarized maps (so that they would not bias the calculation of averages). Fig. 2.7A shows the thresholded partitioning for the same individual shown in Fig. 2.2. The maps used in all subsequent analyses were thresholded by this individualized criterion.

With these thresholded partitions, we recomputed the overall similarity across participants. Compared to before, there was lower topographic agreement across individuals (mean ARI = 0.07, SD = 0.04). The same was true for both PCC (mean = 0.1, SD = 0.07) and mPFC (mean = 0.05, SD = 0.03) separately, although the significance of the differences between areas was preserved (paired permutation, $p < 0.001$; Cohen's D = 1.11). Fig. 2.7B shows the average of the thresholded partitions across all participants, denoting the proportion of times a vertex was affiliated with the DN community. This summary illustrates the common organizational layout of both communities, but also highlights the considerable variability across individuals.

To test the possibility that the higher inter-subject variability in mPFC was driven merely by lower signal quality in the retained vertices, we quantified the temporal signal

to noise ratio (tSNR) for each region, both before and after thresholding. We calculated tSNR using time series that were not demeaned, but were otherwise equivalent to the data originally used. A map of the mean tSNR across individuals can be found in Fig. 2.8. In terms of tSNR variability across vertices within each region, mPFC had overall greater spatial standard deviation both before and after thresholding (mPFC: pre-threshold mean spatial SD =33.96, post-threshold mean spatial SD =30.15; PCC: pre-threshold mean spatial SD =15.28, post-threshold mean spatial SD =14.59). However, mean tSNR after thresholding was significantly higher for mPFC than PCC (mPFC: mean =77.34, SD =13.77; PCC: mean =64.99, SD =10.19; permutation $p < 0.001$, Cohen's D =1.02). This reflected a significant increase in mean tSNR in mPFC as a result of the thresholding step (pre-threshold mean = 66.5, SD = 7.87; paired permutation $p < 0.001$, Cohen's D = 0.97), whereas the mean signal quality in PCC increased only slightly (pre-threshold mean = 64.56, SD = 10.02; paired permutation p-value = 0.04, Cohen's D = 0.04). In short, mPFC had higher overall tSNR, albeit with greater variability across nodes. Applying the thresholding step focused the analysis on vertices with high signal quality.

Test/re-test reliability across days

The relatively high inter-individual variability seen in the aggregate map could reflect at least three factors: (1) measurement noise, (2) dynamic variation in mPFC network organization, and (3) stable patterns of functional organization that differ across individuals. To arbitrate among these possibilities, we examined the test/re-test reliability of thresholded mPFC/PCC community structure across separate days of testing. Insofar as the observed variability reflects individual-specific brain organization, across-day ARI

values should be consistently higher within-individual than between individuals (an example comparison for two individuals is provided in Fig. 2.4). Fig. 2.9 shows pairwise comparisons among ten example subjects for PCC and mPFC separately (left).

Once again, we found low alignment across individuals for PCC (mean = 0.08, SD = 0.06) and mPFC (mean = 0.03, SD = 0.03), but both areas showed comparatively high levels of within-individual agreement (PCC: mean = 0.36, SD = 0.14; mPFC: mean = 0.26, SD = 0.1). We calculated an index of relative specificity by computing the ratio of each individual′s across-day (within-participant) ARI to the mean of all between-participant ARI values involving that individual. The index is expected to take on a value near 1 if partitionings are well aligned across individuals and are subject to a common level of measurement noise. It is expected to exceed 1 insofar as functional network organization is reliable and individual-specific. This index is intended to factor out the potential contributions of measurement noise or dynamic instability, which would introduce variability both across individuals and across days.

Fig. 2.9 shows ARI ratios for PCC and mPFC. A signed-rank test showed evidence for specificity (i.e. ratios >1) in both mPFC (median = 7.45, IQR = 6.08 8.65, V = 5037, $p < 0.001$) and PCC (median = 4.25, IQR = 3.53 5.29, V = 5030, $p < 0.001$). Moreover, the ratios for mPFC were significantly greater than those for PCC when compared in a paired permutation test ($p < 0.001$; Cohen′s D = 0.5). This pattern was mostly unchanged when computed using modularity maximization to detect the communities, showing that the results persisted even without a forced binarization (mPFC: median = 4.46, IQR = 3.55 5.33; PCC: median = 2.85, IQR = 1.94 3.38;

difference: $p < 0.001$; Cohen's $D = 0.52$). These test/retest results suggest that the topographic variability seen in mPFC arose at least in part from stable and subject-specific organizational patterns (examples of these partitionings can be found in Fig. 2.10). We stress that our similarity metric, the ARI, measured the similarity of partitionings in a label-agnostic manner. The greater inter-individual consistency in PCC was therefore not merely an artifact of having used a PCC subregion as the basis for label assignment.

Test/re-test reliability across runs

We extended the analysis of per-day data by examining whether the organization of the DN could be extracted using per-run data only. The duration of each run (approximately 14 min) falls well below a previously suggested stability threshold for fMRI-based modularity estimations, which asymptotes at around 20 mins per individual (Gordon et al., 2017). Nonetheless, high ARI ratios could indicate that the SP algorithm can still obtain information about individual-specific patterns of DN organization from a single run of data.

Run-specific SP results captured unique organizational patterns to some degree, even though the overall levels of agreement decreased (PCC between subjects: mean = 0.04, SD = 0.05; mPFC between subjects: mean = 0.01, SD = 0.02; PCC within subjects: mean = 0.17, SD = 0.14; mPFC within subjects: mean = 0.09, SD = 0.08). Fig. 2.10 shows that even though the community estimates were indeed less reliable within-individuals than those captured using per-day data (and sometimes even failed to produce meaningful partitionings), the layout of DN and non-DN was still observable in many

cases, and was comparable to the organization seen using larger amounts of data. We again computed each subject's ARI ratio in order to quantify the specificity of the partitions, this time using the mean of 6 across-run (within-participant) ARI values in the numerator of the ratio (Fig. 2.9, right).

As before, a signed rank test showed that both regions had ARI ratios significantly greater than 1 (mPFC: median = 6, IQR = 4.14 7.99, V = 4953, $p < 0.001$; PCC: median = 3.51, IQR = 2.4 4.26, V = 4971, $p < 0.001$), and ratios for mPFC were higher than those of PCC (permutation $p < 0.001$; Cohen's D = 0.94). This result further confirms that the intrinsic functional organization of mPFC is uniquely arranged per individual, and provides evidence that information about such patterns can be extracted from relatively small amounts of data.

Correlation versus community detection in mPFC

We next explored the possible advantage of community detection relative to a more conventional seed-based functional connectivity analysis for estimating the individual-specific functional topography of mPFC. We examined whether maps generated with SP were more similar per participant across days than those computed from seed-based correlations. We generated a seed time-series by averaging all vertices in the PCC region of our search space, and calculated its correlation with the activity of each vertex in mPFC. The use of the whole PCC region (instead of just 7 m) was meant to represent a typical approach to seed-based connectivity that relies on the group-average location of canonical DN regions. We compared the map of correlation values in mPFC to the map of unthresholded Fiedler vector values using Spearman correlations

across vertices. Pairwise spatial correlations were calculated among maps computed for each day and method from all individuals. Fig. 2.11A shows that these pairwise comparisons resembled those from the across-day comparisons above, and suggested good alignment between methods, but particularly high agreement within subject and method.

Fig. 2.11B shows the test/re-test reliability across days for patterns derived using community detection, seed-based correlation, and across methods (e.g. Day 1 community detection versus Day 2 seed-based correlation). While both approaches were reliable, community detection displayed a significantly higher median correlation coefficient across days than seed-based correlation (Community: median = 0.77, SD = 0.19; Seed-based: median = 0.63, SD = 0.12; paired permutation $p < 0.001$; Cohen′s D = 0.54). Agreement across methods was fair (median = 0.48, SD = 0.23), signifying that the two approaches identified similar topographic features but also had systematic differences. These findings suggest that graph-theoretic community detection algorithms are advantageous for detecting stable functional topologies, in addition to their other advantages of being data-driven, unbiased and observer agnostic.

Relationship between functional organization and sulcal morphology

Next, we asked if the idiosyncratic organization of the DN corresponded to patterns of sulcal morphology. Several previous studies have provided evidence that sulcal and gyral organization informs the location of functional effects (Amiez et al., 2013; Amiez & Petrides, 2014; Zlatkina, Amiez, & Petrides, 2016). Recent work has suggested that DN regions in individuals lie mostly within sulci in vmPFC (Lopez-

Persem et al., 2019). We sought to reproduce this relationship using the SP communities, and interrogated whether it persisted in PCC and more superior mPFC regions.

Fig. 2.12A shows a qualitative comparison between the thresholded DN and non-DN communities and curvature maps for two individuals. In agreement with findings from Lopez-Persem et al. (2019), the DN community appeared to overlap with the superior rostral sulcus in vmPFC in these individuals, whereas the non-DN community included both gyri and sulci. A similar trend was observable in left PCC, where the DN community traced sulcal layouts and non-DN was more likely to appear in gyri.

We quantified these observations by dividing the search space into three regions: a ventral mPFC area corresponding to the region tested by Lopez-Persem et al. (2019), a superior mPFC area (sup-mPFC) that contained the remaining mPFC regions, and PCC (Fig. 2.12B, left). For each individual, we correlated the unthresholded Fiedler vector with cortical curvature for each combination of region and hemisphere. Negative correlations in this context imply that the DN was found in sulci and non-DN in gyri. Results are shown in Fig. 2.12B. Correlations tended to be slightly negative in vmPFC, both on the left (mean correlation = -0.02, SE = 0.01) and right (mean correlation -0.02, SE = 0.01); the distribution was only significantly different from zero in the right hemisphere, and weakly so (one sample t-test: t = -2.18, p = 0.031, uncorrected; Cohen's D = 0.22). Correlations between FV and curvature were positive and significantly different from zero in sup-mPFC, both on the left (mean correlation = 0.09, SE = 0.01; one sample t-test: t = 6.95, p < 0.001; Cohen's D = 0.69) and right (mean correlation = 0.04, SE = 0.01; one sample t-test: t = 4.49, $p < 0.001$; Cohen's D = 0.45). Correlations in

PCC were negative and significantly different from zero in the left hemisphere (mean correlation = -0.08, SE = 0.01; one sample t-test: t = -6.89, p < 0.001; Cohen's D = -0.69), but significantly greater than zero in the right hemisphere (mean correlation = 0.03, SE = 0.01; one sample t-test: t = 2.19, p = 0.031; Cohen's D = 0.22). The difference across hemispheres was significant (paired t-test: t = -7.07, p < 0.001; Cohen's D = -0.88). These results provide preliminary indications that the association between function and structure is heterogeneous across subregions of the canonical DN.

Alignment of mPFC community structure with a proposed DN sub-network organization

The thresholded partitions we identified had conceptual and topographic similarities to DN sub-networks A and B proposed by Braga and Buckner (2017). We explored the relationship between the two sets of sub-regions by reproducing the previously described seed-based connectivity approach in two of our subjects. In previous work, Braga and Buckner (2017; Braga et al., 2019) manually selected individual vertices in dorsolateral prefrontal cortex (DLPFC) that produced two spatially anticorrelated, interdigitated networks with distinctive patterns in the temporo-parietal junction (TPJ), inferior parietal lobule (IPL), parahippocampal cortex, mPFC, and PCC. We hypothesized that if the SP communities corresponded to one or both of the previously proposed sub-networks, our partitionings should match networks A and B generated by seed-based functional connectivity in these diagnostic areas. For whole-brain functional connectomes from two individuals (100307 and 101006), we evaluated seeds in each diagnostic region that reproduced networks A and B (correlation coefficients thresholded at 0.2), and confirmed their placement based on functional connectivity patterns observed

in the remaining areas. In both individuals, seeds in posterior IPL and TPJ most clearly identified networks A and B, respectively. The whole-brain seed-based functional connectivity maps for the two individuals are juxtaposed with the corresponding community detection results in Fig. 2.13. It is worth noting that a few distinguishing features are missing due to below-threshold correlation values (e.g. default network B in right PCC of 100307).

Visual inspection of these networks showed high similarity between our DN community and the previously reported sub-network A. However, the non-DN community filled areas not covered by either DN-A or DN-B. Since this three-network configuration is at odds with the two- network solution suggested in our previous analyses (i.e. comparison with modularity), we ran additional evaluations to confirm its existence. First, we reproduced the whole-brain k-means clustering analysis (12 clusters, 100 iterations) performed by Braga et al. (2019) using the full time series for two subjects (Fig. 2.14A). In addition to identifying DN networks A and B through this approach (in line with previous findings), we found a third cluster that aligned well with the non-DN community. To understand why modularity maximization did not identify the same three discrete clusters within the search space, we performed a silhouette analysis to determine the ideal number of clusters in our search space (Fig. 2.14B). For all 100 individuals, we ran k-means clusterings restricted to the mPFC/PCC search space with a specified number of clusters ranging from 2 to 5, and computed a silhouette score for each solution (higher silhouette scores indicate a better fit). Scores decreased as the number of clusters increased beyond 2 for all individuals, suggesting that a bisection was indeed the best

solution. Paired permutations comparing silhouette scores across individuals indicated that the difference between two (mean = 0.042, SE = 0.001) and three (mean = 0.032, SE = 0.0009) cluster solutions was significant ($p < 0.0001$, Cohen's D = 0.96). Visualization of a two-cluster k-means revealed a close match with partitionings estimated through modularity, and are comparable to those produced by SP prior to thresholding (Fig. 2.14A, bottom rows). These analyses suggest that SP isolated DN-A and the non-DN community as the dominant opposite signals within our search space, but that DN-B is observable once we take advantage of the continuous information contained in the Fiedler vector.

These results support the idea that the two approaches serve complementary purposes. Whereas Braga and colleagues (2017; 2019) identified subdivisions within the DN, the present community detection approach might be better understood as partitioning DN from non-DN cortex. Furthermore, the results show that the continuous-valued output of our approach provides a method for estimating the location of the three networks in PCC and mPFC that is less computationally demanding than full-brain clustering.

Discussion

A considerable amount of meta-analytic work has been dedicated to characterizing the brain activity patterns associated with psychological processes in medial prefrontal cortex (mPFC), revealing both dissociable and overlapping activation across domains (De La Vega, Chang, Banich, Wager, & Yarkoni, 2016; Hiser & Koenigs, 2018; Kragel et al., 2018). For example, topographic patterns associated with

subjective valuation and with the default network (DN) have been suggested to be indistinguishable in mPFC, with overlap also partially extending to posterior cingulate cortex (PCC) (Acikalin et al., 2017; Bartra et al., 2013; Clithero and Rangel, 2014; Laird et al., 2009). This apparent overlap of task-related effects with DN regions has important implications, as it has motivated theoretical proposals about ways in which these superficially dissimilar domains might involve a shared set of core cognitive processes (Acikalin et al., 2017; Clithero and Rangel, 2014; Northoff and Hayes, 2011).

However, the interpretation of overlap in group-level data depends on the degree to which functional organization is heterogeneous across individuals. Recent studies have shown that heteromodal brain regions have considerable variability in functional connectivity across individuals (Mueller et al., 2013), individual-specific functional topography can be occluded in aggregative estimations (Braga and Buckner, 2017; Gordon et al., 2017; Michalka et al., 2015; Tobyne et al., 2018), and overlap in functional activation can vanish with increases in spatial precision (Woo et al., 2014). These findings suggest that group-level and meta-analysis-level overlap does not necessarily imply overlap in individual brains. To date, our understanding of the individual-level heterogeneity in the functional topography of mPFC has been mostly descriptive (Braga and Buckner, 2017; Braga et al., 2019; Gordon et al., 2017). A strong test of the overlap between task-related effects and DN regions would require a method to reliably and precisely capture the functional topography of mPFC in isolated individuals, as well as a quantitative estimate of the degree of topographic heterogeneity across a large group of individuals.

Here we address these challenges by using spectral partitioning (SP), a graph-theoretic community detection algorithm that efficiently separates a network into two (Fiedler, 1975; Higham et al., 2007; Toker and Sommer, 2019). For each of 100 individuals, we subdivided canonical DN regions into DN and non-DN communities. Restricting our analyses to a general mPFC/PCC search space made it appropriate to use a technique that identified a vertex-wise, binary partitioning that was sensitive to the complex topography of the brain. This contrasts with whole-brain network analyses, which need to allow for multiple sub-networks and which often use parcels that are several orders of magnitude larger than vertices as the units of analysis. Partitioning an individual's brain network through SP has a number of advantages, including identifying communities deterministically, constraining communities to contain a similar number of vertices (i.e. preventing the allocation of most vertices to a single community), providing continuous values that relate to the strength of a node's community affiliation, and the ability to diagnose the connectedness of a network through examination of its resulting eigenvalues (Chung, 1997; Higham et al., 2007). Comparisons with partitionings formed by modularity maximization, which heuristically determines the ideal number of communities (Garcia et al., 2018), as well as a silhouette analysis, suggested the binary partitioning was appropriate.

We found a generalizable pattern across individual partitionings, in which the DN community covered ventral and anterior/superior mPFC and posterior PCC, with the non-DN community concentrated in pregenual ACC and anterior PCC. The precise spatial positioning of this general community structure was highly heterogeneous across

individuals, yet stable across test/re-test evaluations within-individual. The idiosyncrasy in functional topography was particularly pronounced in mPFC, and was identified in both run-based and day-based analyses. Individual-specificity could theoretically arise from a variety of sources. For example, individual variability could be due to shifts in functional organization that are independent of structural features (Conroy et al., 2013; K. H. Nenning et al., 2017), or could relate to the pattern of functional connections with the rest of the brain (Mars, Passingham, & Jbabdi, 2018; Passingham, Stephan, & Kötter, 2002; Tobyne, Sean M. et al., 2018). Alternatively, the functional topography of mPFC could be governed by its underlying sulcal and gyral organization, which has been shown to vary systematically across individuals (Mackey and Petrides, 2014). Our results offer some support for this idea, echoing previous findings that DN is contained within sulci (in particular the superior rostral sulcus) in vmPFC (Lopez-Persem et al., 2019). Structure/function associations were heterogeneous in other regions; DN tended to be located in gyri in more superior mPFC regions, whereas the association differed across hemispheres in PCC. Future studies should further characterize this heterogeneous relationship. Another important goal for future work will be to assess whether the network layout in these regions can also be predicted on the basis of other aspects of brain structure, such as myeloarchitecture (Glasser et al., 2016) or structural connectivity (Osher et al., 2016; Saygin et al., 2011, 2016).

Network-partitioning methods such as SP are data-driven, and therefore provide no labeling information about the resulting communities. We circumvented this issue by independently identifying the DN community based on its coverage of area 7m, a region

in PCC that was preferentially associated with the DN relative to subjective valuation in our meta-analysis. We were able to apply labels derived from this group-level approach on the basis of the topography in PCC, where functional organization was more consistent across individuals. Because each community spanned both mPFC and PCC, the labels extended to mPFC where topography was more heterogeneous.

Our results extend previous work that described individual-specific brain organization. Several recent investigations have identified topographic heterogeneity using a different data aspect ratio than we used here (a small number of individuals and a large number of scanning sessions per individual; Braga and Buckner, 2017; Braga et al., 2019; Gordon et al., 2017). Previous work has also shown that functional correlations among pre-defined cortical parcels are highly stable within an individual (Gratton et al., 2018; Kong et al., 2018). Here we were able to quantify the variability and stability of functional topography in a large sample at a fine, vertex-level spatial granularity, using moderately low amounts of data (down to a single 14 min scan, although estimates based on more data were more reliable). The motivation to subdivide DN also stems from recent work by Kernbach et al. (2018), who identified specialized communication of parcels within DN with the rest of the brain in a large pool of individuals.

In addition to the technical advantages noted above, the SP algorithm offers analytical advantages specific to neuroscience. We found that SP outperformed a traditional seed-based correlation approach in capturing idiosyncratic functional topography. Community detection methods such as SP are stabilized by relying on all pairwise correlations among cortical vertices (rather than correlations with an individual

seed). In addition, we found we could threshold the underlying Fiedler vector on the basis of the temporal stability of SP results. The magnitude of Fiedler vector values has been recently used to characterize the continuous connectivity profile of the insula with the rest of the brain, challenging the notion of discrete parcellations in that region (Tian and Zalesky, 2018). The combination of discrete classification and graded information yielded by SP provides additional flexibility and richness relative to some other clustering algorithms.

The community organization of PCC and mPFC was congruent with DN sub-networks A and B proposed by Braga and Buckner (2017; Braga et al., 2019). The topography of our thresholded DN community closely matched network A, whereas our non-DN community included cortical territory that was not part of either DN network. Subthreshold vertices from the SP communities in turn overlapped with DN-B vertices. Our findings therefore complement the initial identification of DN sub-networks by quantifying the systematic variability of their underlying topography in a larger group of people. Understanding the interaction of networks DN-A, DN-B, and non-DN is an important goal for future research. SP is related to methods that have gained traction recently for distinguishing functional cortical gradients (Huntenburg, Bazin, & Margulies, 2018; Margulies et al., 2016; Tian & Zalesky, 2018). A valuable goal for future work would be to assess whether DN-A and DN-B form part of a gradual information processing sequence, or if their functions can be discretized. Regardless, this set of results collectively suggests that canonical DN regions can be topographically

partitioned into DN and non-DN communities, and that the DN community can in turn be further divided into sub-networks A and B.

Conclusion

Our findings show that the functional topography of mPFC is variable across a large pool of individuals, and that the SP algorithm is a useful tool for identifying individualized topography in a data-driven way. The ability to capture an individual′s functional topography without the need for group priors is clinically relevant, as it could help target the assessment of mPFC activity changes in disorders such as depression and schizophrenia (Hiser and Koenigs, 2018). It will be beneficial for future task-based fMRI experiments to be able to characterize where task-evoked activity is situated relative to an individual′s overall mPFC organization. Our work is relevant to interpreting the overlap of DN regions with task-related brain activity in numerous cognitive domains, including valuation (Acikalin et al., 2017; Shenhav & Karmarkar, 2019), memory (Euston et al., 2012), and self-referential thought (Mitchell et al., 2005). An individualized frame of reference will enhance the ability of future studies to gauge similarities and differences among brain activity patterns associated with diverse psychological domains.

CHAPTER TWO FIGURES

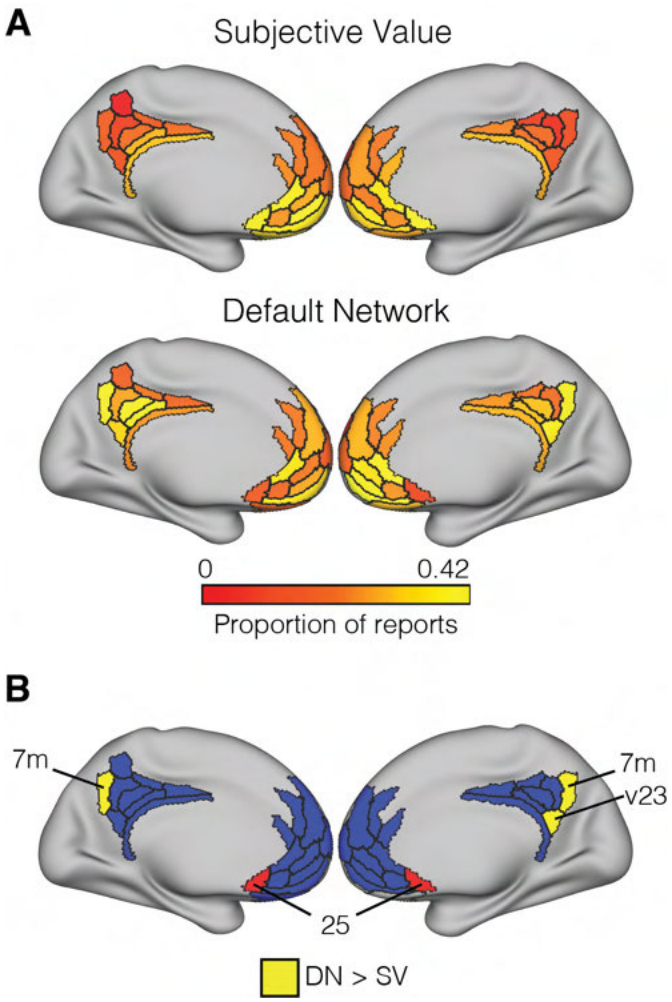

Figure 2.1: Meta-analysis results. **A**: Proportion of times each ROI was reported in the valuation and DN literatures. **B**: Regions identified in permutation-based chi- squared tests contrasting the two literatures. Area 25 (in red) initially appeared to be as associated with valuation, but was not interpreted because the effect was found to reflect carryover from subcortical foci centered in ventral striatum (see text for details). Areas in blue represent the remainder of the search space.

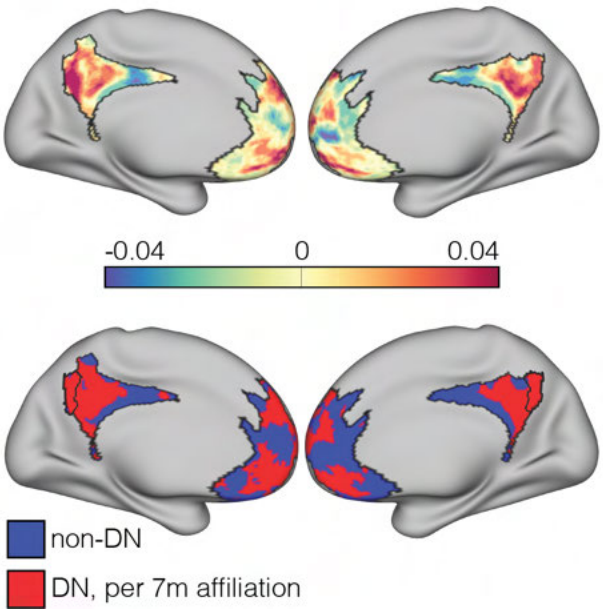

Figure 2.2: Brain partition for an example subject (100307). Fiedler vector values (top) are mapped onto the brain surface, dividing it into positive and negative communities. The bottom brain shows the binarized Fiedler vector, with red areas denoting the DN community (as indicated by coverage of area 7 m, bordered).

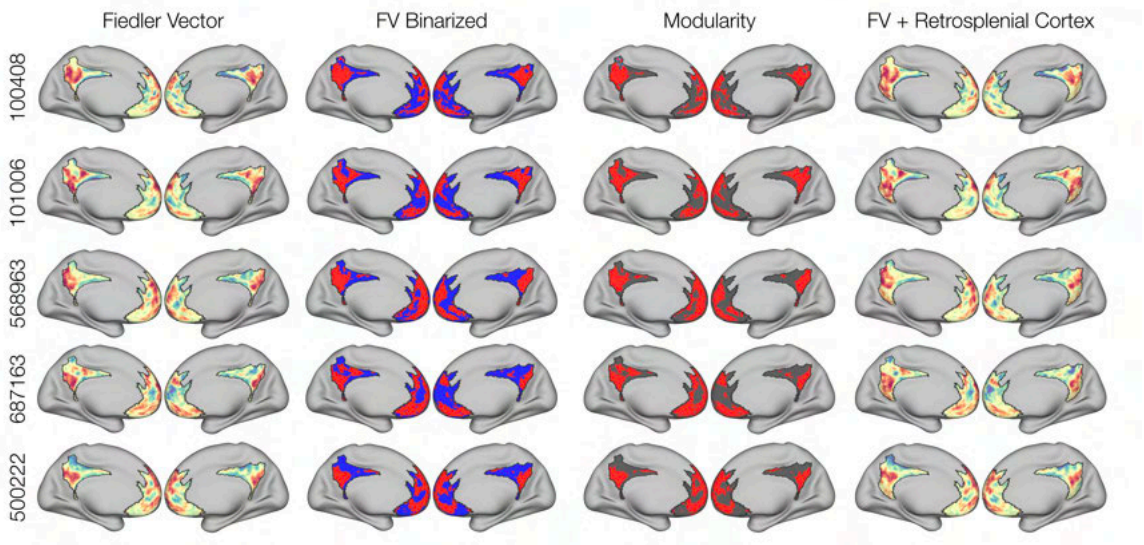

Figure 2.3: Additional examples of individualized partitionings. The first two columns show both Fiedler vector values and binarized communities, respectively. A common organizational principle is visible, even though it shifts topographically across individuals. The organization is also evident when using Modularity (third column), even though some isolated vertices were sometimes placed in a small third community (see subject 100408). The fourth column shows Fiedler vector maps after adding retrosplenial cortex (a common component of canonical DN affected by task-deactivation). The addition of this area to the search space does not alter the original results, and the retrosplenial region tends to be included in the DN community.

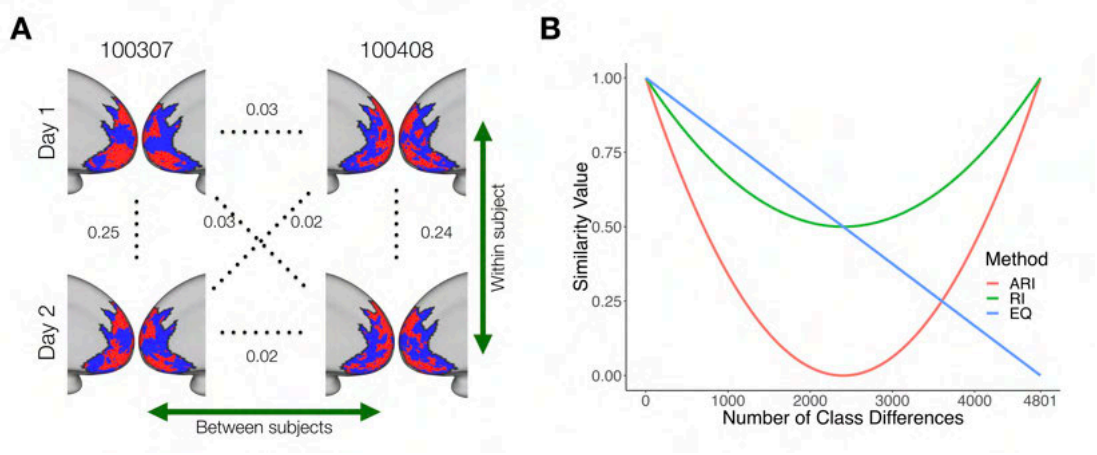

Figure 2.4: Understanding similarities. **A**: Example of an across-day comparison using ARI for two subjects (100307 and 100408). This reflects how qualitatively similar, within-subject partitionings can have relatively small ARI values (here 0.24-0.25), and how partitionings across individuals are much closer to the chance level of zero. **B**: Simulated comparison between two binary partitionings. The allegiance of each node is progressively switched, and the agreement between the new vector and the original one is computed on each change. The x-axis shows the number of nodes switched. Comparing the increasingly dissimilar maps by computing the proportion of equal cluster labelings (EQ) shows the expected linear decrease in similarity. The unadjusted form of the ARI (RI) displays a nonlinearly decaying similarity, and increases after reaching 50% as a result of node pairs once again being grouped in the same/different clusters (making the index label-agnostic). The ARI decays more steeply as a function of increasing dissimilarity, reaching 0 at chance levels. Low ARI values can therefore still occur when there is systematic agreement between partitionings.

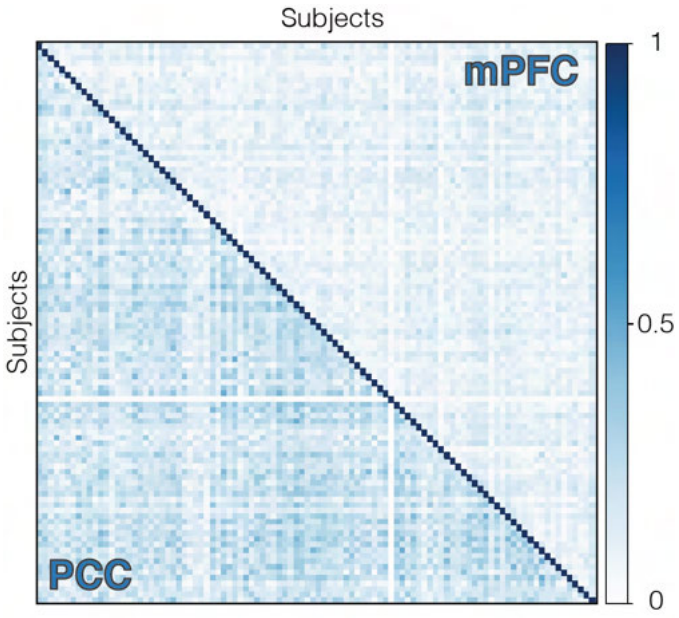

Figure 2.5: Similarity matrix showing ARI values among all subjects for PCC (lower triangle) and mPFC (upper triangle) separately. Functional topographic patterns were better aligned across individuals in PCC than mPFC.

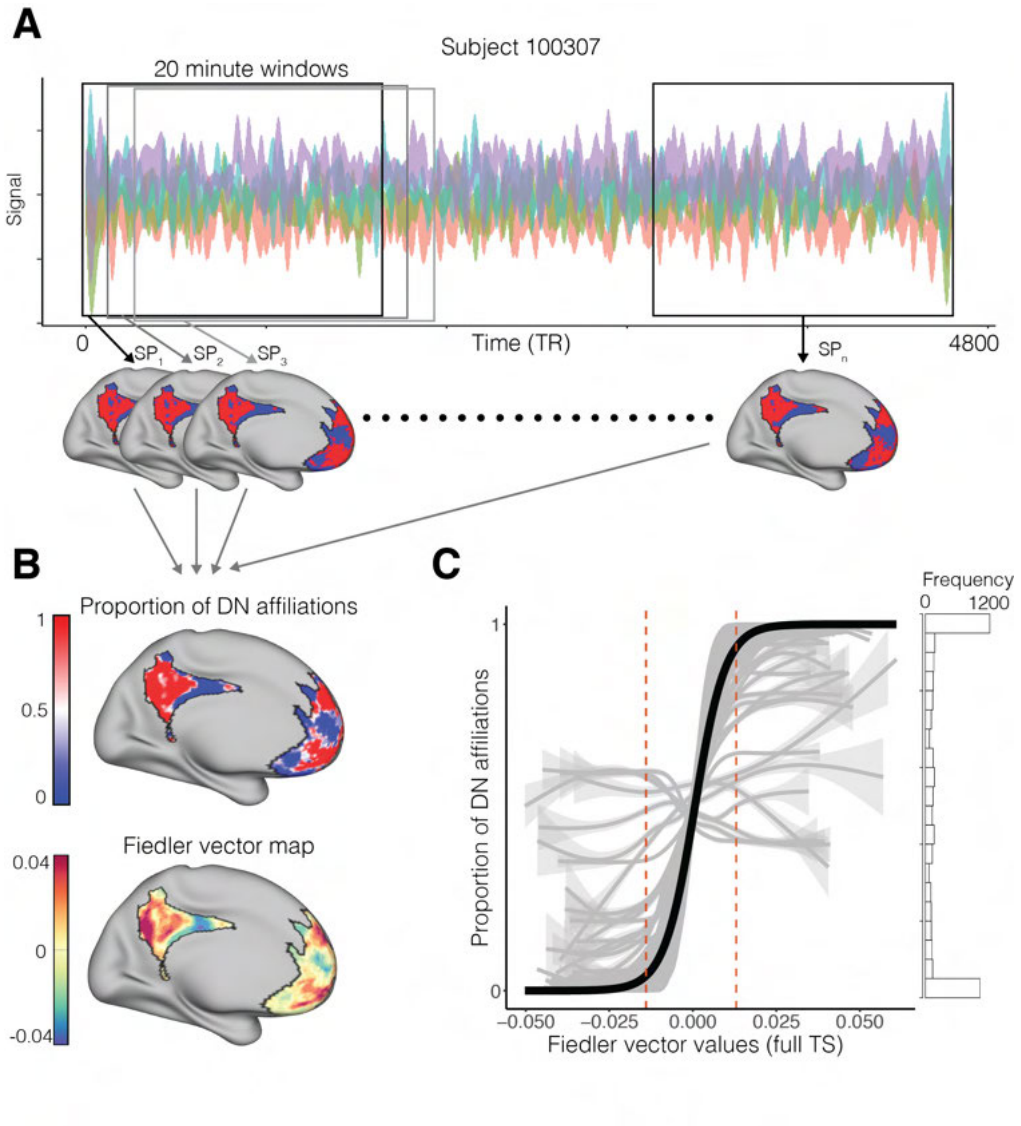

Figure 2.6: **A**: For each individual, we produced partitions for each 20 min sliding window (84 TRs). **B**: Proportion of times each vertex was affiliated with the DN community across windows in one example subject (upper), and the continuous Fiedler vector map for the same subject using their full time series (lower). **C**: Relationship between the magnitude of Fiedler vector values and the proportion of DN affiliations. Grey lines display data for each subject, and the black line shows the fit from a mixed-effects logistic regression. Dashed red lines indicate the mean FV value at which maps were thresholded. The histogram displays the mean frequency distribution of y-axis values.

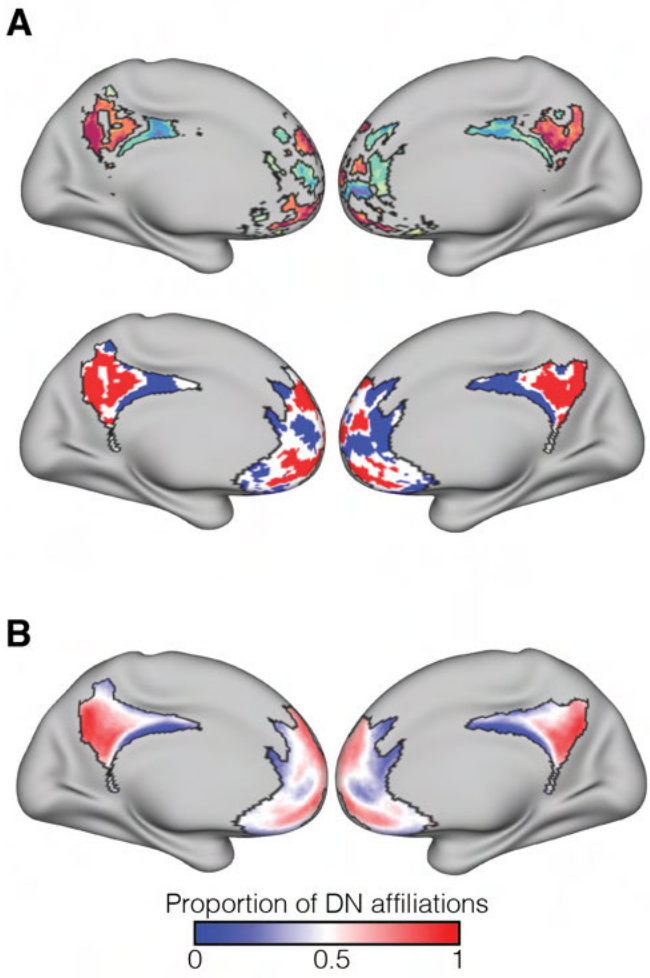

Figure 2.7: A: Thresholded Fiedler vector map for subject 100307 (top), and its binarized form (bottom). Subthreshold values effectively formed a third com- munity of high-variability vertices. **B**: Mean of the binarized maps across all participants, indicating the proportion of DN affiliations per vertex in our sample. Colors represent PCC-based labels ('DN' versus 'non-DN'), which were applied in a subsequent step following the data-driven community-detection analysis and which were necessarily well-aligned in PCC. This aggregate map shows the common organizational principle of the DN and non-DN communities, while also showing the high level of variability in mPFC.

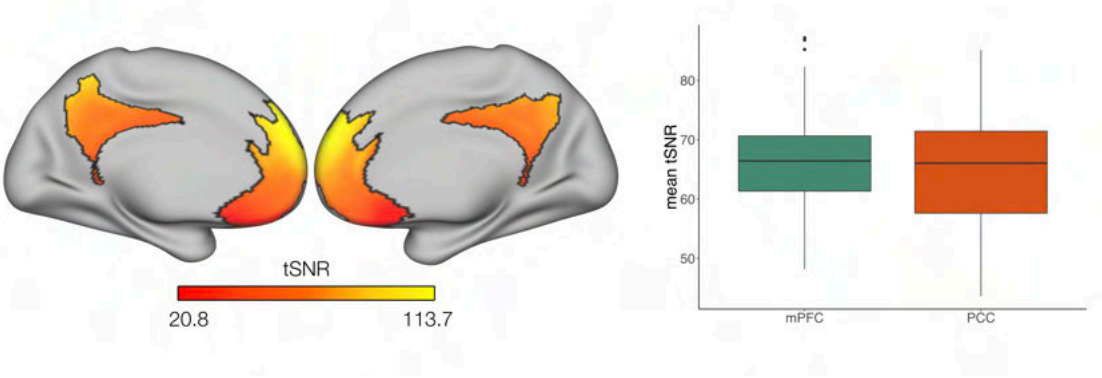

Figure 2.8: BOLD signal quality in the mPFC and PCC search space. **Left**: Surface map displaying the vertex-wise mean tSNR across individuals. **Right**: Mean tSNR for mPFC and PCC across individuals.

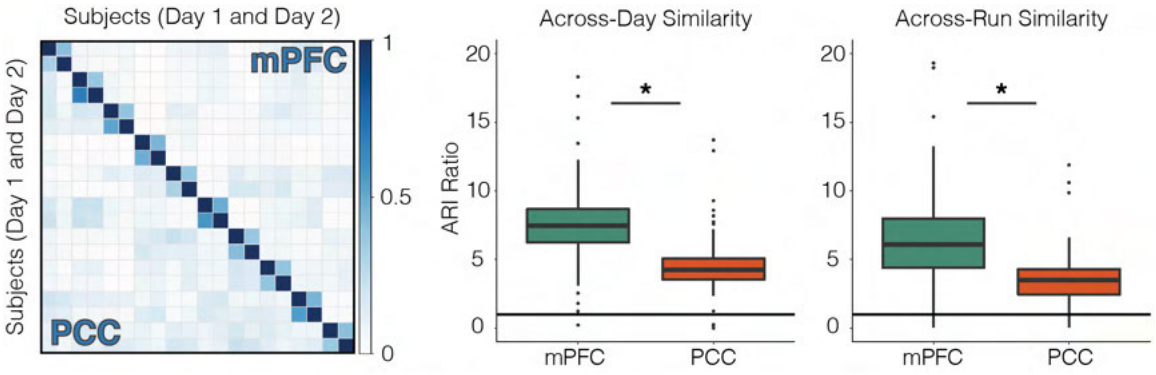

Figure 2.9: Left: Similarity matrix for 10 example participants (2 scanning days each), showing pattern agreement across days and subjects for PCC and mPFC separately. Color scale represents the ARI, which quantifies topographic similarity irrespective of how the communities are labeled. The block-diagonal structure is indicative of test-retest reliability across days within an individual. **Middle**: ratio of within-subject ARI to between-subject mean ARI for all individuals across days suggests idiosyncratic community arrangement for both PCC and mPFC (ratios>1, solid line), with greater subject-specificity in mPFC. **Right**: within-to-between subject mean ARI ratios for run-specific partitionings again show greater subject-specific organization for mPFC.

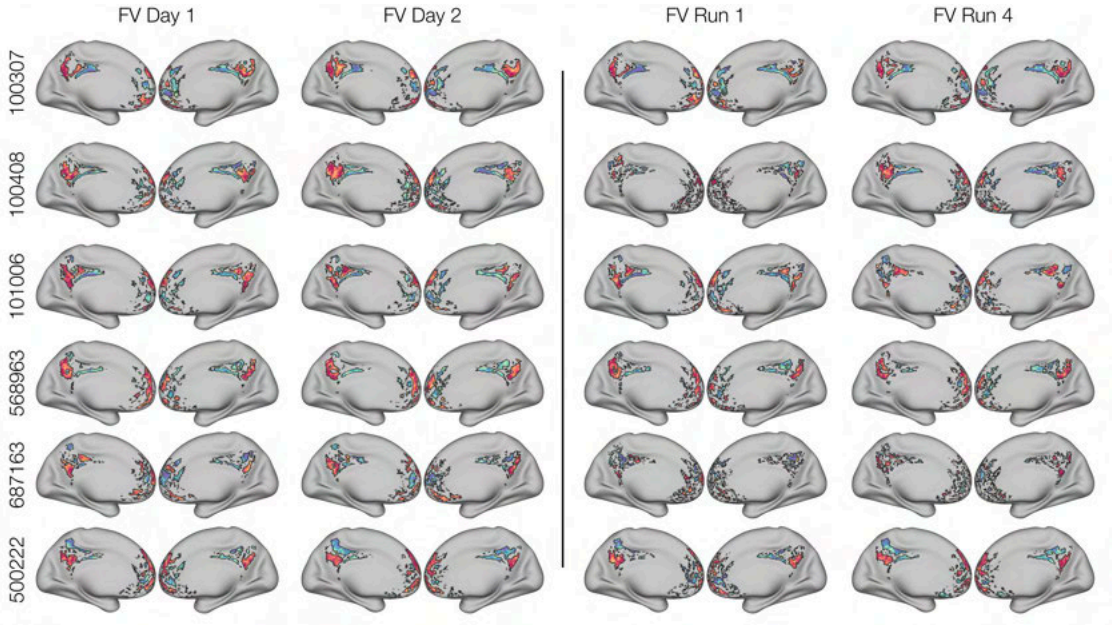

Figure 2.10: Visual examples of idiosyncratic organization across days and runs. **Left**: Thresholded partitionings captured with SP were highly similar across days within individuals, but were topographically distinct across individuals. The organization estimated for each day was generally similar to that captured using the full time series. **Right**: Thresholded partitionings estimated using the first scanning run from day 1, and the last scanning run from day 2. As expected, the considerable reduction in the amount of data used decreased the reliability of the community localization across single scanning runs, which were in some cases irrecoverable (e.g. subjects 100408 and 687163). Even so, the general location of DN and non-DN was still observable in many cases, and is comparable to the organization seen using larger amounts of data. This shows that even with notably noisier estimates and lower within-individual ARI values from working with less data, it was possible to gain information about the general location of DN and non-DN in individuals.

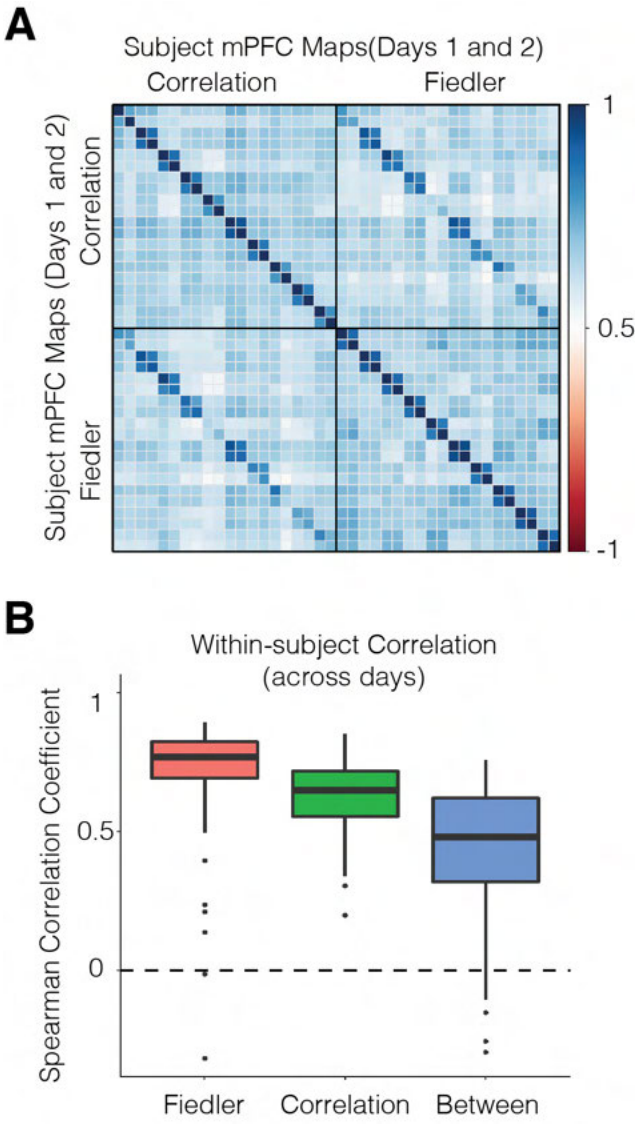

Figure 2.11: A: Correlation matrix comparing the across-day spatial stability of mPFC maps derived from seed-based functional connectivity (using a PCC seed) and the Fiedler vector for 10 example subjects. The top-left quadrant represents seed-based FC maps, and the bottom-right the Fiedler vector, with two single- day-based maps per individual. The upper-right and lower-left quadrants show across-method agreement. **B**: Day 1 vs Day 2 within-subject correlation coefficients for each method, as well as between methods. Community detection through spectral partitioning provided more stable estimates, even though both methods showed good levels of agreement.

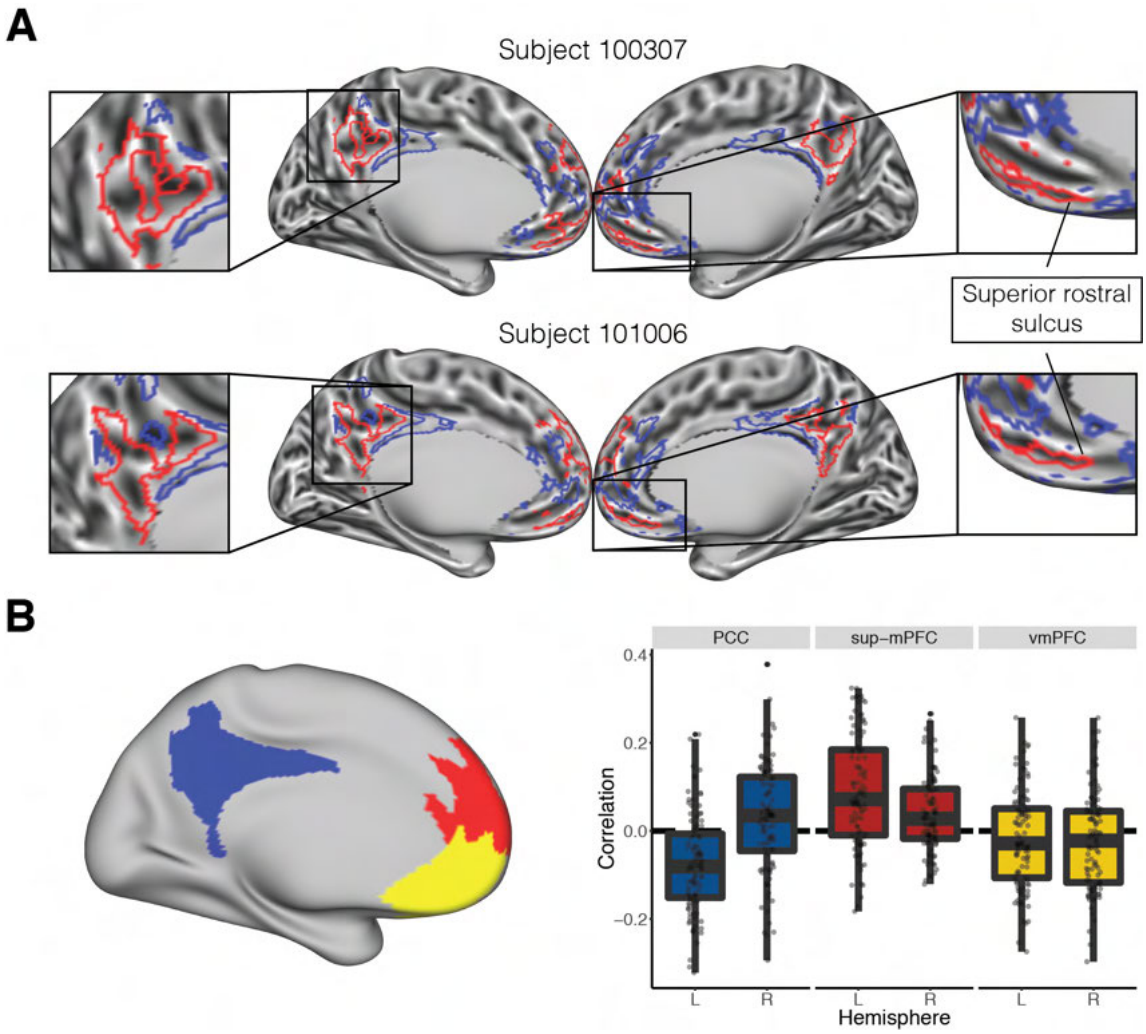

Figure 2.12: A: Correspondence between SP communities and sulcal morphology in two example individuals. Zoomed-in boxes show areas in PCC and ventral mPFC where the DN community appears to follow sulci. The superior rostral sulcus aligned with the DN community, in agreement with findings by Lopez-Persem et al. (2019). **B**: Correlations between cortical curvature and the unthresholded Fiedler values. We divided the search space into three areas (left): vmPFC (yellow), superior mPFC (sup-mPFC, red), and PCC (blue). Boxplots on the right show the distribution of correlation values across individuals for each combination of region and hemisphere, indicating that the DN (defined by positive FV values) was associated with sulci in vmPFC, but with gyri within sup-dmPFC. PCC showed evidence of an opposite association across hemispheres.

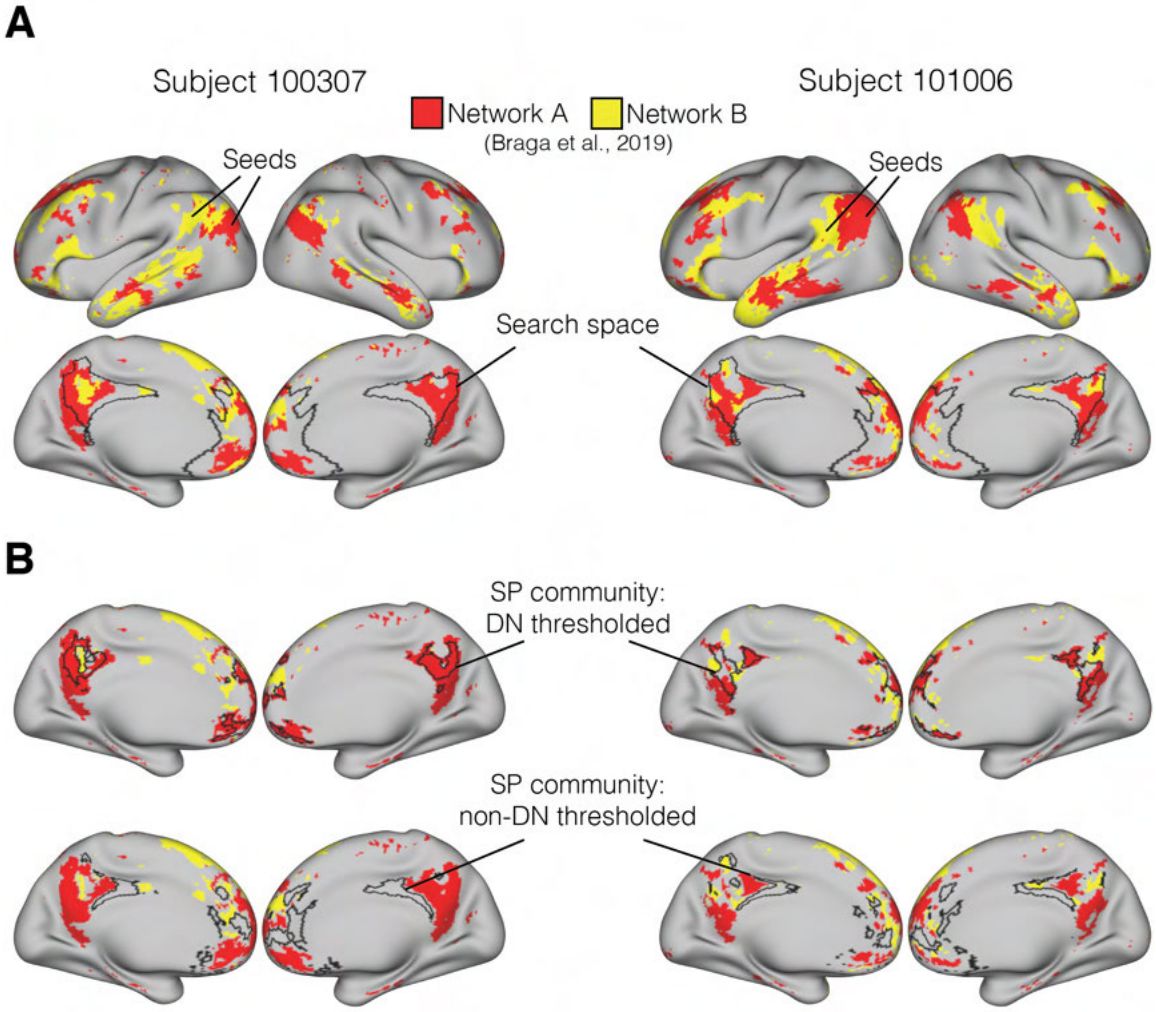

Figure 2.13: Qualitative comparison between DN sub-networks A and B (estimated based on descriptions from Braga et al. (2019)) and SP communities for two individuals. **A**: Whole-brain networks A and B produced by selecting seeds in TPJ, with our community detection search space delineated by black borders. Correlation values are thresholded at 0.2. **B**: thresholded communities (indicated by borders) show strong resemblance between the DN community and network A. The non-DN community covers sections of cortex not associated with either DN sub-network.

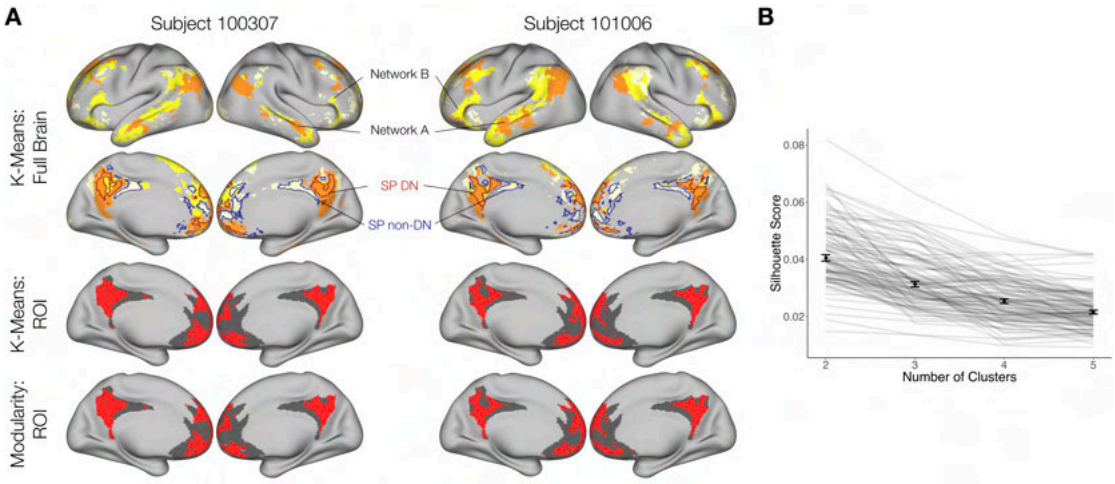

Figure 2.14: Additional checks to confirm the match between our findings and Braga et al. (2019). **A**: DN sub-networks A and B identified through whole-brain k-means clustering were congruent with the existence of a third sub-network, which matched the non-DN SP community (top rows). Limiting k-means to two clusters within the search space aligned with results from modularity (bottom two rows). **B**: K-means based silhouette scores were computed per individual to determine the appropriate number of clusters within the search space (the higher the score, the better). This showed that two networks produced the preferred solution.

CHAPTER TWO TABLES

Hemisphere	mPFC	PCC
Left	10d, 10r, 10v, 25, 9m, a24, d32, OFC, p24, p32, s32	23d, 31a, 31pd, 31pv, 7m, d23ab, PCV, RSC, v23ab
Right	10d, 10r, 10v, 25, 9a, 9m, a24, d32, OFC, p24, p32, s32	23d, 31a, 31pd, 31pv, 7m, d23ab, RSC, v23ab

Table 2.1: Regions that composed the search space. Parcels were visually selected from the Glasser atlas (Glasser et al., 2013), and encompassed zones of cortex that show overlap between DN and valuation (based on Yeo et al., 2011). Note that the number of parcels is asymmetric across hemispheres.

CHAPTER THREE: APPARENT PREFERENCES FOR COGNITIVE EFFORT FADE WHEN MULTIPLE FORMS OF EFFORT AND DELAY ARE INTERLEAVED IN A FORAGING ENVIRONMENT

Introduction

Evaluating whether to pursue a desirable outcome often entails assessing the demands required for obtaining it. In psychology and economics, demands for both time and effort have traditionally been identified as costs that detract from the net subjective value of rewards (Frederick, Loewenstein, & O'Donoghue, 2002; Shenhav et al., 2017; Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006; Westbrook & Braver, 2015). Indeed, decision makers tend to prefer low-effort and immediate rewards over larger rewards that demand longer delays and higher effort (Ainslie, 1975; Hull, 1943; Kool & Botvinick, 2018; Kool et al., 2010; Treadway et al., 2009; Westbrook et al., 2013). However, people still needlessly pursue effortful courses of action in everyday decisions. Emerging experimental and theoretical work has chiefly explained this phenomenon by suggesting a reciprocal relationship between effort and outcomes, whereby exerting effort boosts the perceived value of its outcomes (Hernandez Lallement et al., 2014; Kacelnik & Marsh, 2002; Mochon, Norton, & Ariely, 2012), and positive outcomes imbue value into their preceding effortful behavior (i.e. learned industriousness; Eisenberger, 1992). Though intuitively valid and empirically supported, this explanation leaves open questions about what features of an effortful task can make it attractive in its own right, and how contextual factors mold its perceived costs (Inzlicht, Shenhav, & Olivola, 2018). Understanding what governs the intrinsic costs and rewards

of effort could improve our ability to motivate completion of effortful daily activities, such as schoolwork or physical exercise.

Recent work has proposed that the subjective costs produced by delays and cognitive effort are modulated by the possibility of missing out on better concurrent alternatives, implying that decision makers track the *opportunity cost* of time (OC; Fawcett, McNamara, & Houston, 2012; Kurzban, Duckworth, Kable, & Myers, 2013; Otto & Daw, 2019). Preferences for effortful courses of action could therefore arise from the presence of other types of demands and opportunities in the same choice environment (Kool & Botvinick, 2018). For example, exerting high effort for a reward might be aversive, but the same level of work might be appealing in comparison to the boredom of an equivalent period of passive waiting (e.g. taking the cumbersome way home rather than waiting for a delayed train, even if the resulting time of arrival is the same). This effect of context on effort preferences has been illustrated by experiments on charitable giving, in a which high-effort donation method (such as a run) attracted larger donations than a low-effort method (such as a picnic) when either was available alone (the "martyrdom effect"), but the high-effort event was disfavored when both options were offered as alternatives (Olivola & Shafir, 2013).

Standard models in neuroeconomics and behavioral economics cannot readily address this possibility. The effects on valuation of different types of effort or delay have traditionally been formalized with discounting functions that capture how much less appealing a reward becomes as a function of the associated amount of effort or delay (Ainslie, 1975; Frederick et al., 2002; Green, Fristoe, & Myerson, 1994; Kable &

Glimcher, 2007; Klein-Flügge, Kennerley, Saraiva, Penny, & Bestmann, 2015; Kool et al., 2010; Westbrook et al., 2013). An advantage of relying on these standardized functions is that they can circumvent a key confound when comparing effort and delay, namely that increasingly effortful tasks also tend to take longer to perform. However, although this line of work suggests that discounting functions are similar for delay and effort (Massar, Libedinsky, Weiyan, Huettel, & Chee, 2015; Prevost, Pessiglione, Metereau, Clery-Melin, & Dreher, 2010; Seaman et al., 2018), there is ongoing debate about the shape of the discounting function across types of demands (Arulpragasam, Cooper, Nuutinen, & Treadway, 2018; Białaszek, Marcowski, & Ostaszewski, 2017; Chong et al., 2017; Hartmann, Hager, Tobler, & Kaiser, 2013; Klein-Flugge, Kennerley, Friston, & Bestmann, 2016; Kool & Botvinick, 2018; Prevost et al., 2010), and the degree to which discount rates correlate across cost domains (Seaman et al., 2018; Westbrook et al., 2013). Further, this approach does not directly probe which type of demand people would prefer if given the option, and constrains the understanding of behavioral demands to one in which such demands are bound to reduce the value of prospects.

This evidence sets the stage for examining what people would actually select in environments that intermixed multiple types of demands, including cognitive, physical effort, and delay. *Single choice foraging paradigms* provide a natural way to interrogate preferences among demands, by asking individuals whether the current offer outweighs what could be obtained and experienced during the same time period instead (Constantino & Daw, 2015; Garrett & Daw, 2020; Krebs, Erichsen, & Webber, 1977; Mobbs et al., 2018; Stephens & Krebs, 1986). Such paradigms probe the perceived

opportunity cost of time, which has been hypothesized to underlie the evaluation of behavioral costs (Kurzban et al., 2013). Specifically, foragers must decide whether to spend their limited time by harvesting their currently available resource or by seeking out more profitable (or less costly) alternatives. This could involve, for example, deciding when to leave a patch of depleting richness (e.g. when to quit a job that provides diminishing opportunities for growth), or choosing whether to pursue or forgo offers presented serially (e.g. whether to accept your first job offer or wait for new potential ones). Here we focus on this latter type of foraging, traditionally called prey foraging (Krebs et al., 1977), as it provides a means to examine how choices to accept or reject individual prospects are influenced by contextual knowledge about other potentially obtainable alternatives in the same environment.

Foraging has recently attracted high scientific interest due to its ecological validity, which is rooted in evolutionarily conserved choice behaviors (Hayden, 2018; Mobbs et al., 2018), and offers a number of desirable qualities to address our question. First, it allows for the estimation of optimal, reward maximizing choice patterns on the basis of the opportunity cost of time (Charnov, 1976; Krebs et al., 1977). The opportunity cost depends on the richness of the environment, which can be experimentally manipulated by balancing the amount of time it takes to harvest an offer (*handling time*) versus to search for new ones (*travel time*). We can then interpret observed deviations from such optimal behavior, and examine how foraging behavior changes when the handling time is filled with cognitive effort or physical effort in comparison with unfilled delay (thus disentangling effort from delay duration). Second, in contrast to multi-

alternative economic choice paradigms that often treat choices as independent observations, the sequential nature of foraging encourages individuals to consider both focal and global elements of the decision context. For example, the cost (or value) of effort and delay could globally depend on the availability of alternative demands, impacting the perceived opportunity cost of time. Alternatively, it could be intrinsic and robust to such manipulation, and/or depend on task-specific factors that are adjusted once individuals experience different forms of behavioral costs. Foraging rests on an abundant modeling history that provides computational tools to formalize relevant decision variables (Stephens & Krebs, 1986).

Across two behavioral experiments, we tested the overarching hypothesis that preferences would vary depending on whether individuals faced a single type or multiple types of behavioral demands (i.e. cognitive effort, physical effort, and delay). Experiment 1 used a between-subject design in which participants foraged for rewards of varying magnitudes that imposed varying durations of a single form of demand. Experiment 2 used a within-subject design in which individuals foraged for varying forms of demand in addition to varying reward magnitudes. We leveraged the ecological and computational richness of foraging paradigms to provide new insights on a number of open questions (Inzlicht et al., 2018). First, we examined whether the discounting effects consistently observed in multi-alternative economic choice paradigms persisted when these options were presented serially, and whether the resulting foraging patterns depended on the demands imposed to obtain rewards. The existence of a well-defined optimal, reward-maximizing choice strategy helped us identify when demands boosted or detracted from

the value of rewards (signified by tendencies to overharvest poor offers or underharvest advantageous offers, respectively). Second, we probed whether the perceived cost (or value) of demands was intrinsic or dependent on either task-specific or contextual characteristics. For instance, varying the level of exposure to multiple types of demands across experiments allowed us to assess whether cost evaluation was dependent on the availability of alternative avenues of action. Further, including cognitive demands and physical demands let us examine whether such effects might apply differently to distinct types of effort.

Our results provided mixed support for our pre-registered hypotheses, which are detailed in https://osf.io/2rsgm/. For Experiment 1, we hypothesized that 1) decision makers would integrate reward and timing information in order to approximate optimal foraging strategies, as reflected by acceptance rates; 2) acceptance rates would be higher for trials that demanded physical effort and cognitive effort than for unfilled delay or effortless engagement (based on the apparent value of effort displayed by charitable donors, as described above); and 3) choices would be consistent within trials, reflected by few choice reversals, and acceptance rates for each type of demand would remain stable over time. Confirming our predictions, we found that individuals could approximate optimality by preferring higher rewards and shorter delays, regardless of the richness of the environment. We also found that people in the cognitive effort group accepted more trials than optimal, whereas physical effort and wait groups under-accepted, partially matching our predictions and highlighting that cognitive effort can add value to outcomes. Finally, acceptances (and differences among demands) were both consistent

and stable within each cost condition, in line with expectations. For Experiment 2, we hypothesized that 1) acceptance rates would once again approximate optimality, and would show effects of reward magnitude similar to those seen in Experiment 1; 2) in contrast to Experiment 1, participants would gravitate towards unfilled delays to obtain rewards; and 3) that choices would be consistent and stable over time. As before, decision makers approached the optimal strategy for the foraging environment, resembling the pattern observed in Experiment 1. Contrary to our second and third hypotheses, we found that a preference for cognitive effort was present initially but disappeared over time as decision makers foraged for types of demands in addition to rewards.

Computational modeling results suggested that the normative predictions of traditional foraging models were insufficient to explain our data. Instead, the data supported a model in which demands imposed a bias on the estimated opportunity cost of time, which converged as individuals gained experienced with interleaved types of demands. We hypothesize that this bias arises from modulations of subjective time. For instance, cognitive effort could have become appealing by virtue of compressing the perceived trial time, a perception that could be recalibrated with exposure to alternative demands. The subjective shortening of experienced durations during cognitive effort could be due to the ongoing recruitment of working memory (e.g. preventing individuals from estimating the elapsed time), or because the discrete events within the cognitive task subdivided the time interval. These results expand our understanding of the attractiveness of effort, suggesting potential reward-independent factors that could be leveraged to motivate effort engagement in diverse, everyday scenarios.

Experiment 1: Between-subject Manipulation of Subjective Costs During Foraging

Methods

Participants

The study was preregistered with the Open Science Framework (https://osf.io/2rsgm/registrations). All experimental procedures were approved by the Boston University Institutional Review Board, and written consent was acquired for all participants. For the Experiment 1, we recruited individuals until a desired number of 84 eligible participants was achieved (58 Female, median age = 21, range = 18 - 31; number excluded before reaching goal = 8). The sample size was determined by means of power analysis (ANOVA), using a significance level of 0.05, power of 0.8, an effect size of f = 0.45 (estimated from a pilot study), and three groups (one for each cost type). The resulting per-group sample was 20, which we increased to 21 in order to match three possible block orders. We added an extra group of 21 participants who experienced a minimally effortful condition in order to determine whether effort or pure engagement were driving our results.

Participant data sets could be excluded based on any of the following preregistered criteria: 1) Consent: If they withdrew their participation; 2) Inattentiveness: a catch trial was placed at the end of each experimental block, asking participants to press a key within 3 seconds (time requirement based on pilot study response times). A participant who failed two or more of these checks was excluded and replaced. 3) Improbable choice behavior: The task was structured so that one reward amount must always be accepted. A participant who quit every trial in at least one block was assumed

not to have followed or understood task instructions, or to have disengaged from the task altogether. 4) Performance: Participants were forced to travel if they made 2 mistakes in a cognitive effort trial (see task procedures below), or if they gripped below threshold during physical trials. Any participant with more than 30% forced travels was excluded.

Foraging Task

All experimental tasks were implemented using PsychoPy 2 (v1.85.1, Peirce et al., 2019) on a Macbook Pro laptop. In this task, participants foraged for monetary rewards in an environment in which each trial required either physical effort or cognitive effort for a set period of time (the "handling time"), or an equivalent unfilled delay (Fig. 3.1A). Their goal was to maximize their gains within a fixed amount of time. On each trial a monetary offer was displayed for 2 s, and participants had the opportunity to expend time and/or effort during the handling time in order to earn it. Upon completion of a trial, participants saw a 2 s window displaying the reward obtained, which was followed by a travel time to the next offer. Alternatively, the participant could quit at any point during the handling time by pressing the spacebar on the computer, and immediately start traveling.

Each participant was assigned to one of three types of costs (cognitive effort, physical effort, or delay), and a fourth group of equal size faced an effortless physical task that involved minimal gripping. This latter condition was added to discard the possibility that any differences in acceptances between effort and delay were driven by engagement alone. Each group was unaware that other cost conditions existed. Participants exerted physical effort by gripping a handheld dynamometer (Biopac

Systems, United States) using their dominant hand. Gripping requirements were calibrated at 20% of maximum voluntary contraction (MVC, acquired at the beginning of the session). Cognitive effort entailed switching among Stroop, dot motion coherence, and flanker tasks. This type of task switching is demonstrably effortful (Kool et al., 2010). In the Stroop task, one of three color names was displayed on the screen (red, blue or green), with a font color that was either congruent (e.g. the word red presented in red) or incongruent (e.g. the word red presented in blue). Participants had to select the color of the font (i.e., they had to suppress their tendency to read the word). For the motion coherence task, 100 solid white dots moved on the screen. A fraction of these dots moved coherently to the left or right, while the rest moved in random directions (coherence could be either 30% or 40%, uniformly sampled). Participants had to respond with the direction of the coherent movement. In the flanker task, rows of arrowheads pointed either to the left or the right (maximum of 3 rows, 3 to 13 arrowheads per row). Participants responded with the direction of the center arrowhead, which could point in the same or opposite direction from its neighbors. The tasks were configured so that responses always involved a left or right key press (e.g. for Stroop, two colored circles were presented at each side of the screen). During the handling time, cognitive tasks and their configurations were randomly sampled, and were presented for 1 s followed by a 1 s inter-stimulus interval. Participants were asked to respond within each task's presentation time (e.g. there was a 1 s response deadline). Before the experiment, participants trained in each cognitive task until they correctly performed six consecutive tasks of each kind. While in the handling time, if participants failed to maintain above-threshold gripping or

made two mistakes during the cognitive task, they were forced to travel and missed the reward.

There were three block types, in which handling times of 2, 10, or 14 seconds were paired with travel times of 14, 6, and 2 seconds, respectively. Note that all combinations add up to 16 seconds, meaning that accepted trials took the same amount of time in all conditions, but rejected ones were shorter in environments where the traveling time was shorter. Timing parameters were held constant within each 7-min block. Each block type was presented twice in pseudo-randomized order (total session length = 42 minutes). Reward amounts varied uniformly per trial (4, 8, or 20 cents), with the constraint that every reward was presented twice every six trials. This prevented sequences from being dominated by a single amount during any window of time. Timing information was disclosed at the beginning of each block, and rewards displayed during a 2 s offer window before each trial began. Participants received training prior to the experimental session and were told about all possible timing conditions and reward amounts in order to preclude biases due to experience-dependent learning (Dundon et al., 2020; Garrett & Daw, 2020).

Operationalization of Cost

Foraging theory posits that accepting a delayed reward should depend on the opportunity cost (OC) of time to be incurred in obtaining it, which depends on the richness of the environment (Charnov, 1976; Stephens & Krebs, 1986). In this study, the richness of the environment was manipulated by the length of the handling and travel times, where shorter travel times produced richer environments. Since time combinations

were fixed per block, we calculated each block's optimal accept/reject strategy by computing the expected reward rate (g(*)) from all decision strategies according to the following equation:

$$g(*) = \frac{\sum p_i R_i}{3\text{Time}_{\text{offer}} + (\sum p_i H + \text{Time}_{\text{reward}}) + 3T}, p \in \{0, 1\}$$

where *R* and *p* are the reward, and acceptance probability of offer *i*, respectively, *H* is the handling time for the block, *T* its associated travel time, $\text{Time}_{\text{offer}}$ is the duration of the offer window, and $\text{Time}_{\text{reward}}$ is the duration of the window displaying the reward earned (only presented for completed trials). This gave us the total reward per second attainable in each block as a function of the acceptance threshold (i.e. the lowest amount accepted). Since an ideal forager should always either accept or reject different categories of prey (Krebs et al., 1977), probabilities other than 0 and 1 are not considered. The resulting possible strategies were to accept only 20 cents, accept 8 and 20 cents, or accept everything. Fig. 3.1B shows possible earnings per second for each choice strategy, as well as the lowest amount participants should accept in order to maximize their rewards (circled dots). For example, a participant in a 10 s handling block should accept 8-cent and 20-cent rewards (and reject 4-cent rewards) to maximize their reward rate. Note that accepting every offer was often detrimental to participant earnings.

Analyses

All analyses were performed in R 3.5.2 (R Core Computing Team, 2017). First, we tested whether decision makers integrated delay and reward information. To address the prediction that participants would be more likely to accept higher reward and shorter

handling time trials, we fit a mixed-effects logistic regression to predict trial-wise acceptances (using the lme4 package, Bates et al., 2015), giving a random intercept to each subject to account for biases in participant behavior. We included continuous regressors for handling time and reward, as well as two covariates that probed the influence of recent history on choices. The first was an autoregressive regressor containing the number of consecutive unobtained offers prior to a given trial (unobtained offers could be due to quits or forced travels). The second regressor tracked the sum of the previous *n* offers. We identified the *n* that minimized Akaike's Information Criterion (AIC) among 6 versions of the model (tracing from 2 previous trials up to 7, beyond which the model failed to converge), and reported all coefficients from the winning model.

Our foraging task was configured such that over- and under-accepting were detrimental to total earnings. To confirm this, we fit a general linear model with constant, linear, and quadratic terms to estimate the correspondence between proportion of trials completed (independent variable) and total earnings (dependent variable). A significant quadratic coefficient thus would signal that the task statistics operated as expected. Next, to determine the optimality of the group's decisions, we examined whether each cost type produced a bias to over or under-accept offers 4 and 8 cents (assuming that 20 cents were always be accepted, given our design). The reward-maximizing strategy was to always reject these offers for handling times of 14 s, reject 4 cents and accept 8 cents for 10 s handling times, and accept both for 2 s handling times, yielding a combined optimal proportion of acceptances of 50% (see Operationalization of Costs for details). We

performed a two-sided one-sample chi-squared test of proportions against the null probability of 0.5 for each type of cost. Therefore, a significant difference indicated that participants either over- or under-accepted rewards. This approach was adapted from the analysis preregistered for Experiment 2, as we found it better suited to the question than the original formulation.

Next, we compared acceptance rates across cost conditions. We first performed a one-way ANOVA on the proportion of trials accepted using group as a factor. We then compared the proportion completed among all 4 groups using non-parametric permutation contrasts (6 tests). For each test, on each of 5000 permutation iterations, the group assignment was randomly shuffled without replacement, and the difference in mean acceptance rates across iterations created a putative null distribution. The unpermuted group mean difference was then evaluated against this permuted distribution. The same approach was used for total earnings.

In order to further look at the effect of handling, offer, and cost type, we computed the probability of accepting a trial with a mixed-effects logistic regression. Based on the task structure and our main question, our a priori model of interest included cost condition, handling time, and reward amount as fixed main effects, and a random intercept per subject. Cost condition was modeled with three categorical terms, with the fourth condition as the reference condition. We ran three versions of the model with different reference conditions, in order to test all pairwise differences among the four cost conditions. We then examined whether this a priori model outperformed both simpler and more complex models. We used both AIC and Bayesian Information Criterion (BIC) to

determine the model that minimized the negative log-likelihood while penalizing the addition of parameters. The regressors included in each model were: 1) intercept only; 2) condition only; 3) handling time only; 4) offer amount only; 5) handling time and offer amount; 6) condition, handling time, and reward main effects (a priori model from above); 7) adding a handling-by-reward interaction; and 8) all possible two-way interactions. We predicted that model 6 (the a priori setup) would have the lowest AIC and BIC. Nested models with similar AIC were statistically compared using an analysis-of-deviance. The significance test was computed as the probability of the reduction in deviance, based on a chi-square distribution with degrees of freedom equal to the difference in the number of parameters between models.

Reward-maximizing prey foragers should either accept or reject offers based on their profitability relative to their alternatives (Krebs et al., 1977; Stephens & Krebs, 1986). However, decision makers often curtail persistence based on continuous reevaluation of delayed prospects or lapses in self control (Ainslie, 1975; McGuire & Kable, 2012; McGuire & Kable, 2015; Mischel & Ebbesen, 1970). Therefore, we examined whether participants were consistent (not quitting a trial after it was originally accepted) and stable (similar acceptance rates over time) in their choices. For consistency, we visually examined survival curves indicating at what point during the handling time participants quit a trial (censored points included trial completions and forced travels), in addition to quitting time distributions. We interrogated stability by computing each participant's total proportion of acceptances in the first and second half of the experiment for every block type, comparing the mean proportion of acceptances

using paired permutations (5000 iterations). We then tested whether the observed differences in costs were still present on each half of the experiment by applying winning model from the mixed-effects logistic comparison to each half separately. In each case, the reference cost category was rotated in order to assess pairwise differences among all costs. We evaluated the consistency of the relative costs by visually comparing the patterns of significant differences between experimental halves.

Results

Between-subjects: Decision Makers Integrate Delay and Reward Information

In this experiment, groups faced with different behavioral costs (physical effort, cognitive effort, or delay) chose their preferred strategy to maximize rewards in foraging environments of varying richness. Environmental richness was dictated by the time it took to obtain a reward (the handling time) and the time between trials (the travel time). We hypothesized 1) that participants would approximate reward-maximizing behavior by preferring higher rewards and shorter delays, but that 2) participants confronted with effortful demands would be more likely to accept trials than those faced with passive delay, regardless of handling time and reward amount. The second hypothesis was based on the traditional idea that rewarded work should become appealing (Eisenberger, 1992; Inzlicht et al., 2018).

We performed a mixed effects logistic regression to address the first hypothesis, pooling across the four conditions. Since participants in the effort groups were forced to travel when they failed to perform above threshold (see Methods), we distinguish between acceptances (not explicitly quitting a trial during the handling period) and

completions (successfully obtaining a reward). Even so, the proportion of forced travels was low for cognitive effort (0.05 across participants) and 0 for the physical group. In line with our predictions, larger reward amounts significantly increased acceptance probabilities (β = 0.49, SE = 0.01, $p < 0.001$), whereas longer delays decreased them (β = -0.34, SE = 0.01, $p < 0.001$). Two additional regressors showed that having missed out on consecutive rewards decreased the probability of acceptance (β = -0.06, SE = 0.03, $p < 0.001$), and that participants became more selective when recent offer history was richer (β = -0.01, SE = 0, $p < 0.001$). The counterintuitive reduction in acceptances following consecutive unobtained rewards was driven by the wait group (removing this group from the model yielded a $p > 0.1$), likely because this group accepted the fewest offers. Comparing likelihood fits among lengths of offer history (e.g. sum over 2 or 3 past trials) suggested that participants were influenced by the 6 most recent rewards (AIC = 7682.49), which matched the offer presentation rotation (see Methods).

The next set of analyses examined the hypothesis that both cognitive effort and physical effort groups would uniformly accept more trials than passive delay and easy groups. Fig. 3.2 (left) shows that the cognitive effort group consistently accepted more offers, which resulted in lower earnings (middle). One-way ANOVAs showed that differences among groups were significant both for overall proportion accepted ($F(3, 80) = 6.94$, $p < 0.001$), and total earnings ($F(3, 80) = 8.07$, $p < 0.001$). Post-hoc permutations (5000 iterations) comparing mean acceptance rates between every pair of costs confirmed that acceptance rates of the cognitive effort group (mean = 0.74, SD = 0.13) were higher than those of the physical (mean = 0.6, SD = 0.15; $p = 0.001$, Cohen's D = 1.02) and wait

groups (mean = 0.55, SD = 0.11; $p < 0.001$, Cohen's D = 1.55). It also showed that those faced with the easy task accepted more than those in the wait group, although responses from the easy group were more variable (mean Easy = 0.68, SD = 0.19; $p = 0.009$, Cohen's D = 0.85) (all other $p > 0.05$). The higher acceptance rates produced by the cognitive effort group resulted in significantly lower earnings (mean = 13.55, SD = 0.84) than both physical (mean = 14.03, SD = 0.53; $p = 0.028$, Cohen's D = 0.69) and wait groups (mean = 14.5, SD = 0.58; $p < 0.001$, Cohen's D = 1.32). In addition, we found that the wait group earned more than the physical ($p = 0.01$, Cohen's D = 0.85) and easy groups (mean Easy = 13.94, SD = 0.52; $p = 0.002$, Cohen's D = 1.02). Considering the low proportion of forced travels, performance was an unlikely explanation for the observed differences in choice.

We next examined the optimality of these decisions. The foraging task was configured so that there was a single reward-maximizing strategy per block type (see Operationalization of Cost in Methods), and participants were warned that accepting everything would not maximize their gains. Accordingly, participants who over- or under-accepted earned the least (Fig. 3.2, right; general linear model with quadratic term, F = 25.65, B = -8.99, SE = 2.42, $R^2 = 0.39$, $p < 0.001$). We performed group-level chi-squared tests of proportions to examine whether each cost biased decision makers to under- or over-accept 4 and 8 cent offers, which together should have been accepted 50% of the time. We found a significant over-acceptance of these trials by the cognitive effort group (proportion accepted = 0.59; $\chi^2 = 73.89$, $p < 0.001$), but under-acceptances by the physical (proportion accepted = 0.39; $\chi^2 = 123.81$, $p < 0.001$) and wait (proportion

accepted = 0.31; $\chi^2 = 359.13$, $p < 0.001$) groups (no significant difference for the Easy group). Together, these findings show that even though participants did not always perform optimally, they responded appropriately to the statistics of the environment in each timing condition. Further, they provide evidence that cognitive effort can boost the value of an offer, as reflected by higher acceptances at the expense of earnings.

We next probed whether these results held uniformly across all parameter combinations. Fig. 3.3A shows the mean proportion of acceptances (± SEM) per combination of handling time, reward, and cost, with optimal acceptance rates depicted by the gray circles. Qualitatively, the figure confirms two important predictions. First, participants adapted to the richness of each timing block, gravitating towards optimality regardless of the cost they faced. Second, differences in cost were consistent across blocks, and acceptances were uniformly highest for the cognitive effort group. We tested these observations with a mixed-effects logistic regression, computing the probability of accepting an offer as a function of handling time, reward amount, and demand type (see Methods for details). As before, the model showed significant main effects of handling time (global intercept = -0.81, β = -0.34, SE = 0.01, $p < 0.001$) and offer amount (β = 0.49, SE = 0.01, $p < 0.001$). Comparisons among all conditions are shown on Fig. 3.3B. The matrix of coefficients conveys that the cognitive effort group was significantly more likely to accept any offer than the physical (β = 1.5, $p = 0.005$) and wait groups (β = 2.11, $p < 0.001$), but not the easy group.

These observations could not be more easily explained by simpler or more complex regression models. Fig. 3.3C shows AIC and BIC values for models of

increasing complexity, starting with an intercept-only configuration (see Methods for details). Here we focus on AIC, as both metrics yielded comparable results. As predicted, the a priori model with all main effects performed better than those relying on a single parameter (a priori model pseudo $R^2 = 0.82$), as well as the model with only main effects for handling time and reward amount ($AIC_{a\text{-}priori} = 3416$, $AIC_{handling/reward} = 3427$, $\chi^2(3) = 16.87$, $p < 0.001$).

However, against our expectations, the a priori model was outperformed by models that added an interaction between handling and reward ($AIC_{HR\_interaction} = 3394$, $\chi^2(1) = 24.16$, $p < 0.001$) as well as one that considered all possible interactions ($AIC_{all\_interactions} = 3300$, $\chi^2(7) = 129.58$, $p < 0.001$). Even so, neither of these models offered further interpretable insights on the observed differences across costs, so we restrict our focus to the a priori model containing all main effects. These results further support the hypotheses that handling and reward amounts were similarly integrated across groups, and that the absolute rate of acceptances was affected by the demand faced. However, they offer partial support for the hypothesis that effort would induce higher acceptance rates, which was only true for those confronted with cognitive effort.

Between-subjects: Consistency and Stability

Next, we tested the hypothesis that foragers would exhibit within-trial consistency and across-trial stability in their choices. Participants were deemed consistent if they did not quit after engaging in a trial. To examine this, we plotted survival curves of trial quits across all participants (Fig. 3.4, top-left), with black crosses signaling either completed trials (at 2 s, 10 s, and 14 s) or forced travels. These curves show that most decisions to

quit happened within the first second into the handling time, and participants rarely quit afterwards (proportion of quits after 1 s = 0.04). Physical and easy groups showed slightly lagged responses because the experiment allowed a one-second grace period for participants to begin gripping (marked by the dashed line). This resulted in some individuals opting to wait for that second to indicate an offer rejection (the censor at 4s in the physical group was a completion during a 2 s handling trial that was incorrectly recorded as 4 s by the experimental code). Plotting the quitting time distributions across blocks for the first second of the handling time (ECDFs on Fig. 3.4, top-right) showed that participants in the wait and cognitive groups responded faster as the experiment progressed. These response time observations suggested clarity in participant choices.

Stability was defined as a persistence in acceptance patterns throughout the experimental session. The scatter plot on Fig. 3.4 (bottom-left) shows the overall proportion accepted by each participant before and after the midpoint of the session. Paired permutations showed no differences in acceptances for wait and easy groups ($p > 0.05$), but we found significantly lower acceptances for physical ($p = 0.017$, Cohen's D = 0.27) and cognitive ($p < 0.001$, Cohen's D = 0.56) groups. Fig. 3.5 (left) shows that the reduction in acceptances in effort groups over time was steeper for longer handling times, a hallmark of effortful demands (Treadway et al., 2009). We then replicated the mixed-effects logistic regression from the previous section separately on each experimental half, which confirmed that differences among demands were stable (Fig. 3.5, right). Together, the consistency and stability displayed by participants suggest that decision makers had a clear representation of the environment.

Experiment 2: Within-subjects Comparison of Costs

Methods

Experiment 1 was a between-subjects examination of how cognitive effort, physical effort, and delay would impact prey foraging behavior, and whether canonical discounting effects could be observed under this paradigm. Based on this behavioral baseline, Experiment 2 interrogated which type of demand decision makers would gravitate to when they were allowed to forage for types of demands in addition to rewards. This within-subjects version sought to answer how the presence of alternative avenues would impact the subjective costs of effort and delay, our overarching question.

Participants

We collected data from a new sample of 48 eligible participants (39 Female, median age = 21, range = 18 - 36; number excluded before reaching goal = 6). Sample size was once again determined by means of a power analysis (repeated measures ANOVA), using a significance level of 0.05, power of 0.8, a desired effect size of f = 0.5, and four factors (one for each type of demand, see below for details). The resulting sample size was 45, which we increased to 48 in order to balance the potential order of blocks. We excluded participants using the same criteria as in Experiment 1.

Foraging Task Adaptation

The original between-subjects task from Experiment 1 was modified so that participants foraged for cost types in addition to rewards. Every participant experienced all forms of cost (delay, physical effort, and cognitive effort). There were two block

types, in which delay trials were interspersed with either physical-effort or cognitive-effort trials. This setup prevented participants from having to rapidly switch between response modalities across trials (i.e. keyboard press and handgrip). Each block combination (i.e. physical/wait and cognitive/wait) was experienced three times by each participant in 7-minute-long, interleaved blocks, for a total task duration of 42 minutes. Half of the participants experienced a block sequence that started with the physical block. Timing parameters were matched to the middle condition from Experiment 1 (10 s handling time and 6 s travel time). Participants were informed that this timing combination would be constant prior to beginning the experiment. The upcoming effort type was disclosed at the beginning of each block, and reward/cost offers were displayed at the start of each trial (e.g. "8 cents for physical effort"). Unlike the previous experiment, participants could express their decision to quit during the offer window (instead of during the handling time). Participants were trained in each type of demand until they reached the same criteria as in Experiment 1.

Analyses

The analytical pipeline was mostly preserved from Experiment 1. We first ran tests to determine whether decision makers integrated reward information. We performed a mixed-effects logistic regression with regressors for offer amount, number of consecutive offer misses, and the sum of the *n* previously observed offers (we identified *n* by comparing the AIC of different history lengths). We then confirmed that the under- and over-acceptance were detrimental to total earnings by performing the same quadratic linear model explained above. We tested the group's optimality as before. The optimal

strategy was to reject 4 cent offers and accept 8 cent offers, yielding an ideal acceptance of 50% (not counting 20 cent offers, which should have always been accepted per our design). For these trials, we performed a two-sided one-sample chi-squared test of proportions against the null probability of 0.5 for each type of cost.

To compare acceptances among cost conditions, we first estimated acceptance probabilities using a mixed-effects logistic regression. The model included cost condition and reward amount as fixed main effects, and participant as a random intercept (handling time was not modeled, as participants experienced a single time combination). Cost condition was modeled with three categorical terms, with the fourth condition as the reference condition (two effort conditions, and two delay conditions corresponding to each effort type). We ran three versions of the model with different reference conditions, in order to test all relevant pairwise differences among the four cost conditions. We compared the a priori model to alternative parameterizations using AIC and BIC, varying model complexity as follows: 1) intercept only; 2) cost only; 3) reward only; 4) cost and reward main effects (a priori model); and 5) adding a two-way interaction. We predicted that model 4 would have the last considerable decrease in the negative log-likelihood. Models with similar AIC were formally compared using the analysis-of-deviance approach from Experiment 1.

Next, we probed within-trial consistency and across-trial stability in people's choices. First, we visualized the number of quit responses during the handling time with a survival curve, and plotted the quit response time distribution as a function of block to examine choice consistency over time. Next, we computed the proportion of quits that

occurred during the choice window versus during the handling time. Earlier quits were seen as signifying greater consistency. We then computed each participant's total proportion of acceptances on the first two and last two blocks, and compared them using a paired permutation analysis (5000 iterations). This analysis did not include the third and fourth block because their effort designation was counterbalanced across participants.

Results

Within-subjects: Decision Makers Integrate Reward Information

In Experiment 2, we directly tested an individual's preferences for cognitive effort, physical effort, and delay. We did by modifying the foraging task, such that every participant experienced every type of demand in blocks that interleaved wait trials with either physical or cognitive effort (fixed at 10 s handling and 6 s travel times, to allow for comparisons between experiments). We separately analyzed wait trials from the two block types, labeling the two sets of wait trials according to the type of effort with which they were paired ("Wait-C" for cognitive, and "Wait-P" for physical). We hypothesized that 1) reward decision patterns would match those of the previous experiment; and 2) interleaving effort and delay would produce persistent preferences for wait trials. The second hypothesis was based on the idea that people will gravitate to the easiest way to achieve their goals (Hull, 1943; Olivola & Shafir, 2013).

We first tested the hypothesis that participants would successfully integrate the statistics of the environment. A mixed-effects logistic regression once again showed that larger reward amounts significantly increased acceptance probabilities ($\beta = 0.762$, SE = 0.024, $p < 0.001$), and that participants became more selective if they had recently

observed large offers (β = -0.015, SE = 0.002, $p < 0.001$). Model comparisons suggested that participants were influenced by the 7 most recent rewards (AIC = 4127.63). Unlike Experiment 1, participants were unaffected by the preceding number of consecutive unobtained rewards ($p = 0.43$). As before, a linear model predicting earnings showed that participants who completed too few or too many trials earned the least, in line with the experimental design (Fig. 3.6A; general linear model with quadratic term, F = 11.21, β = -5.53, SE = 2.71, $R^2 = 0.33$, $p = 0.047$). When assessing the optimality of acceptances for 4 and 8 cents (99% of 20 cent offers were correctly accepted), chi-squared tests against the optimal rate of 50% showed that participants under-accepted physical effort trials (proportion accepted = 0.46, $\chi^2 = 6.53$, $p = 0.01$; all other $p > 0.05$). Despite this deviation, participants earned a similar total amount per cost on average (mean Cognitive = 3.13, SD = 0.52; mean Wait-C = 3.56, SD = 0.27; mean Physical = 3.44, SD = 0.25; mean Wait-P = 3.59, SD = 0.24). These results convey that participants in both experiments were similarly influenced by the characteristics of the environment, although participants in the present experiment gravitated more towards optimality.

Next, we addressed the hypothesis that passive waiting trials would be less costly than effortful ones. Even though the reward acceptance pattern at 10 s handling was similar to Experiment 1 (Fig. 3.6B), differences among cost conditions were less evident. Moreover, these differences contradicted our predictions, as cognitive effort remained the least costly among demands. A mixed effects logistic regression confirmed once again that increases in offer amount made acceptances significantly more likely (global intercept = -3.92, $\beta = 0.76$, SE = 0.02, $p < 0.001$), but that the only significant difference

across cost conditions was a higher likelihood to accept cognitive-effort offers than physical-effort offers (Fig. 3.6C; $\beta = 0.38$, $p < 0.001$). Model comparisons suggested that even though a model with all main effects explained a large portion of the variance ($R^2 = 0.82$), there were similar AIC and BIC values for any model including reward amount (Fig. 3.6D). An analysis of deviance among the three models showed that adding cost condition to reward amount significantly improved fit ($AIC_{reward} = 1859$, $AIC_{a\text{-}priori} = 1852$, $\chi^2(3) = 13.24$, $p = 0.004$). However, probing the interaction between these features significantly improved fits from the main effects model ($AIC_{Interaction} = 1830$, $\chi^2(3) = 28.34$, $p < 0.001$), albeit at the cost of interpretability due to choice fluctuations over time (see Consistency and Stability). These initial results contradicted our predictions, suggesting that the availability of alternative demands diminishes preferences among them (rather than promoting wait trials as a way to earn the same reward amounts without the need to exert effort).

Within-subjects: Consistency and Stability

Similar to the between-subject results, a survival analysis displayed that participants made most of their choices quickly within the offer window (Fig. 3.7A). The median percentage of quits during the handling time was 2.61% (SD = 7.67), with only 2 participants quitting over 20% of trials during the handling time. Moreover, the cumulative quitting time distributions show that participant responses during the offer window became faster over time (Fig. 3.7B). Once in the handling time, the overall percentage of forced travels in the effortful trials was low for both cognitive (median = 5.99%, SD = 9.59) and physical (median = 0%, SD = 3.62) conditions. Similar to that of

Experiment 1, these results suggest that decision makers were consistent in their adopted strategies, and were able to perform the tasks well.

However, against our predictions, participants accepted significantly fewer trials on the second half of the experiment, regardless of cost (paired permutations; Cognitive: $mean_{pre} = 0.76$, $SD_{pre} = 0.17$, $mean_{post} = 0.65$, $SD_{post} = 0.21$, $p < 0.001$, Cohen's D = 0.6; Physical: $mean_{pre} = 0.71$, $SD_{pre} = 0.2$, $mean_{post} = 0.65$, $SD_{post} = 0.21$, $p = 0.008$, Cohen's D = 0.27; Wait-C: $mean_{pre} = 0.73$, $SD_{pre} = 0.2$, $mean_{post} = 0.66$, $SD_{post} = 0.22$, $p = 0.004$, Cohen's D = 0.31; Wait-P: $mean_{pre} = 0.72$, $SD_{pre} = 0.2$, $mean_{post} = 0.66$, $SD_{post} = 0.22$, $p = 0.005$, Cohen's D = 0.3; Fig. 3.8). The reduction in acceptances was not driven by a performance decline, as the proportion of forced travels was not significantly different between the first and second half of the experiment (paired permutation with 5000 iterations, $p = 0.16$). Notably, the pattern of choices from the first half of the experiment resembled that from the matching handling/travel block from Experiment 1 (Fig. 3.3A). Mixed effects logistic regressions performed on each half separately confirmed that the relative preference for cognitive effort seen previously was preserved at the beginning of the experiment (Fig. 3.8, top-right, lower triangle), but that it disappeared over time amid an overall decline in acceptance rates as participants gained experience with cognitive effort, physical effort, and both wait conditions (upper triangle). This fading preference for cognitive effort was observed regardless of which block participants experienced first, even though its magnitude appeared to be reduced for those who experienced the physical block first (Fig. 3.8, bottom)

All together, our findings show that the stable preferences for cognitive effort observed during single-demand foraging slowly fade if alternative demands are introduced. This points towards task-specific and outcome-independent features from cognitive tasks that can be recalibrated by new experiences.

Computational Modeling of Foraging Behavior

We developed a computational model to identify cognitive operations that could explain the complex patterns observed in both experiments (e.g. learning, time estimation, etc.). We isolated four key aspects of the behavioral results that an adequate computational model should be able to capture: 1) show deviations in acceptance rates that involved both under- and over-harvesting of offers (e.g. Fig 3.3A); 2) keep track of recent offer history in order to evaluate the relative value of each offer; 3) persistently display higher acceptances for cognitive effort when faced with single behavioral costs (Experiment 1); and 4) initially match this preference for cognitive effort trials when foraging for demands and rewards (Experiment 2), which should decrease to a common rate across costs over time. These conditions excluded a number of popular explanatory candidates that were not modeled, such as learned industriousness, boredom, fatigue, sunk cost effects, or asymmetric learning of environmental richness (see Discussion).

Model Fitting and Model Comparison Procedures

For each model, both population level and individual subject level parameters were fit simultaneously using an Expectation-Maximization algorithm (Huys et al., 2011, 2012). Briefly, it was assumed that, at the population level, each parameter was

distributed according to a Gaussian distribution, defined by a mean and standard deviation; the individual subject parameters were drawn from this distribution. Accordingly, a point estimate for the individual subject parameters were found via maximum a posteriori (MAP) estimation using the population distribution for each parameter as the priors, and the variance in the individual subject parameter estimates were calculated by taking the inverse of the Hessian matrix at the MAP estimate. Next, the mean of the prior was updated by taking the weighted mean of individual subject estimates, and the variance of the prior was updated by taking the weighted variance across the parameter estimates of individual subjects. This process was repeated until the mean and variance of the prior for all parameters converged (a change of less than .1% across successive iterations). Full mathematical details of the algorithm can be found in Huys et al. (2011, 2012), and the algorithm was implemented using a custom written R package, GaussExMax, available on github (github.com/gkane26/GaussExMax).

Model comparison was performed by calculating Bayes Information Criterion at the group level (integrated BIC or iBIC; Huys et al., 2011, 2012; Kane et al., 2019). iBIC penalizes the likelihood of the data given model parameters $p(d|model)$ for model flexibility (the number of parameters, k), and the size of the penalty depends on the number of observations o:

$$iBIC = logp(d|model) + \frac{k}{2}log(o)$$

Laplace approximation was used to find the log marginal likelihood, the likelihood of the data for each subject given the population level distribution over each parameter (Huys et al., 2011; Huys et al., 2012; Kane et al., 2019).

Model Definitions

We began by fitting a base model derived from the normative predictions of the Marginal Value Theorem (MVT; Charnov, 1976), which states that a delayed reward should be accepted if its earning rate surpasses the opportunity cost (OC) of time incurred in obtaining it (see Operationalization of Cost for details). This decision rule can be expressed as

$$\text{Model}_{\text{base}} = R_i - \gamma H_i \tag{1}$$

such that the reward *R* of trial *i* should be accepted if it is larger than the average amount one could earn per second in the environment γ (gamma; the OC) during the handling time H_i, with high gammas promoting selectivity. We fit this and all subsequent models using a softmax function to predict the probability of accepting a given trial

$$P(\text{accept})_i = \frac{1}{1 + e^{-(1/\beta)(\text{Model}_{\text{base}})}} \tag{2}$$

gamma in this base model was computed as an individual's empirical earning rate by the time of each offer, with the free temperature parameter β indexing decision noise. Differences among costs in Experiment 1 could then be due to the impact of each cost on decision noise, which could have also driven the gradual convergence of acceptance rates in Experiment 2. However, it is also possible that individuals updated the estimate of gamma at different rates when faced with different demands. Previous accounts have demonstrated the biases that can impact the learning of the opportunity cost (Dundon et al., 2020; Garrett & Daw, 2020). We therefore tested a second model that allowed γ to evolve over time, adapting a learning rule from Constantino & Daw (2015)

$$\gamma_{i+1} = (1 - (1 - \alpha)^{\tau_i}) \frac{R_i A_i}{\tau_i} + (1 - \alpha)^{\tau_i} \gamma_i \quad (3)$$

where A_i is an acceptance indicator (0 or 1), and τ_i is the duration of the trial. The duration was given by

$$\tau_i = H_i A_i + T_i \quad (4)$$

such that the handling time was counted only for accepted trials. Higher values of the learning rate α (ranging from 0 to 1) resulted in a larger update toward the current trial's earning rate weighted by is duration (first term in the equation). For this adaptive model, we fit the learning rate α and the initial value of γ in addition to the temperature. These modifications allowed for the tracking of recent reward history, while maintaining the same decision rule

$$\text{Model}_{\text{adaptive}} = R_i - \gamma_i H_i \quad (5)$$

where values of γ were now updated on each trial *i* according to Eq 3, rather than by the empirical earning rate.

Another possibility was that rather than impacting the learning rate, covert features specific to each type of demand uniformly affected the way people perceived the opportunity cost of time. If so, biases could remain stable if only one type of demand was experienced (Experiment 1), but adjust over time as biases from other tasks were experienced (Experiment 2). For example, such biases could reflect an individual's global estimation of the richness of the environment, or a focal estimation of the passage of time experienced during performance of each task (e.g. how long 10 seconds actually feel). Altered time durations have occasionally been associated with levels of cognitive

engagement (Csikszentmihalyi, 2014; Wearden, 2016), and have a direct effect on a trial's reward rate. Intuitively, the perceived passage of time could be recalibrated with experience. We modeled these biases by adding a demand-specific term to γ that was gradually integrated across tasks for those participants who foraged among demands. The overall model took the following form

$$\text{Model}_{\text{biased}} = R_i - (\gamma + C_i)H_i \tag{6}$$

where γ was the mean total earning rate per subject (mean = 0.55 cents per second, SD = 0.02). Negative values of C_i would therefore suggest a boost in offer value produced by a demand. At each decision point, C was the cost bias produced by each demand c_{demand} weighted against a global bias estimate c_{global}

$$C_i = c_{\text{global}} w_i + c_{\text{demand}}(1 - w_i) \tag{7}$$

with w_i controlling how early participants started integrating the converging cost c_{global} into their decisions. The value of this weight was given by

$$w_i = \left(\frac{i}{max(i)}\right)^{\alpha_w} \tag{8}$$

where the section in parentheses was the proportion of the experiment completed so far (note that the organization of the experimental session was fully disclosed), with high values of the free parameter α_w denoting a late integration of the global converging bias. Since participants in Experiment 1 experienced a single type of demand, the values of c_{demand} and c_{global} where the same and α_w did not play a role. This resulted in just fitting parameters c_{demand} and β for that dataset. For the within-subjects group, we assumed that the resulting c_{demand} values from the first experiment reflected the actual bias induced

by each demand. We therefore fixed each value of c_{demand} (i.e. $c_{cognitive}$, c_{wait}, and $c_{physical}$) to the mean of the fitted parameters from each demand group from Experiment 1, and fit the global bias c_{global}, the convergence rate α_w, and temperature β. This dynamic allowed for the observed merging of acceptances across demands in Experiment 2, while respecting the stability in the between-subjects study.

Finally, we asked if the biases imposed by each demand could be isolated from a flexible computation of the opportunity cost of time. This fourth model combined the main features from $\text{Model}_{\text{adaptive}}$ and $\text{Model}_{\text{biased}}$, such that γ was estimated on each trial per Eq. 3, with demands imposing distinct and potentially converging biases as described in model 3. The form of this fourth model was

$$\text{Model}_{\text{adaptive + biased}} = R_i - (\gamma_i + C_i)H_i \tag{9}$$

The resulting parameters were compared across cost groups (Experiment 1) and across experiments using ANOVAs, and the c_{demand} biases computed in Experiment 1 were compared against 0 using one-sample t-tests.

Results

The fourth model produced the lowest BIC in both experiments (Fig. 3.9A. Between-subjects = 3552.2; Within-subjects = 1818.57), indicating that the tenets of the Marginal Value Theorem were necessary but not sufficient to explain the observed adaptive behavior. Fig. 3.9B-C show parameter fits from the winning model for both experiments. The learning rate α and the bias c_{demand} were of particular interest, as these parameters distinguished whether different demands globally affected the integration of the environmental richness or focally modulated an offer's perceived cost, respectively.

Analyses of variance showed significant differences among Experiment 1 groups only for c_{demand} ($F(2, 60) = 7.01$, $p = 0.002$), with α being uniformly low for all groups. Pairwise post-hoc permutations confirmed that the cognitive group (mean = -0.08, SD = 0.14) had lower biases than either wait (mean = 0.11, SD = 0.18; $p = 0.001$; Cohen's D = 1.16) or physical (mean = 0.09, SD = 0.21; $p = 0.005$; Cohen's D = 0.95) groups. These biases were significantly higher than zero for the wait group ($t(20) = 2.79$, $p = 0.011$), and lower than zero for the cognitive group ($t(20) = -2.51$, $p = 0.021$). This suggests that rather than reducing costs, cognitive tasks boosted the value of their concomitant offers.

The mean values of c_{demand} from each group were carried over to the within-subjects fitting, and were aggregated over time according to a weighted decay into a final C value. As can be seen on Fig. 3.9C, parameter values were comparable across experiments, pointing to consistency in the way decision makers approached our foraging scenarios. In addition, values of α_w suggested that c_{demand} estimates converged towards the second half of the experiment (mean = 1.75, SD = 2.07), gradually reducing bias by the end of the experimental session (mean = -0.02, SD = 0.06). Model fits provided good qualitative resemblance to the observed dynamics in each experiment. This is illustrated in Fig. 3.9, in which 95% confidence intervals from within-subject fits are overlaid to empirical choices (between-subject fits matched the left side of this figure, as expected).

Discussion

Understanding the sporadic attractiveness of effort remains elusive. In this study, we sought to identify factors underlying this phenomenon, and tested the hypothesis that the perceived cost or value of a demand could depend on the ability to achieve the same

outcome through different means. Using two foraging experiments that exposed individuals to a single or intermixed types of demands, we found ta tendency toward engaging with cognitively effortful prospects when individuals faced a single form of demand, which faded over time when other costs were also experienced.

Our findings provided partial support for our pre-registered hypotheses (https://osf.io/2rsgm/registrations). In accord with our predictions, decision makers in both experiments successfully approximated reward-maximizing strategies, as evidenced by acceptance rates near optimal points. This is in line with previous foraging experiments, where the main deviation from optimality has been a tendency to overharvest (Constantino & Daw, 2015; Wikenheiser et al., 2013). In contrast, our prediction that effort requirements would produce more acceptances only held for the cognitive group, and this tendency for higher acceptances did not switch to passive waiting as predicted for Experiment 2. These hypotheses originally stemmed from prevalent proposals that incentivized work should become more appealing (Eisenberger, 1992; Inzlicht et al., 2018), but that people would opt for the easiest available way to achieve desired outcomes (Hull, 1943; Olivola & Shafir, 2013). Instead, our novel use of foraging paradigms, in addition to our inclusion of types of effort that are often studied separately, highlighted that reward-independent task features can promote the pursuit of effortful prospects. The fact that predicted stable preferences for cognitive effort carried over onto the beginning of Experiment 2 further supported a task-specific modulation. However, the unexpected gradual convergence in acceptances in the second experiment emphasized that these features could be contextually molded.

These choice dynamics ruled out a number of common explanatory candidates. For instance, if the subjective costs of effort or delay were intrinsic, we would not have observed the gradual convergence in acceptances when they were all intermixed in the same environment. This also speaks against a direct role of boredom in our results. The aversiveness of boredom can lead people to attempt any available activity in order to avoid it (Westgate, 2020). This would have likely led to a persistent preference for either effortful condition. Lastly, this gradual convergence in choices rules out the possibility that our sample was mostly composed of demand seekers. Some individuals express a strong need for cognition, which can bias studies of mental labor (Cacioppo et al., 1996; Sayali & Badre, 2019). Even though the prevalence of this trait in our college-level population could have driven the preferences for cognitive effort observed in both experiments, we would have expected such individuals to maintain this preference in the face of alternatives. Another unlikely candidate is fatigue. Even though participants decreased their acceptance rates for effort in both experiments, we would have expected them to increase engagement with wait trials over time in order to preserve earning rates. Further, we do not expect any of our observations to be due to the risk of failing to successfully complete effortful trials, as participants remained accurate throughout both experiments. Finally, we cannot conclude that cognitive effort gained value from its rewarding outcomes, as has been proposed in the past (Eisenberger & Cameron, 1996). This phenomenon could not have benefitted a single type of demand, as cognitive effort, physical effort, and unfilled delay were rewarded at the same rate.

Traditional foraging models, in which a reward is weighed against the earnings foregone by pursuing it (Charnov, 1976; Constantino & Daw, 2015), were also not enough to explain our results. Instead, our results were consistent with a computational model in which each type of demand distinctly biased the estimation of the opportunity cost, and such biases could be gradually recalibrated over time as decision makers gained experience with additional demands. In this formulation, passive delay and physical effort added to the opportunity cost, whereas cognitive effort reduced it. This raises the question of how these biases arise. We hypothesize that demands altered the perceived duration of time. Immersive activities can subjectively speed up the passage of time (Csikszentmihalyi, 2014), and this effect appears to scale with the amount of cognitive processing required by a task (Wearden, 2016). Moreover, adjusting for time sensitivity has been shown to improve fits for delay discounting models (Mckerchar et al., 2009). Therefore, even though participants knew the nominal handling times, cognitive effort might have become valuable by subjectively compressing the perceived durations. In our experiments, shorter subjective handling times would result in higher perceived earning rates, but it is also possible that individuals benefitted from a sense of fast completion. We speculate that this construal of elapsed time could then be adjusted with the experience of alternative demands, as participants learned how long the 10 seconds handling time felt. This could help explain why preferences for cognitive effort were smaller for within-subject participants who started with the physical block, as they had already formed an estimate of the handling time by the time they experienced a subjectively faster cognitive trial. The shorter time sensitivity produced by cognitive

effort could be due to participants counting the number of sub-tasks during pursual (which was always less than the number of seconds in the handling time), or because taxing cognitive resources distracted individuals from estimating the passage of time (Zakay & Tsal, 1989). These hypotheses will have to be experimentally probed in the future.

Relying on foraging paradigms helped us control for common shortfalls in the existing literature. For example, even though engaging with a demanding offer encompasses evaluating, pursuing, and receiving its outcome, investigations of the pursuit and receipt stages have been limited. Many studies have sought to isolate the evaluation stage by presenting individuals with offers that do not come into effect until the end of the experimental session, deliberately setting aside the question of how the ongoing experience of cost and outcome affects subsequent choices (Apps et al., 2015; Chong et al., 2016; Kable & Glimcher, 2007; Massar et al., 2015). However, discrepancies between expected and experienced value are well known (Kahneman & Thaler, 2006), and choosing immediate outcomes could be due to factors other than time preference, such as the risk of being unable to redeem the future rewards, or hoping to find better offers in the near future (Fawcett et al., 2012; Frederick et al., 2002; Sozou, 1998). We circumvented these pitfalls by providing participants with an opportunity to actively experience the demands they accepted, and giving them full knowledge of the available options. Another challenge in comparing effort and delay is that increments in the required effort are often paired with a longer execution time. This issue is often avoided by comparing discounting functions, which yield points at which individuals

faced with increasing costs become indifferent to large rewards (Massar et al., 2015; Prevost et al., 2010). We eliminated this possibility by fixing the handling time for all demands, allowing us to probe people's preferences directly. Moreover, Experiment 1, which more closely resembled the traditional way costs are compared, provided a baseline of foraging behavior that contextualized the findings of the second experiment (a direct test of our hypothesis). Finally, the existence of optimal strategies helped us identify when demands could be attractive, a central motivation for this study. The apparent attractiveness of cognitive effort was reflected both by a higher proportion of acceptances at the expense of earnings, and by the fitted opportunity costs biases discussed above. Of note is that the optimality of the strategies induced by each demand depended on the richness of the block. This can help explain why the wait group earned the most despite having similar acceptance rates to the physical group, as the former was closer to optimality in the richest environment.

A potential limitation of this work is whether our task manipulations were actually costly. For instance, it is possible that participants gravitated towards cognitive demands because the task configuration was not effortful. This could be due to differences in how we calibrated cognitive and physical tasks. For physical demands, we adjusted the gripping threshold to people's maximum strength, whereas we trained participants in the cognitive tasks until they reached criterion. Training might have improved response automaticity and reduced error-related risk. In addition, attributes such as the efficacy to complete a task, or the diversity of its components, have been reported to promote engagement (Bandura, 2010; Eisenberger & Cameron, 1996).

However, hand grips are regularly used to probe physical effort (Chong et al., 2017; Hogan et al., 2018; Prevost et al., 2010), and task switching has been shown to feel cognitively effortful by recruiting cognitive control (Kool et al., 2010; Shenhav et al., 2017). Furthermore, we observed a steeper decrease in acceptances over time at longer handling times for both effort domains, a hallmark of effortful exertion (Treadway et al., 2009).

The present findings have important implications for daily and clinical scenarios. Excessive aversion to delay and effort can manifest as impulsivity and apathy, respectively (Ainslie, 1975; Bonnelle et al., 2016; Chong et al., 2016; Jeffrey L. Cummings, 1993), fluctuations in which are hallmark symptoms of various mental illnesses and substance use disorders (Paulus, 2007). Lack of motivation to exert effort or wait for prospects is a distinguishing factor among mental afflictions like depression and schizophrenia (Barch et al., 2019). Within the former, medicated patients in remission are more sensitive to effort than controls, and this sensitivity is predictive of relapse (Berwian et al., 2020). Encouraging patients to engage in activities has been found to be a helpful treatment for depression (Dimidjian et al., 2006). Tailored reward schedules can successfully achieve this (Eisenberger & Cameron, 1996). However, since depressed patients also show insensitivity to rewards (Huys et al., 2013), identifying what makes effort appealing in its own right can help breach treatment-resistant cases. Our work suggests that subdividing a task can promote its pursuit, although further testing is required. More importantly, it highlights that decision scenarios should be configured to promote pursuit of desired activities, perhaps by removing alternative uses of time. These

same principles can potentially be applied to increasing productivity at work or adherence to physical exercise routines.

Conclusion

In this study, we have examined the factors that motivate engagement with different demands, emphasizing how these can be molded by the decision maker's environment. Unlike existing theories that have focused on the role of incentives and retrospective justifications (Eisenberger & Cameron, 1996; Inzlicht et al., 2018), our work highlights outcome-independent and task-specific features that can motivate the pursuit of effortful avenues, which can be potentially manipulated in order to promote engagement in clinical and daily activities.

CHAPTER 3 FIGURES

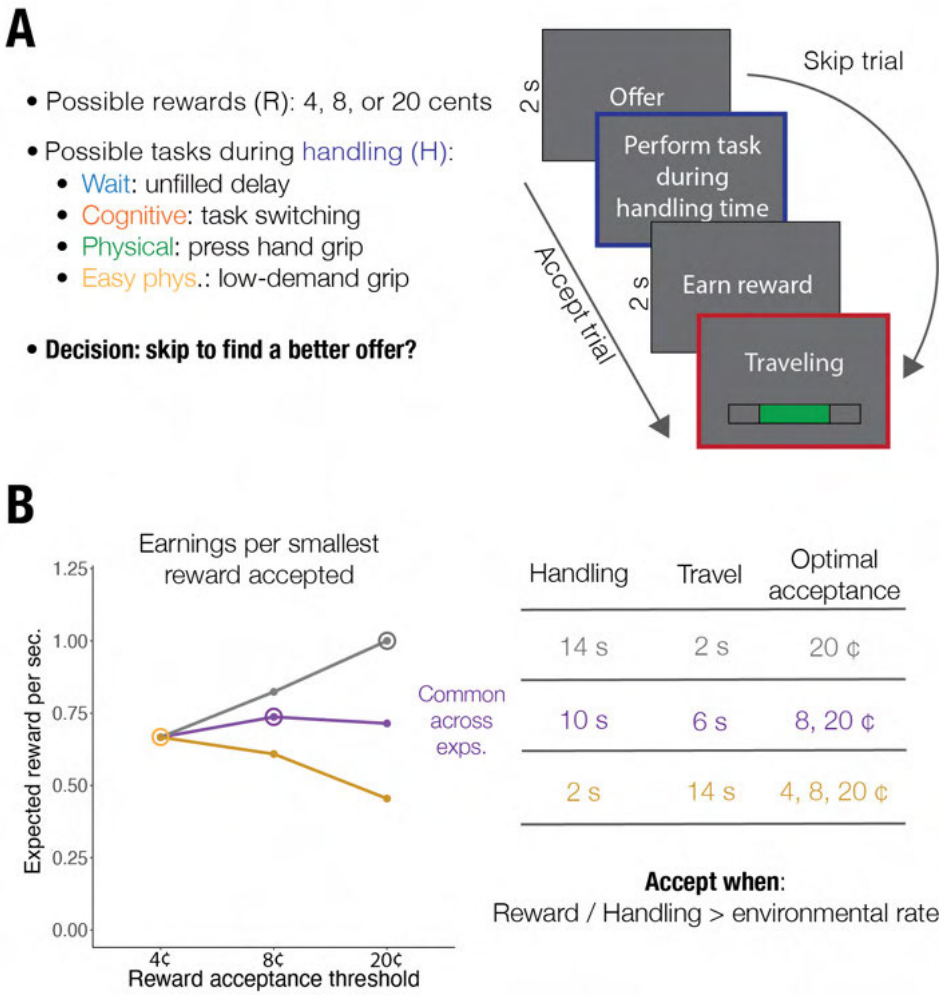

Figure 3.1: Task design and optimal behavior. **A**: General foraging trial structure. On each trial, participants were offered to earn money by sustaining effort or waiting during the handling time (2, 10, or 14 s). The end of a trial was followed by a travel time (handling and travel times always added up to 16 s). Participants could skip unfavorable trials and immediately start traveling to a potentially better offer. In Experiment 1, cost was fixed per participant and handling time varied per block. In the within-subject version, handling time was fixed at 10 s, but a combination of effort and delay trials changed per block. **B**: Possible earnings per second for each acceptance threshold (i.e. the smallest amount accepted) for each handling time. Circles denote the reward-maximizing threshold for each block, which is described in the table. Experiments shared the 10 s handling block, in which it was optimal to skip all 4 cent offers.

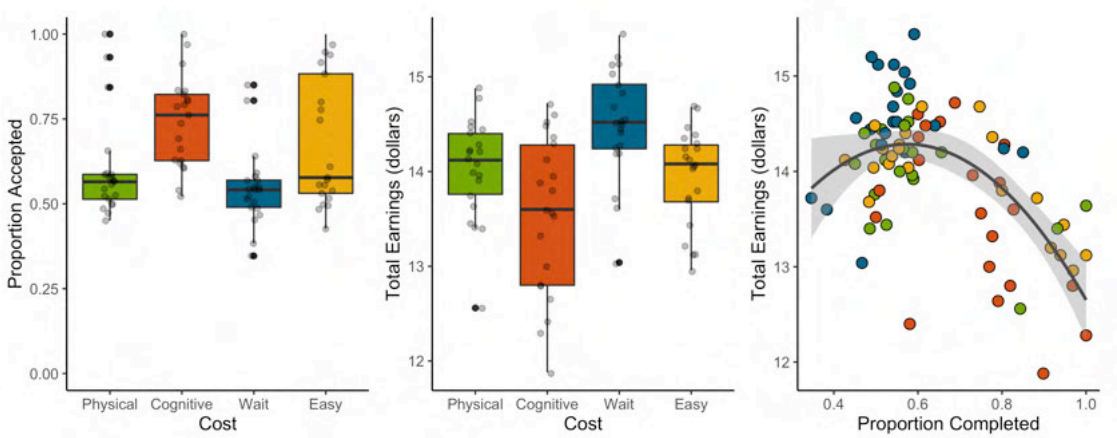

Figure 3.2: Overall choice behavior in Experiment 1. Left: Proportion accepted per cost. Middle: Total number of dollars earned by each group by the end of the experiment. Right: The relationship between proportion accepted and total earned. Consistent with the foraging design, participants who over and under accepted earned the least. These figures suggest that people are willing to pursue cognitively effortful actions at the expense of earnings.

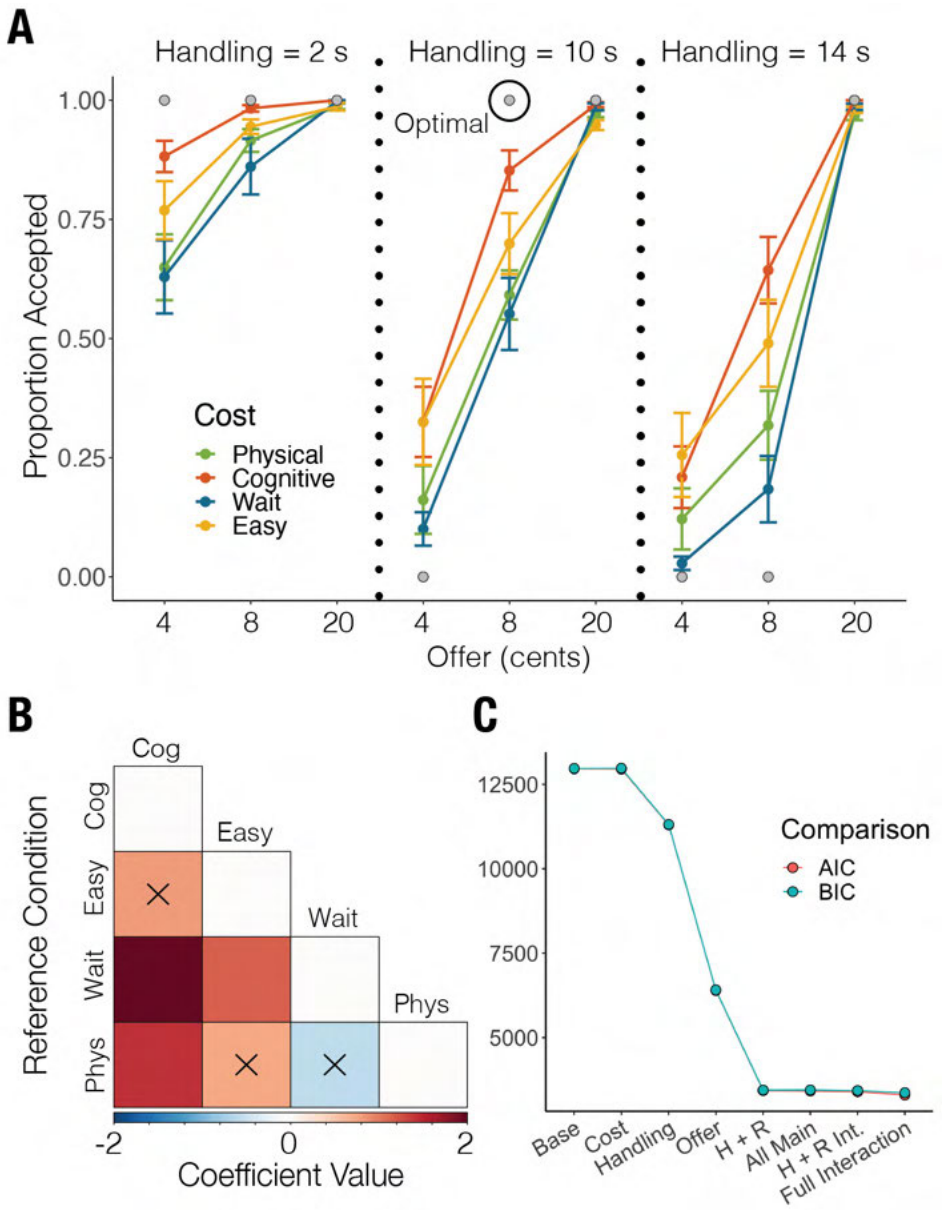

Figure 3.3. Comparing acceptance rates among costs for each reward and handling combination. **A:** Proportion accepted by each group for every combination of handling time and reward. Grey dots indicate the reward-maximizing acceptance for each combination. **B:** Matrix portraying the coefficients (color scale) and significance that resulted from switching the reference cost condition (crosses mark non-significant comparisons). Each entry shows how much more likely a group was to accept compared to the reference condition (row). **C:** AIC and BIC values for a mixed-effects model comparison, ranging from a base intercept-only to a full-interaction model. Both metrics yielded the same values across models.

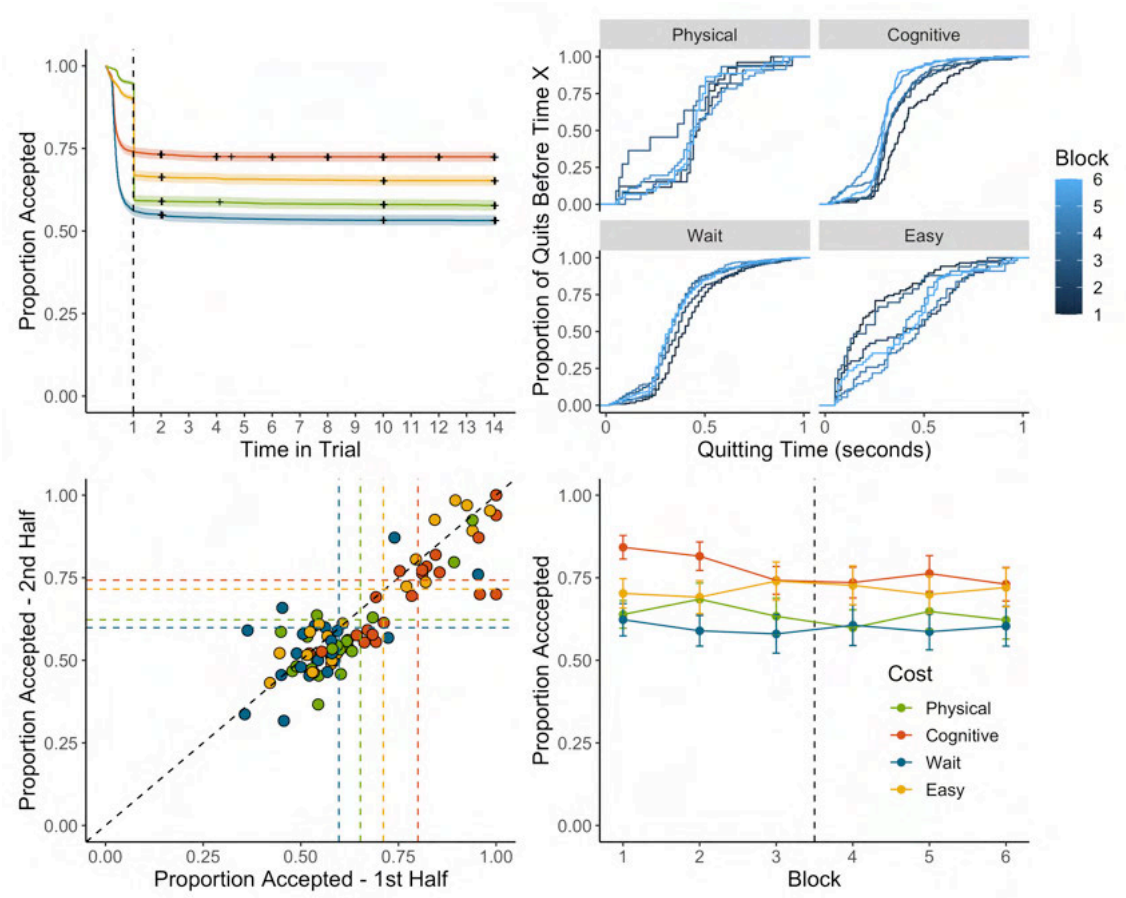

Figure 3.4: Choice consistency and stability in Experiment 1. **Top**: Survival curve (left), and distribution of times when participants quit a trial over time (right). **Bottom**: Proportion accepted in the second half as a function of first half acceptances (left). Dashed lines indicate the mean for each group. On the right, the proportion accepted over time per group. The order of blocks was repeated in the second half (indicated by the black dashed line).

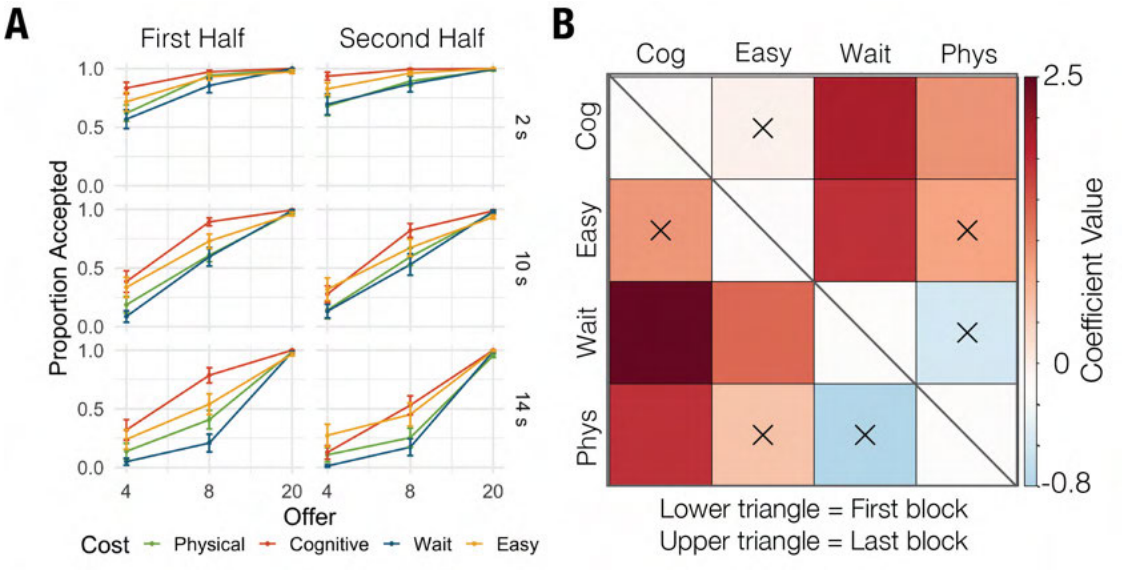

Figure 3.5: Acceptance behavior over time in Experiment 1. **A**: Proportion of acceptances per handling, reward, and half, showing that effortful demands were less frequently accepted over time. **B**: Mixed-effects coefficients denoting the comparison among costs acceptances for each half separately (mirrored along the diagonal). Even though acceptances steadily decreased in the effort groups, the relative preferences were preserved.

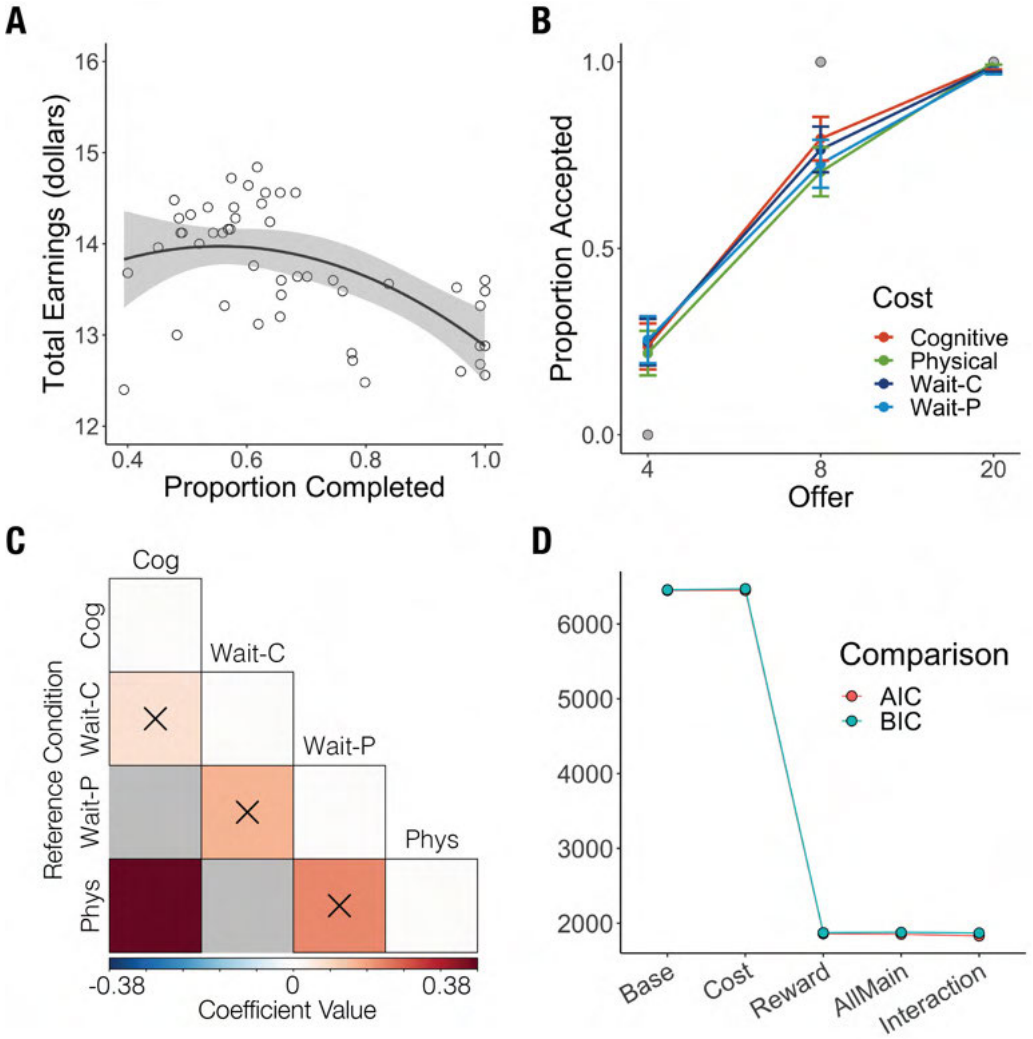

Figure 3.6: Within-subjects acceptance behavior. **A**: Quadratic relationship between trial completions and total earnings. **B**: proportion of acceptances of each reward per cost condition. **C**: Matrix of mixed-effects coefficients showing comparisons among costs (while controlling for reward amount). Comparisons between effort and wait trials from different blocks are omitted, as these are not considered in our original predictions. Crosses mark non-significant comparisons. **D**: AIC and BIC values for models of increasing complexity (a priori = AllMain).

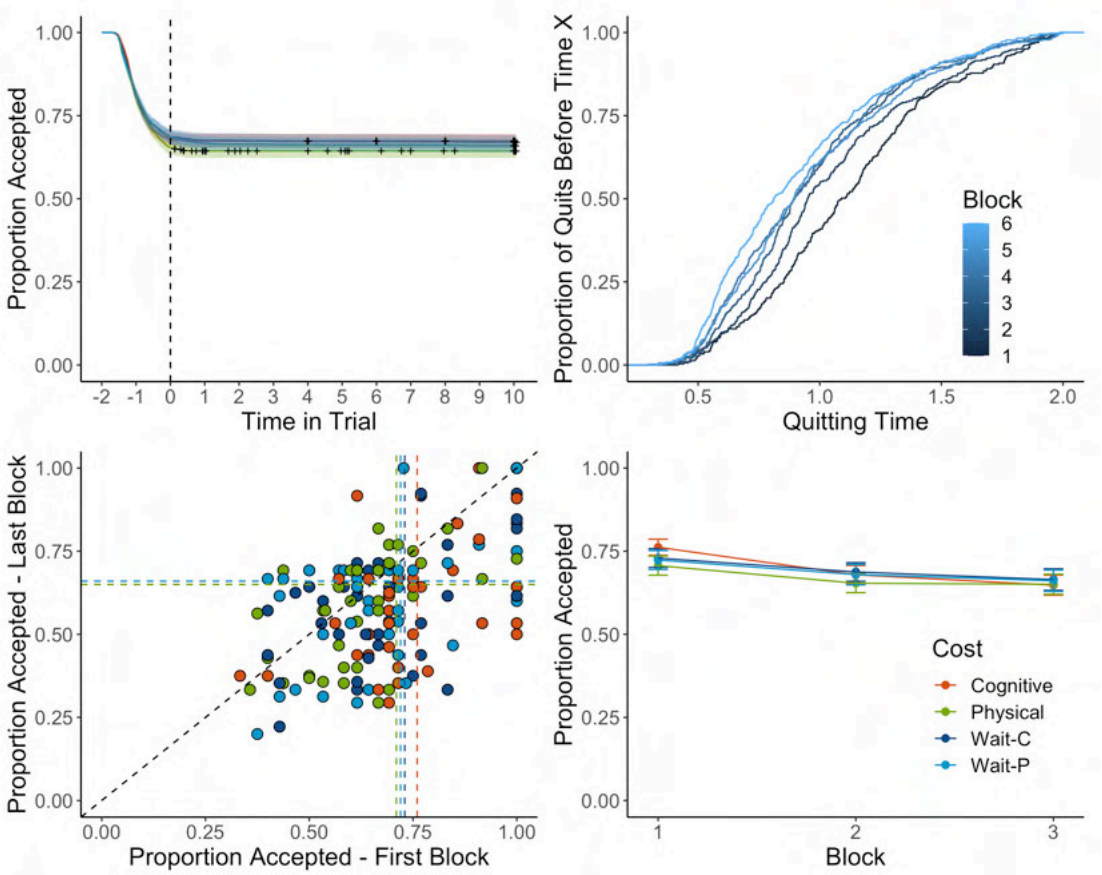

Figure 3.7: Choice consistency and stability in Experiment 2. **Top**: Survival curve (left), and distribution of quitting times over time (right). **Bottom**: Proportion accepted in the second half as a function of first half acceptances (left). Dashed lines indicate the mean for each cost. On the right, the proportion accepted every time participants experienced a given block type.

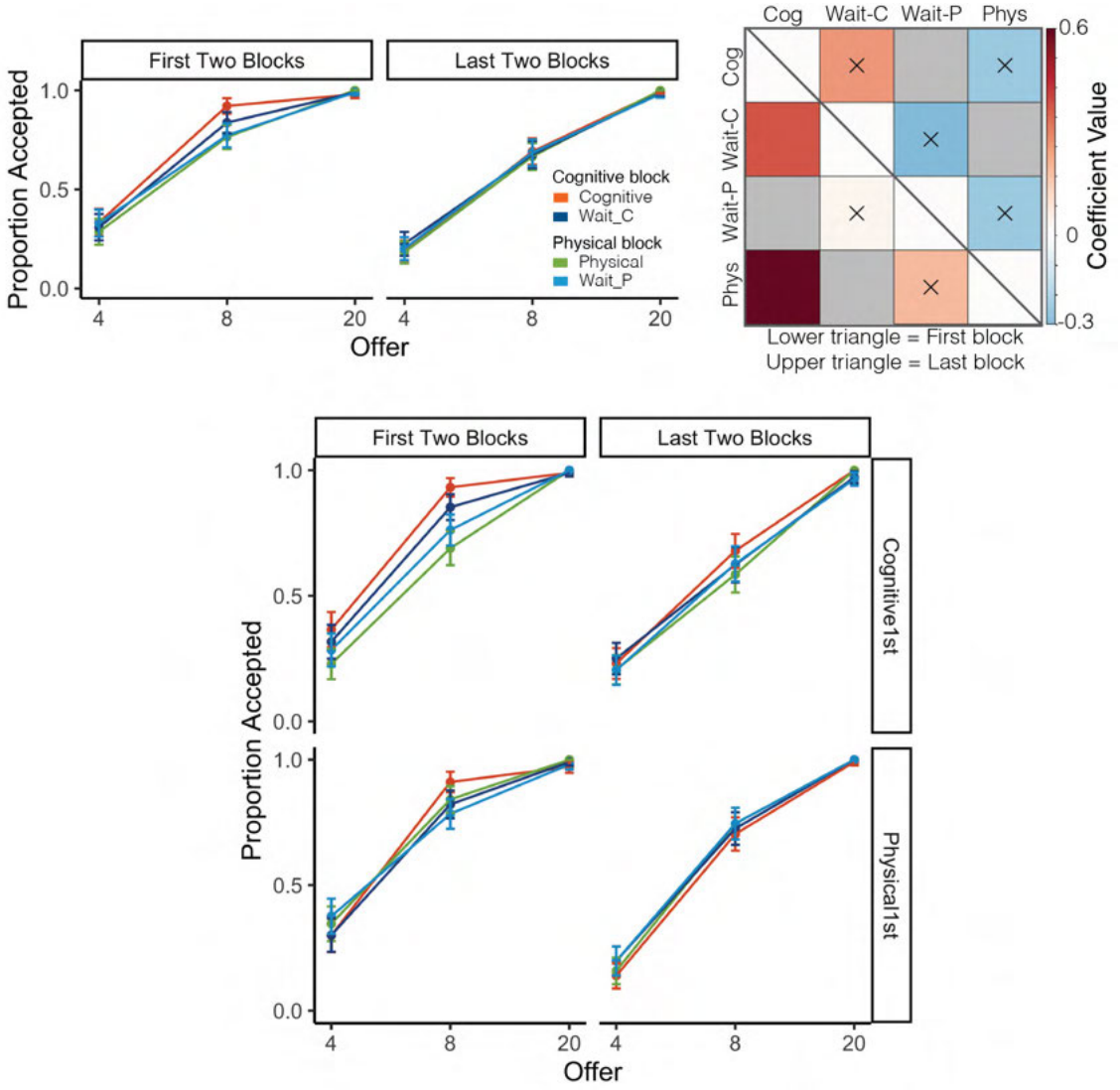

Figure 3.8: Acceptance rates over time in Experiment 2. **Top-Left**: Proportion accepted in the first two and last two blocks of the experiment. An initial preference for cognitive effort trials fades amid a global decrease in acceptance rates. **Top-Right**: Mixed effects coefficient matrix for each half separately. The pattern from the first half resembles what was observed in Experiment 1 (Figure 3). **Bottom**: Block order effects show attenuated differences for those who started with the physical configuration.

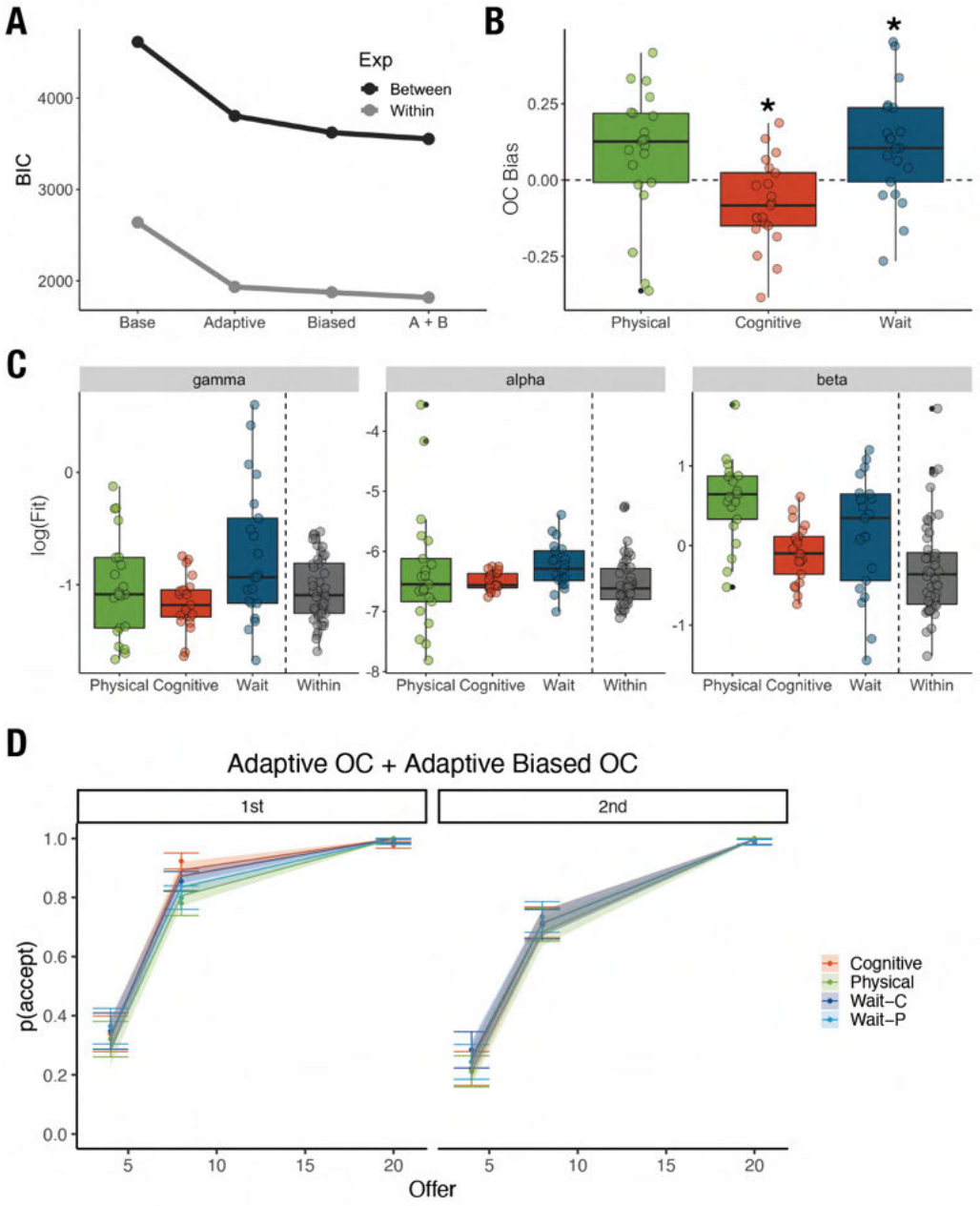

Figure 3.9: Computational modeling. **A**: BIC values for four plausible models of choice behavior for both experiments. Lower values indicate better fit. **B:** Differences in cost biases produced by each demand. Negative values suggest the cognitive group acted as if it perceived its demand as valuable. **C:** Parameter comparison portrays consistent foraging behavior across experiments. **D**: Behavior recovery of the within-subjects adaptive behavior based on the winning model. The model predictions match the initial preference for cognitive effort, and the eventual convergence of acceptances by the end of the session.

CHAPTER FOUR: INDIVIDUALIZED FUNCTIONAL MAPPING SITUATES THE DEFAULT NETWORK AS A FACILITATOR OF ECONOMIC CHOICE

Introduction

Studies of value-based decision making have consistently found valuation effects in regions belonging to the default network (DN; Bartra et al., 2013; Clithero & Rangel, 2014; Laird et al., 2009). Recent meta-analytic work has provided general quantitative support for overlap between valuation effects and DN regions in posterior cingulate cortex (PCC) and medial prefrontal cortex (mPFC) (Acikalin et al., 2017). One possible interpretation of this juxtaposition might be that choice evaluation and the introspective functions of the DN engage a unified set of underlying cognitive and neural operations (Northoff & Hayes, 2011). However, recent studies of individualized functional topography have emphasized that group-level cortical topographic overlap does not necessarily equate to overlap in individual brains. Results reveal high variability in network patterns of association cortex across individuals (Braga & Buckner, 2017; Gordon, Laumann, Gilmore, et al., 2017; Kong et al., 2018; Laumann et al., 2015), and show that this spatial variability is informative about trait-like behavioral characteristics (Bijsterbosch et al., 2018; Kong et al., 2018).

Individual variation in fine-grained network topography poses challenges for the mapping of cognitive phenomena, with group-level aggregations obscuring otherwise dissociable effects at the individual level (Fedorenko et al., 2012; Michalka et al., 2015). It is therefore possible that our understanding of macroscale network functions has been misguided by omitting individual variability. Our results in Chapter 2 showed that the

mPFC and PCC regions of canonical DN could be divided into at least two sub-networks (e.g. DN and non-DN) that were heterogeneously organized across individuals (Toro-Serey, Tobyne, & McGuire, 2020), further highlighting the need to explore any potential functional overlap within this this network at the individual-level. Indeed, a comparable fractionation of canonical DN recently positioned two of these sub-networks as seats of distinct introspective functions, namely theory of mind and episodic projection, that were traditionally grouped together (DiNicola et al., 2020). Based on this idiosyncratic partitioning and the methodology established in our previous work (Toro-Serey et al., 2020), here we ask whether valuation effects are facilitated by the DN or by adjacent sub-networks within mPFC and PCC.

There are two steps to addressing this question. First, we must reconsider whether the spatial configuration of the DN generalizes from rest to tasks at the individual level. The DN was originally described as a distributed set of regions that cohesively deactivated as attention was externally oriented (Laird et al., 2009; Shulman et al., 1997). Early results indicated that this network could be identified either by allowing individuals to introspect during wakeful rest, or by identifying regions that decreased their activity as subjects were confronted with increasingly demanding tasks (Buckner & DiNicola, 2019; Greicius et al., 2003; Raichle et al., 2001). Parallel sets of studies have shown that resting-state networks generally align with task-evoked activity patterns (Fox et al., 2005; Smith et al., 2009; Yeo et al., 2015; Yeo et al., 2011). This has led to proposals that intrinsic functional networks provide a baseline configuration that is partially recruited by task demands (Cole, Bassett, Power, Braver, & Petersen, 2014; Uddin et al., 2019).

However, most studies that have probed the alignment of task-dependent and task-independent networks have employed analyses on aggregated data (Cole et al., 2014; Di, Gohel, Kim, & Biswal, 2013; Smith et al., 2009; Yeo et al., 2015), and topographic comparisons within individuals have either not examined the fine-grained spatial composition of functional networks computed with diverse tasks, or omitted discussion of these within mPFC (Gratton et al., 2018; Kraus, 2020; Tavor et al., 2016). Here we address this gap by assessing whether, and to what extent, the fractionated functional topography of the DN is preserved when measured using rest, task-deactivation, and valuation tasks.

The second step is to interrogate which of the resulting sub-networks displays valuation-sensitive activity profiles. The multifaceted nature of decision making can potentially complicate this endeavor. Even though ventral mPFC (vmPFC) has been proposed to compute the subjective value of prospects in a category-independent manner (Chib, Rangel, Shimojo, & O'Doherty, 2009; McNamee, Rangel, & O'Doherty, 2013), vmPFC valuation activity occurs in the context of a different broader pattern of brain activity depending on the decision stage or context. For example, both the prospective evaluation stage of a choice and the later outcome receipt stage recruit swaths of vmPFC and ventral striatum (VS), but evidence for the engagement of PCC during the outcome receipt stage equivocal across studies (Bartra et al., 2013; Clithero & Rangel, 2014). Even within the prospective evaluation stage, different zones within mPFC seem to consider an item's desirability versus its congruency with the current task goals (Hare et al., 2008; Shenhav & Karmarkar, 2019). This could mean that, depending on which

factors people emphasize, different sub-networks (or different subregions of a network) might drive behavioral choices. Study 2 showed that decision makers can change their attitudes towards well-known decision environments over time, raising the possibility that an individual's underlying network could similarly shift. Therefore, it is worth examining if the same brain networks show value-related activity modulation during the appraisal and outcome of a decision.

In this study, we tested whether valuation and default networks overlap within mPFC and PCC when identified and compared in individuals. Using multi-session task-based and resting resting-state fMRI data, we investigated 1) if data-driven estimates of brain network community structure during a value-based decision-making task were heterogeneous across individuals; 2) if this network's topographic subdivisions estimated from the decision-making task were was at least partially distinguishable from those estimated from resting-state and task-deactivation runs; 3) whether valuation effects could be found in the default or non-default sub-networks; and 4) whether sub-networks displayed unique connectivity profiles with the striatum. In addition, we explored if the functional topography of valuation was reconfigured across sessions depending on how people construed the value-based decision-making task. We hypothesized that resting-state and task deactivation would recover similar individual-specific DN layouts, consistent with previous work (Toro-Serey et al., 2020), and that these would differ from valuation networks as long as individuals focused on value estimation rather than rule-based strategy to guide decisions. Our findings indicated that spatial network configurations in rest and task deactivation indeed followed the same separation of DN

and non-DN seen previously, displaying high topographic idiosyncrasy in mPFC. However, contrary to expectations, we found that the same organization was present during valuation, and that the DN sub-network participated in the evaluation of prospects and outcomes. These results introduce individual-level evidence that functional network layouts are conserved across tasks and the resting-state. The findings offer guidance for future empirical and theoretical investigations of a shared cognitive architecture that could subserve both introspective and evaluative operations.

Experimental Design and Methods

Participants

We recruited 19 healthy volunteers from the Boston University community and the greater Boston area (10 female, average age = 21, age range = 18 - 33) to complete a two-session, multi-task fMRI protocol. Individuals provided written informed consent prior to each experimental session and were compensated for their time, as approved by the Boston University Institutional Review Board (IRB). We screened participants for a history of neurological illness/injury and MRI contraindications, and all were right-handed. One participant was removed due to high motion during scanning (see *Functional Data Preprocessing* for details).

Experimental and Task Design

In each session, participants performed 6-minute runs of an n-back working memory task (3 runs per session), a value-based decision-making task (6 runs per session), and resting fixation (3 runs per session), all coded in PsychoPy 2.0 (Peirce et al.,

2019) running on a Macbook Pro (Fig. 4.1). All tasks were compatible with a single-handed, two-button, MRI compatible response device that was introduced to participants prior to scanning (with the index finger assigned to the left button, and the middle finger assigned to the right button). In the n-back task, a stream of capital letters was presented one at a time in the middle of the screen. For every block of 10 letters, participants had to follow one of three pre-trained rules (with the applicable rule cued at the start of each block). The first, a 0-back rule, asked individuals to press the left button whenever the letter was an X, and the right button otherwise. The second rule was a 1-back, in which participants pressed the left button whenever the current letter was a repeat of the previous one, and the right button otherwise. The final rule was a 2-back, in which participants pressed the left button if the current letter matched the one two letters prior, and the right button otherwise. Within each 10-item block, 5 unique letters were sorted so that 1 to 3 targets of each type appeared, with only one type being relevant in each block. Target rules were presented in a pseudorandomized order during each run in order to avoid immediate repeats (with the order changing per run).

The decision-making task was a card wagering game (Fig. 4.1B; cf. Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Preuschoff, Quartz, & Bossaerts, 2008), in which participants guessed on each trial the direction of the difference between cards drawn from two independent decks (cards ranging from 2 to 9, "shuffled" on each trial to remove history-dependence). After viewing the card from Deck 1, participants wagered whether the card from Deck 2 would be higher or lower. When the second card was revealed, participants gained or lost points according to the size and direction of the

difference between cards. Inter-trial-intervals were jittered to facilitate deconvolution of trial-related BOLD activity. Extreme face values for Card 1 (2 or 9) presented the opportunity for highest gains on average, with an expected value (EV) of 3.5 points (i.e. mean of possible differences), whereas cards with face values in the middle of the range (5 or 6) had a low EV of 0.5. The task separately manipulated offer value and reward prediction error (RPE, the difference between the reward outcome and the EV), which could range from -3.5 to 3.5 points per trial regardless of the value of Card 1. Importantly, participants were explicitly informed that the reward-maximizing strategy was to bet higher if Card 1 was 5 or below, and bet lower if it was 6 or above.

In the fixation task, participants had to lie awake while maintaining their gaze on a cross in the middle of the screen. Participants were told they could let their mind wander, but that they should remain awake with their eyes open and avoid mental activities such as counting. All participants practiced these tasks during a behavioral training session that occurred one or two days before the first scan.

Functional MRI Acquisition

All scans were acquired at Boston University's Cognitive Neuroimaging Center on a 3T Siemens Prisma (Siemens, Erlangen, Germany) fitted with a 64-channel head coil. In each session, anatomical images were acquired using both a T1-weighted multi-echo magnetization-prepared rapid gradient-echo (MPRAGE) sequence (Sagittal; TR = 2530 ms; TE = 1.32 ms, 3.19 ms, 5.11 ms, 7.03 ms; FA = 7°; 0.8 mm isotropic voxels; 224 slices; FOV = 256 x 256 x 180 mm; GRAPPA acceleration 2) and a T2-weighted sequence (Sagittal; TR = 3200 ms; TE = 564 ms; 0.8 mm isotropic voxels; 224 slices;

FOV = 256 x 256 x 180 mm; GRAPPA acceleration 2). Functional T2*-weighted BOLD acquisition parameters were configured to minimize signal dropout and distortion in vmPFC, including oblique axial slices (1.75 mm thickness), TR = 1440 ms, TE = 25 ms, simultaneous multislice (SMS) acquisition = 3, GRAPPA acceleration = 2, and larger in-plane voxels (2.25x2.25x1.75 mm). For precise distortion correction of functional scans, we acquired a matched pair of reference images every 4 runs using spin-echo echoplanar imaging (EPI) with opposite phase-encoding directions.

MRI Data Preprocessing

Each individual's MRI dataset was separately preprocessed using fMRIPrep 1.4.1 (Esteban, Blair, et al., 2018; Esteban, Markiewicz, et al., 2018; RRID:SCR_016216), which is based on Nipype 1.2.0 (Gorgolewski et al., 2011; Gorgolewski et al., 2018; RRID:SCR_002502). Much of the description of the preprocessing steps in the following 3 paragraphs was slightly modified from a boilerplate, copyright-free output automatically generated by the fMRIPrep software.

Anatomical data preprocessing: Each session's T1-weighted (T1w) images was corrected for intensity non-uniformity (INU) with *N4 Bias Field Correction* (Tustison et al., 2010), distributed with ANTs 2.3.1 (Avants, Epstein, Grossman, & Gee, 2008; RRID:SCR_004757). The T1w-reference was then skull-stripped with a Nipype implementation of the *antsBrainExtraction.sh* workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using *fast* (FSL 6.0.1, RRID:SCR_002823, Zhang, Brady, & Smith, 2001). A T1w-

reference map was computed after registration of both T1w images (after INU-correction) using *mri_robust_template* (FreeSurfer 6.0.0, Reuter, Rosas, & Fischl, 2010). Brain surfaces were reconstructed using *recon-all* (FreeSurfer 6.0.0, RRID:SCR_001847, Dale, Fischl, & Sereno, 1999b), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of *Mindboggle* (RRID:SCR_002438, Klein et al., 2017). Volume-based spatial normalization to standard space was performed through nonlinear registration with antsRegistration (ANTs 2.3.1), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c (MNI152NLin2009cAsym) (Fonov, Evans, McKinstry, Almli, & Collins, 2009, RRID:SCR_008796).

Functional data preprocessing: For each of the 24 BOLD runs per subject (across all tasks and both sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A deformation field to correct for susceptibility distortions was estimated based on two EPI references with opposing phase-encoding directions, using 3dQwarp (AFNI 20190100; Cox & Hyde, 1997). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using *bbregister* (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with nine degrees of freedom to account for

distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using *mcflirt* (FSL 6.0.1; Jenkinson, Bannister, Brady, & Smith, 2002). BOLD runs were slice-time corrected using *3dTshift* from AFNI 20190100 (Cox & Hyde, 1997; RRID:SCR_005927). The BOLD time-series were resampled to surfaces on the fsnative space. The BOLD time-series (after slice-timing correction) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series are referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space.

Four confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD) and three region-wise global signals. FD was calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al., 2014). The three global signals were extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (aCompCor; Behzadi, Restom, Liau, & Liu, 2007). Principal components were estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off). Components were calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD to-T1w transformation).

Components were also calculated separately within the WM and CSF masks. For each aCompCor decomposition, the *k* components with the largest singular values were retained, such that the retained components' time series were sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, or combined). Frames that exceeded a threshold of 0.5 mm FD were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using *antsApplyTransforms* (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using *mri_vol2surf* (FreeSurfer). Many internal operations of fMRIPrep use Nilearn 0.5.2 (Abraham et al., 2014; RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.

Next, we applied additional steps to each scanning run similar to those described in previous work (Toro-Serey et al., 2020). First, both native surface and volumetric BOLD time series were converted into the Human Connectome Project’s fs_LR_32k grayordinate space. Then we computed transformation matrices from the standard MNI152Nlin2009cAsymm T1w to MNI152Nlin6thGen using *flirt* (Jenkinson & Smith, 2001), and applied the transformation to all preprocessed BOLD time series. We then combined the surface and subcortical volume using the HCP’s workbench command

tools (Marcus et al., 2011). Second, we performed censoring of high motion volumes (FD > 0.5 mm) by using single-time-point indicator variables to regress out each outlier volume in addition to the two time points adjacent to each outlier. This regression also included the mean gray matter signal, motion parameters (6 dof), and the 5 highest aCompCor components. The residual time series signals were then bandpass filtered for rest scans (allowed range between 0.009 and 0.08 Hz) and high-pass filtered (> 0.01 Hz) for task data. Finally, the BOLD time series were demeaned and concatenated to produce session-specific datasets per subject.

Functional Network Analyses

In order to assess the individual-specificity of task-based functional brain topography, we employed a streamlined version of our previous analytical pipeline, which is detailed in Chapter 2 and in Toro-Serey et al. (2020).

Functional network definition and partitioning. We used the search space from Toro-Serey et al. (2020), including mPFC, PCC, with the addition of striatal regions (Caudate, Putamen, and Accumbens, as indexed in the HCP 32k standard space). Striatal regions were added on the basis of their well-known role in value-based decision making and learning (Haber & Knutson, 2010; Schultz, 2016), but our analyses were mostly concentrated on the ventral striatum (i.e. the Accumbens region from the grayordinate space) due to its established functional relationship to both EV and RPE (Bartra et al., 2013). This addition allowed us to interrogate the degree of functional connectivity of these subcortical regions with the resulting cortical networks. This resulted in a total of

8629 grayordinates selected for analysis. All network analyses were performed in R 3.5.2 (R Core Computing Team, 2017).

Using all grayordinates in the search space, we produced individual-specific functional networks separately for each task and session (2 sessions x 3 tasks per individual), and analyzed each network's community structure using spectral partitioning. The use of spectral partitioning permitted results for all three tasks to be compared in a common format (Fiedler Vector values) even though the inputs differed. Resting-state functional networks were produced by calculating the Pearson correlations of BOLD time series from every pair of grayordinates, resulting in an 8629 by 8629 correlation matrix. The same procedure was used to estimate networks from n-back runs. We estimated networks from decision-making runs by first converting each run's BOLD time series into a "beta series" of per-trial BOLD response amplitudes. The single-trial BOLD amplitudes were estimated using a "least-squares-separate" approach (Mumford, Turner, Ashby, & Poldrack, 2012). For each grayordinate, we performed a series of linear regressions, which each modeled the time series as a function of two variables: one regressor modeling the single event of interest (e.g., presentation of Card 1 on one individual trial), and another modeling the event in all other trials. Prior to fitting, the regressors were convolved with a double-gamma hemodynamic response function (HRF) using the neurosim package (Welvaert, Durnez, Moerkerke, Verdoolaege, & Rosseel, 2011). This produced beta coefficients for each separate trial, which were then concatenated to generate a beta series. In short, this method transformed each grayordinate's raw BOLD time series into a series of coefficients that captured trial-wise

variability. Functional networks were then produced by correlating these beta series between every pair of grayordinates. Separate versions of the analysis used beta-series regressors timed to correspond to either the Card 1 or Card 2 onset event. In that way, the beta-series approach enabled the analysis to focus on variance associated with particular trial events.

As in previous work, we transformed the correlation values from the functional networks to facilitate the construction of the normalized graph Laplacian (i.e. exponentiated Fisher's z), followed by eigen decomposition of the similarity matrix. We focused on the eigenvector associated with the second-to-lowest eigenvalue of each network (the Fiedler Vector, FV). Because of the data-driven nature of the clustering, we oriented the FV values so the sub-network with most coverage of area 7m (i.e. the default sub-network) was positive for all subjects. We confirmed this alignment visually, and flipped the signs of the maps whose DN-network insufficiently covered 7m, but had otherwise DN-like topographic features (e.g. alignment in ventral mPFC with the superior rostral sulcus). We performed this manual alignment on 42 out of the 108 total partitions (18 participants x 2 sessions x 3 tasks). This procedure resolved instances in which opposite labels would have been given to largely overlappig networks from the same individual. Finally, we refined the layout of each network by thresholding the FV, with values above 0.01 assigned to the DN, and those lower than -0.01 to non-DN (all other values were set to 0 to facilitate visual comparisons). The FV was then binarized by sign, with values of 1 denoting a DN affiliation, 0 for non-DN, and 0.5 for everything else (so that calculation of averages were not biased).

Partition validation and comparison. Before comparing networks, we visually validated the individual-specific fixation-based DN layouts by visually comparing them to those produced by resting-state data from Toro-Serey et al. (2020). Next, we compared the networks produced with n-back data to multilevel GLM contrasts of task-deactivation. For each subject and session, BOLD time series in native surface space (fsnative) were smoothed at 5mm full width at half maximum (FWHM), intensity normalized, and high motion volumes (FD > 0.5 mm) were discarded (along with the time point directly before and after). A linear model contrasting 2-back blocks to 0-back blocks was then fit at each vertex using Freesurfer (Fischl, 2012), controlling for head translation and rotation (6 dof) as confounds. The resulting individual maps of two-sided t-values were remapped to HCP's fs_LR_32k to permit the overlay of connectivity- and task-based n-back maps (using the procedure outlined above), and thresholded by an absolute value of 3.

Next, we examined the idiosyncrasy of each network's functional topography, as well as the degree of topographic alignment among networks estimated from different run types. We assembled a similarity matrix that displayed the ARI between partitionings among every combination of individual, session, and run type, sorted by run type and then by individual (the ARI provided a chance-corrected comparison of how many grayordinates were paired together across partitionings). Values directly adjacent to the main diagonal of this matrix yielded session-to-session reliability estimates, and ARI ratios of within-participant (between-session) over between-participant elements served to quantify the degree of idiosyncrasy of these networks in mPFC, PCC, and striatum. The diagonals of the surrounding blocks of the matrix were used to estimate the similarity

across run types for each individual and session (3 comparisons), and paired permutations (5000 iterations) were used to test (1) which network comparison had the highest similarity; and (2) whether such similarity persisted across sessions.

Valuation effects in the individualized DN: We determined the sensitivity of each network to valuation effects in two separate steps. First, using each individual's spatially averaged time course for each network and brain area (mPFC, PCC, and striatum), we plotted the distribution of coefficients from the beta series as a function of the value of each explanatory variable of interest (e.g. RPE values from -3.5 to 3.5) across individuals. Second, we fit mixed-effects linear models to examine whether the mean activity from each fixation-based network was modulated by the specific offers and outcomes presented during the decision-making task. For each individual and session, we used DN and non-DN networks estimated from fixation runs to select grayordinates. We then spatially averaged the network time series from decision-making runs, and concatenated them across sessions. A first model included a constant term, parametric regressors for expected value (EV) and response time (to control for trial difficulty as a confound). A second model included a constant term and a parametric regressor for reward prediction error (RPE). Non-response trials were removed (proportion of skipped trials = 0.07). Regressors contained values at the onset time of the corresponding event (e.g. Card 1 presentation for EV, and Card 2 presentation for RPE), and were convolved with a double-gamma canonical HRF using the neuRosim package (Welvaert et al., 2011). We then concatenated the time series and associated regressors across participants for modeling. In both cases, the mixed effects models were configured with fixed and

random effects for intercept and slopes per subject. This single modeling step emulated the traditional multi-level approach often performed in a two-stage analysis in neuroimaging toolboxes (i.e. estimating coefficient maps per subject and comparing them against the null in group-level tests).

Topography and Behavioral Performance

We performed exploratory analyses to examine whether the degree of covert engagement with the cards task affected the similarity of the valuation and fixation partitionings. For example, a participant who evaluated the expected value given by Card 1 might have responded slower to values in the middle of the range (that is, when there was a high chance of losing points). Conversely, for a decision maker who trusted the provided reward-maximizing strategy, response times (RT) might have become uniform across all values as they gained experience implementing the rule.

We started by examining variability in choice behavior across individuals. First, we determined the accuracy with which participants followed the disclosed reward maximizing strategy (guess that Card 2 will be higher if the value of Card 1 is 5 or below). We performed mixed-effects logistic regressions that predicted the probability of guessing a higher Card 2 value based on Card 1 value. The model had random intercepts per subject, as well as fixed and random effects for card value and session. We then fit a linear mixed-effects model to test the relationship between expected value and response times (log-transformed), with random intercepts for each subject, as well as fixed and random slopes for expected value and session.

Using the results from the behavioral analyses, we evaluated 1) for models whose fixed session coefficients achieved statistical significance, if the absolute value of each subject's random session coefficients correlated with their respective ARI between session 1 and session 2 decision-making runs (i.e. if a person's change in behavior across sessions resulted in session-to-session dissimilarities for decision-making networks); and 2) whether slopes for RT on a given session were correlated with the ARI between networks estimated from fixation runs and from decision-making runs.

Results

In this study, we tested the hypothesis that the valuation network could be distinguished from the default network when accounting for neural and behavioral sources of individual variability. We addressed this question in three levels. First, we evaluated whether the spatial configuration of functional networks underlying valuation differed from the configuration estimated from resting state data and task-deactivation. Second, we probed the existence of valuation effects in the default network. Finally, we explored the connection between behavioral variability and topographic similarities.

Qualitative Validation of Partitionings

Prior to quantification of topographic similarities, we qualitatively evaluated networks in three ways: 1) by comparing group-averaged patterns across tasks; 2) by examining their overlap in individuals; and 3) by assessing whether connectivity and task-deactivation produced similar individualized layouts, as is often assumed in the literature. Steps 1 and 2 were also validated against the topographic patterns observed in previous work (Braga & Buckner, 2017; Toro-Serey et al., 2020).

Fig. 4.2 shows that cortical zones affiliated with the DN across individuals tended to be the same across tasks, although some variability in mPFC was evident (especially for fixation networks). In all cases there was broad coverage of the DN in PCC, although it was concentrated posteriorly near 7m as dictated by our labeling scheme. The DN was consistently present anteriorly in ventral and dorsal mPFC, with pockets of non-DN near pregenual area a24 and towards dorsal anterior cingulate cortex. These group-level patterns resemble the subdivision of canonical DN proposed recently, and match the cortical span of what is usually considered DN (e.g. Yeo et al., 2011). Striatal voxels generally lacked a consistent association with either network across individuals, with the exception of a small zone in ventral striatum that was preferentially affiliated with the DN (median number of DN grayordinates in ventral striatum = 7, IQR = 4 – 17; median for non-DN = 25, IQR = 14 - 46). Visual inspection of this area suggested that this cluster was present in approximately the same coronal slice (y = 66) in all subjects, although there were no voxels that affiliated with the DN more than 50% of the time across individuals. On the other hand, grayordinates in caudate and putamen consistently showed FV values near 0 (median FV values across tasks = -0.003, IQR = -0.004 - -0.002), suggestive of highly variable clustering (see Fig. 2.4). Because of the unreliable clustering both within and across individuals in dorsal striatum, in addition to the consistent (though unbalanced) affiliation in ventral striatum across individuals, we focused subsequent visual evaluations of individual differences on cortical structures.

Examples of individual-level partitionings of the left hemisphere can be found in Fig. 4.3, with the first three columns representing networks computed from different run

types. Partitionings derived from fixation runs shared the general topographic attributes described in group averages above, but with idiosyncratic arrangements that replicated those from Toro-Serey et al. (2020). Thresholded fixation-based DNs (i.e. FV > 0.01, according to the median threshold of our previous work) are overlaid as black outlines on networks estimated from decision-making runs and n-back runs (columns 2 and 3). This juxtaposition showed a considerable degree of similarity between resting and task networks that respected individual heterogeneity, although task-based networks sometimes extended outside fixation outlines in mPFC. For example, the putative DN covered the anterior vmPFC area 10v in both decision-making and n-back runs for subject 770 (blue circle in Fig. 4.3).

Finally, we compared GLM-based and connectivity-based DN estimations. The DN is often described as a set of distributed regions that cohesively reduce their activity during elevated demands for working memory or externally directed attention (Buckner & DiNicola, 2019; Laird et al., 2009). Therefore, we sought to validate the connectivity-based DN with patterns of working-memory-load-related deactivation in individuals. The rightmost column of Fig 4.4 shows individual maps of t-statistics from whole-brain, vertex-wise, two-sided linear regressions comparing 2-back versus 0-back activity patterns during n-back runs, overlaid with the network outline of the DN estimated using the same n-back data (i.e. column 3). As expected, areas that deactivated for higher demands (in blue) were well aligned with the connectivity-based DN in PCC and dorsal mPFC. However, high t-values were at least partially interdigitated with network boundaries in vmPFC. This pattern was clearest for subjects 731 and 770, for whom the

DN outline failed to include some strongly deactivated areas along the suprarostral sulcus. This discrepancy was not due to reliability issues in either GLM (mean within-subject correlation of vmPFC t-maps across sessions = 0.65, SD = 0.23; vmPFC mask defined as in Fig. 2.12B) or connectivity maps (discussed below).

These results visually replicate the idiosyncratic subdivision of canonical DN proposed in Study 1 (which is missed in group-averages), and suggest that this organization is common across tasks. However, they also point at the possibility that topographic patterns of task-deactivation and resting-state functional correlations might not be uniformly coupled across the brain (as is commonly assumed).

Quantitative Network Comparisons

As a first step in quantifying the observed similarities among networks, we computed pairwise ARI values across 108 network parcellations that were estimated independently for each subject, session, and run type. This similarity index gives the chance-corrected proportion of network nodes that were jointly clustered in both partitionings, heavily penalizing discrepancies (see Fig. 2.4 for empirical and simulated examples). All comparisons were performed on binarized networks thresholded at an absolute FV value of 0.01 (see Experimental Design and Methods for details), and all results from decision-making runs were based on beta series aligned with Card 1 onset. The latter condition was implemented because the resulting functional topographies were consistent regardless of the cards event used to compute them (Fig. 4.4B, bottom left). Fig. 4.4A shows a comprehensive view of these similarities for all possible pairings in each cortical region. The visible diagonal structure in every tile of these matrices

indicates that similar subject-specific network patterns were recovered for all run types, as observed qualitatively above.

Next, we used the values from these matrices (in addition to the ventral striatum, not pictured) to quantify the reliability, idiosyncrasy, and cross-task similarity of these networks. The top row of Fig. 4.4B shows session-to-session ARIs per region and task (i.e. the within-subject cells directly adjacent to the main diagonal of the matrix). In spite of the consistent (albeit small) recruitment of ventral striatum by the DN, patterns in this area lacked reliability across tasks (Fixation: median ARI = 0.02, IQR = -0.01 - 0.07; Cards: median ARI = 0.04, IQR = 0.01 - 0.07; N-back: median ARI = 0.02, IQR = -0.01 - 0.08). This could be due to the unstable clustering mentioned above. In cortex, networks estimated with fixation data had the lowest reliability (mPFC: median ARI = 0.16, IQR = 0.04 - 0.23; PCC: median ARI = 0.24, IQR = 0.1 - 0.31), whereas n-back (mPFC: median ARI = 0.3, IQR = 0.19 - 0.45; PCC: median ARI = 0.3, IQR = 0.15 - 0.45) and cards (mPFC: median ARI = 0.35, IQR = 0.31 - 0.4; PCC: median ARI = 0.39, IQR = 0.33 - 0.43) tasks produced very reliable spatial configurations. The reliability of fixation networks was lower than those in previous work (Toro-Serey et al., 2020), although this did not preclude the identification of idiosyncratic patterns discussed next.

We measured each subject's degree of topographic idiosyncrasy by computing the ratio of their within-subject (across-session) similarity over their mean similarity with every other individual for the same run type. These ARI ratios are displayed on the middle row of Fig 4.4B, with values greater than one denoting individual specificity. As expected, the low reliability of VS translated into low ratios (median across tasks =

1.001, IQR = 0.67 - 1.8). However, trends for the fixation networks replicated those from Toro-Serey et al. (2020), such that both mPFC (median = 7.12, IQR = 3.88 - 8.27) and PCC (median = 3.4, IQR = 2.76 – 4.57) yielded ratios significantly greater than one (signed rank tests; $V_{mPFC} = 163$; $V_{PCC} = 167$; both $p < 0.001$), with mPFC being significantly more idiosyncratic than PCC (paired permutation with 5000 iterations, $p < 0.01$; Cohen's D = 0.94). The same idiosyncrasy was observed for networks estimated from decision-making runs (mPFC: median = 6.80, IQR = 6.02 – 8.55; PCC: median = 4.33; IQR = 3.91 – 5.24) and n-back runs (mPFC: median = 6.55, IQR = 4.97 – 7.68; PCC: median = 6.54; IQR = 3.32 – 8.02), although mPFC ratios were significantly higher than PCC only for decision-making runs (paired permutation, $p < 0.001$; Cohen's D = 1.65).

Finally, we assessed the degree of topographic alignment across tasks (i.e. diagonals of the upper-right and lower-left tiles of the similarity matrices), but this time using networks that spanned the whole search space. The boxplot on the bottom right of Fig. 4.4B shows the similarity between every pair of task networks within individuals (per session). The dashed line is the median inter-session ARI from the cards network (0.35), which provides a reference ceiling point to evaluate the comparisons. This plot mirrors previous qualitative observations, suggesting that task-based networks were better aligned. Paired permutations indicated that each comparison was not significantly different between sessions (all $p > 0.1$), so we grouped partitionings across sessions to compare similarities among task networks. This new set of paired permutations confirmed that the ARI between task networks (median = 0.29, IQR = 0.23 – 0.33) was

significantly greater than the ARI between fixation and either cards (median = 0.2, IQR = 0.1 – 0.28; $p < 0.01$; Cohen's D = 0.97) or n-back (median = 0.13, IQR = 0.08 – 0.16; $p < 0.001$; Cohen's D = 1.69).

Collectively, these results provide evidence that 1) functional network topographies underlying wakeful rest, working memory, and decision-making are reliable and idiosyncratic, replicating the relative region specific levels of inter-individual variation observed previously; 2) a common spatial configuration subserves these distinct operations; and 3) tasks yield more stable partitionings than resting-state fixation runs, which could be due to the structure they impose on brain activity or the different network estimation method used for the decision-making runs.

Valuation Effects in the Default Network

The results discussed above indicated that the same sub-division of canonical DN could be identified in data from distinct tasks. This section evaluated which of the resulting sub-networks (DN or non-DN) has a potential functional role in decision making. We did this through two analyses. First, we qualitatively examined each network's distribution of beta-series coefficients as a function of task-related variables at both the of decision and outcome stages of the decision-making task. Second, we formally tested the observed relationships by fitting mixed effects linear models for the mean activity of each sub-network.

Fig. 4.5 shows, for each thresholded network within each cortical region, mean trialwise betas per subject as a function of trial value for four decision-related variables. Ventral striatum is examined separately below. There was an overall tendency for DN

coefficients to be negative, which could reflect task-related deactivation superimposed on valuation effects (and is controlled for in linear models). Plots in the top row present effects expected to occur during the presentation of Card 1. There was a clear decrease of activity in DN as the face value of Card 1 neared the decision boundary (e.g. 5 and 6), with non-DN displaying the opposite pattern. The top-right plot displays this effect as a function of expected value, corroborating that the DN increased its activity as possible earnings grew. The DN's sensitivity to value was somewhat mirrored at the outcome phase (bottom row). The bottom-left plot shows that this network's activity scaled with the magnitude of the resulting earning or loss (losses were rare because participants mostly followed the reward-maximizing strategy). However, net rewards do not take into account that outcomes are sometimes not as good as they could have been, and such violations of expectation are a key influence on activity in the valuation system (Knutson, Fong, Bennett, Adams, & Hommer, 2003). The plot on the bottom right shows that there was a small effect of RPE on overall network activity.

Linear mixed-effects models performed on each subject's averaged sub-network activity time series partially supported these observations (see Experimental Methods and Design for details on the composition of fixed and random effects per model). In order to avoid circularity (Vul, Harris, Winkielman, & Pashler, 2009), we constrained these analyses to grayordinates selected using each subject's whole-brain fixation sub-networks (results were replicated when using n-back networks instead). The first model, which focused on effects at Card 1 onset, showed a significant positive effect of expected value on DN activity (intercept = 4.25, $\beta = 1.20$, SE = 0.48, $p < 0.05$). We also found

significant negative effects for trial onset (β = -5.23, SE = 2.27, $p < 0.05$) and response time (β = -9.11, SE = 2.86, $p < 0.01$) in this network, which are often used to control for general task demands and choice difficulty, respectively. None of these effects were significant for the non-DN network (all $p > 0.1$). A second model showed a decrease in DN activity when Card 2 was presented (intercept = 5.42, β = -13.66, SE = 2.04, $p < 0.001$), but this activity was significantly raised as the total reward earned increased (β = 0.74, SE = 0.24, $p < 0.01$). As before, none of these effects were present in non-DN (all $p > 0.05$). A similar outcome-focused model showed no significant effects for reward prediction error in either network (both $p > 0.1$). Inspections of the random slopes suggested that the effect of RPE on the DN were positive for 13 out of 18 participants. Coefficients for session number were not significant for any model (all $p > 0.05$), indicating that these results were consistent over time.

We also assessed the possibility that valuation effects could be present in the spatial gaps between networks. In our previous work (Toro-Serey et al., 2020), we found that subthreshold areas adjacent to the DN corresponded to Braga and Buckner's (2017) proposed default network B (whereas network A matched what we called the DN). This network has been shown to subserve DN-related cognitive functions separate from those found in network A (DiNicola et al., 2020), raising the possibility that a similar distinction might exist for decision processes. The linear mixed effects models were reproduced on a putative default network B composed of grayordinates with FV values between 0.002 and 0.008 (i.e., positive and under the 0.01 threshold, but above values associated with extreme cluster uncertainty). These exploratory analyses found

significant main effects of expected value ($\beta = 0.33$, SE = 0.14, $p < 0.05$), Card 2 presentation ($\beta = -3.06$, SE = 0.48, $p < 0.001$), and reward amount ($\beta = 0.13$, SE = 0.06, $p < 0.05$), but no effect of RPE ($p > 0.1$). Unlike the decision effects observed in the DN, network B appeared to focus on expected value without being influenced by task demands ($\beta = -1.3$, SE = 0.92, $p > 0.1$) or RT-indexed choice difficulty ($\beta = -1.38$, SE = 0.82, $p > 0.1$).

The addition of this network helped us understand effects in ventral striatum, which were otherwise difficult to evaluate due to the low number of grayordinates affiliated to the DN in this area (see the qualitative section above). Fig. 4.6 shows the beta distributions for ventral striatal voxels associated with each of the three sub-networks. We fitted mixed-effects linear models directly to these distributions in order to test the observed relationships, with fixed and random effects for intercept and slopes per subject. At the outcome phase, activity in both putative default network B and DN-related striatum voxels appeared to be positively modulated by increments in reward and RPE, although these relationships were only statistically significant for network B (Reward: $\beta = 1.18$, SE = 0.36, $p < 0.01$; RPE: $\beta = 0.78$, SE = 0.37, $p = 0.05$). Network B was the only one to present a relationship with expected value, which only trended towards significance ($\beta = 1.6$, SE = 0.87, $p = 0.08$). The cleaner relationships are probably due to the greater number of grayordinates in this sub-network compared to the DN (median = 30, IQR = 12 - 24), but could also point towards a dissociation of network function between cortex and striatum.

These findings extend the view that the default network participates in valuation--at least at the cortical level--with adjacent brain zones showing mixed involvement in choice evaluation and outcome.

Topography and Behavior Interactions

A main aspect of our guiding hypothesis was that accounting for behavioral variability would help us understand functional organization. Specifically, we proposed that people with more stable behavioral patterns over time would also display higher similarity in brain network configurations across sessions, and that disengagement with decision-related values (as indexed by RTs) would result in higher similarity between networks estimated with decision-making runs and those estimated with fixation runs. However, behavioral analyses showed that participants were, for the most part, consistent in their behavior over time.

The disclosed reward-maximizing rule was to bet that Card 2 would be higher if the face value of Card 1 was 5 or lower. Fig. 4.7 (top row) shows the proportion of rule-congruent choices per subject as a function of Card 1 value, indicating that participants mostly adhered to this strategy (mean accuracy = 98%, SD = 0.06). Similarly, the second row of this figure shows that the probability to guess higher adequately switched at values higher than 5, with few deviations (one individual forgot when the switch should happen, but was consistent in their deviated strategy). Indeed, a mixed effects logistic regression showed a significant effect of Card 1 value on the probability of betting high ($\beta = 9.29$, SE = 0.97, $p < 0.001$), but no significant differences across sessions ($\beta = -0.1$, SE = 0.85, $p = 0.9$).

In terms of response times, we predicted that people who actively considered the potential earnings in each trial would spend more time evaluating trials with low expected values (e.g. Card 1 values of 5, 6, etc.). Fig. 4.7 (bottom row) shows this relationship for most subjects, and a mixed-effects linear model confirmed it ($\beta = -0.09$, SE = 0.006, $p < 0.001$). The same model also suggested that people were faster during the second session, even though the relationship was maintained ($\beta = -0.05$, SE = 0.02, $p < 0.05$). However, two participants displayed consistently fast responses regardless of expected value (flat lines in the plot; 639: $\beta = -0.04$; 731: $\beta = -0.004$). These results suggest that even though people consistently followed the optimal rule, their readiness to implement it was somewhat variable across subjects and sessions.

Even though the mixed effects model suggested that people were faster on the second session, the absolute value of the random slopes from this change were not correlated with inter-session (within-subject) similarities of the decision-making networks (Pearson's $r = 0.2$, $t(16) = 0.84$, $p > 0.1$). Another possibility was that response times that did not change as a function of Card 1 values would reflect disengagement with the expected value of a trial, resulting in network patterns during decision-making runs that more closely resembled those during fixation runs. Akin to the reliability test, the random slopes describing the effect of expected value on log-transformed response times (per session) were not correlated with the similarity between card and fixation networks (Pearson $r = 0.19$, $t(34) = 1.16$, $p > 0.1$).

However, these slopes might have been affected by the fact that the main change in response times was non-linear, and mainly visible between the two lowest expected

values. Another way of testing this relationship is by taking the difference in response times between the lowest and highest expected value. This exploratory assessment showed that this delta was significantly correlated with the similarity between fixation and decision-making networks (Pearson r = -0.41, t(34) = -2.65, $p < 0.05$), such that seemingly disengaged choice behavior was associated with greater similarity between fixation networks and decision-making networks. This exploratory assessment points at a possible (though subtle) interaction between functional network topography and behavior, but a larger sample would be needed to confirm such a relationship.

Discussion

Meta-analytic work has suggested that the same brain that comprise the default network also support value-based decision making, promoting theories of a common cognitive architecture that can subserve these two (and additional) phenomena (Acikalin et al., 2017; Euston et al., 2012; Northoff & Hayes, 2011). In this final study, we built upon the individual variability in functional topography and behavior uncovered by the previous chapters to more precisely determine whether the DN facilitates economic choice. In line with previous results, we found that canonical default network regions could be fractionated into at least two networks (DN and non-DN) whose spatial configurations were heterogeneous across individuals, but reliable across sessions within individual. However, contrary to our hypothesis, we found that the same individual-specific functional topography was present across tasks, and that the DN sub-network indeed participated in valuation processes.

Comparing Rest and Task-based Networks

Much work dedicated to understanding the topographic organization of the brain has proposed that, on average, distributed brain areas that participate in the same cognitive operations show correlated spontaneous activity fluctuations in the absence of external stimulation (Fox et al., 2005; Laird et al., 2011; Smith et al., 2009; Yeo et al., 2015). Here we add partial support to this claim, showing that cortical network layouts derived using functional connectivity from different tasks were generally aligned within canonical DN in mPFC and PCC.

Recent work has suggested that this overlap is due to resting-state networks acting as a coarse, default spatial configuration that is partially recruited during engagement with tasks (Cole et al., 2014; Uddin et al., 2019). Our second line of evidence speaks against this idea, both qualitatively and quantitatively. Even when focusing on individuals, we found that the same idiosyncratic sub-division of canonical DN (which was estimated using resting state) was present for all tasks, regardless of the method used to generate them. Further, any discrepancies in topography between domains were due to task networks covering a greater (and similar) extent of cortex than fixation, opposite to the pattern suggested above. Instead, our results converge with recent proposals that functional area boundaries can be reliably (albeit slightly) shifted based on task-induced cognitive states (Salehi et al., 2019). Unlike that work, however, we found that cards and n-back networks were highly similar with each other, though less so with fixation. This could be due to their high inter-session reliability compared to fixation networks, which could stem from the structure imposed by the tasks (Jiang et al., 2020). On the other

hand, the decrease in fixation network reliability compared to Toro-Serey et al. (2020) could be due to our use of a single thresholding value to identify DN and non-DN (versus the customized ones used previously). Regardless, the degree of idiosyncrasy displayed by all networks replicated the pattern observed in the past, with mPFC displaying particularly unique spatial configurations across individuals.

Even though our findings support the idea of a common functional connectivity network organization across tasks, they also provide points of divergence with the existing literature. Many group-level comparisons of task and resting state networks have worked under the assumption that functional connectivity and task-based general linear model analyses are equally capable of retrieving the same networks (Fox et al., 2005; Greicius et al., 2003; Laird et al., 2011; Smith et al., 2009). We made such assumption ourselves when selecting a corpus for the meta-analysis in our own work (e.g. based on Laird et al. 2009; Toro-Serey et al., 2020), and when choosing to validate connectivity layouts with traditional task contrasts. However, our results suggest that this equivalence should be reconsidered. Even though individual partitionings computed using SP were well aligned with task-deactivation GLM contrasts in PCC and dorsal mPFC, we found that the DN sub-network often excluded a region of strong deactivation located along the suprarostral sulcus (i.e. the sulcus parallel and dorsal to the superior rostral sulcus, which was mapped to the DN sub-network in the past; Lopez-Persem et al., 2019). This discrepancy is unlikely to be due to either differences in smoothing (the remainder of the search space was closely aligned across modalities) or lack of reliability (both methods produced high inter-session reliability). Even though this partial apparent decoupling

between methods is beyond the scope of the present work, it warrants further interrogation in future studies.

Valuation Processes in the DN

Against our expectations, we found evidence that the subdivision corresponding to each individual's default network participated in the valuation of both prospects and outcomes, even when controlling for choice difficulty via response times. This relationship appeared to be weaker in PCC than mPFC according to the distribution of single-trial amplitudes, which is supported by the limited effects of expected value often found in this area (Bartra et al., 2013; Clithero & Rangel, 2014). An exception to this trend was the absence of reward prediction error effects in any cortical sub-network. Compared to the ubiquitous presence of prediction errors in ventral striatum (Haber & Knutson, 2010; Hare et al., 2008; Schultz, 2015), the role of the mPFC in encoding reward prediction errors has been inconsistent across studies. Some studies have observed RPE signals in mPFC (Ramnani, Elliott, Athwal, & Passingham, 2004; Ribas-Fernandes, Shahnazian, Holroyd, & Botvinick, 2019), whereas in other cases, the role of mPFC has been relegated to decision values during the appraisal stage (Hare et al., 2008). Other accounts suggest that mPFC provides RPE-related information to the dopamine system only when reward contingencies are uncertain (Starkweather, Gershman, & Uchida, 2018). Since participants in our experiment knew the success probabilities paired with each value of Card 1 (given the disclosed optimal rule), it is possible that mPFC was not always necessary for processing prediction errors (Starkweather et al., 2018), thereby limiting the effects of RPE in the rest of the network.

Even so, considering the effect of expected value observed in the DN, we would have expected more pronounced changes in ventral striatum activity when expectations were violated. It is possible that ventral striatum was active during periods of outcome anticipation that were missed by the beta series deconvolution (Knutson et al., 2003), similar to the ramping up of tonic dopamine observed as a function of reward uncertainty (Fiorillo, Tobler, & Schultz, 2003) and goal approach (Howe, Tierney, Sandberg, Phillips, & Graybiel, 2013). An alternative explanation is that participants focused on the probability of winning or losing points during decision points rather than expected value (i.e. on the positive or negative valence of the outcome; Venkatraman, Payne, Bettman, Luce, & Huettel, 2009). Such probabilities could be computed without making strong predictions about the actual amount they would ultimately earn, but would still represent a modulation of trial value as a the probability of a successful outcome changed (Tom, Fox, Trepel, & Poldrack, 2007). Notably, the focus on expected valence would not preclude the positive effects of reward, which were clear throughout the DN. Future iterations of this work could more directly assess the distribution of value and valence processes across relevant subcortical regions, in particular the amygdala (Berridge, 2019; Joyce & Barbas, 2018).

Part of the difficulty in interpreting the remaining valuation effects in ventral striatum was the inconsistency with which SP assigned voxels to any network, an ambiguity that extended to caudate and putamen. Both group and individual maps showed that the DN (e.g. the cortical network that processed subjective values) recruited a very small ventral portion of the striatum, which is in contrast with the broad and

consistent effects of reward value on striatal activity found in diverse neuroeconomic studies (Bartra et al., 2013; Gregorios-Pippas, Tobler, & Schultz, 2009; Gruber & McDonald, 2012; Haber & Knutson, 2010; Kable & Glimcher, 2007; Levy et al., 2011). Instead, we found that sections of striatum that missed our clustering threshold mirrored the valuation dynamics displayed by the DN in cortex. Toro-Serey et al., (2020) suggested that these grayordinates corresponded to default network B from Braga & Buckner (2017). More recent work suggested that networks A (what we call DN) and B separately process previously-grouped effects of theory of mind and episodic projection (DiNicola et al., 2020), raising the possibility that these networks also process different aspects or stages of decision making. However, under this framework, our results would suggest that the same valuation effects are transferred from cortical DN to subcortical network B instead.

An alternative explanation is that the topography of the striatum is better understood as continuous rather than discrete. Anatomical tracing work in non-human primates has proposed that cortico-striatal projections from prefrontal cortex are gradually organized along a dorsal-ventral axis (Haber & Knutson, 2010). This results in a continuous and diffuse arrangement in this subcortical area, whereby pathways from vmPFC and dmPFC project mostly to ventral and dorsal striatum, respectively. In the present work, every interdigitated sub-network consistently recruited zones of mPFC along this axis, suggesting that signals across cortical sub-networks converged throughout the striatum. We therefore propose that, rather than indexing a specific sub-network affiliation, relatively low FV values in ventral striatum reflect this continuous

organizational scheme, comparable to recent functional mappings of this area that employed similar gradient-based algorithms (Marquand, Haak, & Beckmann, 2017; Tian, Margulies, Breakspear, & Zalesky, 2020a), and in line with the hypothesized role of the striatum in integrating value with additional action signals (Haber & Knutson, 2010). We note that the distinction between an organization of the striatum that is based on functional networks versus anatomical connectivity would be missed in traditional voxel-wise modeling approaches, as they normally just highlight which units present valuation signals, and not their relationship within and across networks.

Behavioral Influences on Functional Topography

Recent studies have uncovered a relationship between both structural and functional brain topography and unique behavioral phenotypes (Bijsterbosch et al., 2018; Kong et al., 2018; Mansour L, Tian, Yeo, Cropley, & Zalesky, 2020). More nuanced individualized work has further shown that the boundaries of functionally coherent subregions shift from task to task (Michalka et al., 2015; Salehi et al., 2019). Study 2 offered evidence that the same decision environment can produce time-varying choice behavior based on features that individuals covertly integrate (i.e., the subjective passage of time). Based on this evidence, we explored whether network layouts could be reconfigured depending on the degree of behaviorally indexed engagement with a simple and transparent decision-making task. However, the majority of our participants displayed consistent rule-consistent choice behavior across sessions, precluding this possibility. None of the results from our planned behavioral analyses were correlated with changes in network alignment, either across sessions or across run types. Instead, the

only piece of evidence came from an exploratory analysis across individuals, which suggested that individuals whose response times were similar for lowest and highest expected values (i.e., less strongly modulated by proximity to the decision boundary) displayed network topography during the decision-making task that more closely resembled configurations from resting fixation runs. However, response times can follow distributions like those observed here for non-economic judgments (Moyer & Landauer, 1967), limiting our ability to relate this phenomenon to valuation. Further, visual inspection failed to identify systematic topographic features underlying the variation in similarities. While we cannot make strong inferences based on current findings, we hope that future studies will continue to explore this possible interaction with more tailored experimental designs.

Limitations

The present study faces a few limitations and outstanding questions. First, it is unclear why DN and valuation networks overlapped in individuals, but were segregable in PCC at the meta-analytic level (Toro-Serey et al., 2020). We saw that expected value patterns were weaker in PCC than mPFC, resembling the pattern observed across studies (Bartra et al., 2013; Clithero & Rangel, 2014). This could be related to the mismatch in individuals between task-derived and connectivity-derived network topography. A second potential limitation is that we did not discriminate between modes of decision values. Even though regions like vmPFC have been known to process value in a category-independent manner (Chib et al., 2009), previous work has highlighted that different sections of mPFC are recruited for the evaluation of goal-congruency, offer

value, and relative value among options (Frömer, Dean Wolf, & Shenhav, 2019; Hare et al., 2008; Shenhav & Karmarkar, 2019). Whether these are facilitated by distinct sub-networks, or by different sections of the same network, should be assessed in the future. Finally, we are presently limited to fully endorse the hypothesis that the functional connectivity between prefrontal cortex and striatum is organized along a dorsal-ventral gradient rather than discrete functional networks in individuals, a hypothesis that has been proposed on the basis of cortico-striatal anatomical projections (Haber & Knutson, 2010). This limitation is mostly due to our scanning protocols, which were optimized for cortical signal acquisition. Even so, future work could test this hypothesis by expanding the search space towards dorsal anterior cingulate and orbitofrontal cortex, areas where the anatomical projection gradient appears to shift the most (Haber & Knutson, 2010). Another possibility would be to slice each individual's sub-networks along this dorsal-ventral axis in mPFC, and correlate the mean activity of each sub-network within each slice with activity of striatal voxels. A gradual organization would result in changes to the correlation patterns in striatum across slices, whereas a discrete organization would exhibit differences in correlation patterns between DN and non-DN in each slice. Regardless, the structural basis of this organizational scheme, and its recent functional support at the group-level utilizing a similar family of algorithms as the one employed here (Marquand et al., 2017; Tian, Margulies, Breakspear, & Zalesky, 2020b), add confidence to this hypothesis.

Conclusion

The work presented here advances our understanding of the functional organization of the brain in two ways. First, to our knowledge, it demonstrates for the first time that the same individual-specific network configuration observed during rest underlies task engagement (in contrast to previous comparisons performed at the group or meta-analytic level). Second, it situates the DN as a cortical network that facilitates the implementation of economic choice, particularly within mPFC, and presents supporting evidence for the role of the striatum as a hub that can combine diverse sources of cortical information with value in order to produce actionable behavior. The topographic heterogeneity of functional networks in mPFC might help explain why a recent multi-site replication attempt of well-known valuation effects failed to survive diverse statistical correction pipelines (Botvinik-Nezer et al., 2020), bringing attention to the possibility of false negatives in previous group-level results. The present findings strengthen the rationale for theorizing about a common architecture that subserves distinct types of cognitive functions (Euston et al., 2012; Northoff & Hayes, 2011).

CHAPTER 4 FIGURES

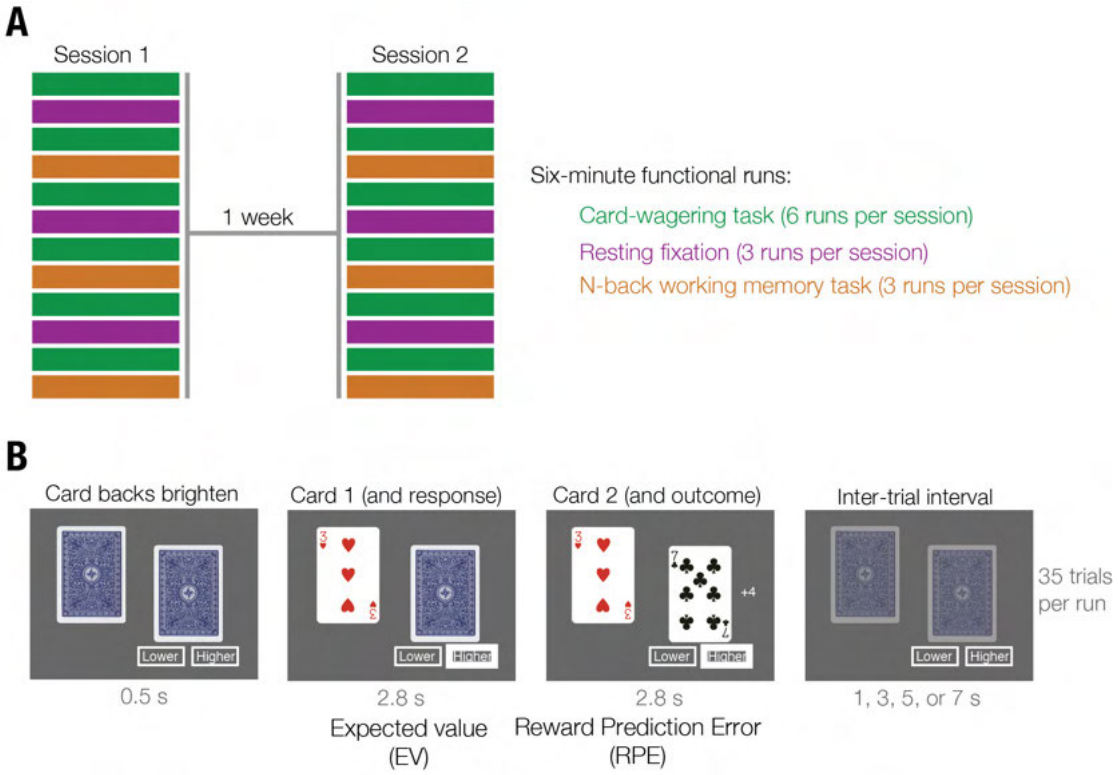

Figure 4.1: Task structure. **A:** Each of 18 participants was scanned in three paradigms across two sessions separated by a week. The order of functional scans was preserved across sessions. Participants practiced each task a day or two before Session 1. **B:** Card wagering task. On each trial, a card from Deck 1 was revealed, and participants had to guess whether the top card from Deck 2 would be higher or lower in value. Participants gained or lost points according to the difference in cards. This task manipulated expected value and reward prediction error separately, and participants were told about the reward-maximizing strategy (bet higher for Card 1 face values of 5 or lower).

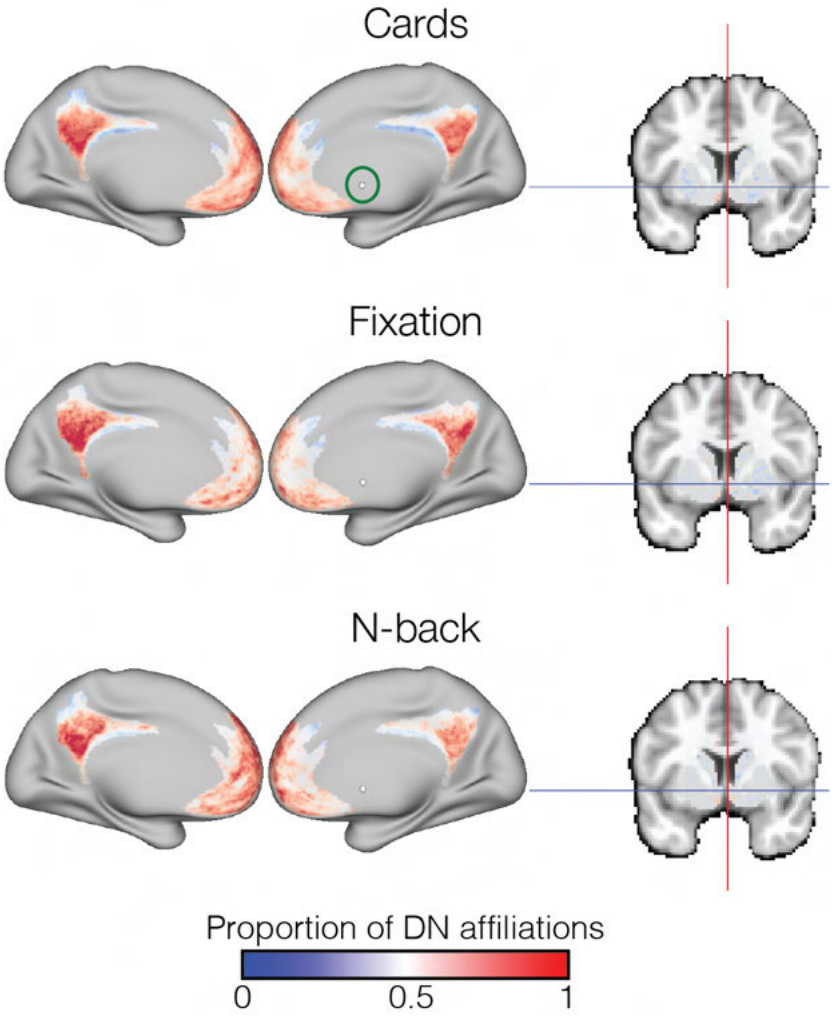

Figure 4.2: Proportion of DN affiliations across individuals in each task. Each subject's Fiedler vectors from the graph Laplacian are binarized according to the sign of their values, and oriented so that the DN label is given to the network with most grayordinates covering area 7m. The dot in the green circle shows the point corresponding to the coronal volumetric slice (MNI coordinate y = 66). Cards networks were computed on the basis of activity at Card 1 onset (similar results were obtained using Card 2 onset; see Fig. 4.4B). All tasks yielded similar spatial network configurations in cortex, although fixation-based maps tended to be more variable. A small region in ventral striatum appeared to be consistently affiliated with the DN, although no voxel was affiliated with either network more than 50% of the time across individuals.

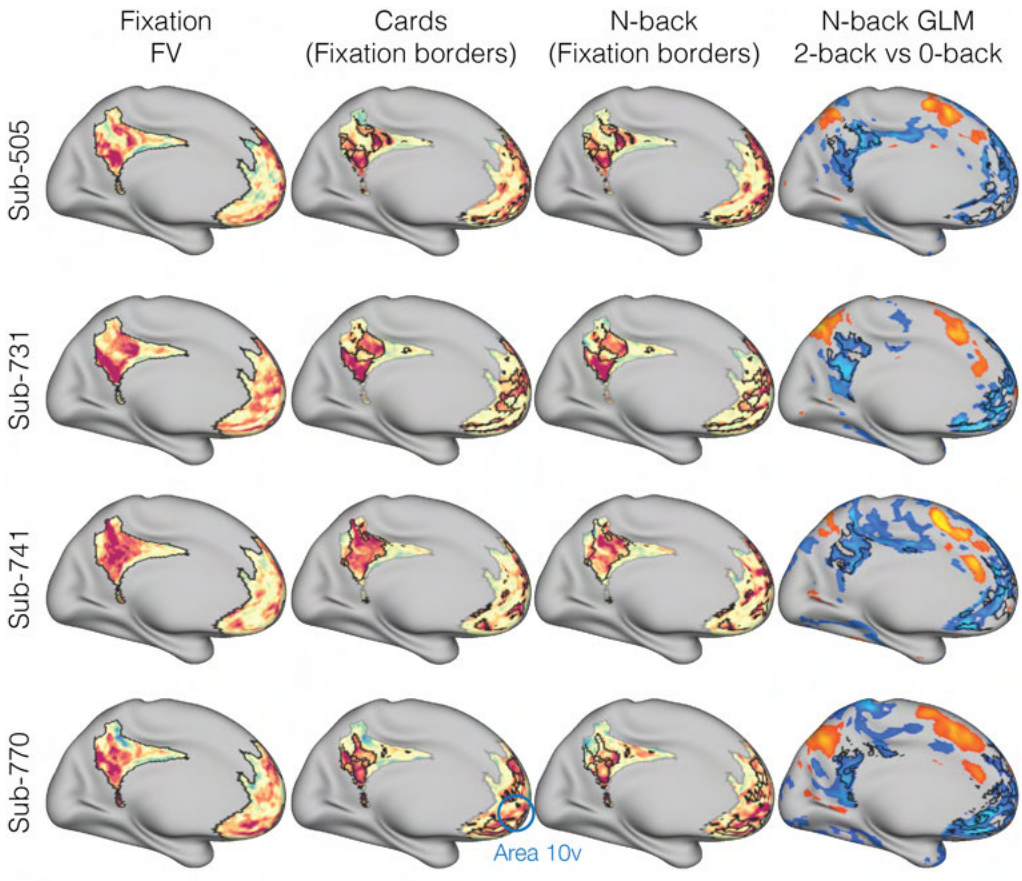

Figure 4.3: Examples of individual network arrangements per task (same session). Column 1 shows a replication of the networks from Toro-Serey et al. (2020) using the present resting dataset. Columns 2 and 3 display task-related networks, with a black outline depicting the DN from the respective fixation partitionings. There was good overall alignment among networks. Column 4 shows the overlap between connectivity-based and whole brain GLM-based patterns of n-back activity (two-sided t-maps thresholded at an absolute value of 3, with blue denoting task-based deactivation). There was good alignment in PCC and dorsal mPFC, but not in vmPFC.

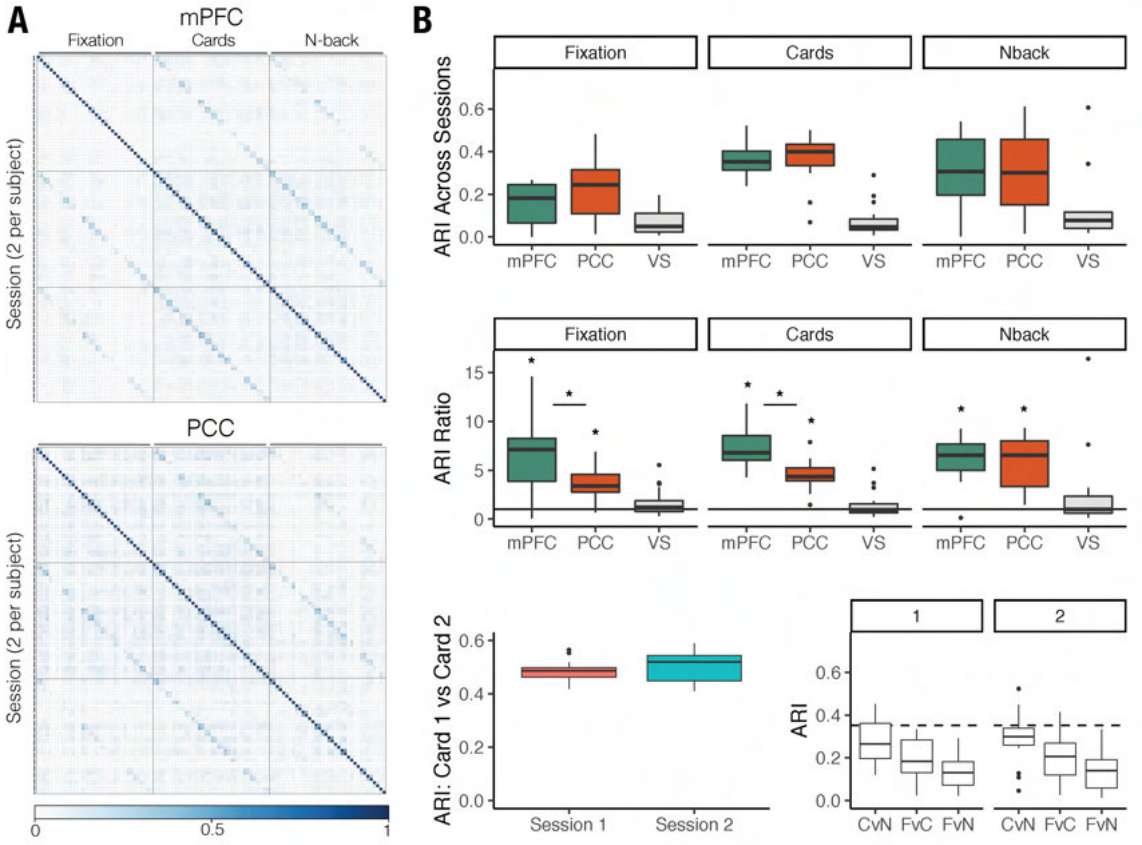

Figure 4.4: Quantitative similarities among networks. **A:** Similarity matrices of ARIs between every pair of run type, individual, and session. Each of the 9 tiles represents a within- or across-task comparisons. Visible diagonal structures within each tile suggest a subject-unique layout underlying all tasks. **B:** Inter-session reliability for each network and region (top row), ratios that represent the degree of subject-specificity of each layout (middle row), similarity between networks estimated from decision-making runs using using either Card 1 or Card 2 onset-related activity (bottom left), and similarities among task networks within individuals for each session (bottom right; CvN = Cards versus N-back; FvC = Fixation versus Cards; FvN = Fixation versus N-back; dashed line is the median inter-session ARI for the decision-making networks). It appeared that an idiosyncratic network topography subserved all cognitive operations (especially in mPFC), and that runs with task structure (in contrast to resting fixation) yielded more reliable network estimates.

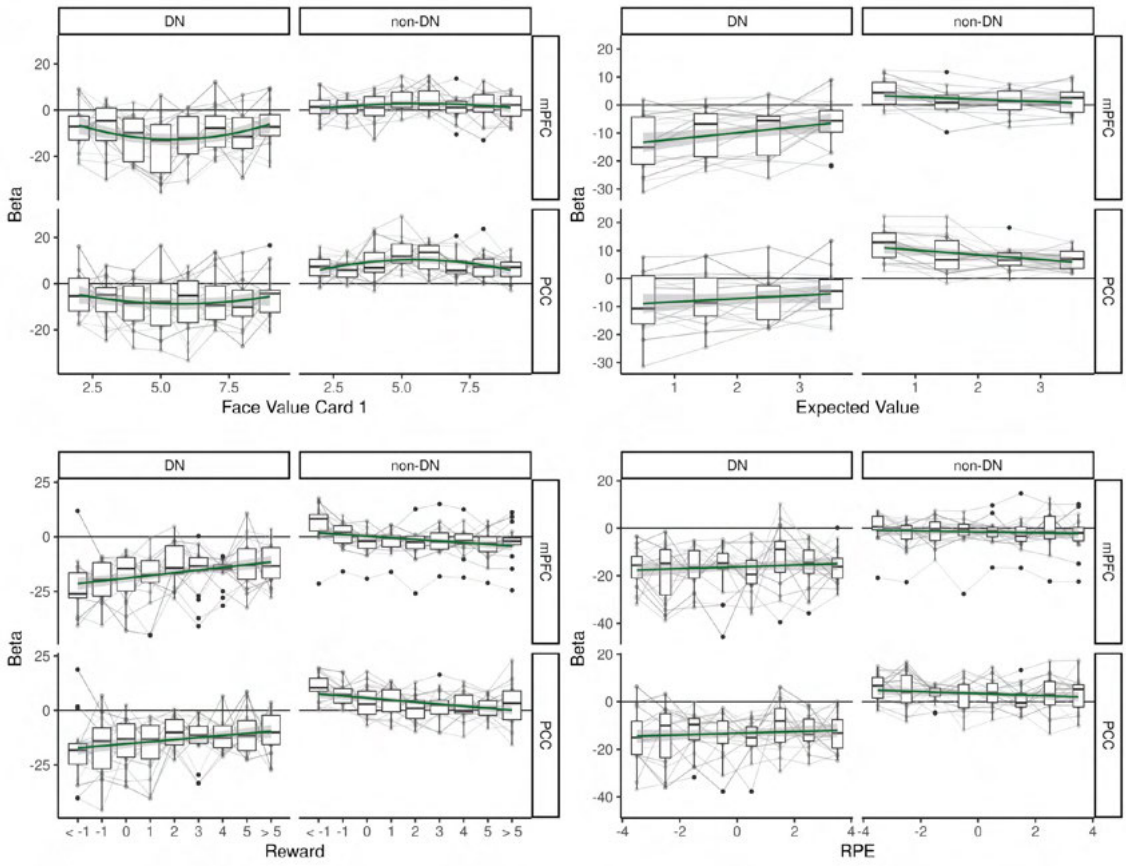

Figure 4.5: Valuation effects per network and region. Plots show the distribution of single-trial coefficients from each individual's beta series, separated by region and thresholded network. Top row shows effects pertaining to the presentation of Card 1, whereas the bottom shows outcome-related effects. These distributions suggest a general role of the DN in valuation. Linear mixed effects models on mean network activity of both mPFC and PCC corroborated effects of EV and reward magnitude in the DN.

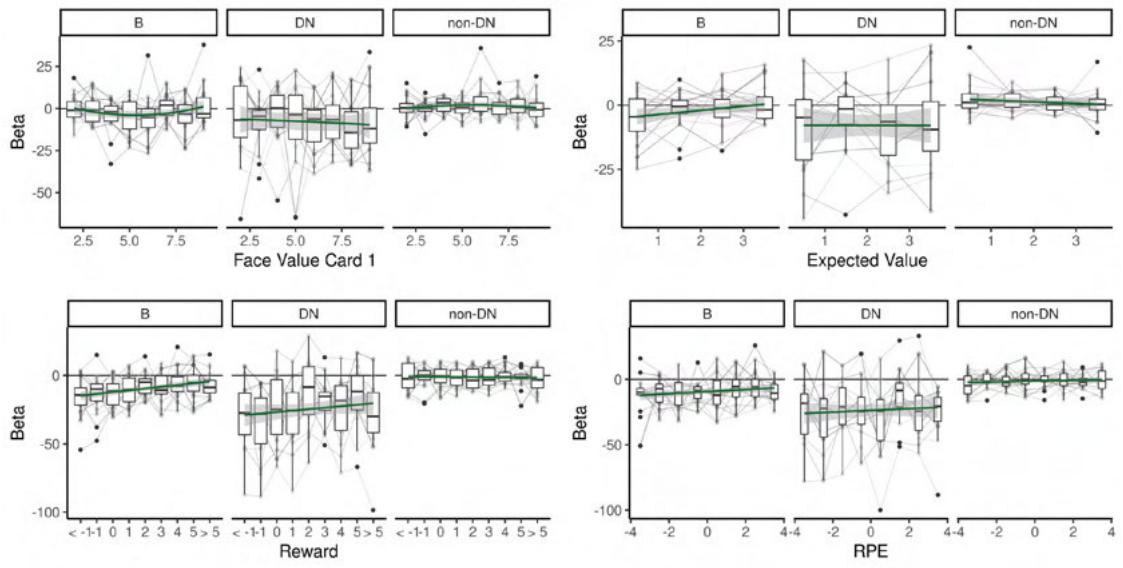

Figure 4.6: Valuation effects in ventral striatum. Beta distributions similar to those in Fig. 4.5 are plotted for striatal voxels associated with the DN and non-DN as well as the putative default network B (computed based on findings from Toro-Serey et al., 2020). Unlike in cortex, the DN-related striatal voxels appeared to participate mainly during the reward phase of each trial, whereas network B-related voxels presented canonical value-sensitive patterns during both decision and outcome stages. Linear mixed-effects models fitted to these distributions only corroborated the effects of reward and RPE on putative network B. The noisy patterns in DN are likely due to the small number of grayordinates affiliated with this network within this region across individuals (median = 7).

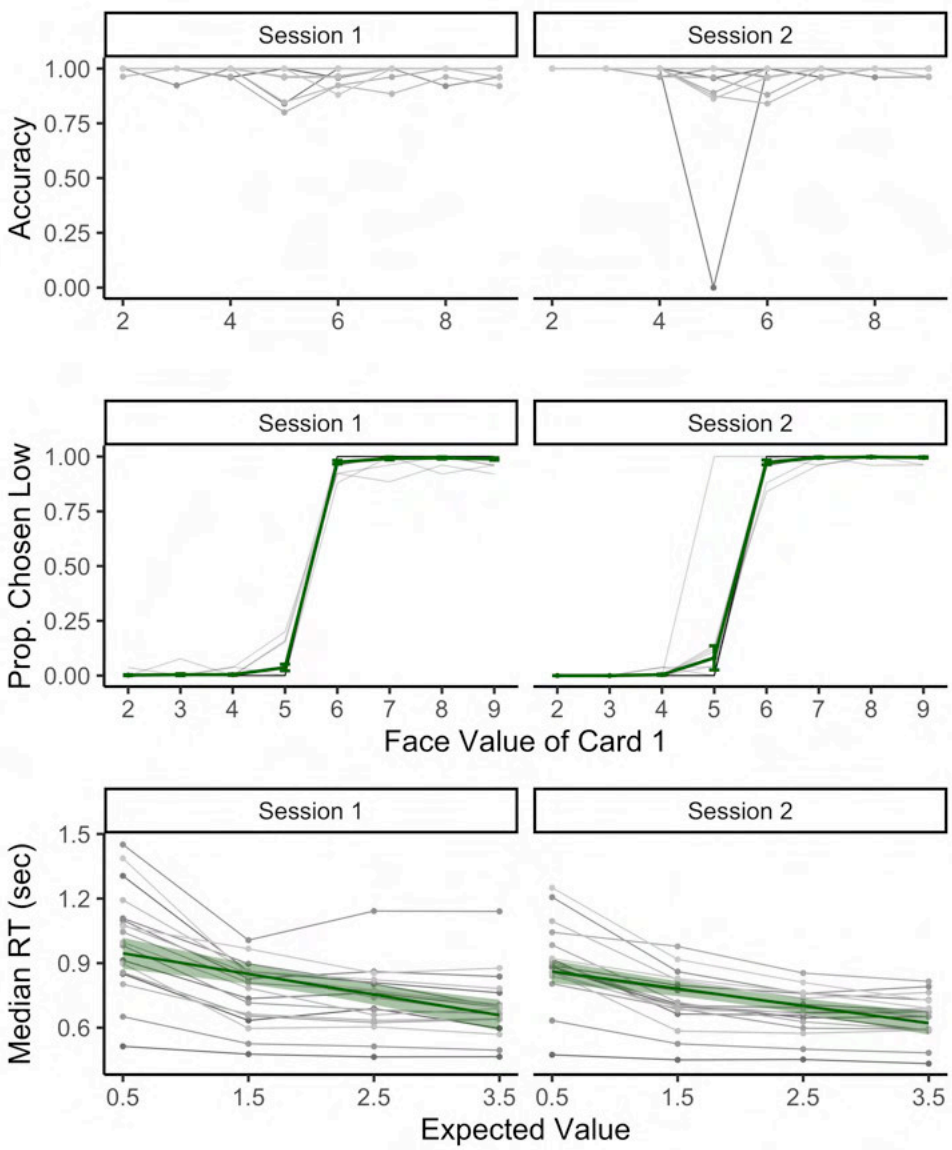

Figure 4.7: Choice behavior in the cards task per session. **Top**: proportion of trials in which participants followed the optimal rule. **Middle**: Probability of guessing that the card from the second deck would be lower. Participants readily followed the optimal rule in both sessions, with the exception of one person. **Bottom**: Response time as a function of the expected value given by Card 1. Participants were consistently faster in responding to trials with greater expected earnings.

CHAPTER FIVE: SUMMARY AND DISCUSSION

Restatement of Original Goals

The overarching goal of this dissertation was to advance our understanding of the neural basis of human decision making by examining within-person and between-person variation in functional brain network topography and choice behavior. I pursued this objective through three behavioral and neuroimaging studies. The first study sought to establish a procedure to robustly estimate an individual's functional brain networks, and to quantify their degree of topographic idiosyncrasy within medial cortical regions of the default network that consistently overlap with subjective value effects (PCC and mPFC). A second study examined potential sources of choice fluctuations by examining task-context-related variation in attitudes towards different types of behavioral demands. The final study explored the potential interaction of these types of variability to determine whether value-sensitive regions could be disentangled from the default network across multiple task contexts at the individual level.

Summary of Findings

The first study, described in Chapter 2, sought to advance our understanding of the individual heterogeneity of the DN within mPFC and PCC, quantifying that heterogeneity for the first time. First, we employed a cortical surface-based meta-analysis to identify a parcel in human PCC that was more strongly associated with the DN than with valuation effects. We then used resting-state fMRI data and a data-driven network clustering algorithm, spectral partitioning, to fractionate mPFC and PCC into "DN" and "non-DN" subdivisions in individual participants (n = 100 from the Human Connectome

Project). The spectral partitioning algorithm identified individual-level cortical subdivisions that varied markedly across individuals, especially in mPFC, and were reliable across test/retest datasets. Further, we found that the resulting eigenmaps from this algorithm contained information about the stability of the sub-network label at each cortical location, giving rise to a third interdigitated sub-network. Within these partitionings, an individual's DN showed a region-dependent alignment with sulci and gyri throughout the search space. Finally, we validated these spatial layouts against recently proposed subdivisions of canonical DN identified in a small sample, and demonstrated the advantage of spectral partitioning in robustly capturing these layouts compared to seed-based functional connectivity. These results highlighted the need to ascertain whether distinct cognitive functions engage common or distinct mPFC subregions at the individual level, and established an analytical pipeline to do so.

Chapter 3 explored the factors that can underlie changes in people's choices over time, even when decision environments are well known. To do this, we embedded mental and physical effort in a "prey selection" foraging task, which required decision makers not only to evaluate the magnitude and delay of a focal prospective reward, but also to estimate the general opportunity cost of time. In two experiments, independent sets of participants collected rewards that required equivalent periods of cognitive effort, physical effort, or unfilled delay. Decision makers were able to approximate reward-maximizing strategies unique to each foraging environment, regardless of the combination of time, reward, and cost. Even so, when individuals faced only one type of cost (Experiment 1; between-subjects design), cognitive effort uniformly produced the

highest acceptance rate compared to trials with an equivalent period of either physical effort or unfilled delay. Conversely, a similar higher acceptance rate for cognitive effort trials disappeared over time amid an overall decline in acceptance rates when individuals could choose which demand to pursue (Experiment 2; within-subject design). A combination of foraging and reinforcement learning models could account for the results by positing that each type of demand imposed a bias on the estimation of the opportunity cost incurred in pursuing offers. These biases, which are potentially attributable to a distortion in the perceived passage of time, were gradually averaged away when multiple forms of demand were interleaved.

In Chapter 4, we built upon the individual and context-related variability identified in Chapters 2 and 3 to ascertain the degree of overlap between valuation and default networks within mPFC, PCC, and striatum across multiple task contexts. Individuals were scanned in two sessions while performing an n-back task, wakeful resting, and a decision-making task that orthogonalized expected value with reward prediction error. Using the methodology established in Chapter 2, we first found that the same idiosyncratic organization discovered in Chapter 2 was recoverable from all three tasks, as well as from both the evaluation and reward stages of the decision-making task. Furthermore, within each partitioning, the default network persistently displayed canonical valuation effects in mPFC and PCC (except for reward prediction error), whereas non-DN showed an opposite activity pattern. Regions within striatum were not directly linked to any singular sub-network. Participants consistently adhered to a reward-

maximizing strategy, resulting in negligible evidence that behavioral variability could affect the topographic alignment across tasks.

Discussion

A diverse set of higher order psychological phenomena, including economic judgments, engage overlapping swaths of mPFC (De La Vega et al., 2016; Hiser & Koenigs, 2018; Kragel et al., 2018), prompting the question of whether these are truly distinguishable functions. The neuroscientific literature has addressed such issues by organizing juxtaposed functions into macroscale functional networks with nodes distributed throughout the brain, with individual networks implicated in broad domains of cognition and behavior (Power et al., 2011; Yeo et al., 2011). The network that encompasses a majority of anterior mPFC, the default network, has been suggested to serve self-referential thought, prospective thinking, theory of mind, and general mnemonic functions (Buckner et al., 2008; Buckner & DiNicola, 2019; Gusnard et al., 2001; Schacter et al., 2007; Yeshurun, Nguyen, & Hasson, 2021). In neuroimaging studies, the default network typically shows an increase in activity when attention is internally oriented, and a decrease during increasingly demanding externally oriented tasks (Laird et al., 2009). Even though the introspective role assigned to the default network is conceptually cohesive with many of its associated cognitive operations, it is less obviously compatible with the increases in activity observed in mPFC as a function of increasingly valuable prospects and outcomes during decision making (Bartra et al., 2013; Clithero & Rangel, 2014). Yet, the apparent overlap of valuation effects with the DN within mPFC has been confirmed at the meta-analytic level (Acikalin et al., 2017),

and has led to theories that attempt to unite these seemingly disparate functions under a common cognitive architecture (Euston et al., 2012; Northoff & Hayes, 2011).

Examination of anatomical projections suggests that broad functional networks are composed of closely intertwined circuits that perform similar, but nonetheless distinguishable, parallel operations (Goldman-Rakic, 1988). This view of cortical organization has found renewed support with the advent of improved noninvasive imaging techniques and a recent focus on individualized mapping. For instance, functional connectivity analyses customized to each individual have shown that canonical frontoparietal, dorsal attention, and default networks can be subdivided into interdigitated sub-networks that spatially vary across people (Braga & Buckner, 2017; Braga et al., 2019), with each sub-network participating in distinct functions (DiNicola et al., 2020). Similar refinements have been identified for sensory-biased working memory processes within dorsolateral prefrontal cortex (Michalka et al., 2015; Tobyne et al., 2017), and for distinguishing language-specific effects from general cognitive-load-related modulation in Broca's area (Fedorenko et al., 2012). This idiosyncratic, fine-scaled organization of cerebral cortex had previously been obscured by the statistical need to average functional data with low signal-to-noise ratio across individuals, and its identification has called for a reconsideration of the degree to which cognitive functions truly overlap throughout cortex. The work presented here advances our knowledge on this matter by demonstrating that an individual's spatially refined default network is indeed modulated by a number of valuation-related factors.

A key goal of this dissertation was to explore the sources of individual variability that could make the brain's functional network structure difficult to identify. Until recently, the comparison of individual-specific functional network arrangements had been qualitatively performed using a few highly sampled participants (Braga & Buckner, 2017; DiNicola et al., 2020; Gordon, Laumann, Gilmore, et al., 2017; Laumann et al., 2015), and some studies purposefully omitted discussion of personalized functional topography in ventral mPFC due to BOLD signal artifacts commonly found in this area (Kraus, 2020; Tavor et al., 2016). Chapter 2 bridged these gaps by first validating a data-driven method to quantify the idiosyncrasy of an individual's functional brain networks, and then demonstrating in a large pool of individuals that the spatial partitioning of canonical DN is particularly idiosyncratic in mPFC. Our approach provided evidence that these partitionings could be captured without the need to extensively scan each individual (an extended sampling procedure proposed by Gordon et al., 2017), making it amenable to the data ratio that is more characteristic of task-based neuroimaging studies (i.e., more subjects with up to a few hours of scanning). This was fundamental in order to achieve the results presented in Chapter 4, which replicated these initial findings using resting-state scans and extended them to task-based networks.

A second potential challenge to high-precision functional mapping stems from behavioral variability. Recent work has suggested that the functional boundaries of brain regions are flexible and vary with the cognitive states induced by diverse task demands (Michalka et al., 2015; Salehi et al., 2019). Within the realm of valuation and decision making, people are prone to change their attitude towards repeated decisions in everyday

life, depending on aspects of the context. This ecological consideration raised the possibility that even for well-practiced tasks, the extent of recruitment of a given region of cortex could vary depending on how individuals construe their decision environment. These dynamics have often been investigated in the context of learning, and have tended to focus on the evolving connectivity within a network rather than its topographic organization (e.g. Bassett, Yang, Wymbs, & Grafton, 2014). Capitalizing on the fluctuating tolerance for effortful tasks experienced in everyday choices, Chapter 3 demonstrated that preferences could evolve over time even when the decision environment was fully known. Even though foraging scenarios like the one employed in this work do usually require decision makers to learn the richness of adaptive environments (Constantino & Daw, 2015; Dundon et al., 2020; Garrett & Daw, 2020; Stephens & Krebs, 1986), our participants were able to quickly and consistently estimate reward-maximizing strategies based on the information we gave them. Within-subject choice variability appeared to arise from the way in which individuals construed task elements that were not experimentally manipulated (i.e. the hypothesized subjective passage of time). We found no strong evidence that similar choice fluctuations played a role in the individualized mappings from Chapter 4 (which also provided participants with comprehensive knowledge of their options), but it remains possible that alternative decision-making tasks would evoke stronger shifts in cognitive state that could impact individualized functional mapping.

Our findings contribute to ongoing discussions of rigor and reproducibility in cognitive neuroimaging. In one recent multi-site study, 70 analysis teams produced

inconsistent fMRI results that ultimately failed to replicate a well-known valuation effect in mPFC (Tom et al., 2007), even though they used the same large dataset (Botvinik-Nezer et al., 2020). It is possible that the highly interdigitated DN and non-DN sub-networks identified here, which showed either opposite or orthogonal functional patterns in Chapter 4, overlapped across individuals and reduced the statistical power needed to replicate the effects. This phenomenon could also imply the existence of considerable false-negatives in the field. These issues could be confronted by adopting inter-subject alignment procedures that rely on functional characteristics in addition to structural features (Conroy et al., 2013; Guntupalli et al., 2018), some of which utilize gradient-like algorithms like the one employed here (Nenning et al., 2020; Nenning et al., 2017).

The work presented here is also pertinent to a number of clinical applications. At the neural level, the idiosyncratic network layouts exhibited by both task and rest networks could pose a problem for treatments that rely on either intra- or trans-cranial brain stimulation. Recent work has suggested that personalizing targets for such stimulation, either by adjusting stimulation to the spectral composition of reward signals or by identifying connectivity-based topography, can improve therapeutic outcomes for depression and obsessive-compulsive disorders (Cash, Cocchi, Lv, Fitzgerald, & Zalesky, 2020; Cash, Fitzgerald, Cocchi, & Zalesky, 2021; Grover, Nguyen, Viswanathan, & Reinhart, 2021; Riva-Posse et al., 2014; Scangos, Makhoul, Sugrue, Chang, & Krystal, 2021). Understanding the extent to which cognitive functions can be isolated in regions of psychiatric relevance (such as the reward circuitry examined here) will continue to add precision to our treatments. From a behavioral perspective, motivating patients to engage

in activities has been suggested as a treatment for severe depression (Dimidjian et al., 2006). However, the attenuated sensitivity to rewards experienced by this population presents a challenge for outcome-based approaches to incentivized behavior (Eisenberger & Cameron, 1996; Huys et al., 2013). Chapter 3 provided clues as to what outcome-independent features of a decision's environment can make certain types of effortful activities appealing. These characteristics, such as the availability of alternative activities or the segmentation of a task into discrete components, could benefit treatment-resistant patients as well as healthy individuals that fail to pursue otherwise desirable daily activities.

Future Directions

The neuroimaging findings in this dissertation were purposefully focused on a set of brain regions within the canonical default network that exhibit valuation effects (mPFC and PCC). Even though this circumscribed search space helped us confront the computational demands of fine-scaled connectivity mapping, it also constrained some of our interpretations. For instance, our claim that mPFC topography is particularly idiosyncratic could be strengthened in future work by comparing this region's heterogeneity to additional heteromodal or sensory areas. Regions such as dorsolateral prefrontal, posterior parietal, and lateral temporal cortex also contain default network regions and have been targets for qualitative examination (Braga & Buckner, 2017; DiNicola et al., 2020; Tavor et al., 2016), but to our knowledge have yet to receive the type of quantitative assessment undertaken here. Analyzing those areas would let us probe whether the same sub-network that tracked valuation in mPFC and PCC (Chapter

4) also overlaps with subregions that encode the decision costs of effort and self-control in dlPFC (Chong et al., 2017; Kool, McGuire, Wang, & Botvinick, 2013; McGuire & Botvinick, 2010). Finally, expanding the search space would permit a more formal test of the graded striatal organization suggested in Chapter 4. Even though this connectional gradient has already been functionally observed at the group-level using resting state data (Marquand et al., 2017), it remains to be seen whether this scheme persists within individuals. More orbital, caudal, and dorsal regions of mPFC than those examined here would be needed to provide a complete picture of the circuitry suggested by anatomical data (Haber & Knutson, 2010).

Additional considerations for our neuroimaging results pertain to the nature of decision processes. In Chapter 4, we failed to observe the hypothesized interaction between an individual's behavior and functional network topography. This was perhaps due in part because behavior was highly stable and consistent across sessions and across individuals, and also because the small participant sample was not conducive to establish such associations. A potential future approach would be to scan individuals as they performed the within-subjects version of the foraging task from Chapter 3, which evoked a systematic pattern of behavioral change over time. Another lingering question relates to the potential different encoding of different sources of value during decision making. Even though the appraisal of diverse prospects appears to rely on a common neural currency computed by vmPFC (Chib et al., 2009), some studies have proposed that different subregions of mPFC process the value of concrete and abstract prospects (i.e. food versus money, respectively; Clithero et al., 2014), and their desirability or goal-

congruency (Hare et al., 2008; Shenhav & Karmarkar, 2019). Future individualized work should examine whether these distinct valuation signals fall within the same or different interwoven networks.

At the behavioral level, Chapter 3 investigated the characteristics underlying dynamic preferences for different behavioral costs. The potential interpretation of our results in terms of variation in the subjective passage of time remains a hypothesis in need of further investigation. Future iterations of the same foraging experiment could adapt the configuration of each effortful task in order to test this idea directly. For example, if participants estimated the elapsed duration of a trial by counting each sub-task within cognitive effort trials rather than by tracking time continuously, it should be possible to influence acceptance rates by manipulating the number of discrete sub-tasks performed during a fixed handling time for either cognitive or physical effort trials. Alternatively, if the biases were produced by an inability to direct cognitive resources to the estimation of elapsed time (Zakay & Tsal, 1989), increasing working memory load during the handling time should reduce the OC biases from our computational model. This could be done, for example, by assigning a different level of n-back to each participant. These manipulations could identify specific features of the decision context that could motivate effortful behaviors in both healthy and clinical populations.

Conclusion

Recent neuroimaging work has highlighted the need to consider how different sources of individual variability can impact the localization of function in the brain, and whether distinct cognitive operations are indeed juxtaposed at the neural level. Building

upon this framework, this dissertation provided further evidence that, within individuals, the default network can be spatially refined along with at least one other interwoven network in mPFC and PCC, and that this arrangement is particularly idiosyncratic in mPFC. Further, it demonstrated that some canonical valuation effects are localized to each person's default network, providing support for theories that unite choice evaluation and the introspection-related processes associated with the default network as a common cognitive domain. The heterogeneous nature of functional topography and decision behavior displayed here invites reevaluation of the neural localization of additional decision-related processes at the individual level. This understanding can, in turn, potentially help craft personalized medical interventions for psychiatric disorders, many of which involve the brain's reward circuitry.

BIBLIOGRAPHY

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., … Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014

Acikalin, M. Y., Gorgolewski, K. J., & Poldrack, R. A. (2017). A coordinate-based meta-analysis of overlaps in regional specialization and functional connectivity across subjective value and default mode networks. *Frontiers in Neuroscience*, *11*(JAN), 1–11. https://doi.org/10.3389/fnins.2017.00001

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496. https://doi.org/10.1037/h0076860

Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., & Procyk, E. (2013). The location of feedback-related activity in the midcingulate cortex is predicted by local morphology. *Journal of Neuroscience*, *33*(5), 2217–2228. https://doi.org/10.1523/JNEUROSCI.2779-12.2013

Amiez, C., & Petrides, M. (2014). Neuroimaging evidence of the anatomo-functional organization of the human cingulate motor areas. *Cerebral Cortex*, *24*(3), 563–578. https://doi.org/10.1093/cercor/bhs329

Andrews-Hanna, J. R., Reidler, J. S., Huang, C., & Buckner, R. L. (2010). Evidence for the Default Network's Role in Spontaneous Cognition. *Journal of Neurophysiology*, *104*(1), 322–335. https://doi.org/10.1152/jn.00830.2009

Arulpragasam, A. R., Cooper, J. A., Nuutinen, M. R., & Treadway, M. T. (2018). Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(22), E5233–E5242. https://doi.org/10.1073/pnas.1800444115

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 353–364. https://doi.org/10.1038/nn.4502

Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2014). Learning-Induced Autonomy of Sensorimotor Systems. *Nature Neuroscience, 18*(5). https://doi.org/10.1038/nn.3993

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method ({CompCor}) for {BOLD} and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionalitu reduction and data representation. *Neural Computation*, *15*, 1373–1396. https://doi.org/10.1162/089976603321780317

Berridge, K. C. (2019). Affective valence in the brain: modules or modes? *Nature Reviews Neuroscience*, *20*(4), 225–234. https://doi.org/10.1038/s41583-019-0122-8

Białaszek, W., Marcowski, P., & Ostaszewski, P. (2017). Physical and cognitive effort discounting across different reward magnitudes: Tests of discounting models. *PLoS ONE*, *12*(7), 1–25. https://doi.org/10.1371/journal.pone.0182353

Bijsterbosch, J. D., Woolrich, M. W., Glasser, M. F., Robinson, E. C., Beckmann, C. F., Van Essen, D. C., … Smith, S. M. (2018). The relationship between spatial configuration and functional connectivity of brain regions. *ELife*, *7*, 1–27. https://doi.org/10.7554/eLife.32992

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

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., Van Dijk, K. R. A., Polimeni, J. R., Eldaief, M. C., & Buckner, R. L. (2019). Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *Journal of Neurophysiology*, jn.00808.2018. https://doi.org/10.1152/jn.00808.2018

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Buckner, R. L., & DiNicola, L. M. (2019). The Brain's Default Network: Updated Anatomy, Physiology, and Evolving Insights. *Nature Reviews Neuroscience*. https://doi.org/10.1038/s41583-019-0212-7

Burgess, G. C., Kandala, S., Nolan, D., Laumann, T. O., Power, J. D., Adeyemo, B., … & Barch, D. M. (2016). Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. *Brain Connectivity*, *6*(9), 669–680.

Cash, R. F. H., Cocchi, L., Lv, J., Fitzgerald, P. B., & Zalesky, A. (2020). Functional Magnetic Resonance Imaging–Guided Personalization of Transcranial Magnetic Stimulation Treatment for Depression. *Journal of the American Medical Association Psychiatry, 78(3):*337-339. https://doi.org/10.1001/jamapsychiatry.2020.3794

Cash, R. F. H., Fitzgerald, P. B., Cocchi, L., & Zalesky, A. (2021). Personalized connectivity-guided DLPFC-TMS for depression : Advancing computational feasibility , precision and reproducibility. *Human Brain Mapping.* 2021, 1–18. https://doi.org/10.1002/hbm.25330

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, Vol. 9, pp. 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, *29*(39), 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009

Chong, T. T. J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biology*, *15*(2), 1–28. https://doi.org/10.1371/journal.pbio.1002598

Chung, F. R. K. (1997). *Spectral Graph Theory* (Vol. 92). American Mathematical Soc.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E, 70(6)*, 1–6. https://doi.org/10.1103/PhysRevE.70.066111

Clithero, J. A., & Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, *9*(9), 1289–1302. https://doi.org/10.1093/scan/nst106

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238–251. https://doi.org/10.1016/j.neuron.2014.05.014

Conroy, B. R., Singer, B. D., Guntupalli, J. S., Ramadge, P. J., & Haxby, J. V. (2013). Inter-subject alignment of human cortical anatomy using functional connectivity. *NeuroImage*, *81*, 400–411. https://doi.org/10.1016/j.neuroimage.2013.05.009

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, *15*(4), 837–853. https://doi.org/10.3758/s13415-015-0350-y

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4–5), 171–178. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, *1695*.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999a). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Dale, A. M., Fischl, B., & Sereno, M. I. (1999b). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

De La Vega, A., Chang, L. J., Banich, M. T., Wager, T. D., & Yarkoni, T. (2016). Large-Scale Meta-Analysis of Human Medial Frontal Cortex Reveals Tripartite Functional Organization. *The Journal of Neuroscience*, *36*(24), 6553–6562. https://doi.org/10.1523/JNEUROSCI.4402-15.2016

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*(6), 3072–3077. https://doi.org/10.1152/jn.2000.84.6.3072

Di, X., Gohel, S., Kim, E. H., & Biswal, B. B. (2013). Task vs. rest-different network configurations between the coactivation and the resting-state brain networks. *Frontiers in Human Neuroscience*, *7*(SEP), 1–9. https://doi.org/10.3389/fnhum.2013.00493

DiNicola, L. M., Braga, R. M., & Buckner, R. L. (2020). Parallel distributed networks dissociate episodic and social functions within the individual. *J Neurophysiol*, *123*(3), 1144–1179. https://doi.org/10.1152/jn.00529.2019

Dundon, N. M., Garrett, N., Babenko, V., Cieslak, M., Daw, N. D., & Grafton, S. T. (2020). Sympathetic involvement in time-constrained sequential foraging. *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/https://doi.org/10.3758/s13415-020-00799-0

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*(2), 248–267. https://doi.org/10.1037/0033-295X.99.2.248

Esteban, O., Blair, R., Markiewicz, C. J., Berleant, S. L., Moodie, C., Ma, F., … Gorgolewski, K. J. (2018). fMRIPrep. *Software*. https://doi.org/10.5281/zenodo.852659

Esteban, O., Markiewicz, C., Blair, R. W., Moodie, C., Isik, A. I., Erramuzpe Aliaga, A., … Gorgolewski, K. J. (2018). {fMRIPrep}: a robust preprocessing pipeline for functional {MRI}. *Nature Methods*. https://doi.org/10.1038/s41592-018-0235-4

Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, *76*(6), 1057–1070. https://doi.org/10.1016/j.neuron.2012.12.002

Fawcett, T. W., McNamara, J. M., & Houston, A. I. (2012). When is it adaptive to be patient? A general framework for evaluating delayed rewards. *Behavioural Processes*, *89*(2), 128–136. https://doi.org/10.1016/j.beproc.2011.08.015

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*(21), 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011

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 of the United States of America*, *110*(41), 16616–16621. https://doi.org/10.1073/pnas.1315235110

Feher da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-020-0905-y

Fiedler, M. (1975). A Property of Eigenvectors of Nonnegative Symmetric Matrices and its Application to Graph Theory. *Czechoslovak Mathematical Journal*, *25*(100), 619–633.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete Coding of Reward Dopamine Neurons. *Science*, *299*(March), 1898–1902. https://doi.org/10.1126/science.1077349

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–

207. https://doi.org/10.1006/nimg.1998.0396

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*. https://doi.org/10.1016/S1053-8119(09)70884-5

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, *40*(2), 351–401.

Fritsch, A. (2012). *mcclust: Process an MCMC Sample of Clusterings. R package version 1.0.* Retrieved from https://cran.r-project.org/package=mcclust

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-12931-x

Garcia, J. O., Ashourvan, A., Muldoon, S., Vettel, J. M., & Bassett, D. S. (2018). Applications of Community Detection Techniques to Brain Graphs: Algorithmic Considerations and Implications for Neural Function. *Proceedings of the Institute of Electrical and Electronics Engineers*, *106*(5), 846–867. https://doi.org/10.1109/JPROC.2017.2786710

Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*(1), 1–12. https://doi.org/10.1038/s41467-020-16964-5

Ghasemian, A., Hosseinmardi, H., & Clauset, A. (2019). Evaluating Overfit and Underfit in Models of Network Community Structure. *Institute of Electrical and Electronics Engineers Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2019.2911585

Gkantsidis, C., Mihail, M., & Zegura, E. (2003). Spectral analysis of Internet topologies. *Twenty-second Annual Joint Conference of the Institute of Electrical and Electronics Engineers Computer and Communication Societies (IEEE Cat. No. 03CH37428)*, San Francisco, CA, USA, 2003, pp. 364–374 vol 1. https://doi.org/10.1109/INFCOM.2003.1208688

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

Goldman-Rakic, P. S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. *Annual Review of Neuroscience*, *11*, 137–156. https://doi.org/10.1146/annurev.ne.11.030188.001033

Gordon, E. M., Laumann, T. O., Adeyemo, B., Gilmore, A. W., Nelson, S. M., Dosenbach, N. U. F., & Petersen, S. E. (2017). Individual-specific features of brain systems identified with resting state functional correlations. *NeuroImage*, *146*, 918–939. https://doi.org/10.1016/j.neuroimage.2016.08.032

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., … Dosenbach, N. U. F. (2017a). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791-807.e7. https://doi.org/10.1016/j.neuron.2017.07.011

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., … Dosenbach, N. U. F. (2017b). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791-807.e7. https://doi.org/10.1016/j.neuron.2017.07.011

Gorgolewski, K. J., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics*, *5*, 13. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Esteban, O., Markiewicz, C. J., Ziegler, E., Ellis, D. G., Notter, M. P., … Ghosh, S. (2018). Nipype. *Software*. https://doi.org/10.5281/zenodo.596855

Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., … Petersen, S. E. (2018). Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. *Neuron*, 439–452. https://doi.org/10.1016/j.neuron.2018.03.035

Green, L., Fristoe, N., & Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, *1*(3), 382–389. https://doi.org/10.3758/BF03213979

Gregorios-Pippas, L., Tobler, P. N., & Schultz, W. (2009). Short-Term Temporal Discounting of Reward Value in Human Ventral Striatum. *Journal of Neurophysiology*, *101*(3), 1507–1523. https://doi.org/10.1152/jn.90730.2008

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253–258. https://doi.org/10.1073/pnas.0135058100

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., … Smith, S. M. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*, *95*, 232–247. https://doi.org/10.1016/j.neuroimage.2014.03.034

Grover, S., Nguyen, J. A., Viswanathan, V., & Reinhart, R. M. G. (2021). High-frequency neuromodulation improves obsessive–compulsive behavior. *Nature Medicine*. https://doi.org/10.1038/s41591-020-01173-w

Gruber, A. J., & McDonald, R. J. (2012). Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior. *Frontiers in Behavioral Neuroscience, 6.* https://doi.org/10.3389/fnbeh.2012.00050

Guntupalli, J. S., Feilong, M., & Haxby, J. V. (2018). A computational model of shared fine-scale structure in the human connectome. *PLoS Computational Biology*, *14*(4), 1–26. https://doi.org/10.1371/journal.pcbi.1006120

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4–26. https://doi.org/10.1038/npp.2009.129

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience*, *28*(22), 5623–5630. https://doi.org/10.1523/JNEUROSCI.1309-08.2008

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting

of monetary rewards by physical effort. *Behavioural Processes*, *100*(October 2013), 192–196. https://doi.org/10.1016/j.beproc.2013.09.014

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Hayden, B. Y. (2018). Economic choice: the foraging perspective. *Current Opinion in Behavioral Sciences*, Vol. 24. https://doi.org/10.1016/j.cobeha.2017.12.002

Hernandez Lallement, J., Kuss, K., Trautner, P., Weber, B., Falk, A., & Fliessbach, K. (2014). Effort increases sensitivity to reward and loss magnitude in the human brain. *Social Cognitive and Affective Neuroscience*, *9*(3), 342–349. https://doi.org/10.1093/scan/nss147

Higham, D. J., Kalna, G., & Kibble, M. (2007). Spectral clustering and its use in bioinformatics. *Journal of Computational and Applied Mathematics*, *204*(1), 25–37. https://doi.org/10.1016/j.cam.2006.04.026

Hiser, J., & Koenigs, M. (2018). The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making, Social Cognition, and Psychopathology. *Biological Psychiatry*, *83*(8), 638–647. https://doi.org/10.1016/j.biopsych.2017.10.030

Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, *500*(7464), 575–579. https://doi.org/10.1038/nature12475

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, *2*(1), 193–218. https://doi.org/10.1007/BF01908075

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). *Functional magnetic resonance imaging* (3rd ed.). New York, NY: Oxford University Press.

Hull, C. L. (1943). Principles of Behavior. In R. M. Elliott (Ed.), *Principles of Behavior, An Introduction to Behavior Theory*. New York, NY: Appleton-Century.

Huntenburg, J. M., Bazin, P. L., & Margulies, D. S. (2018). Large-Scale Gradients in Human Cortical Organization. *Trends in Cognitive Sciences*, *22*(1), 21–31. https://doi.org/10.1016/j.tics.2017.11.002

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Computational Biology*, *7*(4). https://doi.org/10.1371/journal.pcbi.1002028

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3). https://doi.org/10.1371/journal.pcbi.1002410

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The Effort Paradox: Effort Is Both Costly and Valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Jiang, R., Zuo, N., Ford, J. M., Qi, S., Zhi, D., Zhuo, C., … Sui, J. (2020). Task-induced brain connectivity promotes the detection of individual differences in brain-behavior relationships. *NeuroImage*, *207*(November 2019), 116370. https://doi.org/10.1016/j.neuroimage.2019.116370

Joyce, M. K. P., & Barbas, H. (2018). Cortical connections position primate area 25 as a keystone for interoception, emotion, and memory. *The Journal of Neuroscience*, *38*(7), 2363–17. https://doi.org/10.1523/JNEUROSCI.2363-17.2017

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625–1633.

Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *ANIMAL BEHAVIOUR*, *63*, 245–250. https://doi.org/10.1006/anbe.2001.1900

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341. https://doi.org/10.1037/0003-066X.39.4.341

Kane, G. A., Bornstein, A. M., Shenhav, A., Wilson, R. C., Daw, N. D., & Cohen, J. D. (2019). Rats exhibit similar biases in foraging and intertemporal choice tasks. *ELife*, *8*, 1–22. https://doi.org/10.7554/elife.48429

Kernbach, J. M., Yeo, B. T. T., Smallwood, J., Margulies, D. S., Thiebaut de Schotten, M., Walter, H., … Bzdok, D. (2018). Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants. *Proceedings of the National Academy of Sciences*, (November), 201804876. https://doi.org/10.1073/pnas.1804876115

Klein-Flugge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. *Journal of Neuroscience*, *36*(39), 10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016

Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral Modeling of Human Choices Reveals Dissociable Effects of Physical Effort and Temporal Delay on Reward Devaluation. *PLOS Computational Biology*, *11*(3), e1004116. https://doi.org/10.1371/journal.pcbi.1004116

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., … Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *NeuroImage*, *18*(2), 263–272. https://doi.org/10.1016/S1053-8119(02)00057-5

Kong, R., Li, J., Orban, C., Sabuncu, M. R., Liu, H., Schaefer, A., … Yeo, B. T. T. (2018). Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cerebral Cortex*, (May 2018), 2533–2551. https://doi.org/10.1093/cercor/bhy123

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(December), 899–908. https://doi.org/10.1038/s41562-018-0401-9

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision Making and the Avoidance of Cogntive Demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198.Decision

Kool, W., McGuire, J. T., Wang, G. J., & Botvinick, M. M. (2013). Neural and Behavioral Evidence for an Intrinsic Cost of Self-Control. *PLoS ONE*, *8*(8), 1–6. https://doi.org/10.1371/journal.pone.0072626

Kragel, P. A., Kano, M., Van Oudenhove, L., Ly, H. G., Dupont, P., Rubio, A., … Wager, T. D. (2018). Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nature Neuroscience*, 1. https://doi.org/10.1038/s41593-017-0051-7

Kraus, B. (2020). Network variants are similar between task and rest states. *NeuroImage*, *118370*, 117743. https://doi.org/10.1016/j.neuroimage.2021.117743

Krebs, B. Y. J. R., Erichsen, J. T., & Webber, M. I. (1977). Optimal prey slection in the great tit (*Parsus major*). *Animal Behavior*, *25*(2), 30–38.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(06), 661–679. https://doi.org/10.1017/S0140525X12003196

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., & Fox, P. T. (2009). Investigating the Functional Heterogeneity of the Default Mode Network Using Coordinate-Based Meta-Analytic Modeling. *Journal of Neuroscience*, *29*(46), 14496–14505. https://doi.org/10.1523/JNEUROSCI.4004-09.2009

Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., … Fox, P. T. (2011). Behavioral Interpretations of Intrinsic Connectivity Networks. *Journal of Cognitive Neuroscience*, *23*(12), 4022–4037. https://doi.org/10.1162/jocn_a_00077

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85. https://doi.org/10.1137/0701007

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M. Y., … Petersen, S. E. (2015). Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron*, *87*(3), 658–671. https://doi.org/10.1016/j.neuron.2015.06.037

Levy, I., Lazzaro, S. C., Rutledge, R. B., & Glimcher, P. W. (2011). Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. *Journal of Neuroscience*, *31*(1), 118–125. https://doi.org/10.1523/JNEUROSCI.3214-10.2011

Lopez-Persem, A., Verhagen, L., Amiez, C., Petrides, M., & Sallet, J. (2019). The human ventromedial prefrontal cortex: sulcal morphology and its influence on functional organization. *The Journal of Neuroscience*, *39*(19), 2060–18. https://doi.org/10.1523/JNEUROSCI.2060-18.2019

Mackey, S., & Petrides, M. (2014). Architecture and morphology of the human ventromedial prefrontal cortex. *European Journal of Neuroscience*, *40*(5), 2777–2796. https://doi.org/10.1111/ejn.12654

Mansour L, S., Tian, Y., Yeo, B. T. T., Cropley, V., & Zalesky, A. (2021). High-resolution connectomic fingerprints: mapping neural identity and behavior. *NeuroImage, 229*. https://doi.org/10.1016/j.neuroimage.2020.117695

Marcus, D. S., Harwell, J., Olsen, T., Hodge, M., Glasser, M. F., Prior, F., … Van Essen, D. C. (2011). Informatics and data mining tools and strategies for the human connectome project. *Frontiers in Neuroinformatics*, *5*(June), 1–12. https://doi.org/10.3389/fninf.2011.00004

Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., … 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

Marquand, A. F., Haak, K. V, & Beckmann, C. F. (2017). Functional corticostriatal connection topographies predict goal-directed behaviour in humans. *Nature Human Behavior, 1,* 0146. https://doi.org/10.1038/s41562-017-0146

Mars, R. B., Passingham, R. E., & Jbabdi, S. (2018). Connectivity Fingerprints: From Areal Descriptions to Abstract Spaces. *Trends in Cognitive Sciences*, *22*(11), 1026–1037. https://doi.org/10.1016/j.tics.2018.08.009

Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *NeuroImage*, *120*, 104–113. https://doi.org/10.1016/j.neuroimage.2015.06.080

McGuire, J. T., & Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 7922–7926. https://doi.org/10.1073/pnas.0910662107

McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition*, *124*(2), 216–226. https://doi.org/10.1016/j.cognition.2012.03.008

McGuire, J. T., & Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nature Neuroscience*, *18*(5), 760–769.

McKiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., & Binder, J. R. (2003). A Parametric Manipulation of Factors Affecting Task-induced Deactivation in Functional Neuroimaging. *Journal of Cognitive Neuroscience*, *15*, 394–408. https://doi.org/10.1162/089892903321593117

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, *16*(4), 479–485. https://doi.org/10.1038/nn.3337

Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., & Somers, D. C. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. *Neuron*, *87*(4), 882–892. https://doi.org/10.1016/j.neuron.2015.07.028

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*(2), 329–337. https://doi.org/10.1037/h0029815

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315.

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: insights from ethology. *Nature Reviews Neuroscience*, 1. https://doi.org/10.1038/s41583-018-0010-7

Mochon, D., Norton, M. I., & Ariely, D. (2012). Bolstering and restoring feelings of competence via the IKEA effect. *International Journal of Research in Marketing*, *29*(4), 363–369. https://doi.org/10.1016/j.ijresmar.2012.05.001

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgment of numerical inequality. *Nature*, *215*, 1519–1520.

Mueller, S., Wang, D., Fox, M. D., Yeo, B. T. T., Sepulcre, J., Sabuncu, M. R., … Liu, H. (2013). Individual Variability in Functional Connectivity Architecture of the Human Brain. *Neuron*, *77*(3), 586–595. https://doi.org/10.1016/j.neuron.2012.12.028

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, *59*(3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076

Nenning, K.-H., Xu, T., Schwartz, E., Arroyo, J., Woehrer, A., Franco, A. R., … Langs, G. (2020). Joint embedding: A scalable alignment to compare individuals in a connectivity space. *NeuroImage*, *222*, 117232. https://doi.org/10.1016/j.neuroimage.2020.117232

Nenning, K. H., Liu, H., Ghosh, S. S., Sabuncu, M. R., Schwartz, E., & Langs, G. (2017). Diffeomorphic functional brain surface alignment: Functional demons. *NeuroImage*, *156*(December 2016), 456–465. https://doi.org/10.1016/j.neuroimage.2017.04.028

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*, 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Noyce, A. L., Cestero, N., Michalka, S. W., Shinn-Cunningham, B. G., & Somers, D. C. (2017). Sensory-biased and multiple-demand processing in human lateral frontal cortex. *Journal of Neuroscience*, *37*(36), 8755–8766. https://doi.org/10.1523/JNEUROSCI.0660-17.2017

Olivola, C. Y., & Shafir, E. (2013). The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *Journal of Behavioral Decision Making*, *26*(1), 91–105. https://doi.org/10.1002/bdm.767

Osher, D. E., Saxe, R. R., Koldewyn, K., Gabrieli, J. D. E., Kanwisher, N., & Saygin, Z. M. (2016). Structural Connectivity Fingerprints Predict Cortical Selectivity for Multiple Visual Categories across Cortex. *Cerebral Cortex*, *26*(4), 1668–1683. https://doi.org/10.1093/cercor/bhu303

Otto, A. R., & Daw, N. D. (2019). The opportunity cost of time modulates cognitive effort. *Neuropsychologia*, *123*(May 2018), 92–105. https://doi.org/10.1016/j.neuropsychologia.2018.05.006

Passingham, R. E., Stephan, K. E., & Kötter, R. (2002). The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience*, *3*(8), 606–616. https://doi.org/10.1038/nrn893

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., … Petersen, S. E. (2011). Functional Network Organization of the Human Brain. *Neuron*, *72*(4), 665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014a). 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., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014b). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*(Supplement C), 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Prelec, D., & Loewenstein, G. (1991). Decision Making over Time and under Uncertainty : A Common Approach. *Management Science*, *37*(7), 770–786.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, *28*(11), 2745–2752. https://doi.org/10.1523/JNEUROSCI.4286-07.2008

Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L., & Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *Journal of Neuroscience*, *30*(42), 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-

10.2010

R Core Computing Team. (2017). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.r-project.org/

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Ramnani, N., Elliott, R., Athwal, B. S., & Passingham, R. E. (2004). Prediction error for free monetary reward in the human prefrontal cortex. *NeuroImage*, *23*(3), 777–786. https://doi.org/10.1016/j.neuroimage.2004.07.028

Reddan, M. C., Wager, T. D., & Schiller, D. (2018). Attenuating Neural Threat Expression with Imagination. *Neuron*, *100*(4), 994-1005.e4. https://doi.org/10.1016/j.neuron.2018.10.047

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Ribas-Fernandes, J. J. F., Shahnazian, D., Holroyd, C. B., & Botvinick, M. M. (2019). Subgoal- and Goal-related Reward Prediction Errors in Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *31*(1), 8–23. https://doi.org/10.1162/jocn_a_01341

Riva-Posse, P., Choi, K. S., Holtzheimer, P. E., McIntyre, C. C., Gross, R. E., Chaturvedi, A., … Mayberg, H. S. (2014). Defining Critical White Matter Pathways Mediating Successful Subcallosal Cingulate Deep Brain Stimulation for Treatment-Resistant Depression. *Biological Psychiatry*, *76*(12), 963–969. https://doi.org/10.1016/j.biopsych.2014.03.029

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Salehi, M., Greene, A. S., Karbasi, A., Shen, X., Scheinost, D., & Constable, R. T. (2019). There is no single functional atlas even for a single individual: Functional parcel definitions change with task. *NeuroImage*, 116366. https://doi.org/10.1016/J.NEUROIMAGE.2019.116366

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*,

90, 449–468. https://doi.org/10.1016/j.neuroimage.2013.11.046

Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D. E., & Saxe, R. R. (2011). Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nature Neuroscience*, *15*(2), 321–327. https://doi.org/10.1038/nn.3001

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., … Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Scangos, K. W., Makhoul, G. S., Sugrue, L. P., Chang, E. F., & Krystal, A. D. (2021). State-dependent responses to intracranial brain stimulation in a patient with depression. *Nature Medicine*. https://doi.org/10.1038/s41591-020-01175-8

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657–661. https://doi.org/10.1038/nrn2213

Schiller, D., Levy, I., Niv, Y., LeDoux, J. E., & Phelps, E. A. (2008). From fear to safety and back: Reversal of fear in the human brain. *Journal of Neuroscience*, *28*(45), 11517–11525. https://doi.org/10.1523/JNEUROSCI.2265-08.2008

Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*, (95), 853–951.

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32. https://doi.org/10.1038/nrn.2015.26

Seaman, K. L., Brooks, N., Karrer, T. M., Castrellon, J. J., Perkins, S. F., Dang, L. C., … Samanez-Larkin, G. R. (2018). Subjective value representations during effort, probability and time discounting across adulthood. *Social Cognitive and Affective Neuroscience*, *13*(5), 449–459. https://doi.org/10.1093/scan/nsy021

Shenhav, A., & Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. *Scientific Reports*, *9*(1), 172320. https://doi.org/10.1038/s41598-019-38927-7

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*(1). https://doi.org/10.1146/annurev-neuro-072116-031526

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common Blood Flow Changes across Visual Tasks: II.

Decreases in Cerebral Cortex. *Journal of Cognitive Neuroscience*, *9*(5), 648–663. https://doi.org/10.1162/jocn.1997.9.5.648

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., … 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 LP – 13045. https://doi.org/10.1073/pnas.0905267106

Spreng, N. (2012). The fallacy of a "task-negative" network. *Frontiers in Psychology*, *3*(MAY), 1–5. https://doi.org/10.3389/fpsyg.2012.00145

Starkweather, C. K., Gershman, S. J., & Uchida, N. (2018). The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty. *Neuron*, *98*(3), 616-629.e6. https://doi.org/10.1016/j.neuron.2018.03.036

Stephens, D. W., & Krebs, J. W. (1986). *Foraging Theory*. Princeton University Press.

Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 216–220. https://doi.org/10.1126/science.aad8127

Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020a). Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience*. https://doi.org/10.1038/s41593-020-00711-6

Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020b). Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience*, *23*(11), 1421–1432. https://doi.org/10.1038/s41593-020-00711-6

Tian, Y., & Zalesky, A. (2018). Characterizing the functional connectivity diversity of the insula cortex: Subregions, diversity curves and behavior. *NeuroImage*, *183*, 716–733. https://doi.org/10.1016/j.neuroimage.2018.08.055

Tobyne, Sean M., Somers, D. C., Brissenden, J. A., Michalka, S. W., Noyce, A. L., & Osher, D. E. (2018). Prediction of individualized task activation in sensory modality-selective frontal cortex with 'connectome fingerprinting.' *NeuroImage*, *183*, 173–185.

Tobyne, S. M., Osher, D. E., Michalka, S. W., & Somers, D. C. (2017). Sensory-biased attention networks in human lateral frontal cortex revealed by intrinsic functional connectivity. *NeuroImage*, *162*(August), 362–372. https://doi.org/10.1016/j.neuroimage.2017.08.020

Toker, D., & Sommer, F. T. (2019). Information Integration In Large Brain Networks.

PLoS Computational Biology, *15*(2). https://doi.org/10.1371/journal.pcbi.1006807

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518. https://doi.org/10.1126/science.1134239

Toro-Serey, C., Tobyne, S. M., & McGuire, J. T. (2020). Spectral partitioning identifies individual heterogeneity in the functional network topography of ventral and anterior medial prefrontal cortex. *NeuroImage*, *205*. https://doi.org/10.1016/j.neuroimage.2019.116305

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the "EEfRT"? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS ONE*, *4*(8), 1–9. https://doi.org/10.1371/journal.pone.0006598

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Uddin, L. Q., Yeo, B. T. T., & Spreng, R. N. (2019). Towards a Universal Taxonomy of Macro - scale Functional Human Brain Networks How Many Functional Brain Networks Are A fundamental construct in neuroscience is the definition. *Brain Topography*, *32*(6), 926–942. https://doi.org/10.1007/s10548-019-00744-6

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., … Yacoub, E. (2012). The Human Connectome Project: A data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., & Huettel, S. A. (2009). Separate Neural Mechanisms Underlie Choices and Strategic Preferences in Risky Decision Making. *Neuron*, *62*(4), 593–602. https://doi.org/10.1016/j.neuron.2009.04.007

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on Psychological Science*, *4*(3), 319–324. https://doi.org/10.1111/j.1745-6924.2009.01132.x

Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H., & Van Snellenberg, J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *NeuroImage*, *45*(1 Suppl), S210–S221. https://doi.org/10.1016/j.neuroimage.2008.10.061

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, *19*(8), 1302–1314. https://doi.org/10.1016/j.neunet.2006.03.005

Wang, D., Buckner, R. L., Fox, M. D., Holt, D. J., Holmes, A. J., Stoecklein, S., … Liu, H. (2015). Parcellating cortical functional networks in individuals. *Nature Neuroscience*, *18*(12), 1853–1860. https://doi.org/10.1038/nn.4164

Welvaert, M., Durnez, J., Moerkerke, B., Verdoolaege, G., & Rosseel, Y. (2011). neuRosim: An R Package for Generating fMRI Data. *Journal of Statistical Software*, *44*(10), 1–18. Retrieved from http://www.jstatsoft.org/v44/i10/

Westbrook, A., & Braver, T. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 395–415. https://doi.org/10.3758/s13415-015-0334-y

Westbrook, A., Kester, D., & Braver, T. S. (2013). What Is the Subjective Cost of Cognitive Effort? Load, Trait, and Aging Effects Revealed by Economic Preference. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0068210

Westgate, E. C. (2020). Why Boredom Is Interesting. *Current Directions in Psychological Science*, *29*(1), 33–40. https://doi.org/10.1177/0963721419884309

Woo, C.-W., Koban, L., Kross, E., Lindquist, M. A., Banich, M. T., Ruzic, L., … Wager, T. D. (2014). Separate neural representations for physical pain and social rejection. *Nature Communications*, *5*(May), 1–12. https://doi.org/10.1038/ncomms6380

Yeo, B. T. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., … Chee, M. W. L. (2015). Functional specialization and flexibility in human association cortex. *Cerebral Cortex*, *25*(10), 3654–3672. https://doi.org/10.1093/cercor/bhu217

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*, 1125–1165. https://doi.org/10.1152/jn.00338.2011

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, *22*(3), 181–192. https://doi.org/10.1038/s41583-020-00420-w

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*,

20(1), 45–57. https://doi.org/10.1109/42.906424

Zilles, K., Palomero-Gallagher, N., & Amunts, K. (2013). Development of cortical folding during evolution and ontogeny. *Trends in Neurosciences*, *36*(5), 275–284. https://doi.org/10.1016/j.tins.2013.01.006

Zlatkina, V., Amiez, C., & Petrides, M. (2016). The postcentral sulcal complex and the transverse postcentral sulcus and their relation to sensorimotor functional organization. *European Journal of Neuroscience*, *43*(10), 1268–1283. https://doi.org/10.1111/ejn.13049

CURRICULUM VITAE