University of California
Santa Barbara

Network-Based Investigations of Human Functional Brain Dynamics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Kimberly J. Schlesinger

Committee in charge:

Professor Jean Carlson, Chair
Professor Scott Grafton
Professor Everett Lipman
Professor Boris Shraiman

December 2017

The Dissertation of Kimberly J. Schlesinger is approved.

Professor Scott Grafton

Professor Everett Lipman

Professor Boris Shraiman

Professor Jean Carlson, Committee Chair

October 2017

Network-Based Investigations of Human Functional Brain Dynamics

by

Kimberly J. Schlesinger

Acknowledgements

Kimberly Schlesinger was supported during the completion of this work by the National Science Foundation Graduate Research Fellowship Program under grant DGE-1144085, the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office, and the Worster Fellowship.

Elizabeth Davison acknowledges additional support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1656466 and the Francis Robbins Upton Fellowship in Engineering. Danielle Bassett acknowledges support from the John D. and Catherine T. MacArthur Foundation, the Army Research Laboratory and the Army Research Office through contract numbers W911NF-10-2-0022 and W911NF-14-1-0679, the National Institute of Mental Health (2-R01-DC-009209-11), the National Institute of Child Health and Human Development (1R01HD086888-01), the Office of Naval Research, and the National Science Foundation (#BCS-1441502, #BCS-1430087, and #PHY-1554488). Jean Carlson acknowledges additional support from the David and Lucile Packard Foundation.

The content of the information in this dissertation does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Finally, Kimberly Schlesinger would like to thank the following people for their contributions to this work: Jean Carlson, Liz Davison, Ben Turner, Charles Lieou, Dani Bassett, Ann Hermundstad, Scott Grafton, Mike Miller, Chantal Nguyen, Matthew Cieslak, Stella Von Meer, Joshua Mueller, Eric Jones, Sean Stromberg, Kuang Wei, Alberto Busetto, Jeff Teeters, Brian Mitchell, Brian Lopez, Mary-Ellen Lynall, Ben LaRoque, Stacy Copp, Keith Fratus, and John Bushnell.

Curriculum Vitæ

Kimberly J. Schlesinger

Education

2017	Ph.D., Physics (Expected), University of California, Santa Barbara.
2014	M.A., Physics, University of California, Santa Barbara.
2011	B.S., Physics, University of Notre Dame.
2011	B.A., English, University of Notre Dame.

Publications

K.J. Schlesinger, B.O. Turner, S.T. Grafton, M.B. Miller, J.M. Carlson. Targeted Node Removal for Improved Resolution of Communities in Dynamic Networks. PLOS ONE. In Production.

C. Nguyen, **K.J. Schlesinger**, and J.M. Carlson. Collective Decision Dynamics in Group Evacuation. Proceedings of the Summer Computer Simulation Conference (SCSC), 10 July 2017.

E.N. Davison, B.O. Turner, **K.J. Schlesinger**, M.B. Miller, S.T. Grafton, D.S. Bassett, and J.M. Carlson. Individual Differences in Dynamic Functional Brain Connectivity Across the Human Lifespan. PLoS Computational Biology 12 (11), e1005178, 23 November 2016.

K.J. Schlesinger, B.O. Turner, B.A. Lopez, M.B. Miller, and J.M. Carlson. Age-dependent Changes in Dynamic Modular Organization of the Human Brain. NeuroImage 146, 741-762, 1 February 2017. Epub 3 September 2016.

E.N. Davison, **K.J. Schlesinger**, D.S. Bassett, M.E. Lynall, M.B. Miller, S.T. Grafton, and J.M. Carlson. Brain Network Adaptability Across Task States. PLoS Computational Biology 11 (1), e1004029, 8 January 2015.

K.J. Schlesinger, S.P. Stromberg, and J.M. Carlson. Coevolutionary Immune System Dynamics Driving Pathogen Speciation. PLoS ONE 9(7), e102821, 23 July 2014.

W.J. Gannon, W.P. Halperin, C. Rastovski, **K.J. Schlesinger**, J. Hlevyack, M.R. Eskildsen, A.B. Vorontsov, J. Gavilano, U. Gasser, and G. Nagy. Nodal Gap Structure and Order Parameter Symmetry of the Unconventional Superconductor UPt_3. New Journal of Physics 17, 023041, 13 February 2015.

C. Rastovski, **K.J. Schlesinger**, W.J. Gannon, C.D. Dewhurst, L. DeBeer-Schmitt, N.D. Zhigadlo, J. Karpinski, and M.R. Eskildsen. Persistence of Metastable Vortex Lattice Domains in MgB_2 in the Presence of Vortex Motion. Physical Review Letters 111, 107002, 4 September 2013.

P. Das, J.M. Densmore, C. Rastovski, **K.J. Schlesinger**, M. Laver, C.D. Dewhurst, K. Littrell, S.L. Budko, P.C. Canfield, and M.R. Eskildsen. Field dependence of the superconducting basal plane anisotropy of $TmNi_2B_2C$. Physical Review B 86 (14), 144501, 1 October 2012.

P. Das, C. Rastovski, T.R. OBrien, **K.J. Schlesinger**, C.D. Dewhurst, L. DeBeer-Schmitt, N.D.Zhigadlo, J. Karpinski, and M.R. Eskildsen. Observation of Well-Ordered Metastable Vortex Lattice Phases in Superconducting MgB_2 Using Small Angle Neutron Scattering. Physical Review Letters 108, 167001, 20 April 2012.

Presentations

K.J. Schlesinger and J.M. Carlson. Identification, Validation, and Control of Networked Neuronal Systems. Observability and Controllability of Network Dynamics Workshop, Mathematical Biosciences Institute, Columbus, OH. 11 April 2016. Invited talk.

K.J. Schlesinger, B.O. Turner, B.A. Lopez, M.B. Miller, and J.M. Carlson. Age-dependent community dynamics and brain system organization in human functional brain networks. Society for Neuroscience Annual Meeting, Chicago, IL. October 2015. Contributed poster.

K.J. Schlesinger, S.P. Stromberg, and J.M. Carlson. Dynamics of Coevolution and Branching in the Immune System. American Physical Society March Meeting, Denver, CO. 5 March 2014. Contributed talk.

Abstract

Network-Based Investigations of Human Functional Brain Dynamics

by

Kimberly J. Schlesinger

The human brain is a complex system in which interactions of billions of neurons give rise to behavior. fMRI allows researchers to measure the functional activity of the working brain, allowing both the localization of specific functions within the brain and the investigation of multivariate patterns of functional activation. These patterns have been found to correspond both to short-term brain states such as focused attention or daydreaming, and to characteristics such as age or disease. Functional patterns also show substantial variation across individuals. Understanding the correspondence of distributed functional activity to these various factors is an ongoing research area.

Network science is a valuable tool for representing complex brain function, providing a framework for quantifying multivariate activity as a network of interactions. Here, we build upon recent advances in dynamic network science, using time-evolving networks to investigate how the organization of brain dynamics is related to demographics and brain states.

We use *hypergraphs* to analyze brain network dynamics during different cognitive tasks and the transitions between them. We identify the presence of *hyperedges*, groups of functional interactions that fluctuate coherently in strength over time both within and across brain states. We develop metrics to quantify the variation of hyperedge structure between tasks and across individuals. We find that the spatial location of hyperedges is relatively consistent across individuals, serving as a signature of a cognitive task, while hyperedge size exhibits variation across individuals but remains consistent between tasks.

We also investigate the variation of brain dynamics across the human lifespan, using both hypergraphs and dynamic clusters, or *communities*, of brain regions with similar activity. We find significant relationships between age and dynamic organization: younger subjects tend to have larger hyperedges, as well as less fragmented and more coherent communities, and their brain regions tend to switch between communities less often. Further, the dynamics of different cognitive brain systems respond differently to aging.

Finally, we propose and evaluate a method of targeted node removal during the data-driven detection of communities, using synthetic and fMRI-derived networks to show that the method can improve identification of multi-scale community structure, and help to resolve key features of community dynamics.

Contents

Chapter 1

Introduction

1.1 Mapping Human Brain Function: A Brief History

For hundreds of years, humans have been working to understand exactly how the brain's structure and function give rise to the fascinating range of behaviors of which we are capable. The earliest investigators focused mainly on anatomical studies, developing a detailed understanding of the physical structure of the human brain through post-mortem studies. However, when it came to uncovering relationships between structure and function, scientists and doctors were hampered by among the greatest challenges in neuroscience: they lacked the ability to observe and measure electrical activity in the living, working brain, without causing damage or destruction to either the body or the mind of the patient.

In 1848, Phineas Gage, a railroad construction foreman, suffered a gruesome accident in which an explosion drove a sharp metal tamping rod entirely through his head. Despite severe brain damage to his left frontal lobe, Gage fully recovered from the accident,

but was reported to have notably altered personality traits. This famous case was used afterwards to support the theory that some functions were localized within the brain, and that damage to one brain area could selectively impact certain abilities and behaviors while leaving others unchanged [1]. It also typified what became an important trend in the study of brain function in the 19th and early 20th centuries: because researchers could not ethically measure or interfere with function in a healthy human subject's brain, many key breakthroughs came from studying cases of accidental brain injury, disease, or medically necessary intervention.

There are many classic examples of this approach, which led to an understanding of brain function pieced together from a patchwork of tragic but informative case studies. Not long after Gage's case, Paul Broca and Karl Wernicke localized aspects of language by studying patients with aphasia who had experienced brain lesions in specific cortical regions in the left cerebral hemisphere [2]. Lobectomies led to many influential case studies, including that of the famous patient H.M., whose extreme episodic memory loss was studied extensively after most of his hippocampus was removed as a treatment for severe epilepsy [3]. Roger Sperry and Michael Gazzaniga made groundbreaking discoveries about the lateralization of brain function by studying split-brain patients whose corpus collosum had been severed for medical reasons [4, 5].

In the 20th century, advances in medical imaging made less invasive measurements of human brain function possible. Electroencephalography uses electrodes placed on the scalp to record electrical activity from the brain, and electrocorticography does the same with electrodes placed inside the skull during brain surgery. PET scans detect gamma rays emitted by an injected radionuclide, which correlate with oxygen flow in the brain. None of these cause extensive damage or injury, but they are not without risks (ECoG, PET), or have extremely limited spatial resolution (EEG) [2]. Thus, the focus largely remained on medically necessary imaging, for which these risks could be justified.

1.2 fMRI: A Window into the Healthy Brain

After functional MRI was discovered in 1990, it quickly because an invaluable tool, as it finally provided a way to move beyond a disease-centric study of brain function. For the first time, researchers could non-invasively image the activity in healthy, living brains at a high spatial resolution.

fMRI works by measuring changes in blood flow coupled to neural activity. When neural activity occurs in an area of the brain, a hemodynamic response is triggered: blood flow to that area increases temporarily, and the blood oxygen levels are depleted and then replenished over several seconds, producing a relative change in the ration of oxygenated to deoxygenated blood in the region. Because these two types of blood have differing magnetic properties, this also creates a local change in the MR signal, called the blood-oxygen-level-dependent (BOLD) contrast. By tracking the BOLD signal during an MRI scan, researchers can non-invasively map blood flow corresponding to neural activity across the entire brain. The BOLD time series is typically sampled on the order of once per second, from each of over 100,000 voxels; voxels are on the order of a millimeter in size and contain around 100,000 to 1,000,000 neurons on average [6].

Over the past 25 years, fMRI has allowed researchers to non-invasively form maps of brain activity in healthy populations. Due in part to the unprecedented spatial resolution of the technique, a major focus of the research done with fMRI over its first few decades has been localizing specific functions in the brain. Typically, controlled experiments and statistical hypothesis testing are used to determine whether any voxels in the brain show significantly altered BOLD signal between the baseline activity and the function of interest, across a population of participants. Despite challenges in the pre-processing of the noisy BOLD signal, and the statistical intricacies of controlling for multiple sources of noise and thousands of comparisons, such studies have revolutionized our ability to

understand the spatial patterns of activity in the brain, and highlighted anatomical areas that are generally associated with specific functions and behaviors [6].

1.3 Dynamic Networks and Human Brain Function

While much of the success of fMRI has come in localizing specific brain functions, it is widely accepted that many complex abilities and behaviors require the dynamic integration of several processes and areas across multiple regions of the brain, and that higher-level cognition likely utilizes distributed representations of concepts and memories. In order to probe the distributed organization of human brain activity, there has been growing interest in analyzing fMRI images with methods that can take their multivariate properties into account, and associate behaviors with distributed patterns rather than simply voxel-by-voxel activation [7].

In recent years, network science has emerged as a method well suited to answering these problems. Networks, or graphs, provide a well-studied mathematical framework for describing complex systems, including brain function. A network model allows us to distill the information contained in thousands of interactions between brain areas into graph-based metrics that capture important local and global properties of the brains organization. These metrics can then be used to quantify, classify, and compare these properties across conditions and individual brains, often providing insights into fundamental organizational principles underlying the complex activity we observe [7–9].

Functional brain networks derived from fMRI data are represented as a collection of brain regions, or nodes, and connections between pairs of brain regions, or edges. Typically, brain regions are groups of spatially adjacent voxels defined by anatomical boundaries within the brain. Each pair of brain regions has an edge between them, which is given a weight that represents the similarity of the BOLD signals in the two regions

over time. If two regions show highly correlated activity during a scanning session, their connecting edge will have a higher weight than that between two regions with unrelated activity.

Based upon the nodes and their weighted edges, several widely used graph metrics can be used to capture properties of the functional brain network and characterize its organization. Previous work on human functional brain networks has established consistent patterns in the organization of brain activity. For example, these networks are typically modular – i.e., they can be partitioned into distinct clusters of strongly similar brain regions [8, 10] – and show an organization of a few highly connected core regions surrounded by more a sparsely connected periphery [11]. There are also similarities between functional brain networks and the anatomical white matter tracts that physically support communications between brain regions [12, 13].

Recent advances in network science have enabled the extension of these methods to dynamic networks, which can capture the important changes in brain activity and its organization over the course of a single fMRI experiment. A dynamic network is composed of a sequence of static networks, each representing the edge weights or similarities between brain regions in a single time window within the whole experiment. To include enough samples for a reliable estimate of the edge weights, each time window might span anywhere from a minute of time to entire scanning sessions of an hour or more. Many metrics can be applied to quantify the organization and time evolution of these dynamic networks – some simple extensions of existing methods for static networks, and others designed specifically for temporal dynamics. Properties of dynamic functional networks in the human brain have been shown to predict learning in a simple motor task [14]. However, many questions remain unanswered about the factors that inflence brain dynamics and the methods used to measure and describe them.

1.4 Our Approach: Dynamic Networks, Cognitive Tasks, and Age

In this dissertation, we use dynamic functional networks, derived from fMRI measurements of human brain activity, to investigate how brain dynamics relate to behavior and demographic factors. We apply two complementary methods of describing network dynamics: *dynamic communities*, or clusters of brain regions with strongly similar dynamics that can change their makeup over time [15]; and *hypergraphs*, which group regions whose edges all evolve together with correlated dynamic profiles [16]. For each method, we develop metrics that capture different aspects of brain dynamics, and quantify how they vary based on both short-term cognitive states and more lasting characteristics such as age and other demographics.

We use two complementary data sets for these investigations. First, we use a data set in which participants in early to middle adulthood complete a set of multiple cognitive tasks, calling for attention, memory, and unfocused thought. This data set is used in Chapter 2 and is referred to as the multi-task data set in Chapters 3 and 5. It is used to probe how brain dynamics depend on brain state, and how they change as the brain transitions between brain states. Second, we use a data set in which participants over a wide range of ages (18-75 years) perform the same memory-based task. This data set is used in Chapter 4 and is referred to as the age-memory data set in Chapter 3 and the single-task data set in Chapter 5. It is used to understand how brain dynamics depend on demographic traits, with a special focus on age.

1.4.1 Analyzing Task-Switching Dynamics with Hypergraphs

In Chapters 2 and 3, we use the graph-theoretic hypergraph to understand and describe brain task-switching dynamics. Hypergraphs group dynamic edges between brain region pairs with other edges that change together with similar dynamic profiles, over the course of a specific cognitive task or across several tasks. These groups are known as hyperedges. First, we ask how these groups of dynamically related edges and their corresponding brain regions are distributed, both within and across cognitive tasks. We find that several significant groups, or hyperedges, exist, and we find clear differences in the spatial distribution of these groups between tasks that reflect task-related areas.

Next, we statistically quantify how these hypergraph dynamics vary between individuals based on demographics and other measured traits. We find that in the multi-task data, age is a mild predictor of a metric describing the size of hyperedges, although hyperedge size is very consistent across cognitive tasks within a single individual. In the single-task age-memory data, this correspondence between age and hyperedge size is also seen, demonstrating significant changes in brain dynamics with age.

1.4.2 Finding and Interpreting Evolving Clusters of Brain Activity

Next, inspired by the previous idea of age being an important mediator of brain dynamics, we consider the effect of age on more straightforward brain activity clusters, or dynamic communities. We use a common algorithm to detect clusters based on fMRI data, by maximizing a quality metric known as *modularity*. We then apply and develop metrics for quantifying the changes in these communities over time. We find that age does affect the number of communities found, and the propensity of brain regions to flexibly switch between communities over the course of the task.

Finally, we find that dynamic networks often make the process of resolving functional brain communities very difficult, but we propose a strategy of removing strongly associated brain systems from the functional network, in order to resolve other systems more clearly. We devise a series of synthetic networks to systematically test this strategy, and show that these especially coherent brain systems (e.g. visual cortex) can mask dynamic organization in both these synthetic networks and in the real brain.

Chapter 2

Hypergraph Analysis of Human Functional Brain Dynamics

2.1 Introduction

An essential characteristic of the human brain is the ability to transition between functional states in synchrony with changing demand. A central focus in neuroscience involves quantifying this adaptability and understanding the underlying brain organization that supports it. Several studies have accomplished this with functional MRI techniques by delineating changes in regional blood-oxygen-level-dependent (BOLD) signal associated with different cognitive tasks, or between task states and task-free (resting [17, 18]) states [19, 20]. However, this approach, which examines the magnitude of brain activity alone, is unable to completely describe the complex correlation structure linking spatially segregated neural circuits. In particular, while providing crucial insight into the spatial structure and anatomical distribution of functional activity and how it differs between task and resting states, these methods are not well suited to probe the intrinsic organization of the dynamics of task-driven transitions between cognitive states, or co-evolving

associations among brain regions throughout a particular task.

Recent advances in network science provide tools to represent and characterize the functional interactions between brain regions forming cognitive systems. In this formalism, brain regions are represented as network nodes and functional connections (estimated by statistical similarities between BOLD signals [21]) are represented as network edges [7, 22]. These approaches enable the statistically principled examination of large-scale neural circuits underlying cognitive processes, and have enabled quantitative comparisons *between* circuits [23, 24]. Indeed, a growing literature provides evidence that individual tasks may elicit specific functional connectome configurations [25], while maintaining a relatively stable functional backbone reminscent of the connectome configuration evident in the resting state [26].

Nevertheless, these studies have focused on examining task or cognitive states as separate and independent entities, and tools to quantify how brain networks reconfigure between these task states remain significantly underdeveloped. Initial efforts to examine reconfiguration properties of brain networks have focused on quantifying properties of dynamic functional connectivity at rest [27]. A relatively few studies have begun to examine reconfiguration properties during task states [28–32] or across a series of brain states accompanying behavioral change [11, 14, 33, 34]. These studies have robustly demonstrated that functional connectome patterns change during task execution, and that individual differences in these reconfiguration properties have implications for task performance [11, 14, 28, 33].

In this chapter, we ask a complementary set of questions that focus on sets of functional connections rather than on the entire functional connectome pattern. We ask whether sets of functional connections evolve independently within or across brain states, or whether they evolve cohesively, each set controlled by a common regulatory driver. To answer this question, we employ recently developed dynamic network science methods

to estimate brain functional networks in one-minute time intervals as 86 participants engage in four task states: a task-free resting state, an attention-demanding state, and two memory-demanding states. We treat the evolving patterns of functional connectivity as temporal, or dynamic, networks [11, 14, 29, 30, 34, 35] and estimate the pairwise correlation between the strengths of functional interactions over time in order to identify groups of functional interactions which display similar changes in strength within and across task states. These groups of network edges with similar dynamic patterns, known as *hyperedges*, have been used to quantify the co-evolution in functional brain networks over the course of a learning task [16]. Our goal is to adapt this dynamic network science method to investigate the organization of evolving functional correlations both within and between task-specific cognitive states, using hyperedges as a measure of co-evolution. We hypothesize that overall, functional interactions between brain regions especially important for particular tasks are likely to be grouped in hyperedges with interactions between regions used strongly in other tasks, capturing co-evolution between task-specific functional networks as they turn off or on together when switching tasks. Furthermore, we expect that those functional correlations that link sets of brain regions whose coordination is crucial to a particular task will be more likely to co-evolve significantly during that task alone.

In this chapter, we demonstrate the existence of hyperedges driven by significant co-evolution within groups of functional interactions, both within and across task states. We develop novel network diagnostics to characterize hyperedges according to their structure, anatomy, and task-specificity. These analyses provide a unique window into the adaptability of the brain as it transitions between states and offer quantitative statistics for the comparison of such adaptability across subject cohorts.

2.2 Methods

2.2.1 Ethics Statement

Informed written consent was obtained from each subject prior to experimental sessions. All procedures were approved by the University of California, Santa Barbara Human Subjects Committee.

2.2.2 Tasks

Subjects engaged in a resting-state (task-free) period, as well as three separate tasks designed to engage different cognitive skills and task-specific brain networks: two separate functional runs of the same attention-demanding task, a memory task with lexical stimuli, and a memory task with face stimuli.

During the resting-state period, participants were asked to lie still with their eyes open and look at a blank screen. The attention task (Figure 2.1) required subjects to view sequences of visual stimuli on a screen, with the goal of detecting the presence or absence of a target stimulus in each of several test displays. Before each test display, subjects were presented with a cue arrow whose color and direction provided probabilistic information on whether and where the target stimulus might appear. The test display was then flashed for approximately 50ms, after which the subjects were required to choose whether or not the target stimulus had appeared. In both memory tasks (Figure 2.1), 180 previously studied stimuli and 180 novel stimuli were presented to the subjects, who were asked to determine whether each stimulus was "old" or "new" – i.e., whether it had been previously studied. As in the attention task, the memory tasks included probabilistic cues: each stimulus was shown either in a particular color (lexical stimuli) or bordered by a color (face stimuli) which provided subjects with the probability that the stimulus

was novel. Face stimuli were drawn from a variety of online faces databases [36–41]. For additional experimental details, see [12], [42], and supplemental information therein.

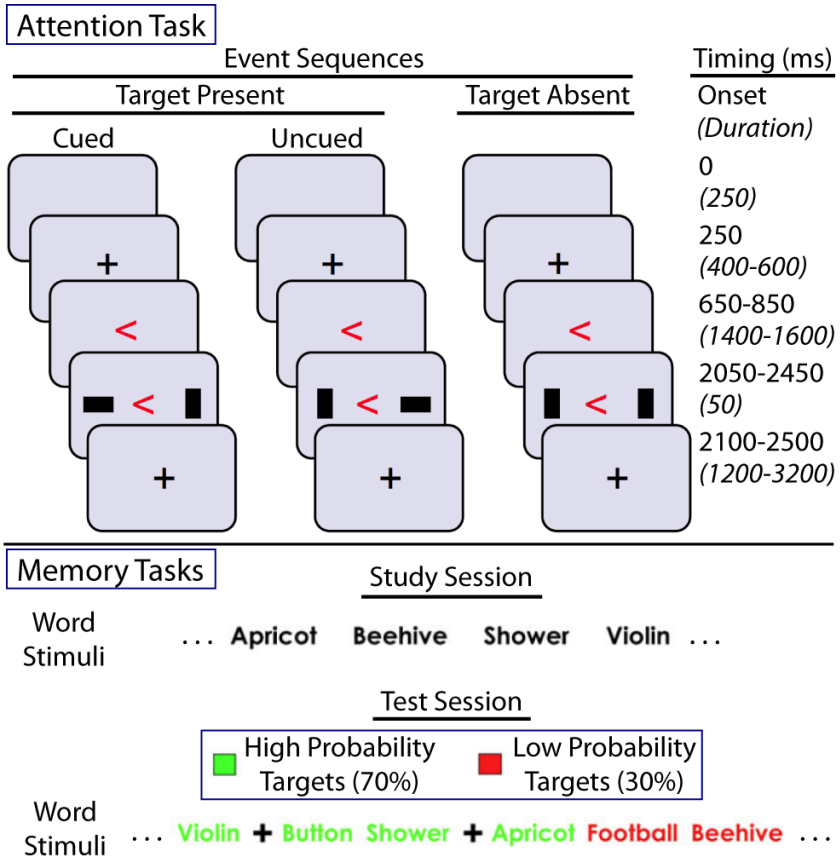

Figure 2.1: **Task Setup: Top panel:** Setup of a single trial sequence in the attention-demanding task. Here, the target stimulus is a horizontal rectangle on either side of the center cross. In each trial sequence, the cross is presented, followed by a cue (arrow) giving probabilistic information about whether and where the target stimulus wil appear, and finally by the stimuli, displayed for approximately 50ms. The target will either appear as cued, appear in the uncued location, or not appear at all; subjects are required to choose which of these possibilities occurred. **Bottom panel:** Setup of the memory-demanding tasks (same format for word and face memory). In the study session, subjects are presented with a sequence of stimuli. During the test session, another sequence of stimuli is presented; subjects are required to distinguish whether each test stimulus is novel or identical to a stimulus from the study session. Colors of lexical stimuli and colored borders of face stimuli (not pictured) indicate the probability that the test stimulus has been seen before.

2.2.3 Imaging

MRI data was acquired at the UCSB Brain Imaging Center from 116 healthy adult participants using a phased array 3T Siemens TIM Trio with a 12 channel head coil. Functional MRI data was taken while each participant engaged in the four tasks described above. This analysis combines two separate functional runs of the same attention task [12]. The sampling period (TR) was 2s for the rest and attention tasks and 2.5s for both memory tasks. In addition to functional data, a three dimensional high-resolution T1-weighted structural image of the whole brain was obtained for each participant.

2.2.4 Image Analysis

Structural MRI acquisition and pre-processing

Structural scans were intensity-corrected, skull-stripped, normalized, segmented and parcellated (as described below) using Freesurfer v.5.0.0 cortical reconstruction all with default settings, accessed via the Connectome Mapping Toolkit v.1.2.0 [43]. The starting atlas was the updated Lausanne2008 multi-scale atlas [44]. For each subject, parcellations containing 83, 129, 234, 463 and 1015 regions were generated, covering cortical grey-matter regions, the thalamus, caudate, putamen, pallidum, accumbens area, hippocampus, amygdala and brainstem. The highest-resolution parcellation of 1015 regions was not investigated further, since a large number of regions contained very few or no voxels when the atlas was downsampled into fMRI space.

Functional MRI pre-processing and time series analysis

Preprocessing was performed using FSL v5.0 [45–47], AFNI v. 2011 12 21 1014 http://afni. nimh.nih.gov [48] and Matlab (2013, The Mathworks, Natick, MA). Functional MRI scans were preprocessed as follows. FSL programs MCFLIRT [49] and fsl

motion outliers were used to correct for head motion and derive a volume-by-volume measure of head motion: framewise displacement. Framewise displacement (FD) is calculated as the sum (in mm) of rotational and translational displacements from volume N to N+1 [50]. Next, we performed slice timing correction (AFNI 3dTshift), auto-masked to obtain a brain-only fMRI image (AFNI 3dAutomask), and smoothed the time series at each voxel (AFNI 3dDespike with default parameter settings). Despiking has been shown to reduce the motion-related distance dependent bias in correlation estimates [51]. Each voxel's time series was then detrended with respect to framewise displacement using AFNI 3dDetrend. This uses linear regression to remove variability related to the nuisance regressor, framewise displacement, at each voxel. Runs were only included in the analysis if mean framewise displacement for the run was less than 0.25mm per frame; this led to 73 fMRI runs (of 763 total runs) being excluded from this analysis. Registration proceeded as follows: a participant's time-averaged fMRI image was aligned to their structural T1 scan using FSL FLIRT boundary-based registration [49, 52], and the inverse of this transformation was applied to all subjects parcellation scales (generated in structural space). Parcellations were downsampled into EPI (AFNI 3dfractionize, voxel centroid voting, requiring 60% overlap), and the mean signal across all the voxels within a given brain region was calculated to produce a single representative time series. The data was not spatially smoothed at any stage.

Creation of a hybrid atlas

We sought to create an atlas with low inter-individual and cross-brain variability in the amount of fMRI data acquired per region. Many existing atlases use parcellations that have roughly equal region sizes as measured on structural MRI scans [53]. However, downsampling the atlas from structural MRI voxels to fMRI voxels, along with inhomogeneous fMRI signal-loss, mean that this does not produce equally sized regions

in functional MRI space. To mitigate this, we generated a 'hybrid' atlas by choosing those regions from various scales of the Lausanne2008 atlas that minimized cross-brain and intra-subject variability in region size. The intra-subject size variability was quantified by the coefficient of variation, defined for each region i as $100\sigma_i/\mu_i$, where μ_i is the mean size of region i over all subjects and σ_i is the standard deviation. Starting with the scale 234 atlas, an iterative process was used to decrease intra- and intersubject variability in region size. Where a region had very few voxels (mean size $<$ 25th percentile), or high variability in size across subjects (coefficient of variation $>$ 30%), it was tentatively exchanged for a region from the next highest resolution atlas, effectively combining the initial region with other higher-resolution regions subsumed under the same anatomical heading. If this combination of regions decreased the inter-subject or within-subject variability in region size, the combined region was retained. If not, the initial poor quality region was rejected from the "hybrid atlas". This was repeated until no further combinations of regions could decrease intra- and inter-subject variability while retaining neuroanatomically sensible groupings. Regions were excluded from the analysis altogether if there were fMRI runs in which no data was acquired in that region (frontal pole, entorhinal cortex and temporal pole), or if the inter-subject coefficient of variation was greater than 30% (this applied to 7 of the 8 inferior temporal regions; 1 of the 8 middle temporal regions; 2 of 8 fusiform regions; 1 of the 6 caudal middle frontal regions, and 1 of the 14 precentral regions). Table 2.1 lists the 194 regions identified by this hybrid atlas. This approach considerably reduced intra-subject variability in region size as well as reducing the inter-subject variability at problematic outlier regions, while minimizing the amount of data that had to be excluded from analysis.

2.2.5 Functional Connectivity

Specific frequencies of oscillations in the BOLD signal have been associated with different cognitive functions. We focus our investigation on low frequency (0.06-0.125 Hz) oscillations in the BOLD signal that have proven useful for examining resting [54,55] and task-based functional connectivity [14]. The task-related oscillations are posited to be specific to this frequency range, possibly due to a bandpass-filter-like effect from the hemodynamic response function [56]. We apply a Butterworth bandpass filter to isolate frequencies in the (0.06-0.125 Hz) range [57].

To construct a functional brain network, we use the 194 region hybrid atlas, where each region contains a roughly equal number of voxels. These 194 regions represent the network nodes. The x, y, and z positions of each node are given by the centroid of the voxels which comprise the node. Edge weights in the functional brain network are computed by taking Pearson's correlations between the filtered time series within a defined time period for each pair of nodes [58].

2.2.6 Time Windows for Temporal Network Construction

Dynamic networks are constructed by taking the filtered time series in temporal windows of 60 seconds and computing a $N \times N$ adjacency matrix of nodal correlations for each time window, where $N = 194$ is the number of nodes. Each of these $N \times N$ adjacency matrices represents the functional network over the 60 seconds in question. From this set of networks, we extract the edge weight time series by considering the correlation strength in each sequential network. We let $E = N(N-1)/2 = 18721$ be the total number of edges between the 194 nodes and construct an $E \times E$ adjacency matrix $\mathbf{X}$, where X_{ab} gives the Pearson correlation coefficient between the time series of edge weight for edges a and b. The entries of the $E \times E$ adjacency matrix represent pairs of

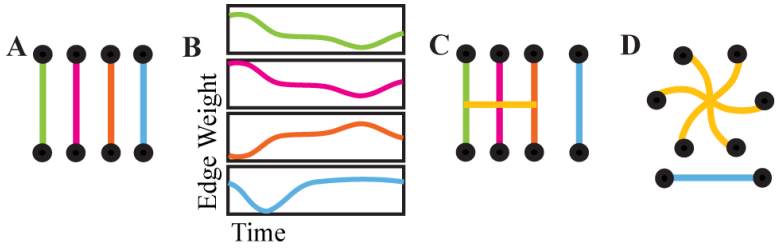

Figure 2.2: **Hyperedge Construction:** A schematic illustration of the method used to identify hyperedges. We begin with a set of node-node edges (A) and their time series (B), of which three [green, pink and orange traces, (B)] exhibit strong pairwise temporal correlations. These edges are cross-linked (C) by temporal covariance in edge weight time series, and thereby form a hyperedge (D) of size three on six nodes. The final [blue] edge forms a singleton, an edge which is not significantly correlated with any other edges.

edges with correlated weight time series [16].

We consider a range of temporal window lengths from 40 to 120 seconds and find that our results for hyperedge size and spatial distributions are robust to changes in window length in this range. Because the TR varies between the memory tasks and the rest and attention tasks, windows of equal time length include different numbers of data points in different segments of the experiment. To ensure this does not affect our analysis, we conduct an analysis with the number of data points per window held constant, and obtain very similar results (see Figure 1 in Appendix A).

2.2.7 Hyperedge Construction

The cross-linked network structure, which contains information about groups of edges with similar time series (hyperedges), is extracted from the edge-edge correlation matrix $\mathbf{X}$ [16]. Figure 2.2 provides a schematic illustration of the process of determining the cross-linked structure of a network. To exclude entries of $\mathbf{X}$ that are not statistically significant, we threshold $\mathbf{X}$ by evaluating the p-values for the Pearson coefficient R for each edge-edge correlation using a false discovery rate correction for false positives due

to multiple comparisons [59]. If the p-value for an entry X_{ij} satisfies the false discovery rate correction threshold, we set $\xi_{ij} = R(i,j)$ for our thresholded matrix ξ. We set the thresholded entry of all other elements X_{ij} to zero. We binarize this thresholded matrix and obtain ξ'_{ij}, where

$$\xi'_{ij} = \begin{cases} 1, & \text{if } \xi_{ij} > 0; \\ 0, & \text{if } \xi_{ij} = 0. \end{cases} \tag{2.1}$$

Each connected component in ξ represents a hyperedge, a set of edges that have significantly correlated temporal profiles. The groups of nodes in Figure 2.2(D) are examples of such connected components. A single hyperedge may include any number of edges between one (a singleton) and $E = N(N-1)/2$ (the system size); these edges may be spatially clustered or at disparate locations throughout the brain. The set of all hyperedges defined in ξ produces an individual hypergraph.

This hypergraph technique builds on recent trends in the wider field of network science. First, identifying groups of network edges that share similar properties, rather than the groups of nodes that have traditionally been the focus of community detection methods, has been recently shown to provide more intuitive representations of overlapping nodal communities and hierarchical structure [60–62]. Second, the idea of identifying functional groups based on the temporal patterns of their interactions has proven useful [62, 63]. Hypergraphs provide a straightforward method, both edge-based and intrinsically dynamic, of identifying and analyzing temporal patterns in network organization. In this work we focus on functional networks in the human brain, but the hypergraph-related diagnostics introduced below are easily generalizable to a broad variety of dynamic networked systems.

2.2.8 Hypergraph Diagnostics

We use several methods to extract statistical features from individual hypergraphs and across the set of subjects.

Hyperedge size: We define the size, $s(h)$, of a hyperedge h, as the number of edges contained in it,

$$s(h) = \sum_{i,j \in h} \xi'_{i,j}, \tag{2.2}$$

where the sum is performed over the upper triangular elements of ξ', and ξ' is the binarized edge-edge adjacency matrix defined above. Hyperedges with $s(h) = 1$ are singletons, which display no significant correlation between that edge and any other in the network. These singletons are excluded from further analyses. Additionally, we compute the cumulative hyperedge size distribution across all subjects in the study.

Hyperedge node degree: We define the hyperedge degree of a node to be the number of hyperedges that contain that node. We examine the hyperedge node degree distribution as a spatial distribution over the subjects as a group to understand characteristic hyperedge properties.

Co-evolution network: We construct a "co-evolution network" to consolidate hypergraph results into a single graph that illustrates where hyperedges are most likely to be physically located over an ensemble of individuals. Figure 2.3 illustrates a schematic of our construction. We begin by defining the matrix, $\mathbf{C}$, of probabilities that edges are included in a hyperedge over a set of hypergraphs. Again, nodes correspond to brain regions and connections correspond to inter-region associations, but here the weight of a connection joining nodes i and j is the matrix entry $\mathbf{C}_{i,j}$. The resulting static network encompasses the dynamics of hyperedge activity, with connection weight corresponding to the probability that the two nodes are co-evolving over all of the hypergraphs consid-

ered. In later sections, we refer to co-evolution connection "strength," which we define as the magnitude of the probability matrix entry corresponding to that connection.

2.2.9 Task-Specific Classification

Previous work identified regions with task-specific activity in rest, attention, and memory tasks [12]. Further understanding of the regions that have a correlation structure unique to one task provides insight into network structure differences between tasks. To investigate the task-specific hyperedge structure, we first group hyperedges that exhibit a significantly higher correlation within one task into task-specific sets. If a hyperedge is significantly correlated in two or more tasks, it is excluded from the task-specific hypergraphs. The task-specificity of hyperedges is calculated by comparing the correlation within a single task to the correlation over the same time length with time points chosen randomly from other tasks. This permutation test uses a Bonferroni correction for false positives due to multiple comparisons [64]. Task-specific hypergraphs are then used to construct task-specific hyperedge size distributions, hyperedge node degree distributions, and co-evolution networks.

To quantitatively probe the differences in spatial organization of dynamic functional co-evolution networks for the four tasks, we investigate two summary metrics that show significant variation across tasks. Choice of these measures is primarily motivated by observed coarse differences in co-evolution network structure.

The first "length-strength" metric is the Pearson correlation coefficient, R, between the strength of a connection in the co-evolution network and Cartesian distance between the two nodes linked by the connection (physical length). The Cartesian distance is computed by taking the x, y, and z coordinates of each node and calculating the square root of the differences squared. The length-strength metric identifies a geometric property

of the network, as well as a coarse estimate of the length of the strongest connections. Furthermore, connection length is related to network efficiency [65, 66], so differences in this measure could indicate varying levels of functional network efficiency corresponding to task states.

The second "position-strength" metric is the Pearson correlation coefficient, R, between the strength of the co-evolution network connection with the average anterior-posterior position of the two nodes. A measure of anterior-posterior position for each connection was found by taking the average y position of the two nodes in the connection. Identifying the location of strong co-evolution network connections along the anterior-posterior y axis provides a measure of where hyperedges are physically present in task states. Both the structural core [44] and a dynamic functional core area, comprised of sensorimotor and visual processing areas [11], are located in the posterior, so nodes in these regions have negative y values. A larger negative position-strength value corresponds to a higher probability that hyperedges are active in these core areas.

The length-strength and position-strength metrics are evaluated for significance by comparing the correlation between length or position and connection strength to the same correlation performed on randomly chosen co-evolution connections. Again, the Bonferroni correction is performed to eliminate false positives due to multiple comparisons.

In Section 2.3, we discuss how these metrics reveal quantitative differences between task-specific networks. A more detailed analysis of the overlap between hyperedge co-evolution networks and relevant cognitive processing regions is also presented. In this analysis, we describe how delineated areas of higher hyperedge activity consistently correspond to recognized centers of task-specific activity.

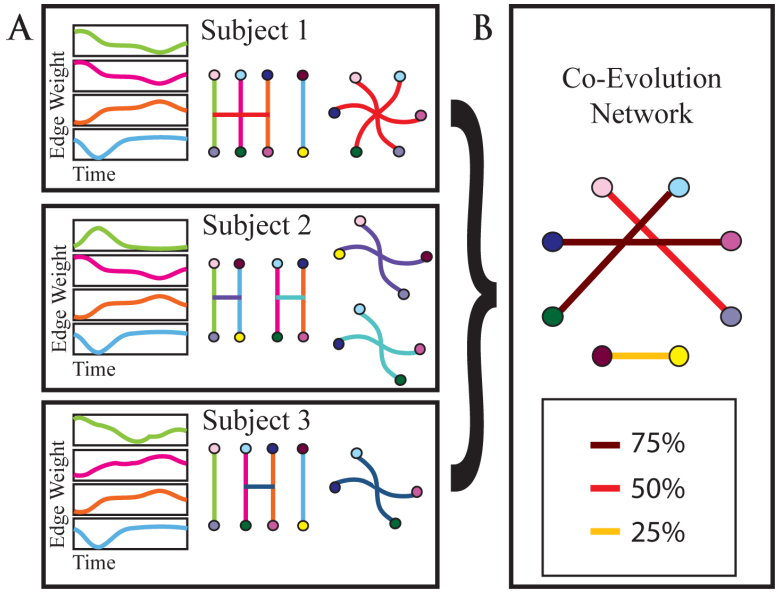

Figure 2.3: **Schematic Construction of the Hyperedge Co-evolution Network:** In (A), we analyze edge time series and group edges exhibiting similar temporal profiles into a hyperedge (as in Figure 2.2). Here, node colors are used to indicate individual nodes and the edge color indicates distinct edges. We construct hypergraphs for each subject and find the matrix $\mathbf{C}$ of probabilities that two nodes are in the same hyperedge over all subjects and hyperedges. In (B), this matrix is used to create a co-evolution network, where the weight for an edge connecting nodes i and j corresponds to the entry $\mathbf{C}_{i,j}$.

2.2.10 Null Models

In this analysis, we compare our results with two statistical null models based on measures for dynamic networks [35]. Hyperedges are formed from correlated edge time series; consequentially the null overall model randomly shuffles each edge time series over all experiments. This null model is designed to ensure that the hyperedges identified in our analysis can be attributed to the dynamics of the system, rather than some overall statistical property of the data set.

The other null test we perform, which we will refer to as the null within-task model, reorders each edge time series within each task, keeping tasks distinct. This is constructed in order to determine whether there are specific differences in the data between tasks.

2.3 Results

We compile the results from the hypergraph analysis for each of the subjects and combine these results to obtain a size distribution, anatomical node degree distribution, and co-evolution network for the group. We then divide the data into task-specific hypergraphs and perform the previously mentioned analyses on the task-specific hypergraphs.

2.3.1 Hypergraph Analysis and Statistics

We construct a hypergraph for each individual and examine the cumulative distribution of hyperedge sizes ($s(h)$ from Equation 2.2), shown in Figure 2.4. There is a distinct break in the slope between two branches of the distribution occurring at a size of approximately 100 edges, which we use to distinguish between "large" and "small" hyperedges. The total number of small hyperedges appears to roughly follow a power law with an exponent of approximately -2.5. The number of large hyperedges peaks around the maximum size, with relatively few in the middle range from 100 to 1000 edges. In Figure 2.4, the sharp drop off in the distribution at large hyperedge sizes reflects the system size limitation on hyperedge cardinality.

There is a distinct partition in all individual frequency versus sizes distributions; one or two "large" hyperedges ($s(h) > 100$), and many "small" hyperedges ($s(h) < 100$) that peak at the smallest size. A subject with relatively small maximum hyperedge size has hundreds of edges in this largest hyperedge, as well as multiple "small" hyperedges. The corresponding hypergraph of a subject with a maximum hyperedge near the system size is strongly dominated by the largest hyperedge, which contains almost all edges in the brain.

The null overall model shuffles the data over all tasks. There are no hyperedges greater than size one, so the results from this null model are not depicted in Figure

2.4. These singletons signify no significant correlation with other edges. As a result, we performed no further analysis on this null model. The fact that no significant hyperedges were found in the null overall model validates the statistical significance of our results.

The null within-task model shuffles the data but ensures that task data stays within the same task. The size distribution of hyperedges from the null within-task model is shown in Figure 2.4. The shape of the two distributions is similar, although the null within-task model has fewer hyperedges in the large regime and there are more singletons than in the original data. This indicates there is co-evolution structure across tasks because this structure corresponds to changes in edge states between two or more tasks. For example, if groups of edges have an overall high correlation in one task and a significantly lower correlation in another, it would induce a hyperedge across the tasks regardless of how the within-task time series are shuffled.

Examining the cumulative hyperedge size distribution provides information about the network topology but does not supply descriptive spatial information. Next, we quantify which anatomical locations in the brain participate in hyperedges, identifying differential roles in task-induced co-evolution. Figure 2.5A depicts the hyperedge node degree on a natural log scale. The densest regions are located in posterior portions of the cortex, primarily in visual areas, while a second set of dense regions is located in the prefrontal cortex.

We construct a co-evolution network, as illustrated schematically in Figure 2.3, where connection weight corresponds to the probability that two nodes participate in the same hyperedge. In Figure 2.5B we present this co-evolution network over all individuals and all tasks. The graph includes sparse long-range connections between regions that are densely connected. Within the strongest 1% of connections, the high degree of bilateral symmetry indicates that corresponding nodes in the left and right hemispheres have a high likelihood of being placed together in a hyperedge. Dense areas of the graph include

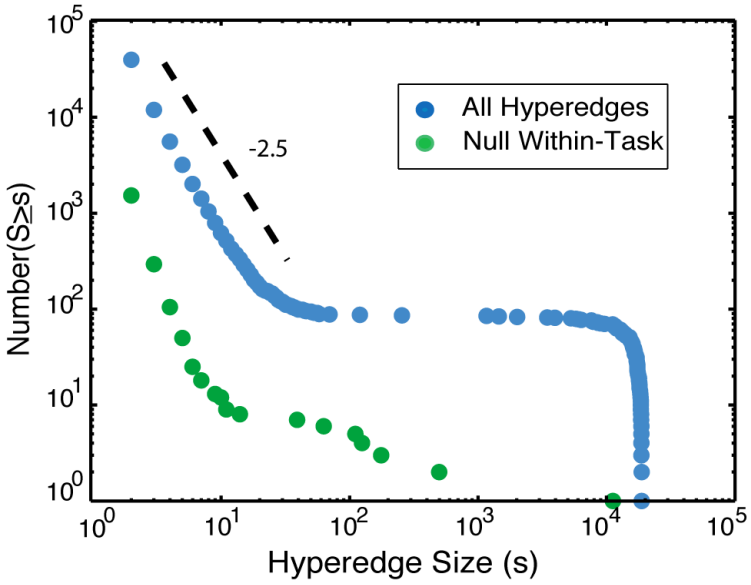

Figure 2.4: **Hyperedge Size Distribution:** In the cumulative frequency distribution of hyperedge sizes, the small hyperedges appear to roughly follow a power law with an exponent of approximately -2.5, while the large group is concentrated near the maximum size. In the null overall model, there are no non-singleton hyperedges. Results for the null within-task model, where the data is shuffled within each task, are in green.

primary visual areas, portions of prefrontal cortex, and primary motor cortex.

2.3.2 Task-Specific Hyperedges

The hypergraph algorithm groups together edges with significantly similar temporal behavior. However, this basic classification does not distinguish whether the correlation is present throughout the edge time series, or whether highly correlated sections of the time series drive the selection. We compute the average within-task edge correlation for each hyperedge and find that in some cases, strong edge correlation spans the tasks, while in other hyperedges, a strong correlation between edges within one task drives the hyperedge. An example of this task-specific correlation structure can be seen in

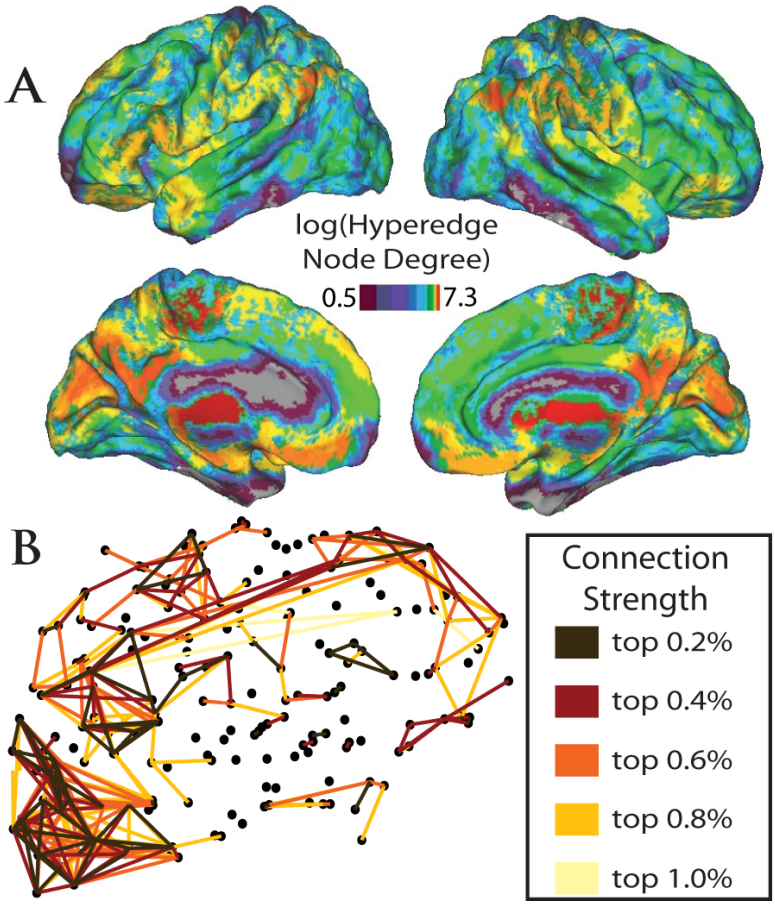

Figure 2.5: **Hyperedge Node Degree and Co-evolution Network:** In (A), we show hyperedge node degree on a natural log scale. The cumulative number of hyperedges at each node over all individuals is plotted on the brain, where higher values at a node correspond to more hyperedges that include the node. (B) depicts a sagittal view of the co-evolution network. The edge strength represents the probability that the edge will be in a hyperedge over all individuals. Edge color corresponds to threshold percentage value, where only the top 1% of co-evolution probabilities are shown. Within this 1%, brown connections correspond to the highest 0.2% of probabilities, red connections correspond to 0.2% to 0.4%, orange connections correspond to 0.4% to 0.6%, gold connections correspond to 0.6% to 0.8%, and yellow connections correspond to 0.8% to 1%.

Figure 2.6. In the average within-task correlation on the left, there is a stronger average correlation in the word memory task than in any other task. Furthermore, the edge time series in the first hyperedge indicates it is driven mainly by a correlation within the word memory task.

To investigate this further, we construct task-specific co-evolution networks, composed of hyperedges with significantly stronger average correlation in one task than the others (see Methods). To identify these task-specific hyperedges for each task, we perform a permutation test on the edge weight time series, as described in Methods, and compare the total correlation within the task to the expected values. If a hyperedge displays significant edge correlation (determined by the Bonferroni correction on the p-values from the permutation test) in only one task, we label it as a task-specific hyperedge. Hyperedges with two or more tasks exhibiting significant correlation are not included in the task-specific hypergraphs.

Figure 2.7 illustrates the size distributions of all the task-specific results alongside the overall hyperedge size distribution. The sizes and spatial distributions of single task-driven hyperedges vary across tasks and incorporate significant information about functional network organization with respect to changing cognitive states. Attention has the greatest number of task-specific hyperedges, followed by face memory, word memory, and rest. In the small regime, the tasks follow a similar distribution. There are fewer large attention and rest hyperedges, while the face memory task closely mimics the overall distribution. The distinction in the distributions indicates that the tasks can be characterized by differing complexities of edge co-variations.

The spatial distributions of hyperedge node degree in each task, along with task-specific co-evolution networks, are shown in Figure 2.8. The rest hypergraph has the least activity in posterior regions of the cortex, both in the hyperedge node degree plot and co-evolution network. In the attention network, long connections connecting the

front and back of the brain distinguish it from the rest network. Furthermore, the concentration in the occipital lobe is larger in the memory co-evolution networks than in the rest or attention networks. We characterize these observed differences with two statistics, which are described in more detail in Methods. The length-strength metric is a correlation between connection length and strength in the co-evolution network. The position-strength metric is a correlation between connection position (anterior-posterior) and strength. The results of this analysis over the full unthresholded co-evolution network are in Figure 2.9. All correlation values are negative, indicating that, in all tasks, stronger connections in the co-evolution network are located in posterior portions of cortex and are physically shorter.

We compare these values across tasks by performing pairwise permutation tests to determine which networks have statistically different properties. Figure 2.9 depicts the p-values from these tests, where the horizontal axis represents the statistic being tested and the vertical axis corresponds to the task being tested against. The black squares in this figure represent significant values, which are summarized in the following list:

1. The rest task has a significantly less strong position-strength correlation than the word and face memory tasks. This confirms the observation that the rest co-evolution network is less likely than the memory networks to have strong connections in posterior regions of the cortex.

2. The attention task is less strongly correlated than the word memory task, as measured by the position-strength metric and the rest task in terms of the length-strength metric. Thus, the attention co-evolution network is less likely than word memory to have strong connections in the posterior, and less likely than the rest network to have strong connections that are short.

3. The word memory task has a weaker length-strength correlation than the rest and

attention tasks. Thus, strong connections in the word memory co-evolution network are less likely be short than they are in attention and rest networks.

These results delineate significant differences in co-evolution network structure between the tasks, confirming that the hypergraph analysis is a useful method for distinguishing between task states. Additional features of the task-specific co-evolution networks are described in more detail below.

Rest

Rest-specific hyperedges are primarily represented in the "small" range of the size distribution in Figure 2.7. Although it is difficult to distinguish in Figure 2.7 due to the logarithmic scale, the rest task also has the lowest number of task-specific hyperedges. Consequently, its spatial hyperedge node degree distribution in Figure 2.8A has the lowest overall magnitude across task states. The areas with the highest degree of hyperedge activity are in the posterior portions of the brain, a configuration that is consistent across tasks. This suggests there is an underlying pattern of hyperedge generation centered in the occipital lobe.

The rest-specific co-evolution network is highly clustered in the most probable 0.2% of co-evolution pairs, as visualized in Figure 2.8B. High probability clusters occur in areas including the inferior parietal lobule, superior frontal gyrus, precuneus, and posterior cingulate cortex. Although the rest network displays clustering at this highest threshold of probability, lower thresholds show very little structure; the top 1% of connections shown in Figure 2.8B is far more randomized in rest than in the other task-specific co-evolution networks. There is relatively little lateral symmetry and few visible "core" areas with high hyperedge node degree.

The negative length-strength correlation for connections in the co-evolution network

is significantly stronger for the rest task than the word memory task. This indicates that the strongest connections in the rest-specific co-evolution network are short, reflecting the initial observations in Figure 2.8B. The rest co-evolution network also has the smallest negative correlation between connection position and strength, which the permutation test (Figure 2.9B) confirmed to be significantly smaller than the word or face memory tasks. This means that the strongest rest-specific hyperedges are less likely to be located in the posterior of the brain than the strongest hyperedges specific to either memory task, a result again consistent with Figure 2.8B.

Attention

Overall, there are more hyperedges associated with attention than any other task, although this is difficult to visualize in Figure 2.5. The attention-specific hypergraph consists almost exclusively of small hyperedges. This lack of large hyperedges may account for the increased disorganization in the co-evolution structure at lower probability thresholds observed in both rest and attention co-evolution networks in Figure 2.8B.

The spatial hyperedge node degree distribution for the attention task (Figure 2.8A) appears qualitatively similar to the rest task, with a few areas of increased degree in the occipital lobe, and with overall larger hyperedge node degree values corresponding to the greater overall number of attention-specific hyperedges compared to rest.

The co-evolution structure specific to the attention task (Figure 2.8B) has a higher degree of bilateral symmetry than the rest network, and has fewer strong connections in the occipital lobe than either memory task. There are multiple prefrontal cortical regions that are likely to cohesively evolve with several other nodes. Regions of high clustering in the most probable threshold include the lateral parietal and occipital lobes, the superior frontal cortex, and dorsal parietal cortex.

Numerous strong connections between rostral and caudal brain regions are another

feature of the attention-specific co-evolution network. The negative length-strength correlation in the attention co-evolution network is significantly less strong than in the rest task, consistent with the observation that the attention network has strong connections that reach across the brain (Figure 2.8B). Additionally, the attention task has a significantly weaker position-strength correlation than the word memory task, likely driven by the strong attention co-evolution connections in the prefrontal cortex.

Memory for Words

The word memory-specific hyperedge size distribution includes more large hyperedges than rest or attention, although it is not as close to the overall distribution as the face memory distribution.

The spatial hyperedge node degree distribution for the word memory task has high node degrees in similar brain areas to the other task-specific distributions. There is a marked increase in node degree of regions in the parietal lobe from rest and a decrease in degree of regions in the occipital lobe from attention (seen in Figure 2.8A).

In the word memory co-evolution network in Figure 2.8B, the strongest connections are highly clustered in the occipital or frontal lobes, with few connections to nodes in between, while the strength and number of bilateral links is diminished compared to the attention task. The negative length-strength correlation of connections in this network is the weakest for the word memory co-evolution network, and significantly weaker than in the rest or face memory tasks. As in the attention task, this is consistent with the many connections between the occipital and frontal lobes visible among the strongest links in the word memory co-evolution network (Figure 2.8B).

Memory for Faces

There are more large hyperedges significantly correlated in the face memory task than any other task-specific group. The task-specific size distribution closely resembles the overall distribution in the large regime, indicating that a significant portion of all large hyperedges are driven by correlations in the face memory task.

The face memory-specific hyperedge node degree values are consistently the largest across the brain. This is primarily due to the many large hyperedges specific to the face memory recognition task. In the word memory and attention degree distributions, there are areas of higher hyperedge node degree in the parietal lobe and occipital lobe, respectively, but the face memory degree distribution is more evenly dispersed over the brain.

The structure of the face memory-specific co-evolution network, shown in Figure 2.8B, is most dense in the occipital lobe, consistent with the visual nature of the task. There are several strong connections from the occipital lobe to other brain regions, specifically in the prefrontal cortex and frontal lobe. While the structure looks similar to that of the word memory co-evolution network, the strong cluster of face memory co-evolution connections in the occipital lobe has fewer strong connections and less nodes involved overall than the corresponding word memory co-evolution network cluster, but more strong connections to a few particular nodes. Compared to word memory, the face memory-specific network also displays fewer strong connections in the frontal lobe but more strong connections among regions in the dorsal attention network. In addition to the properties discussed in previous sections, the face memory co-evolution network has a strong negative position-strength correlation, indicating that the strongest connections tend to be in the posterior of the brain.

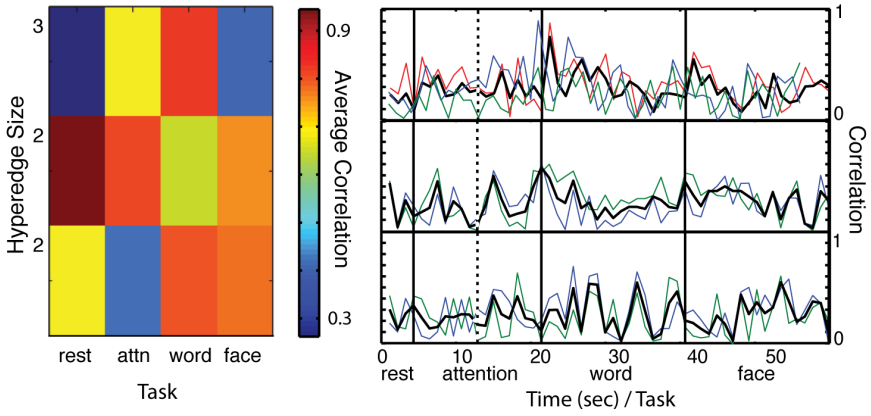

Figure 2.6: **Task-specific Hyperedges: Left:** Average hyperedge correlation in each task for three hyperedges (where hyperedges with small sizes are chosen for illustrative purposes). **Right:** Correlation (absolute value) time series for the same three hyperedges. The colored lines represent each edge, while the black line is the average edge time series. Each time point represents the static network over 60 seconds, and the attention task is broken into two sections because two separate iterations of the same task were combined in this analysis. These results display the task-specificity of hyperedges, where significant correlations in the hyperedge are restricted to one task. For example, the first hyperedge is word-specific because there is a much stronger average correlation in the word task than in any other task.

2.4 Discussion

Progress in understanding functional brain network topology provides significant insight into broad neuroscience questions regarding the brain's organization and ability to effectively transition between cognitive states. Quantifying complex network dynamics in the brain will further understanding in these areas and has promising applications to behavioral adaptation and learning [11, 14, 34]. We apply hypergraph analysis, a tool from dynamic network science, to functional brain imaging data in order to determine co-evolution properties of the brain as subjects perform a series of tasks. A previous application of this method to neuroscience uses hypergraphs to analyze how functional network structure changes over a long term learning task [11]. The learning experiment considers hypergraphs constructed over 6 weeks of training while subjects acquire a new

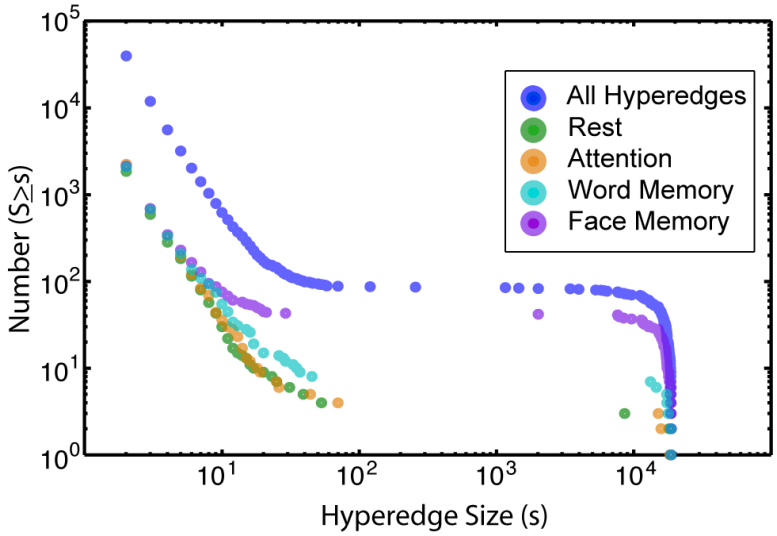

Figure 2.7: **Task-specific Hyperedge Size Distributions:** Cumulative frequency distribution as a function of hyperedge size for all task-specific groups. The results are compared to the overall distribution of hyperedges (dark blue), previously illustrated in Figure 2.4. There are fewer large hyperedges attributed to attention and rest tasks, while the memory tasks have a greater number of large task-specific hyperedges.

motor skill, while our analysis compares hypergraphs over three different tasks performed within an interval of hours. Our analysis shows that hypergraphs are a useful tool for investigating shorter time scales and differentiating between task-specific networks.

Instead of analyzing the time-dependent behavior of groups of nodes, the hypergraph investigation considers the edge weight time series, where edges with statistically significant similarities in their temporal profiles are grouped into hyperedges. This approach is advantageous because it considers all edges, regardless of correlation strength, unlike previous methods which focus exclusively on strong correlations [12, 13]. The use of a data-driven analysis also allows us to investigate the dynamic changes in brain function over a series of tasks without prior assumptions of the structure of the connectivity network. This is a significant advantage over methods that characterize task states based on their differences with respect to the rest network [19, 20]. A comparison between the

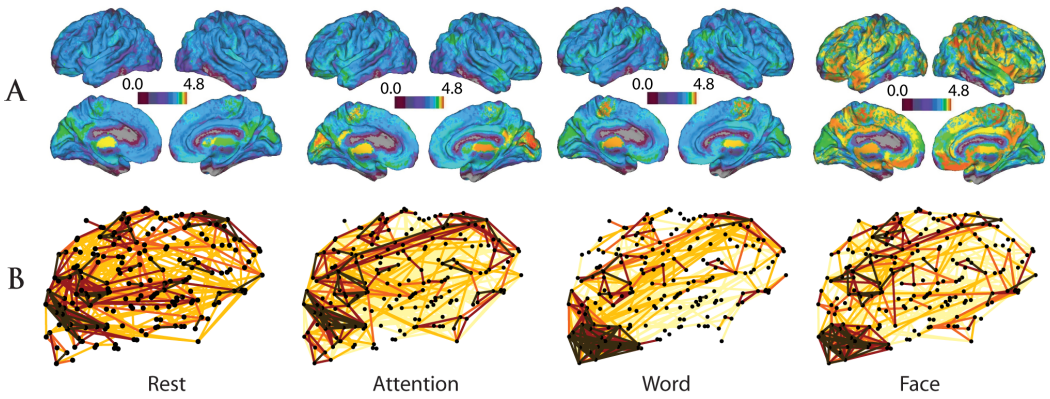

Figure 2.8: **Task-specific Co-evolution Networks and Hyperedge Node Degrees:** (A): Distribution of task-specific hyperedge node degree on the brain. Here, the log of the total number of hyperedges containing each node is represented on the brain. The color scale represents the log of hyperedge node degree as in 2.5A, although here the range of values is from 0 to 4.8. (B): Co-evolution networks for each task. Edge strength corresponds to the probability that a hyperedge will contain the edge over all individual hypergraphs. Color represents a threshold in percentage value, with the scale given in Figure 2.5B, and the top 1% of co-evolution probabilities are shown. Once again, the top 2 % of probabilities are brown, red indicates the top 0.2% to 0.4% of connections, orange indicates the top 0.4% to 0.6% of probabilities, gold indicates the top 0.6% to 0.8% of probabilities, and yellow indicates the top 0.8% to 1% of probabilities.

hypergraph analysis and these methods in a future analysis could reveal how the concentration of hyperedges varies in known task-positive or task-negative areas and determine whether this variation has an effect on task performance.

2.4.1 Hypergraph Statistics and Structural Metrics

We demonstrate the existence of hypergraph structure in functional brain dynamics and statistically characterize the hyperedge distributions in comparison to appropriate null models. Shuffling the time series over all time produces no significant hyperedges, while shuffling within each task results in a size distribution that resembles the overall size statistics in shape, but with far fewer hyperedges. The distinct differences between the

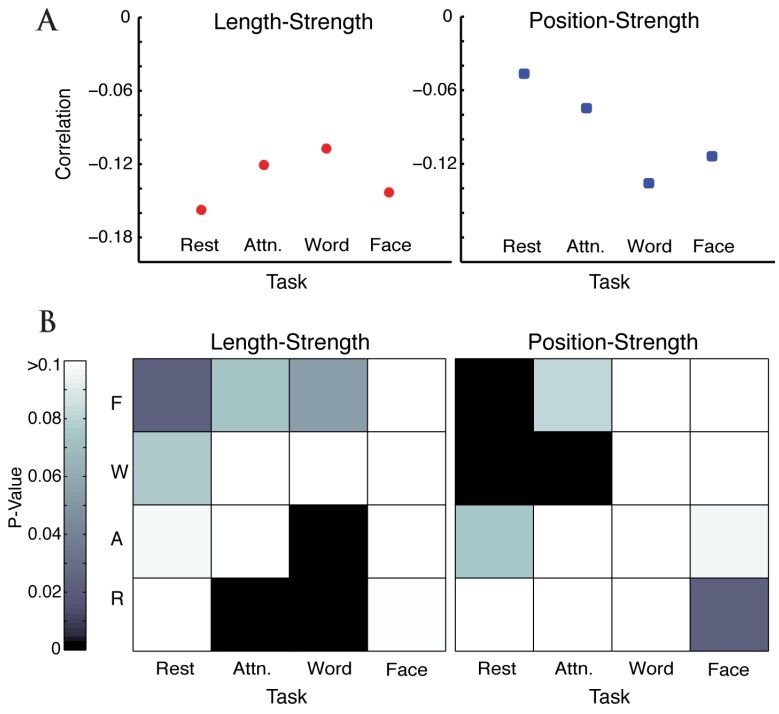

Figure 2.9: **Task-specific network statistics:** Values for the position-strength metric (blue) and the length-strength metric (red) for the four tasks are depicted in (A). (B) shows p-values for the pairwise statistical permutation test between tasks, where black denotes a significant value after a Bonferroni correction for multiple comparisons. Values are obtained for length-strength and position-strength metric. For example, on the y position plot in (B), attention-word is significant. Referring back to (A), we see that this implies the difference in the y position-strength correlation between the attention and word tasks is statistically significant.

two null models and our results based on the original time series establish the significance of our findings. Furthermore, the existence of hyperedges after the within-task shuffling indicates the presence of activity in some edges that is differentiated between tasks. Since there are fewer large hyperedges after the within-task shuffling, we can also confirm that there are hyperedges caused by edge dynamics within tasks. This chapter primarily concentrates on hyperedges correlated within a particular task, but future analyses to understand the properties of hyperedges that are grouped due to other general properties would supplement our results.

Table 2.1: **Brain Regions**

Region Name	**L**	**R**		**Region Name**	**L**	**R**
lateralorbitofrontal	2	2		pericalcarine	1	1
parsorbitalis	1	1		lateraloccipital	5	5
medialorbitofrontal	1	1		lingual	2	3
parstriangularis	1	1		fusiform	3	3
parsopercularis	2	2		parahippocampal	1	1
rostralmiddlefrontal	5	6		inferiortemporal	1	0
superiorfrontal	9	8		middletemporal	3	4
caudalmiddlefrontal	3	2		bankssts	1	1
precentral	7	6		superiortemporal	5	5
paracentral	1	1		transversetemporal	1	1
rostralanteriorcingulate	1	1		insula	2	2
caudalanteriorcingulate	0	1		thalamusproper	1	1
posteriorcingulate	2	2		caudate	1	1
isthmuscingulate	1	1		putamen	1	1
postcentral	7	5		pallidum	1	1
supramarginal	5	4		accumbensarea	1	1
superiorparietal	7	7		hippocampus	1	1
inferiorparietal	5	6		amygdala	1	1
precuneus	5	5				

Anatomical locations of the 194 brain regions used as network nodes in the hyperedge analysis, including the number of regions in left and right hemispheres in each brain area.

The hyperedge size distribution is comprised of "small" and "large" hyperedges, where the size distribution of the small hyperedges follows a power law and the large hyperedges peak at the system size. We explore the overall spatial hyperedge distribution by constructing a hyperedge node degree plot, and find that the majority of the most densely connected nodes lie in the posterior portions of the brain. To better observe spatial hyperedge properties, we develop a co-evolution network, where connection weights correspond to the probability that a hyperedge will include the connection. The top 1% of connections in the network with the highest probability of inclusion in a hyperedge are most concentrated in the occipital lobe and prefrontal cortex. These are expected areas

of hyperedge concentration, consistent with the visual nature of the tasks, as well as the coordination of quick decision making and the selection of specific motor responses.

2.4.2 Task-specificity and Anatomical Placement

We find there are hyperedges that are more correlated in one task and hyperedges that have a distinct profile across the tasks. Our results suggest that edges with a high probability of inclusion in task-specific hyperedges are often found in previously identified brain areas associated with the corresponding tasks, as discussed in detail below, confirming that the approach captures relevant information about task networks. In some cases, brain regions expected to show strong co-variation in a certain task are not included among the strongest connections of that task-specific co-evolution network; we also discuss examples of this in detail below. Repeating the analysis and grouping hyperedges that are significantly correlated in two tasks might lend insight into whether brain systems relevant to a certain task contain hyperedges that are correlated in another task and thus are rejected from our task-specific analysis.

In all tasks, stronger connections in the co-evolution network tend to be located in posterior portions of cortex and to be physically shorter. The higher probability of posterior edges to be included in hyperedges is consistent with the identification of a core set of highly structurally connected regions centered in the posterior of the brain, thought to play an important role in integrating large-scale functional connectivity [11, 44]. The tendency of strong connections to be physically shorter suggests high efficiency in task-specific co-evolution networks. This may reflect efficient wiring properties associated with minimal wiring for rapid processing and low energy expenditures found in structural brain networks and shared by some other biological and technological networked systems [67].

Rest

Resting-state brain activity contains correlated patterns that comprise a default mode network, a system that is engaged during internal cognition [68,69]. Certain brain regions active at rest are consistently deactivated during goal-oriented tasks, indicating that they comprise a functional mode that is rest-specific [17].

Our result that rest has fewer specific hyperedges than the attention or memory tasks could be a result of the specificity of correlated resting state regions, or a simplicity intrinsic to resting state function that does not necessitate more concerted efforts involving numerous brain regions [26]. In addition, we see a relative randomization and asymmetry in the spatial co-evolution distribution of rest-specific hyperedges, as well as a relative lack of long, strong connections; these results may correspond to a diminished need for efficient processing in a task-free environment.

Dense areas of the co-evolution network with high probabilities of being in rest-specific hyperedges include brain regions traditionally associated with the resting state. The inferior parietal lobule, superior frontal gyrus, precuneus, and posterior cingulate cortex have been identified as integral components of the default mode network; in addition, the posteromedial cortex, which includes the precuneus and posterior cingulate cortex, plays an important role in awareness [70–72].

Attention

Two attention systems exist in the human brain: a "top-down" network controls goal-directed attention, while a "bottom-up" group of brain regions detects and orients attention to relevant sensory stimuli that are generally novel or unexpected [73, 74]. Our task probes the former, as subjects are asked to focus on repetitive stimuli in a controlled environment. This requires an "executive control network," a bilateral dorsal system that

governs guided attention and working memory [75]. The relatively high degree of bilateral symmetry and the dorsal concentration of connections observed in the attention-specific co-evolution network suggests a higher probability for connections within this executive control network to co-evolve with other edges during the attention task.

Specifically, we observe regions of high clustering among the strongest connections in the attention-specific co-evolution network in the lateral parietal and occipital lobes, superior frontal cortex, and dorsal parietal cortex, areas known to be involved in attention networks. Parietal and frontal areas are involved in attention control and localization, specifically in visual attention tasks [73, 76]. Activation of the superior frontal cortex occurs in attention tasks, especially those that involve a shift to peripheral locations in the visual field [77, 78]. The dorsal parietal cortex also performs a central role in the executive control network: patients with lesions in the dorsal parietal cortex have shown significant impairment in goal-directed attention tasks [79].

Strong connections in the attention co-evolution network are more likely to be long than those in rest, corresponding to the high probability that long rostral-caudal edges will be included in hyperedges (visible in Figure 2.8B). This may reflect a greater need for coordination between prefrontal executive control regions and regions in the occipital lobe during the attention task. In addition, strong attention-specific co-evolution connections are less likely to be located in the posterior of the brain than those specific to word memory; this could indicate that the attention task state has less reliance on core visual regions than the word memory task state.

Memory for Words

Our results for the word memory-specific and the face memory-specific hypergraphs were similar in several ways. Both displayed many more "large" hyperedges than the rest or attention tasks, suggesting that some aspect of the memory tasks requires dynami-

cally coherent evolution over much of the brain. We speculate that this variation in the task-specific size distributions may correspond to the cognitive complexity demanded by the tasks, with the more involved memory tasks requiring more coordination between different cognitive networks and functions, and therefore producing more large hyperedges. This possibility could be further tested by examining hyperedge size variation across tasks specifically designed to vary in complexity.

Visual orthographic and face processing have a common reliance on central vision [80] and share neural circuitry [81]. The resemblance of the co-evolution networks for the two tasks, especially when compared with the very different graph structure of the attention and rest networks, indicates a similarity in the hypergraph representation of the memory tasks. This in turn signifies a correspondence in brain dynamics specific to memory. The task-specific analysis identifies hyperedges that show a significant correlation in only one task, so there is no overlap in these co-evolution networks.

Existence of a dedicated visual word processing network has been a topic of frequent discussion in neuroscience. The visual word form area (vWFA), located in the occipito-temporal cortex, is consistently activated by orthographic stimuli [82] and is invariant to changes in case, size, font, or type of visual stimulation [83, 84]. The vWFA has also been shown as functionally linked to the dorsal attention network in resting state fMRI data, indicating that it fulfills a complex cognitive role [85].

In the word memory-specific co-evolution network, the vWFA is highly connected, but there is minimal strong structure in dorsal attention areas, which we would expect to see in a functional connectivity analysis [85]. This can be explained by our methodology of selecting task-specific hyperedges. If edges in the dorsal attention network have similar co-evolution properties within the word memory and attention tasks, they will not be identified as task-specific edges.

Memory for Faces

Face recognition in humans requires a complex network distributed throughout the visual cortex that includes extended connections branching to other cortical regions [86]. The majority of visual processing occurs in the occipital lobe, located in the posterior of the brain. Functional MRI studies have identified multiple regions in the occipital cortex that respond more strongly to faces than other visual stimuli, indicating that the cognitive processes involving facial recognition are highly specialized [87, 88]. The especially dense concentration of connections in the occipital lobe at the highest probability levels of the face memory-specific co-evolution network is consistent with this.

The face perception system is composed of multiple bilateral regions; the lateral symmetry observed in the face memory-specific co-evolution network is consistent with this structure [86]. An aspect of the co-evolution network that breaks this symmetry is the right fusiform gyrus, which is strongly connected to other areas in the occipital lobe by high probability co-evolution pairs. A region in the fusiform gyrus, the fusiform face area (FFA), has been found to be selectively active in whole human facial perception, and the right FFA in particular has been found to have the most salient response to faces, with damage to the region severely impairing face recognition [89, 90]. The high probability of co-evolution between the right fusiform gyrus and other regions in this task-specific hypergraph is consistent with our expectation that regions involved in the memory of faces in particular (as opposed to words) are most likely to be included in face memory-specific hyperedges.

The co-evolution networks for both memory tasks show a significantly higher hyperedge probability in visual areas than the attention and rest tasks, and the differences in structure indicate that the hypergraph representation of memory tasks is significantly different from rest or attention. The marked differences in hyperedge statistics between

task states in our task-specific analysis suggest hypergraphs as a measure of functional network changes due to task states. With measures derived from the hyperedge analysis, we can begin to quantitatively probe the mechanisms of functional switching between tasks and gain insight into how distinct features of the network evolve in synchronized patterns.

2.4.3 Methodological Considerations

Because they consider both strong and weak edges with no thresholding, hypergraphs are well-suited for identifying groups of brain regions that, for example, initially have uncorrelated activity but become more correlated in synchrony (or vice-versa), as we expect task-associated cognitive networks to do as a result of switching between tasks. In order to extract these dynamic patterns, the hypergraph technique considers strong and weak edges equally, ignoring any offset between the average correlation strengths of different edge time series. This is intended to provide a complementary method to the common thresholding approach of separating or ignoring network edges with correlation strengths weaker than some critical value [12, 13]. Since weak edge connectivity has been shown to contain functionally relevant and predictive information in various contexts, retaining these edge weights is desirable [55, 91, 92]. There is also evidence that mean edge correlation values can be driven by non-biological artifacts such as head motion, even after applying standard motion-correction techniques [33]; by remaining indifferent to edge weight offsets, a hypergraph analysis avoids this concern.

In applications where the overall correlation strength of network edges is nevertheless important, it may be useful to supplement the dynamic information given by a hypergraph analysis with a measure that retains this edge weight information. Efforts to make quantitative comparisons between the hypergraph analysis and other dynamic

graph theoretical methods in the context of the human brain are ongoing. We are currently investigating whether dynamic community detection on weighted brain networks, a node-based analysis which relies on edge correlation strength, provides complementary information to the hypergraph analysis.

Because we choose a linear measure to compute correlations between edge weight time series, our analysis as presented here does not account for time lag in these correlations. However, our framework could be extended to nonlinear measures that include time-lag information.

It is important to note that our method of computing a dense matrix of edge-edge correlations and thresholding according to significance does not necessarily identify direct conditionally-dependent correlations between time series, or correlations that represent the underlying structural connectivity of the brain. As with any method that infers a network structure from correlation data simply by thresholding, we expect many of these correlations to be indirect. For example, a significant correlation between two edge weight time series may occur because both edges are being controlled by a third, more central edge – and not because the two edges are directly connected either causally or structurally. In this sense, the edge-edge correlation structure does not capture relations that necessarily reflect the underlying control structure or the physical architecture of the brain. Our hyperedge analysis moves the focus away from such indeterminate dyadic relationships, considering only groups of all edges that share similar dynamic patterns without any intra-group organization or structure.

It is also possible, as in any fMRI analysis, that edge-edge correlations arise from task-induced indirect drivers, such as visual stimuli. Two regions that are both activated by a visual stimulus may show strong functional connectivity with one another in a single time window. Moreover, such regions may show similar changes in functional connectivity over time if their activation profiles to the stimulus evolve similarly during the

experiment. As with any measurement of functional connectivity based on the Pearson correlation coefficient [93], a common and robust measurement of functional connectivity, such indirect drivers of functional connectivity are not distinguished from other more direct drivers of communication or interaction.

We observe a significant amount of individual variability in the hypergraph properties of interest. In this chapter, we have completed a group-level analysis and focused on investigating task-related differences in hypergraph structure. However, individual variability may be related to differences in cognitive ability and provide additional insight into the role of hyperedges in task performance, a possibility which we address in Chapter 3.

2.4.4 Conclusion

In this chapter, we use hypergraph analysis to identify significant co-evolution between brain regions in task-based functional activity and develop new tools to summarize the spatial patterns of these co-evolution dynamics over the group of subjects. By isolating task-specific hyperedges, we quantify significant differences between the spatial organization of co-evolution dynamics within different tasks. This hypergraph analysis adds a crucial perspective to previous treatments of task-based brain function, describing temporal similarities between spatially segregated neural circuits by specifically examining the organization of connections that co-evolve in time. It provides a promising approach for understanding fundamental properties of task-based functional brain dynamics, and how individual variation in these properties may correspond to differences in behavior and task performance.

Chapter 3

Individual Differences in Human Brain Hypergraph Properties

3.1 Introduction

Functional connectivity (FC) analyses based on fMRI data are effective tools for quantifying and characterizing interactions between brain regions. Many approaches borrow methods from the field of graph theory, in which FC is used to build graphs that model the brain as a complex network, treating brain regions as nodes and using functional connections (pairs of nodes with significantly related BOLD signal dynamics) to determine the edge structure of the network [7, 21]. Individual differences in both underlying FC and the complex network structure resulting from graph theory approaches have been investigated for a variety of task states, developmental stages, and clinical diagnoses [94–96].

Certain characteristics of FC have been found to vary consistently over the course of normal human aging. The loss of clear segmentation between neural systems is widely reported: many intrinsic functional connectivity networks in the brain tend to become

less internally coherent with age, and the functional differences between these intrinsic networks generally become less pronounced [97–99]. These changes are most commonly reported in the default mode network (DMN) [100–106], although they have also been observed in other networks, including those associated with higher cognitive functions [100, 102, 105–107]. In addition, inter-network connectivity between the DMN and other regions of the brain has been found to increase, diminishing the ability to discriminate between networks based on FC [104, 106]. There are some intrinsic functional networks, however, that show no changes or even increased intra-network connectivity with age, such as sensory networks [101, 103, 105].

The bulk of studies on age-related changes and other individual differences in FC, including those that use methods from complex networks and graph theory to represent FC patterns, are performed using static FC analysis, which represents the similarities of brain region activity (or some other measure of concordance) aggregated across an entire data set. In the present investigation, we build upon recent advances in network science to study individual differences in human brain activity and behavior from a dynamic network science perspective [108]. Dynamic functional connectivity (DFC) extends FC to examine how functional organization evolves over time [27, 109], allowing investigation of the changes in FC during the course of a cognitive task or scanning session. Efforts to probe the dynamics of functional brain networks have revealed that functional structure reconfigures over time in response to task demands [29, 31, 32, 110, 111] and spontaneously at rest [27, 112]. DFC methods have also been used to inform understanding of individual differences related to aging. In particular, dynamic community structure was found to vary significantly with age [113] and amplitude of low-frequency fluctuations of FC (ALFF-FC) was used to show age-dependent changes in the dynamics of interactions between networks [114]. Both studies imply that functional dynamics should be considered when investigating how aging affects brain network organization.

To address this, we use hypergraph analysis, a method from dynamic graph theory, to examine individual differences in DFC network structure in fMRI data acquired as subjects perform cognitively demanding tasks. The method is based on a generalization of standard graph theoretical techniques. In particular, by defining the standard node-node FC graph in successive temporal epochs, we construct a set of edge timeseries—that is, a vector of how the edge changes over time. The edge-edge DFC graph is constructed by treating these edge timeseries analogously to the node timeseries in the first step, and computing the relationship between every edge pair. Finally, we focus on "hyperedges," which are connected components of the absolute valued edge-edge DFC graph (described in more detail in Section 3.2) [16]. To contextualize hypergraph analysis, we define the graph theoretic elements used to construct hypergraphs as follows:

Node: As in the FC literature, nodes denote brain regions, or groups of voxels.

Edge: Also corresponding to the FC literature, edges denote correlations in activity between pairs of nodes over time. A node-node graph $G = \{V, E\}$ on N nodes will have $\binom{N}{2}$ edges, because each pair is considered. Unlike the majority of FC analyses, the edges are not thresholded for significance in the hypergraph analysis.

Links: Links denote significant correlations in activity between pairs of edges over time. An edge-edge graph $G' = \{V', E'\}$ on $\binom{N}{2}$ edges will have $\binom{\binom{N}{2}}{2}$ possible links, but tends to be sparse in practice.

Hyperedge: A hyperedge denotes a group of links connecting two or more edges with significantly correlated temporal profiles. Hyperedges are the simplest form of link community, since they are simply the connected components of the edge-edge graph $G' = \{V', E'\}$, where V is the set of edges and E is the set of links.

Hypergraph: A hypergraph is a set of hyperedges.

The hypergraph analysis is a simple first step toward understanding the structure of functional dynamics. Hyperedges are the connected components of the edge-edge

graph, and so avoid the introduction of additional unconstrained parameters, unlike many common FC and DFC methods such as community detection.

The groups of brain regions that comprise hyperedges are not necessarily strongly active or strongly interconnected brain regions. Rather, correlations in the dynamic connectivity of these regions are the defining characteristics that determine hyperedge structure. As a result, hyperedge analysis is able to identify groups of dynamic connections that change from strong to weak (or *vice versa*) cohesively together over time, providing complementary information to other DFC methods that focus on only the strongest node-node correlations, such as dynamic community detection [11, 35, 113]. Note that our choice of hyperedge metrics, as opposed to any other graph theoretic measure, is due to the simplicity of the hyperedge. Although it is beyond the scope of the present investigation, other graph properties of the edge-edge graph are likely to provide insight into dynamic brain network structure along other relevant dimensions. Nonetheless, hyperedges have some appealing intuitive validity in terms of the neural properties they might uncover—that is, in defining collections of nodes (or more technically, edges) on the basis of their similar dynamics.

In Chapter 2, we demonstrated that hyperedges discriminate between diverse task states in a group-level analysis of an fMRI data set spanning four tasks, which we refer to as the "multi-task" data set [111]. We also observed notable variation in descriptive hypergraph measures across individuals. However, given the level of abstraction involved in the construction of the hypergraph, an important first question is whether the method is able to capture well-known phenomena. In this chapter, we investigate the relationship between the variability in hypergraph cardinality and other individual difference measures. We develop and employ hypergraph measures that capture individual differences in functional brain dynamics to determine correspondences between dynamics and specific demographic and behavioral measures. In the multi-task data set,

we find that hypergraph cardinality—the number of distinct hyperedges within a subject's hypergraph—exhibits marked variation across individuals. At the same time, we find this measure is consistent within individuals, across overall hypergraphs and those associated with specific tasks.

To elucidate the drivers of this striking variation in hypergraph metrics observed across subjects, we explore systematic relationships between hypergraph cardinality and individual difference measures spanning distinct domains such as demographics, cognitive strategy, and personality. In the multi-task data set, we find a suggestive relationship between hypergraph cardinality and participant age. This relationship is confirmed with an independent analysis of a data set with participants who range in age from 18 to 75, which we refer to as the "age-memory" data set. We report a strong positive relationship between age and hypergraph cardinality: older participants are significantly more likely to have a larger number of distinct hyperedges in their hypergraph. This agrees with the widely reported phenomenon of the loss of cohesion within intrinsic functional brain systems, because an increase in the number of distinct hyperedges linking various brain regions points to interconnections between functional groups evolving in time [104, 106]. Thus, the hypergraph method agrees with previous descriptions of age-related brain changes, while capturing information about dynamics that adds a novel dimension to previous studies. The results of this chapter further recommend the hypergraph as a useful tool in studying structure in dynamic functional connectivity.

3.2 Methods

3.2.1 Ethics Statement

Informed written consent was obtained from each participant prior to experimental sessions for the multi-task and age-memory experiments. All procedures were approved by the University of California, Santa Barbara Human Participants Committee.

3.2.2 Background and Multi-Task Methods

Multi-Task Experimental Design

Participants were scanned at rest (task-free) and while engaging in three distinct tasks designed to elicit distinct cognitive functions: an attention-demanding task, a memory task with lexical stimuli, and a memory task with face stimuli. Participants were instructed to lie still and look at a blank screen for the duration of the rest period. During the attention task, participants were instructed to attend to sequences of images on a screen and detect the presence or absence of a target stimulus in designated test displays. Prior to the test display, a cue arrow provided probabilistic information on whether and where the target stimulus might appear. The test display was flashed for approximately 50 ms, after which participants chose whether or not the target stimulus had been present. Attention trials were separated by inter-stimulus intervals between 1200ms and 3200ms. In both memory tasks, participants were presented with 180 previously examined stimuli and 180 novel stimuli and were asked to discriminate between the two. The stimuli in the word and face memory tasks were 1.5 s or 1 s in duration, respectively, with a 1 s inter-stimulus interval. The memory tasks also included probabilistic cues indicating the probability that the stimulus was novel. For additional experimental details, see [12] and [42].

After completing the scans described above, the following individual difference measures were obtained for study participants: self-reported demographic information, self-reported state of mind (including physical and mental comfort) information, results from the Beck Depression Inventory II [115], tests for cognitive style (Santa Barbara Learning Style Questionnaire [116], Object Spatial Imagery Questionnaire [117], The Need for Cognition Questionnaire [118], Verbalizer-Visualizer Questionnaire [119], Card Rotation and Paper Folding Tests [120]), personality tests (Big Five Inventory [121] BIS/BAS scales [122], and PANAS mood assessment [123]). More individual difference measures were also collected, but do not match the individual difference measures collected from subjects in the age-memory study.

Image Acquisition and Processing

The MRI data were acquired from 116 participants at the UCSB Brain Imaging Center using a phased array 3T Siemens TIM Trio with a 12 channel head coil. In addition to functional data, a three dimensional high-resolution T1-weighted structural image of the whole brain was obtained for each participant. Functional MRI data were collected from 116 healthy adult participants over the four states described above in a block design format. Due to various sources of attrition, only 77 participants completed the functional scan and accompanying survey of detailed in [42]. The sampling period (TR) was 2 s for the rest and attention tasks and 2.5 s for both memory tasks (TE = 30ms, FA = 90). The rest task consisted of 146 acquired brain volumes (for a total of 292 s), each attention task consisted of 240 scanned volumes (480 s), and each memory task consisted of 540 volumes (1350 s).

The functional data is parcellated into regions using a "hybrid" adaptation of the multi-resolution Lausanne2008 atlas registered to MNI space [44] in order to apply the hypergraph analysis. This 194 region "hybrid" anatomical atlas minimizes variability in

region size between subjects and brain regions [111].

The functional data are preprocessed using FSL [47], AFNI [48] and Matlab [124]. Head motion was corrected for with MCFLIRT and voxelwise despiking was performed with AFNI 3dDespike [49]. Other preprocessing steps include non-brain removal with AFNI 3dAutomask, slice-timing correction with AFNI 3dTshift, and additional motion artifact correction with AFNI 3dDetrend. Additionally, each participant's time-averaged fMRI image is aligned to their structural T1 scan using FSL's FLIRT with boundary-based registration [49,52]. The inverse of this transformation is applied to all participants' parcellation scales (generated in structural space) and parcellations are down-sampled into functional space with AFNI 3dfractionize. The mean signal across all voxels within a given brain region is calculated to produce a single representative time series. An integer number of minute-long intervals is taken from the beginning of the time series for each task, and these are concatenated to produce a single time series, 3840 s in length, for each brain region.

Construction of Temporal Graphs

For each subject, we construct a dynamic graph model of brain function that accounts for changes in connectivity over time. Each of the $N = 194$ brain regions in the hybrid atlas is a node in the graph. The BOLD signal time series from each brain region is bandpass filtered to obtain data in the 0.06-0.125 Hz frequency range that contains task-related brain activity [14, 54–56]. Node-node adjacency matrices of size $N \times N$ are constructed by taking Pearson's correlations between each pair of the $N = 194$ nodes for each consecutive 60 s window of the bandpassed time series. Each node-node adjacency matrix represents the state of the functional network over that period of 60 s. Previous work found robust hyperedge size and spatial distributions for temporal window lengths between 40 to 120 s, and we choose a 60 s length consistent with previous analyses [111].

Furthermore, the variation in TR between tasks and corresponding variation in number of data points in the 60 s windows was shown to have minimal effects on the size and spatial hyperedge distributions [111].

Given the duration of each scan, this windowing yields four rest, 18 attention, 18 word memory, and 18 face memory node-node adjacency matrices. The set of node-node adjacency matrices, one for each one-minute segment, represents the dynamic functional connectivity graph; each edge, or pairwise connection between nodes, has an edge weight time series describing its temporal evolution across time windows, as depicted in Figure 2.2 (B).

Hypergraph Construction

Hyperedges are groups of edges that have related temporal profiles, so hypergraph structure is determined from the correlations between the time-evolving weights of edges [16] (See Figure 2.2 for a schematic illustration of hypergraph construction). These are represented in an edge-edge adjacency matrix $\mathbf{X}$, of size $E \times E$, where $E = N(N-1)/2$ is the total number of possible edges in one time window of the DFC graph. Each entry in $\mathbf{X}$ is given by the Pearson correlation between the corresponding pair of edge weight time series in the DFC graph. The p-values from these correlations are thresholded by a false discovery rate correction, which is more sensitive than other corrections for multiple comparisons and is thus effective for such neuroimaging network analyses [59]. When the correlation between edges i and j is significant ($p < 0.05$), we set $\xi_{ij} = X_{ij}$, to form the thresholded matrix ξ. All other elements of ξ are set to zero. We binarize this thresholded

matrix and obtain ξ'_{ij}, where

$$\xi'_{ij} = \begin{cases} 1, & \text{if } \xi_{ij} \neq 0; \\ 0, & \text{if } \xi_{ij} = 0. \end{cases} \tag{3.1}$$

Each connected component in the thresholded edge-edge correlation matrix ξ' – that is, each set of edges with correlations between any two edges in the set but no significant correlation with edges in any other set – forms a hyperedge. Taken together, all hyperedges in ξ form a hypergraph. Since the edge weight time series are never thresholded and both high and low edge weights are preserved, hypergraphs provide information about edge dynamics without restricting the analysis to strong correlations in regional time series. Hypergraphs are constructed from significant positive and negative correlations to incorporate a broad definition of whether two edges are "related." By definition, this approach precludes pairs of hyperedges that are anti-correlated. An alternative mechanism for constructing hypergraphs would treat positive and negative correlations separately, but here we only consider absolute valued relationships.

In this analysis, we consider a system with 194 nodes, and $\binom{194}{2}$ possible edges. Hyperedges are collections of edges, which results in a high dimension for this system. As such, we illustrate an schematic hypergraph on a smaller number of nodes for visualization purposes, seen in Figure 3.1.

Our results are compared with a null model designed to ensure that hyperedges identified in our analysis can be attributed to system dynamics, rather than overall statistical properties of the data [35]. To destroy temporal correspondences between edges but retain the mean and variance of each edge weight time series, the null model randomly reorders each edge time series individually and calculates correlations between the reordered edges.

Once hypergraphs are identified for each individual in the multi-task data set, hyperedges are classified according to whether the correlation in a cognitive state (i.e., rest or one of three cognitive tasks) is significant compared to a permutation null model over all states [111]. The hyperedges that satisfy these requirements are denoted as task-specific hyperedges, which we combine to form task-specific hypergraphs.

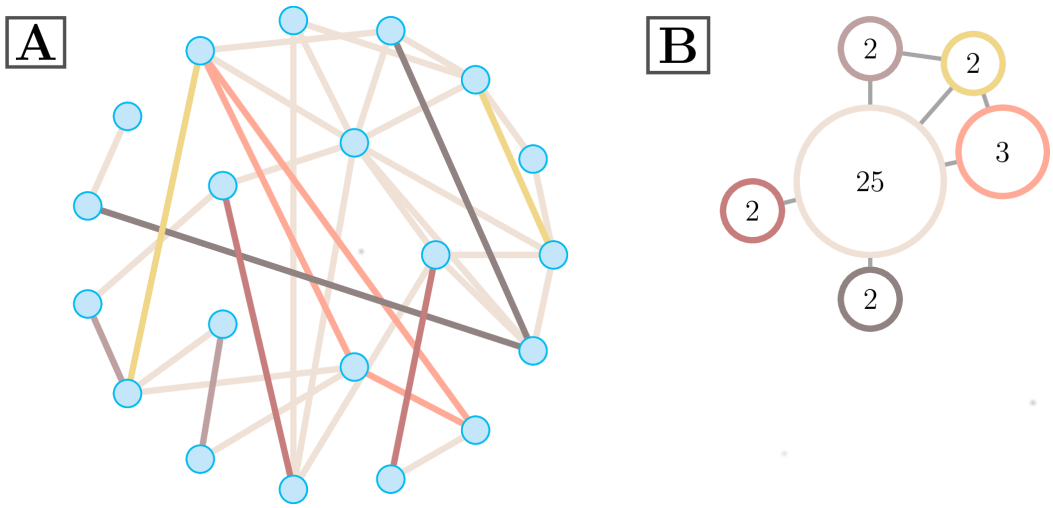

Figure 3.1: **Example ξ':** Two depictions of a representative hypergraph on 20 nodes. The hypergraph with singletons removed is shown in (A), where edge color represents hyperedge assignment. The size of a particular hyperedge is the number of edges in it, as in Equation 3.2. This illustrative hypergraph is comprised of six distinct hyperedges of various sizes. An alternative hypergraph representation is depicted in (B), where colors directly correspond to (A). Each hyperedge in the hypergraph is represented by a single vertex. Vertex labels correspond to hyperedge size, and edges are drawn when hyperedges connect to at least one common node.

Hypergraph Metrics

In this analysis, we examine several complementary measures on individual hypergraphs and focus on one of the most straighforward and transparent of these measures to extract meaningful information from the overall hyperedge distribution.

Hyperedge size: The size, $s(h)$, of a hyperedge h, is defined by

$$s(h) = \sum_{i,j \in h} \xi'_{i,j}, \tag{3.2}$$

where the sum is over the upper triangular elements of ξ', the binarized edge-edge adjacency matrix defined above. This is equivalent to the number of edges that are designated as part of this hyperedge. An illustration of a hypergraph with hyperedges of varying sizes can be seen in Figure 3.1.

Singletons: Singletons are hyperedges with $s(h) = 1$, edges with no significant correlation with any other edge in the graph. We exclude singletons from the following analyses.

Hypergraph cardinality: The cardinality of an individual hypergraph is the number of non-singleton hyperedges present in the hypergraph. The cardinality of the representative hypergraph in Figure 3.1 is six.

Hyperedge node degree: The hyperedge degree of a node is the total number of hyperedges that contain that node.

Task-specific hyperedges: Hyperedges that exhibit a significantly higher correlation within one particular task are grouped into task-specific sets. The sets are calculated by using a permutation test to compare the correlation between edge time series for groups of edges in hypereges in a single task to the same correlation with edge time series data chosen randomly from all tasks. A Bonferroni correction for false positives due to multiple comparisons is employed to select task-specific hyperedges using the most stringent requirements [64].

Regression Procedure

To investigate possible correlates of variability in individual hypergraph metrics, we perform a series of regression analyses. In each analysis, we use the hypergraph metric as the dependent variable and factors representing individual difference measures from the psychometric tests as the independent variables.

Behavioral data categorization: Behavioral and performance data for the multi-task study consist of 231 measures, while there are 115 measures for the age-memory study participants. There are 42 individual difference measures common to both studies, which we group into five categories, given in Table 3.1. These categories are comprised of differing numbers of individual difference measures, which are summarized in Table B.1.

Category	**Factors**	**Information Retained**
Performance	2	91.41%
Demographics	2	92.62%
State of Mind	3	80.45%
Cognitive Factors	4	77.64%
Personality	6	77.79%

Table 3.1: **Information retained for multi-task study:** Categories, number of factors for each, and how much overall variance from the multi-task individual difference data was retained for each category. Each category represents a subset of the 42 individual difference measures and the factors represent a percentage of the variance contained in the category for the multi-task data.

Singular value decomposition: Once the individual difference measures have been categorized, we demean all measures and perform a singular value decomposition (SVD) separately for each category. We choose the minimum number of factors from the SVD for each category that retain at least 75% of the variance across the category of measures from the multi-task study. Results from this process are presented in Table 3.1.

R^2 *change:* The number of factors retained is not constant across categories, so we implement an adapted multivariate hierarchical regression [125, 126] to establish the comparative informativeness of each category. To assess the explanatory power of a given

category, all factors in that category are held out for a "control" regression, and the difference in model R^2 between this reduced model and the full model is denoted as the contribution for that category. This corresponds to repeatedly performing a hierarchical regression with each category computed last, which gives a conservative estimate for the amount of variance attributable to the category [126].

Significance test: To determine the significance of the regression coefficients, we use the p-values from t-tests on each multiple regression performed. The Bonferroni procedure for correcting for false positives due to multiple comparisons is used to adjust the t-test p-values over all regressions performed in this study [64]. We employ the Bonferroni correction for multiple comparisons in all regression analyses because it is the most stringent test for significance.

3.2.3 Age-Memory Methods

The majority of the methods are identical to those discussed for the multi-task data set. Below, we point out aspects that differ between the two analyses.

Age-Memory Experimental Design

The word memory task in the age-memory study is constructed similarly to the word memory task in the multi-task data set. In addition to the memory task, participants completed a resting state scan and diffusion-tensor imaging, which we do not analyze further. Participants did not complete the face memory or attention tasks described in the first data set. The BOLD data were acquired while adult participants performed a recognition memory task with probabilistic cues. Prior to the scanning session, the participants studied 153 common English words, which were mixed with 153 novel lexical stimuli during the task. Participants were asked to determine whether the stimuli

were studied or unstudied, with font color cues indicating whether the word had a 70% probability or a 30% probability of having been previously studied [127].

Image Acquisition and Processing

Functional and structural data were collected from 126 healthy participants engaged in the word memory task. All functional data was acquired with a 3T Siemens TIM Trio MRI system with a 12-channel head coil. Scans consisted of T2*-weighted single shot gradient echo, echo-planar sequences sensitive to BOLD contrast (TR = 1.6 s; TE = 30 ms; FA = 90) with generalized autocalibrating partially parallel acquisitions (GRAPPA). Subjects were scanned performing the task, with a total of 948 brain volumes acquired (1516.8 s). In additon to the functional scans, high-resolution anatomical scans were performed for each participant using an MPRAGE sequence (TR = 2.3 s; TE = 2.98 ms; FA = 9; 160 slices; 1.1 mm thickness). Study participants also underwent behavioral assessments and psychological testing. Functional data from 31 participants were excluded due to technical issues, metal screening issues, claustrophobia, attrition, or lack of a complete individual differences survey. The results presented here are from 95 participants with usable functional and individual difference data.

The functional data are preprocessed using FSL [47], AFNI [48], and Matlab [124]. Preprocessing includes head motion correction (MCFLIRT) [49], non-brain removal (BET) [128], high-pass temporal filtering ($\sigma = 50$s), spatial smoothing, and grand mean intensity normalization (FEAT) [129]. Each voxel's time series is further denoised using a nuisance regression. The nuisance regression includes regressors for the six motion correction terms returned by MCFLIRT, their temporal derivatives, and the mean signal time series from the cerebrospinal fluid. The denoised data is registered to MNI space using FLIRT [130, 131]. The T1 scan is first registered to the MNI template (12 df affine transformation), the functional data are registered with the T1 image (6 df affine

transformation, trilinear interpolation), and the transformations are combined. As in the multi-task study, the mean BOLD signal across all voxels within a given brain region is calculated to produce a single representative time series.

Construction of Temporal Graphs

Time series are demeaned and concatenated across the three functional runs of the word memory task to produce a single time series for each brain region. DFC graphs are constructed here analogously to the multi-task study, with one key difference. In the age-memory analysis, we remove a single node-node adjacency matrix (i.e., a single time window) from the beginning and end of each functional run. This is to counteract edge effects from processing and ensure continuity across runs. We address this choice further in Section 3.4.3 and Appendix B.

Regression Procedure

The regression procedure is similar to the analysis performed on the multi-task data. The individual difference data is kept in the common format, where only the 42 measures common to both studies are used and the categories are the same. Furthermore, the R^2 change and significance tests are calculated as above.

Singular value decomposition: We demean all measures and perform a singular value decomposition (SVD) on the combined multi-task and age-memory data separately for each category. This differs from the multi-task analysis, where we only consider the variance retained over the multi-task data. We choose the minimum number of factors from each SVD that retain at least 75% of the variance across both studies. Results from this process are presented in Table 3.2.

Category	Factors	Information Retained
Performance	1	87.18%
Demographics	1	86.14 %
State of Mind	3	77.09%
Cognitive Factors	3	81.25%
Personality	4	78.56%

Table 3.2: **Factors common to the multi-task and age-memory trials:** Categories, number of factors assigned to each, and how much of the overall variance was retained in each category. Each category represents a subset of the 42 individual difference measures and the factors represent a percentage of the variance contained in the category.

3.3 Results

As mentioned above, the hyperedge method has been applied to the multi-task data set in a previous study [111]. Here, we first recapitulate the key findings from that investigation and provide results of exploratory analyses that motivate the followup analyses on the age-memory data set. We then present results from the age-memory analysis.

3.3.1 Summary of Prior Results

A previous study of the multi-task data, detailed in Chapter 2, identified measures that capture significant differences in population-level hypergraph structure across tasks [111]. Furthermore, extensive variation was observed in several hypergraph measures, including hypergraph cardinality, across individuals. These results emphasize that hypergraph structure can be used to differentiate between task states and motivates our investigation of the correspondence between hypergraph structure and individual difference measures.

Figure 2.7 depicts the empirical cumulative hyperedge size distributions for all hyperedges found across all subjects in the multi-task data set. As a null test, we shuffle the data over time and find no hyperedges of size greater than one. There is a rough

power law for the smaller sizes ($s < 100$), followed by a gap in the distribution from about 100 to 1000 and a sharp drop at the system size ($s = \binom{194}{2} = 18721$). The shape of the distribution is due to the consistent hypergraph structure across individuals; the majority of subjects in this study have a hypergraph composed of one large hyperedge and many small hyperedges. While this characteristic structure is common to most subjects in the study, the size of the largest hyperedge varies across individuals. This size is closely related to the hypergraph cardinality, defined as the number of hyperedges in a hypergraph, a measure which also exhibits large variation.

Figure 2.7 also depicts task-dependent differences in the cumulative size distributions of task-specific hyperedges. Memory-specific hyperedges tend to be more numerous than those specific to the rest and attention tasks. However, the total number of task-specific hyperedges for any task is at least ten times fewer than the total number of hyperedges. Our strict definition of task specificity includes only hyperedges specific to a single task and discards those associated with more than one task. This approach is conservative, and likely leaves some meaningfully task-related hyperedges unclassified. However, it reduces the dimension of the task-specific results, and provides greater confidence that any hyperedges classified as task-specific are indeed providing truly task-driven information due to coherence within that task alone, rather than coherence due to an unrelated driver that is common to several tasks.

There are significant differences in the spatial organization of task-specific hyperedges over all individuals, visualized in Figure 2.8A, in Chapter 2. The plots depict task-specific hyperedge degree across the brain for each of the four tasks. In addition to the differences in magnitude between word memory and the other tasks, the locations of high hyperedge concentration vary with task.

These significant differences in hypergraph structure between the tasks confirm that hypergraph structure varies between task states. However, persistent variability in hyper-

graph measures across individuals indicates that the hypergraph method reflects innate differences beyond the current task state. The work presented here follows this line of inquiry, beginning with an analysis of individual differences in the multi-task data set.

3.3.2 Multi-Task Results: Individual Differences

Here, we illustrate and quantify the wide variation in hypergraph measures across individuals in the multi-task data. In brief, we identify a particular measure, hypergraph cardinality, that demonstrates large variance across all individuals but is consistent within individuals. Following this, we investigate relationships between the variation in individual difference measures and the variation in hypergraph cardinality. The results from this study are not statistically significant due to the limited variation in individual difference measures and strict corrections for multiple comparisons. However, we report a marginally significant result relating demographics and word-memory hyperedge cardinality that motivates further analyses on the age-memory data set.

Individual Variability and Consistency in Hypergraph Metrics

Although our previous study focused on group-level properties of hypergraphs across tasks, notable individual differences in functional dynamics were also seen [111]. Here, we confirm those preliminary observations by investigating the hypergraph cardinality measure and finding that it displays extreme variations across subjects in the multi-task data set, as shown in panel (A) of Figure 3.2. These individual variations in hypergraph cardinality span several orders of magnitude.

Despite this large variation between participants, hypergraph cardinality follows a consistent pattern within each participant across tasks. Panel (B) of Figure 3.2 depicts individual measures of hypergraph cardinality for hyperedges specific to each task, with

subjects sorted by rest hypergraph cardinality. Within participants, the task-specific hypergraph cardinality is consistent across task states and follows the distribution for rest-specific hyperedges, which further emphasizes the consistency of hypergraph cardinality within individuals.

Consistent hypergraph cardinality within participants over all tasks indicates that there are characteristics specific to individuals that drive hypergraph properties, even in designated task-specific hypergraphs. These patterns imply the existence of driving influences on hypergraph structure that are independent of performance on a specific task. To investigate this further, we examine how individual difference measures from demographic and behavioral data relate to hypergraph cardinality.

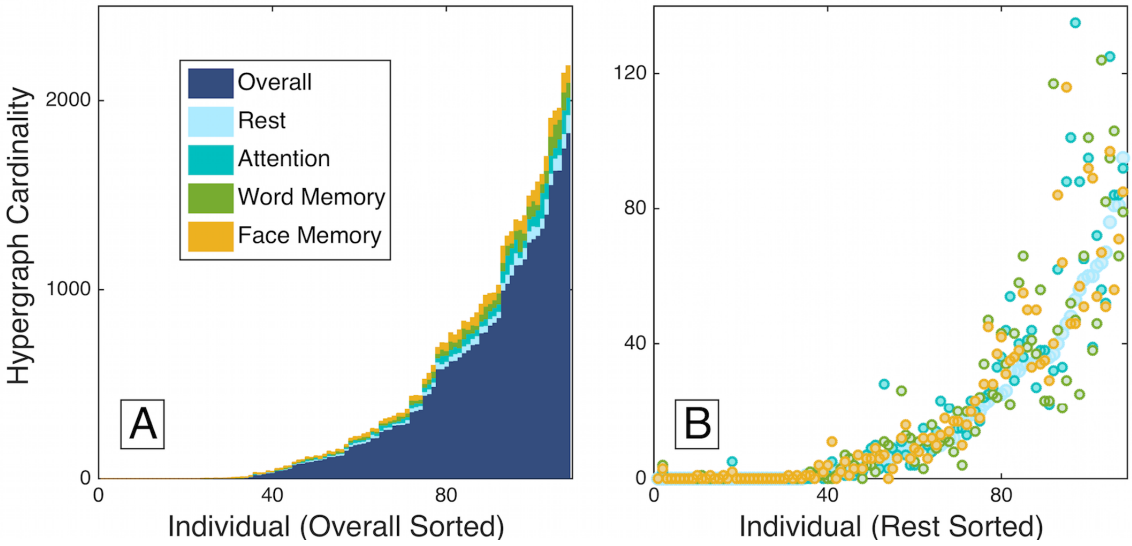

Figure 3.2: **Individual variability:** Hypergraph cardinality for individual overall multi-task hypergraphs is shown in panel (A), sorted by increasing overall cardinality. Individual task-specific hypergraph cardinality is shown atop the overall cardinality for comparison, and is also sorted by increasing overall cardinality. Panel (B) depicts the cardinality for task-specific hyperedges, sorted by rest cardinality. The number of hyperedges across tasks is fairly consistent within individuals, in contrast to the range of hyperedge number across individuals.

Drivers of Individual Variability

To investigate possible sources of the large variation in hypergraph cardinality seen above, as well as to quantify the extent of the consistency of hyperedge cardinality across tasks, we perform a series of multiple regression analyses on the multi-task data, as described in Section 3.2.2.

First, using the cardinality of task-specific hypergraphs as the dependent variable, we perform a regression analysis for each non-resting task (attention, word memory, and face memory) that includes the cardinality of the rest-specific hypergraph and the factors shown in Table 3.1 as independent variables. Table 3.3 gives the R^2 change values and p-values associated with the rest predictor for each task-specific regression. In all three tasks, the rest predictor alone significantly explains the variance in task-specific hypergraph cardinality. This confirms and quantifies our observation in Figure 3.2 that hypergraph cardinality is consistent across each individual's task-specific hypergraphs—i.e., it is trait-like. The individual difference measures used as independent variables are not significant after the Bonferroni correction for multiple comparisons over all tests. However, including the rest-specific hypergraph cardinality, which is closely linked to overall hypergraph cardinality, as an independent variable in the regression accounts for the variation across individuals that is consistent across tasks.

	Attention	**Word Memory**	**Face Memory**
R^2 **change**	0.72	0.58	0.68
p**-value**	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$

Table 3.3: **Rest regression R^2 values:** R^2 values for the regression between rest-specific hyperedge cardinality and hyperedge cardinality for each of the other three tasks.

To identify possible drivers of this individual variation, we perform another regression analysis, using the individual difference measures from Table 3.1 as independent variables and overall hypergraph cardinality as the dependent variable. Figure 3.3 depicts the R^2

changes from this analysis for each category of factors. The t-test identifies no factors with significant correspondence to hypergraph cardinality, but we observe that the demographics category has the largest R^2 change. The t-test p-value for one of the factors in the demographics category is $p < 0.05$ and is by far the lowest p-value in this stage of the analysis. However, due to our stringent requirements for correcting for multiple comparisons and the number of tests we performed, this correlation is not statistically significant. The marginally significant demographics factor has a loading of -0.95 for the age measure and -0.31 for the years of education measure; the loading for sex and handedness demographic measures are comparatively negligible, with magnitudes less than 0.02.

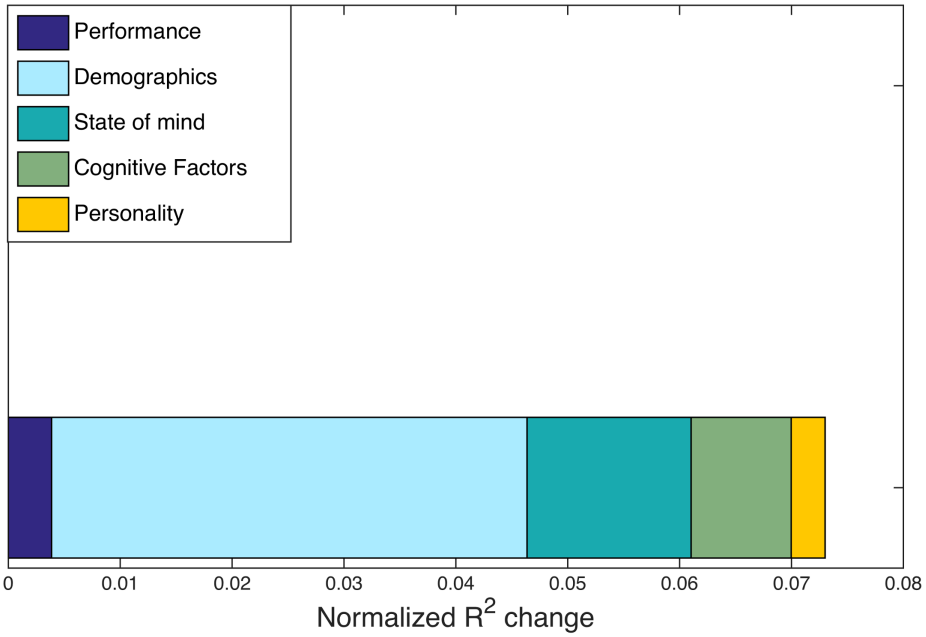

Figure 3.3: **Multi-task R^2 changes:** Normalized R^2 changes with respect to hypergraph cardinality are shown for individuals in the multi-task study. R^2 changes are calculated from the regression procedure outlined in Section 3.2.2, with five distinct categories common to the multi-task and age-memory studies. The largest normalized R^2 change is from the demographics factor, but no factors exhibit a signficant correspondence with hypergraph cardinality.

Summary of Multi-task Results

On the basis of our previous results applying hyperedge analysis to this data set in Chapter 2, which hints at substantial variability across individuals in hypergraph structure (Figure 2.7), we carry out several regression analyses designed to identify individual drivers of this variability. There were two key results. The first result is that overall and task-specific hypergraph cardinality show notable variation between subjects, but remarkable consistency within subjects for all tasks (Figure 3.2).

The second key result from this exploratory analysis is the finding of a marginally significant relationship between the demographics category and hyperedge cardinality. Limits to the explanatory power of the multi-task data set may be determined by limited variation in some demographic measures – particularly the small range (27–45) and standard deviation (4.24) in subject age, which poorly represents the ages observed in the entire population. We thus extend our analysis to a complementary data set collected on a longer study of the word memory task with participants aged 18–75, with a standard deviation of 22.11. In the next section, we report the results of our independent analysis of this age-memory data set, which confirm the relationship between age and hypergraph cardinality suggested by the multi-task results.

3.3.3 Age-Memory Results

To supplement the findings from the multi-task data set, we perform a parallel set of analyses on the age-memory data set. The data set includes participants with ages ranging from 18 to 75, a range three times larger than the range of ages in the multi-task study. Furthermore, the age-memory study uses an almost identical task to the multi-task word-memory task. In this section, we combine hypergraph results for all participants in the age-memory data set and obtain a distribution of hyperedge size

over all participants with similar features to the hyperedge size distribution from the word-memory task of the multi-task data. We then identify and test specific drivers of individual variation in hypergraph cardinality for the age-memory study participants. We find a strong correspondence between age and hypergraph cardinality that confirms the preliminary result from the multi-task study.

Hypergraph Statistics

The cumulative size distribution of hyperedges for all individuals in the age-memory study is depicted in blue in Panel (A) of Figure 3.4. To compare these age-memory hyperedges with the word memory portion of the multi-task study, we identify a new set of hyperedges using only the portion of the multi-task functional time series recorded during the word-memory task for each subject; the distribution of sizes for these hyperedges are plotted in pink. Note that these new word-memory hyperedges from the multi-task data are fundamentally different from the "word memory-specific" hyperedges depicted in Figure 2.7. The "word memory-specific" hyperedges are those hyperedges computed over all tasks, but classified to be driven by correlations in the word memory task alone. In contrast, the new word-memory hyperedges in Figure 3.4 are found by using just the word-memory subset of the multi-task data, with no further classification applied.

The distributions of sizes are similar at smaller size scales, but differ somewhat at larger size scales. There are many more hyperedges close to the system size in the age-memory task, while the word-memory hyperedges from the multi-task data set tend to be smaller. The length of the multi-task word-memory time series is shorter than the age-memory time series, which may contribute to this effect [132]. To investigate the size distributions without the effect of full-brain hyperedges, we remove the largest hyperedge from each subject's hypergraph and plot the resulting distribution in Panel B of Figure 3.4. With this adjustment, the distribution of age-memory hyperedge sizes has

a striking agreement with the size distribution of hyperedges constructed from the multi-task word memory data. In both distributions, there is power law behavior for small sizes, similar to that observed in Figure 2.7. Furthermore, the distributions without the largest hyperedges are almost identical; the power of the fit to multi-task word memory data is -2.21 and the intercept is 7.91×10^4, while the power of the fit to the age-memory data is -2.37 and the intercept is 1.46×10^5.

We construct a null model, as detailed in Sections 2.2.10 and 3.2.2, by temporally shuffling the data and find no hyperedges with size greater than one, indicating that the hyperedges identified in the unshuffled data are capturing statistically significant aspects of brain dynamics. In addition, the close correspondence between these two distributions of word-memory hyperedges suggests that the analysis captures aspects of brain dynamics that are robust across imaging sessions and populations.

The inter-subject variability in multi-task hypergraph cardinality spanned several orders of magnitude and followed consistent patterns within subjects for differing cognitive states. We compare the individual hypergraph cardinality for the age-memory and multi-task word-only studies in Figure 3.5. In the age-memory data, hypergraph cardinality ranges from 0 to 1817, which is a similar range of variability as that observed for the complete overall multi-task data set in Figure 3.2. There are 79 subjects with nonzero hyperedge cardinality, indicating that significant non-singleton hyperedges are present in less than two thirds of the subjects. For the remaining analyses, we only consider the 79 subjects with nonzero hypergraph cardinality. For the overall hypergraphs, hypergraph cardinality ranges from 0 to 1832. The maximum hypergraph cardinality for the multi-task word-only data is 1408, which is markedly less than that observed for the age-memory data and may be a result of the shorter time series for the multi-task word task. The presence of near-system size hyperedges, which may also be due to the shorter multi-task word time series, affects hypergraph cardinality by resulting in hypergraphs

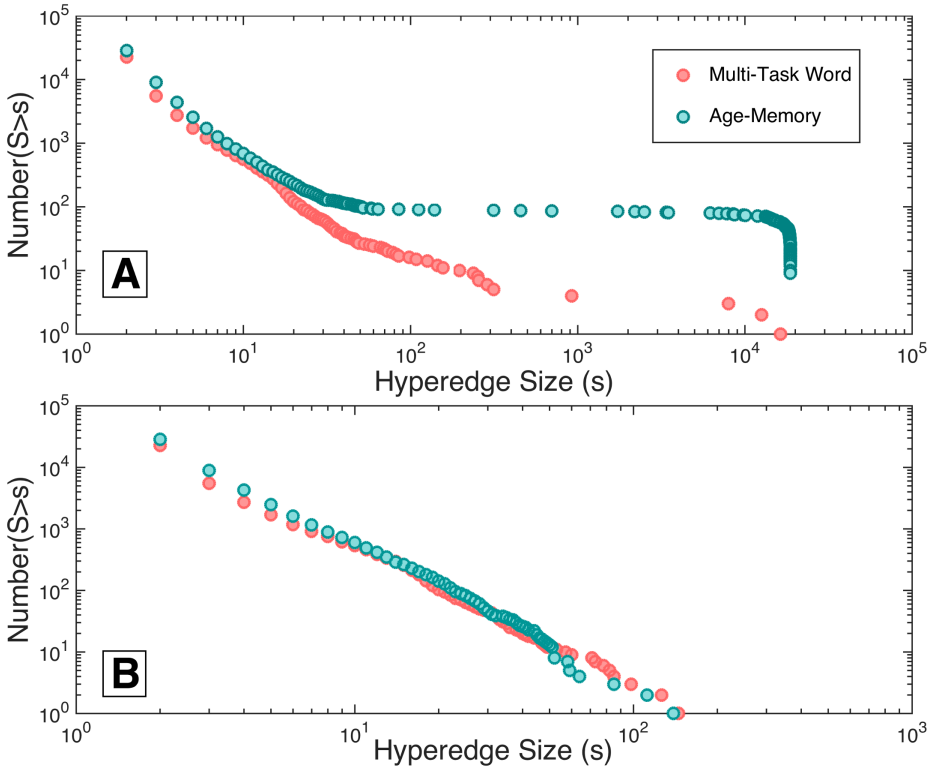

Figure 3.4: **Comparison of cumulative size distribution:** Panel (A) depicts the cumulative distribution of hyperedge sizes over all individuals in the age-memory study compared with the sizes of the set of hyperedges constructed from only the word-memory task of the multi-task data set. Differences in the number of large hyperedges can be observed between the two tasks, but both contain a similar number of hyperedges over all individuals. Panel (B) illustrates the cumulative distribution of sizes for all individuals in both studies with the largest hyperedge for each individual subject removed. When this is done, the distributions overlap and are well described by a power law with close alignment in slope and magnitude across studies.

with cardinality near one.

Age-Memory Hypergraph Correspondence With Age

Having confirmed that hypergraph composition is similar for the multi-task word study and the age-memory study, we investigate whether the individual variability in hypergraph cardinality seen in Figure 3.5 corresponds to individual difference factors for the age-memory study.

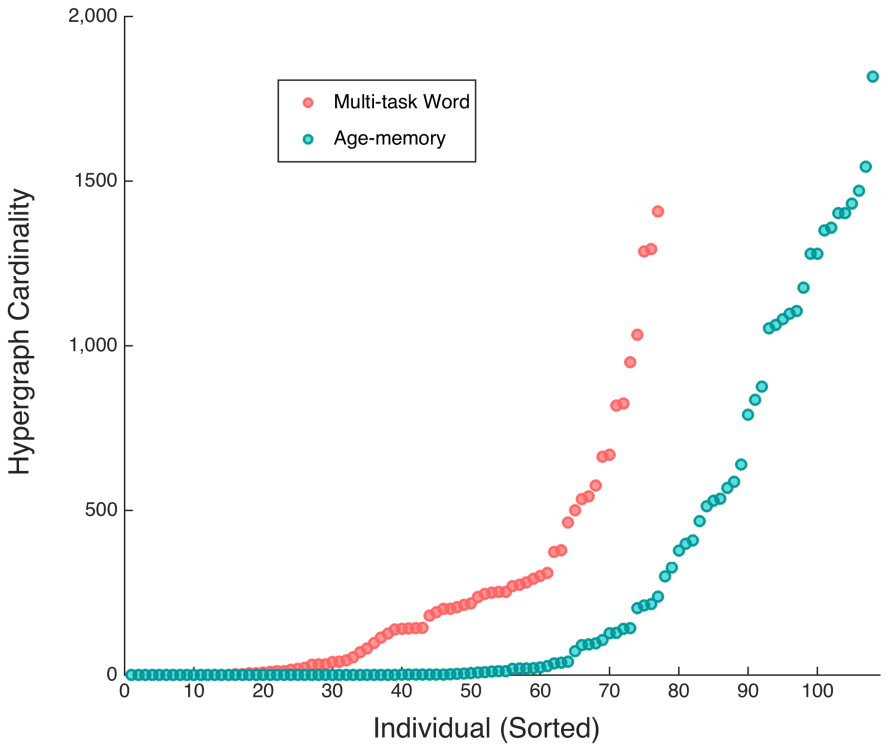

Figure 3.5: **Sorted hypergraph cardinality:** Increasing hyperedge cardinality for individual multi-task word-only and age-memory hypergraphs. The variability for both studies is similar to the variability in multi-task overall hypergraph cardinality, depicted in Panel (A) of Figure 3.2. The range of hypergraph cardinalities for subjects in the word-only data is smaller than either the overall multi-task study or age-memory study.

We perform a multiple regression on the 12 factors distributed across five categories in Table 3.2. Head motion has been found to induce correlations in FC analyses [50], and a previous study using this data found a significant correlation between age and amount of head motion during the experiment [127]. To ensure that excessive head motion is not contributing to our result in any way, we include head motion (operationalized as the average relative movement as computed by MCFLIRT) as a predictor in this regression.

The overall R^2 value for the multiple regression analysis was 0.3452, indicating that the predictors explain about a third of the variance in the overall data. After a Bonferroni correction for multiple comparisons across all regression studies included in this chapter

[64], the demographics factor is the only significant predictor of hyperedge cardinality. The normalized R^2 changes for hypergraph cardinality can be seen in Figure 3.6; the demographics factor has the largest normalized R^2 change and the only significant p-value ($p < 0.005$) in the regression. These results correspond with the marginal result from the multi-task data set, where the demographics factor is a marginally significant predictor.

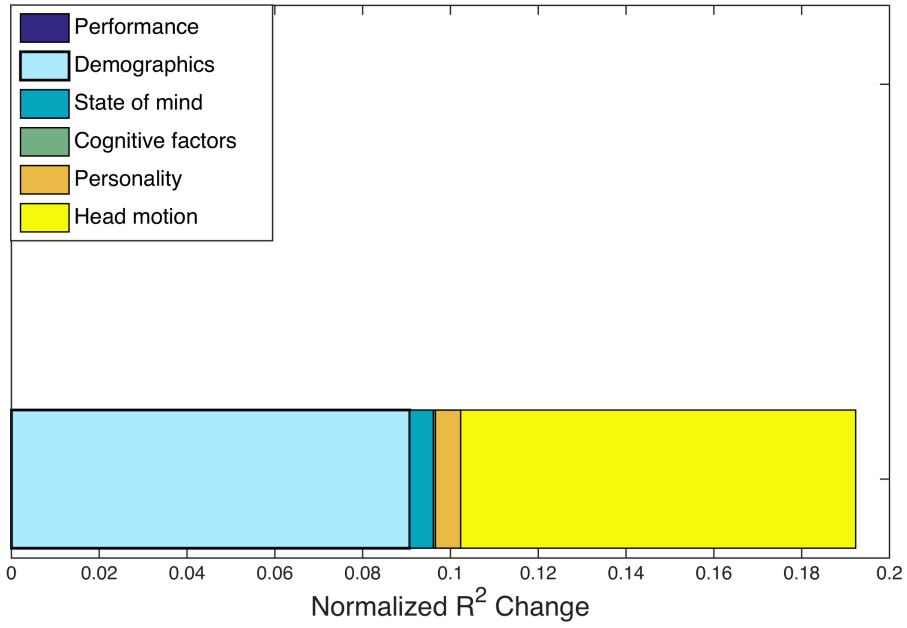

Figure 3.6: **Age-memory R^2 changes:** Normalized R^2 changes with respect to hypergraph cardinality across individuals in the age-memory study. The largest normalized R^2 changes are from the demographics factor and head motion measure, but the demographics factor is the only significant predictor of hypergraph cardinality. In this figure, prediction significance is denoted with a bold outline. The composition of R^2 changes for the age-memory task is consistent with that seen for the multi-task data in Figure 3.3, in that the normalized R^2 change is largely due to the demographics factor.

Much of the variation in the demographics factor (73.5%) is directly attributable to age. We attempt to isolate the specific relationship between age and hypergraph cardinality by performing a separate regression. In this regression, hypergraph cardinality is the dependent variable and the independent variables are age and head motion. The

relationship between age and hypergraph cardinality is significant, with the t-test p-value well below the Bonferroni correction over all regression analyses presented in this work, at $p < 0.001$.

This is a positive relationship, indicating that older individuals tend to have higher hypergraph cardinality, while younger participants tend towards lower hypergraph cardinality. An illustration of this correspondence between hypergraph cardinality and age is presented in Figure 3.7. As age increases, the number of hyperedges in a participant's hypergraph increases as well. We verify that this relationship holds beyond this particular study by reintroducing the word-memory data from the multi-task study and performing a correlation between hypergraph cardinality and age over both studies. Age and hypergraph cardinality have a Spearman correlation coefficient of $\rho = 0.32$, and the p-value for this correlation, $p < 10^{-5}$, is significant when we use the Bonferroni correction over all analyses presented in this chapter.

Spatial Distribution of Hyperedges

Given the positive relationship between age and hypergraph cardinality, we next identify how the spatial organization of hyperedges reflect the increase in cardinality. We group subjects from the age-memory study into three age ranges based on the age-memory task data distribution: 18 years old (39 subjects), 25-33 years old (34 subjects), and 60-75 years old (35 subjects). For each set of subjects, we calculate the average hyperedge node degree for each region and depict them on the brain in Figure 3.8. The plots for the two younger populations exhibit few differences, although there is a slight increase in degree for the middle population. Hypergraphs in the oldest population exhibit higher hyperedge node degree across the brain, although regions of relatively high hyperedge node degree are consistent with those in the other populations.

Thus, the increased cardinality is due to global changes, with regions of relative high

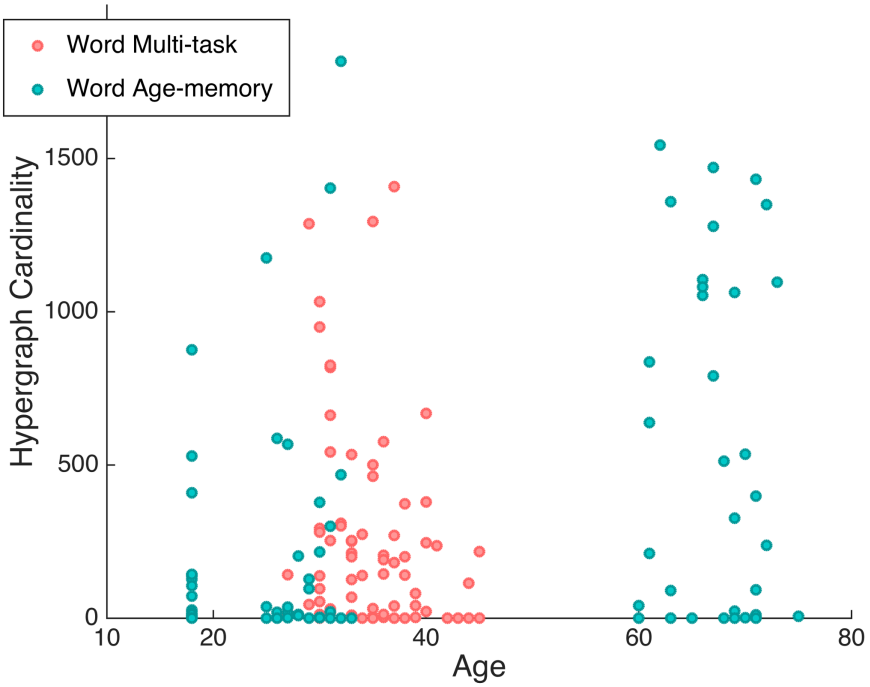

Figure 3.7: **Hypergraph cardinality and age:** Hypergraph cardinality is shown as a function of age for the age-memory data set (blue) and word memory task of the multi-task data set (pink). Three distinct age groups are present for the age-memory data, while the multi-task ages overlap with the middle age-memory group. The correspondence between increasing age and larger hypergraph cardinality can be observed, where few older subjects have low hypergraph cardinalities, but the majority of the youngest subjects have cardinalities lower than 500.

and low degree that are stable across ages. The number of hyperedges that include each region increases by about a factor of five over the age range studied. We conduct a paired t-test for each brain region in the youngest and oldest populations and found that each region has a significantly higher hyperedge node degree in the oldest population.

3.4 Discussion

Improving our understanding of the drivers of individual differences in functional brain imaging data can give insight into the dynamic mechanisms that lead to individual

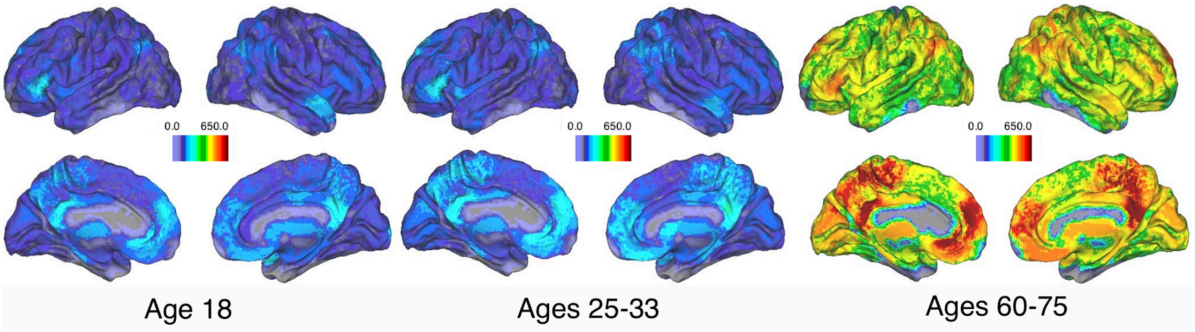

Figure 3.8: **Spatial distribution of hyperedges for three age groups:** Average hyperedge node degree for three discrete age groups in the age-memory study. Regions of relative high node degree are consistent across the three groups, but the overall node degree is about five times larger in the group with ages from 60–75. This corresponds to previous observations of increasing cardinality with age and illustrates how the increase in cardinality is spread across the brain.

behavior. Dynamic FC has been used over groups to explain changes in the brain attributed to individual differences in learning [11, 14, 34]. Hypergraphs in particular have been used to analyze how long-term learning impacts the functional network structure [11] and how the brain switches between cognitive states [111]. A previous DFC study found task-dependent hypergraph properties at the level of the group, indicating that hypergraphs can be used to describe how functional dynamics differ between tasks [111]. Here, we develop new hypergraph metrics to investigate individual differences in hypergraph structure and possible drivers of these variations. Our primary goal in the present investigation is to continue validating the hypergraph approach by demonstrating its ability to reproduce a well-known phenomenon in the FC literature.

Hypergraphs are constructed from correlations between edges, providing a method of analysis complementary to static and dynamic graph theoretic methods including dynamic community detection and ALFF-FC [113, 114]. In this method, hyperedges are defined as connected components of the edge-edge graph. A natural extension of the hyperedge formalism would be to perform edge-based community detection on the edge-edge adjacency matrix, which would further partition the connected components

of the edge-edge graph [60, 61]. Similarly, any graph theoretic measure that can be computed on the standard node-node graph—clustering coefficient, assortativity, global efficiency, *et cetera*—can be computed with respect to the edge-edge graph, although the interpretation would of course be quite different.

The hypergraph method provides a rigourous graph theoretical formalism to study network dynamics. Throughout this study, we investigate hypergraph cardinality as a dependent variable. However, future investigations should be performed to determine whether hypergraph cardinality is a useful independent variable with predictive power.

3.4.1 Disparate Sources of Variability in Hypergraph Structure

As we showed in the Multi-Task Analysis, the hypergraph cardinality varies widely across individuals, but is consistent between task states. Previous work on the multi-task data set found that the probability for hypergraphs to appear in a particular network configuration over individuals was significantly different depending on task state [111]. Consistent spatial organization rules for each task existed at the level of the group. There were similarities in the spatial arrangement of hyperedges in the brain for differing tasks, but certain properties were found to vary significantly between tasks. Brain areas in the occipital lobe in particular were highly likely to participate in the hypergraph across individuals and across tasks, likely due to the visual nature of most of the cognitive tasks studied.

Here, we study hypergraph cardinality, which displays high variability across individuals and consistency across tasks within individuals (Figure 3.2). This indicates that hypergraph cardinality serves as an individual signature of a subject's brain dynamics. The similarities across subjects in the spatial distributions of hypergraphs described in [111] capture information orthogonal to the information summarized by hypergraph

cardinality. For example, there are some individuals for whom the visual brain regions are linked by many hyperedges, and some for whom those same regions are linked by relatively few hyperedges, but these regions are more likely than others to be included in hypergraphs in the majority of subjects. This suggests that, for some subjects, brain regions tend to be more dynamically integrated in general, with co-varying functional relationships across many brain circuits; in other subjects, connectivity dynamics are more fragmented across the brain.

The high degree of variability in hypergraph cardinality across subjects and consistency within subjects, combined with the significant differences in spatial hyperedge arrangement across tasks, indicate that hypergraphs are a useful analysis tool for investigating both individual and task-based differences in brain function in a variety of settings. At the same time, hypergraphs can provide a view of dynamic patterns that complements other commonly used DFC methods. For example, many FC methods exclusively investigate the structure of strong correlations in functional data [13,35,133,134]; hypergraph analysis captures information about both strongly and weakly correlated dynamics and how sets of brain regions transition between them [16].

Although they are highly informative, many of the hypergraph metrics we study here are representative measures that greatly reduce the dimension of the hypergraph and only reveal a small part of the information contained in its structure. Further development of methods to utilize more of the information that hypergraphs provide will allow characterization of the consistency of particular hyperedges and dynamic modes, an understanding of which are important for behavior, or influenced by demographics or disease. Future work is also needed to further quantify the spatial differences in hypergraph arrangement across both individuals and tasks, to clarify the extent of overlap between the two types of information, and to determine whether the individual variability in cardinality can be mapped to individual spatial differences in hypergraph structure.

3.4.2 Relationship Between Age and Changes in DFC Networks

FC studies have established clear trends associated with aging, including a decrease in connectivity within functional networks and an increase in connectivity across different functional networks in resting and task states [106, 135–138]. Many of these studies have considered resting-state FC, because the absence of task stimulus provides a simple and reliable setting for comparison between subjects [139], although recent studies have successfully used FC networks to study various cognitive proceses [140]. The default mode network (DMN) and similar resting-state analyses may miss functional changes evoked by task states; while the DMN FC decreases with age, task-related sensorimotor network FC has been shown to increase with age [103,105]. Similarly, FC in memory tasks shows increased segmentation with age [141]. Extending these analyses to incorporate the dynamics of functional interactions is a necessary step towards quantifying individual changes in functional brain dynamics associated with age.

Several efforts have been made to capture individual age-related differences with methods from dynamic FC. Dynamic community structure and amplitude of low-frequency fluctuation of FC were both found to be strongly correlated with age, illustrating that functional dynamics are closely linked with aging [113, 114]. In the dynamic community detection analysis, functional communities were found to be more fragmented with age, which agrees with the hypergraph cardinality result presented here [113]. A multi-scale community detection analysis uncovered similar fragmentation with age for small scales [142]. Our finding that hypergraph cardinality also increases with age aligns with this result and provides further information based upon its ability to capture higher-order dynamic patterns across larger ensembles of brain regions. Not only do the functional similarities of communities of brain regions themselves become less distinct as humans age, but the temporal profiles of these functional similarities also become less integrated

across brain regions. The agreement of this result with known age-related changes in FC [97–99, 104, 106] demonstrates the ability of hypergraph methods to capture and quantify major brain changes. Moreover, since the hypergraph analysis is not limited to strong correlations, our analysis further suggests that age is related not only to the organization of functional activity in groups of brain regions with strongly coherent activity, but also to the coordination between groups of regions that transition from being strongly to weakly correlated over time (or *vice versa*).

The reported correspondence between age and hypergraph cardinality is significant in the age-memory data set, but our analysis did not include data that could verify this relationship for cognitive tasks other than the word memory task. Although memory is a cognitive ability known to decline with age in many individuals, it is unlikely that the specific task studied in the age-memory data set drives this result. Rather, the consistency of hypergraph cardinality across tasks seen in the multi-task data set in Figure 3.2(B) suggests that similar hypergraph cardinalities may be found during other tasks in data sets with higher age variability, and that the relationship between age and cardinality is unlikely to depend primarily on the behavioral task. Further investigation is needed to determine whether individual differences in hyperedge structure have any significant relationship to behavioral or cognitive performance on any particular task.

3.4.3 Methodological Considerations

Atlas-Based Variations: In this chapter, we use the hybrid atlas described in Section 2.2.4 for all analyses. However, several studies have reported variation in graph topology with differing choices of atlas in both structural [53, 143, 144] and functional imaging studies [145, 146]. It is possible that use of a variety of atlases may produce variation in the results presented here, which could be verified in future work.

Edge Effects in Task Concatenation: In this chapter, we investigate dynamic functional connectivity changes across multiple cognitive tasks and two separate imaging data sets. In order to capture changes across tasks in the multi-task data set, we concatenate the time series for all tasks, as in [111]. In our analysis of the age-memory data, we concatenate time series from three functional runs of the word memory task, and remove time windows from the ends of the time series of each task to reduce edge effects. Edge effects appear to be confined to the data points adjacent to the beginning and end of each run, but we remove the full $N \times N$ adjacency matrix to ensure we are not including any edge effects in the analysis. The resulting change in the cumulative size distribution is depicted in Figure B.1. With the edge blocks removed, there are fewer system-size hyperedges and more small hyperedges.

Figure B.1 includes a comparison with another method for treating edge effects. In this case, the time series data for each of the three tasks is filtered separately before concatenation. This approach dramatically reduces the number of hyperedges. If filtering is responsible for introducing edge effects that drive hyperedges, the number of hyperedges are likely to increase when we employ this method. Instead, only 13 subjects had non-singleton hyperedges. We choose to not analyze these results further because there are too few subjects with hyperedge data.

Two further efforts to understand the effects of concatenating across functional runs on the cumulative size distribution are depicted in Figure B.2. In the trial-by-trial analysis, we performed the hypergraph method separately on each edge time series (10 data points each) for the three trials. Only 30 subjects have significant non-singleton hyperedges in at least one of the three trials and the number of large hyperedges is much lower than the original result. This decrease may be a result of our removal edge effects, but it is likely the shorter task length is driving the difference, as we discuss in the next section. To explicitly investigate the effect on the size distribution caused by each tran-

sition, we also split the time series data into three sets of 18 edge time series data points. The first includes the transition between the first and second trials, the last includes the transition between the second and third trials, and the middle includes both transitions. These distributions are also plotted in Figure B.2. We see that the overall number of hyperedges is greater than both the original age-memory hypergraph over all individuals, which is driven by a decrease in the number of system-size hypergraphs in the 18-split analysis. The distributions for all three follow similar patterns, indicating there is not a large discontinuity in the pattern of the distribution when we include both transitions.

Edge Time Series Length in Hypergraph Construction: When we construct hypergraphs from the much shorter single task measurements within the multi-task data set, the number of large hyperedges is greatly reduced, with fewer hyperedges in the population near the system size (see Panel A of Figure 3.4). We see a similar effect when we compare the distributions seen in Figure B.2 for the split data sets. The trial-by-trial hypergraphs contain fewer hyperedges overall and far fewer system-size hyperedges than the 18-split hypergraphs. However, this increase is not driven by inclusion of the transitions alone, since the middle 18-split hypergraph contains approximately half the number of system-size hyperedges when compared to the full analysis. Since both hypergraphs are constructed across both transitions, this indicates that the edge time series length is more influential to population-level hypergraph properties than concatenation.

Further work is needed to elucidate the relationships between hyperedge size and the overall length and composition of the data set. Additionally, it remains to be determined whether there is an analogue to the scan length proposed for reliable FC estimates [132]; an edge time series length that ensures minimal fluctuations in the size distributions for longer scans. However, the very close correspondence between small-size hyperedges found during the word memory task in both data sets suggests that these hyperedges are capturing important characteristics of the dynamics within this task that are robust

across imaging sessions and populations.

3.5 Conclusion

Here, we have shown that the considerable differences in functional connectivity dynamics across individuals are closely linked with age. The hypergraph method is presented as an analysis tool that captures information about group-level similarities that differ between task states as well as individual differences that are consistent within individuals, across tasks. Further investigation into a single hypergraph metric (hypergraph cardinality) that varies across individuals uncovers a significant relationship between hypergraph cardinality and age. Specifically, there are a greater number of hyperedges in older individuals' hypergraphs, suggesting that there are more small groups of regions with cohesively evolving dynamics and indicating a loss of coherence across larger, spatially distributed intrinsic functional connectivity networks. This complements widely reported relationships between FC and human aging by providing new insight into how FC activity and the co-evolution of FC activity are altered with increasing age, including the loss of large groups of co-evolving brain regions in older individuals. The correspondence with and extension of classic FC results to new dynamic regimes, along with the unique capacity of hypergraphs to probe multiple dimensions of both strong and weak dynamic variability, show that hypergraph analysis is a valuable tool for understanding age-related changes and other individual differences in dynamic brain function.

Chapter 4

Age-dependent Community Structure and Dynamics in Human Brain Function

4.1 Introduction

Humans experience notable changes in cognitive ability and behavior as they age, often in situations involving memory encoding, memory retrieval, and executive control functions [147–151]. Over the past few decades, advances in brain imaging have made it possible to observe and quantify neural changes associated with advanced age. One of the most widely-reported phenomena associated with aging is the loss of segregation between neural systems: many networks become less internally coherent, while at the same time they become more similar to other networks. This result has been reported using a number of methodological approaches, including whole-brain ICA [100], whole-brain parcel-based functional connectivity methods [103–105, 107, 138] as well as similar analyses confined to a subset of systems [106, 152], whole brain voxel-wise analyses [101], and

seed-based methods [153] (for reviews, see [97–99]). Moreover, these changes have been tracked longitudinally within participants [106], have been shown to affect various properties theoretically associated with the efficiency and efficacy of information processing in the brain [154, 155], and have been associated with behavioral effects [106, 154].

Although the dominant change associated with aging is one of decreased intra-network connectivity and increased inter-network connectivity, this pattern varies across networks. The loss of intra-network connectivity is found most consistently in the default mode network (DMN), even among those studies that consider brainwide connectivity [100–106]. Some studies also report similar decreases in networks associated with higher cognitive functions [100, 102, 105–107]. However, other networks consistently show no change, or even an increase in intra-network connectivity, especially those associated with sensory functions [101,103,105]. Similarly, connectivity between the DMN and other networks tends to increase (or, equivalently, the uniqueness of the networks decreases) [104, 106].

In parallel with this line of research on how the brain's functional architecture changes with age, a largely separate effort has sought to extend connectivity methods by accounting for the fact that the brain is not static (for a review, see [109]). To the contrary, this work has demonstrated that patterns of connectivity are quite variable [156], which can be characterized as constituting a series of transitions between fairly well-defined brain states [157]. It has been proposed that the greatest variability occurs in regions that serve to connect fairly well-segregated systems [112], and that a small set of networks may modulate the organization across a large number of others [158]. The time-resolved approach adds yet another dimension for investigating age-related effects; for example, [114] reports increased variability in connectivity across networks including DMN and cerebellum, and decreased variability between those two and within the cingulo-opercular network, as a function of age.

Having established these aging-related changes in functional connectivity—along with some general principles of dynamic connectivity—in the resting state, an obvious next question is how the results differ during task performance. "Task-free" paradigms dominate studies of functional connectivity. Incorporating a task could affect connectivity, including its relationship with age and its dynamics, in a number of ways. For instance, compensatory strategies employed by older—but not younger—adults could drive the connectivity profiles of the two groups even further apart; alternatively, the presence of an extrinsic input could impose structure on the systems that have become homogenized in older adults. Indeed, [159] demonstrates widespread changes in the relationship between age and connectivity across resting and task scans, with the largest effects being a weakening in the age–connectivity relationship during tasks compared with rest. Likewise, connectivity between and within networks could change as participants learn, change strategies, or even simply become fatigued.

For the present study, we used a memory task that incorporated a strong element of cognitive control. In particular, after studying a list of items, participants were presented with the studied items, along with novel (unstudied) items, and instructed to indicate whether each item was studied or not. Items occurred in one of two contexts: a "liberal" context indicating that each item in that context was likely to have been studied (70% of items were studied items) or a "conservative" context indicating that each item was unlikely to have been studied (30% of items studied). In the face of imperfect memory evidence, participants must exert cognitive control—adjusting the criterion they use to endorse an item as studied—in order to perform well on this task. Given that the domains of memory and cognitive control are fundamental in human cognition, and are associated with changes over the lifespan [160], this task is an appealing choice for studying how the brain's architecture changes with age when *not* at rest. Previous results with this task revealed wide individual differences in adaptability [42], and implicated a network

of regions including lateral prefrontal and lateral posterior parietal cortex in performing this task [161].

Although the brain regions associated with the performance of this task are well documented, these results are derived from the standard mass-univariate GLM analysis of BOLD data, and therefore give little basis for predictions in terms of network-level dynamics. In fact, by definition, these existing results assume stationarity and consider each voxel as independent. Even results derived from methods that explicitly model the spatiotemporal nature of brain activity (e.g., ICA) would require a theoretical framework in order to define regions of interest in the context of how network dynamics relate to other factors, such as age. Thus, there remains a gap in understanding of the neural processes related to performance of this task on the level of dynamic interactions between large-scale brain regions and networks. Our current understanding of these processes, based on existing theories and results, is specified on a very different level from the target of our current investigation. Our goal in this chapter is to apply a data-driven analysis method to investigate the dynamics of these regions and networks, which allows us to uncover age-related changes at scales at which it is difficult to make specific hypotheses based upon existing literature.

We apply a dynamic community detection method to quantify several higher-order aspects of task-based functional connectivity and their dependence on age. This method and other network science approaches have proved successful in distilling the information in fMRI data into intuitive, descriptive, and predictive network characteristics [14, 33, 55, 111, 162–164]. While previous results suggest that static community structure will meaningfully differ on a group level between older and younger adults at rest [137], we ask whether the dynamic changes in these communities are affected by age during task-based cognition, and how such effects vary across individual participants. We quantify the size and number of functional brain communities, the degree to which brain regions

flexibly switch between communities, and the association of the community structure with known intrinsic functional connectivity networks or systems, in order to determine whether these systems are differentially involved in age-related changes.

4.2 Materials and Methods

4.2.1 Participants

126 participants were recruited from the UCSB and Santa Barbara communities and scanned at the UCSB Brain Imaging Center. 22 subjects were not included in this analysis due to technical issues, metal screening issues, claustrophobia, and attrition. The 104 participants assessed here came from three separate age groups: 35 adolescents (age 18, 18 female), 34 young adults (ages 25-33, mean age 28.5, 16 female), and 35 older adults (ages 60-75, mean age 67.2, 18 female). All subjects had a history of normal memory ability for their age, and a Mini-Mental State Examination score of 27 or above [127] [165]. All subjects gave informed written consent prior to experimental procedures and were paid for their participation. All procedures were approved by the University of California, Santa Barbara Human Subjects Committee.

4.2.2 Stimuli and Procedure

Subjects performed a recognition memory task designed to test memory for words and to measure how participants strategically use probabilistic information as a supplemental guide to memory [42]. During a study session (which occurred inside the scanner, immediately before beginning scanning), subjects were asked to memorize a list of 153 common English words. Subjects were then scanned during three consecutive test sessions, each consisting of 102 trials (each spanning a single TR) in which subjects were shown a word

and asked to indicate whether they had seen that word in the study session, interspersed with 214 blank jitter TRs.

The test session word stimulus in each trial was presented with probabilistic information about the correct response; words of one color (blue or orange, counterbalanced across subjects) had a 70% probability of having been presented in the study session, and words of the other color a 30% probability. These probability contexts were presented in a blocked fashion, such that the probability context changed every 5–7 trials. Half of the trials in each functional run (51 words) were studied, while the other half were unstudied. A schematic of the task design is shown in Figure 4.1 [42]. For more information on the details of the procedure, see ref. [127].

4.2.3 Imaging Acquisition and Preprocessing

Subjects were scanned with a 3T Siemens TIM Trio MRI system with a standard 12-channel head coil. Functional data were collected with a T2*-weighted echo-planar sequence (30 interleaved slices, 3 mm thickness, 3×3 mm in-plane resolution; TR = 1.6 s; TE = 30 ms; FA = 90) with generalized autocalibrating partially parallel acquisitions (GRAPPA). A high-resolution anatomical image was collected at the beginning of the scanning session for each participant using an MPRAGE sequence (TR = 2.3s; TE = 2.98 ms; FA = 9°; 160 slices; 1.1 mm thickness). Additionally, diffusion-tensor imaging and resting state fMRI scans were acquired but are not considered further here.

The data were preprocessed using FEAT v6.0, part of FSL [47]. Preprocessing included motion correction with MCFLIRT, non-brain removal with BET, spatial smoothing (FWHM = 5mm), high-pass temporal filtering ($\sigma = 50$s), and grand mean intensity normalization. The mean relative motion across all TRs, averaged across functional runs, was also recorded for each subject. It has previously been established that motion

varies reliably with age [127], so all subsequent analyses are conducted with mean motion partialed out.

The data were then processed further using a nuisance regression with the following regressors: the six relative motion correction terms returned by MCFLIRT, their temporal derivatives, and the mean signal timecourse from cerebrospinal fluid (generated by segmenting the high resolution T1 image, thresholding the CSF probability image at 0.9, and taking an average over all in-mask voxels). The model also included regressors for each probability context block, which were modeled as a boxcar of duration equal to the context, convolved with an HRF (gamma model, phase = 0s, standard deviation = 3s, mean lag = 6s), plus temporal derivatives. To generate the final denoised data, we took the residuals of this model with respect only to the motion and CSF regressors. Finally, the denoised data were registered to MNI space using FLIRT. First, the high resolution T1 image was registered to the MNI template (12 *df* affine transformation), then the functional data were registered with the high resolution image (6 *df* affine transformation, trilinear interpolation), and the transformations were combined.

4.2.4 Creating Dynamic Brain Networks

In order to investigate the large-scale network structure of brain activity, a dynamic network was constructed separately from each subject's measured functional activity. Each network is defined as containing n nodes, treated as constant over time. Here these nodes consist of the $n = 194$ regions of a "hybrid" anatomical atlas, an adaptation of the multi-resolution Lausanne2008 atlas minimizing cross-brain and inter-subject variability in region size [44, 111]. This atlas was registered to MNI space. Node-specific time series from each functional run were generated for each node by averaging the time series of all voxels within the node [111].

Each network has $e = n(n-1)/2$ edges, each with a real-valued, non-negative connection weight that may change dynamically over time, taking on a new value in each of T sequential time windows spanning the experiment. The weight of an edge between nodes i and j in a given time window l, denoted A_{ijl}, is defined as the mean low-frequency (0.06-0.125 Hz) wavelet coherence between the BOLD time series of i and j within time window t [166]. Edge weights are always valued in $[0, 1]$. In this study, we investigate two separate time window sizes. We focus primarily on windows containing 52 time samples (18 windows in total), with each window representing approximately 80 seconds or 1.3 minutes. We also analyze for comparison the results from more temporally coarse-grained windows containing 316 time samples (3 windows in total), with each window representing approximately 500 seconds or 8.4 minutes, the length of a single functional run of the experiment. Note that these window sizes are both significantly longer than individual trials (each of which contains one word stimulus and one decision, and lasts for approximately 2 seconds), as well as blocks of trials belonging to each of the two probability conditions (which contain about six trials each and are approximately 12 seconds long); we make this choice in order to ensure that each window contains sufficient time sampling statistics to provide a reliable estimate of the coherence or edge weight within that window [14, 33, 111].

4.2.5 Detecting Dynamic Community Structure

To study the time-evolving modular structure of these networks, we identify distinct communities, or sets of brain regions with strong intra-set functional coherence, and quantify how these communities change over time. For each subject's dynamic network, a community partition is determined by maximizing the multislice modularity Q, a function indicating the quality of the modular structure of a given partition of the network,

in comparison to that expected of a randomized "null" network [15]. The multislice modularity of a network partition is given by

$$Q = \frac{1}{2\mu} \sum_{ijlr} \left\{ \left(A_{ijl} - \gamma_l P_{ijl} \right) \delta_{lr} + \omega_{jlr} \delta_{ij} \right\} \delta \left(g_{il}, g_{jr} \right). \tag{4.1}$$

Here, the Kronecker delta $\delta \left(g_{il}, g_{jr} \right)$ is equal to 1 when the community assignment of node i in window l (g_{il}) is the same as the community assignment of node j in window r (g_{jr}); otherwise, it takes a value of 0. A_{ijl} is the edge weight between nodes i and j in time window l, as defined above; P_{ijl} is the corresponding edge weight in the null network, with a spatial resolution factor γ_l determining the relative weight of the null model within in each time window (see Appendix C for details). Thus, the first term in brackets provides a positive contribution to Q, for each pair of nodes assigned the same community in the same time window, proportional to the difference between the actual edge weight between the pair and that in the weighted null model. The second term in the brackets includes a time resolution factor ω_{jlr} for each node j and each pair of time windows l and r (see Appendix C for details). This term provides a positive contribution of ω to Q, for each node j and each pair of time windows l and r, when j is assigned to the same community in both time windows. μ is a normalizing factor given by $\mu = \frac{1}{2} \sum_{jl} \kappa_{jl}$, where $\kappa_{jl} = c_{jl} + k_{jl}$, $c_{jl} = \sum_r \omega_{jlr} = \omega T$, and $k_{jl} = \sum_i A_{ijl}$, or the weighted degree of node j in time window l [15].

This multislice modularity is larger for community partitions that group together nodes with comparatively strong pairwise edge weights (as compared to the null network) within each window, and that group more nodes in the same community as themselves across multiple time windows. In this study we use the Newman-Girvan null model, in which $P_{ijl} = \frac{k_{il} k_{jl}}{2m_l}$ and $2m_l = \sum_{ij} A_{ijl}$; this commonly used choice of null model treats the measured edge weights as randomly distributed within each window while

preserving the node degree distribution. We maximize Q over partitions with a Louvain-like locally greedy algorithm [15, 167]. Due to the stochasticity of the algorithm and the expected high degeneracy of solutions near the maximum value of Q, we use a community consensus procedure to distill a statistically representative partition from an ensemble of 100 solutions [35].

4.2.6 Brain Network Community Structure Diagnostics

Basic community structure

We use several measures to quantitatively describe the dynamic community structure of each network, and to compare subjects' networks to each other.

The first set of metrics involves the number of distinct communities in a subject's brain network. The number of *dynamic communities* is evaluated over the entire dynamic network, and counts each community that appears in at least one time window. Communities which stretch over several time windows, but are associated together under the same community label, are counted as only one dynamic community. (Note that the community detection algorithm automatically identifies communities in each window with those that have similar membership in other time windows, and assigns them the same label; this self-identification is enforced to an extent controlled by the strength of the inter-window coupling parameter ω). In contrast, the number of *static communities* is evaluated within each time window separately, counting each distinct, dynamically detected community appearing within that time window once, regardless of whether that community spans multiple time windows. Because communities may appear or disappear between windows, the dynamic community number and the static community numbers need not be equivalent.

The *flexibility* f of a node i within a network is defined as the number of times that

node switches communities between all distinct pairs of time windows, normalized by the total possible number of switches:

$$f(i) = \frac{1}{T(T-1)} \sum_{t \neq t'} \left[1 - \delta(g_{it}, g_{it'})\right]. \tag{4.2}$$

Here, t and t' both run from 1 to T, the total number of time windows; $\delta(g_{it}, g_{it'})$ equals 1 if node i is assigned to the same community in time window t and time window t', and 0 otherwise. A node with high flexibility changes communities in every or nearly every time window and has a flexibility at or near 1, while a node with low flexibility may remain in the same community in all windows and have a flexibility of 0. We further define the *community flexibility* as the mean flexibility of all nodes in a particular community.

This method of computing "categorical" flexibility compares nodes between every possible pair of time windows, in contrast to "time-ordered" flexibility, which compares only time-adjacent windows. While many applications of categorical flexibility are used to compare communities across categories or tasks, we use categorical flexibility here to emphasize the consistency of nodes across long time windows which are statistically identical in terms of task design (for 500-second windows) or nearly so (for 80-second windows), without an assumed change in brain dynamics over time in the experiment. All results reported in this chapter are essentially unchanged when using time-ordered flexibility (see Appendix C).

Comparing communities to functional systems

To understand how the community structure of this data corresponds to known functional brain systems, we compare the community partitions to a basic *functional system partition* of the nodes. Based on the primary functional roles of different anatomical brain areas as reported in the literature, and as detailed in [168] and [33], each node

is assigned to one of ten functional systems: auditory, cingulo-opercular, default mode, dorsal attention, fronto-parietal, somatosensory, subcortical, ventral attention, visual, and other. These systems have been distilled using a network-based clustering approach [169] and used to describe and quantify system-specific functional interactions in the brain [24, 26, 170]. The relationship of these functional systems to the community structure is described by the following quantitative metrics.

The *recruitment coefficient* of a given node is a measure of the consistency with which that node is grouped in the same community as other nodes within its own functional brain system. It is given by

$$R(i) = \frac{1}{n(s_i) - 1} \sum_{j \neq i} P(i,j)\delta(s_i, s_j), \tag{4.3}$$

where $\delta(s_i, s_j)$ equals 1 if the system of node i (denoted s_i) and the system of node j (denoted s_j) are the same, and 0 otherwise; $n(s_i) = \sum_j \delta(s_i, s_j)$, or the number of nodes in system s_i; and $P(i,j)$ is the frequency with which node i and node j are grouped in the same community [33, 168]. Specifically, $P(i,j)$ is computed as the observed proportion of instances (i.e. time windows or modularity-optimizing partitions) in which i and j are placed in the same community.

We further define the *self-recruitment* Ψ of a given system S as the average recruitment coefficient of all nodes in the system, given by

$$\Psi(S) = \frac{1}{n(S)(n(S) - 1)} \sum_{ij, i \neq j} P(i,j)\delta(s_i, S)\delta(s_j, S). \tag{4.4}$$

This measures the extent to which nodes in system S are cohesively grouped together in the same community [33, 168].

Correcting for Mean Relative Motion

As mentioned above, subject age is correlated with mean relative motion in these data (Spearman's $\rho = 0.48$, $p = 1.80 \times 10^{-6}$), as found previously in ref. [127]. Thus, all subject-wise correlations presented here are performed with mean relative subject motion partialed out – i.e., each correlation variable was first regressed separately on mean relative motion, and we assessed the correlation between the residuals of these regressions, to ascertain the extent of their relationship that could not be explained by motion. Since subject age is significantly correlated with mean relative motion, it is possible that motion also affects the correspondence measures of community dynamics and age, and potentially other performance and demographic measures as well, due to the broad and non-uniform distribution of ages in our sample. Appendix C provides further details on the results of this correction.

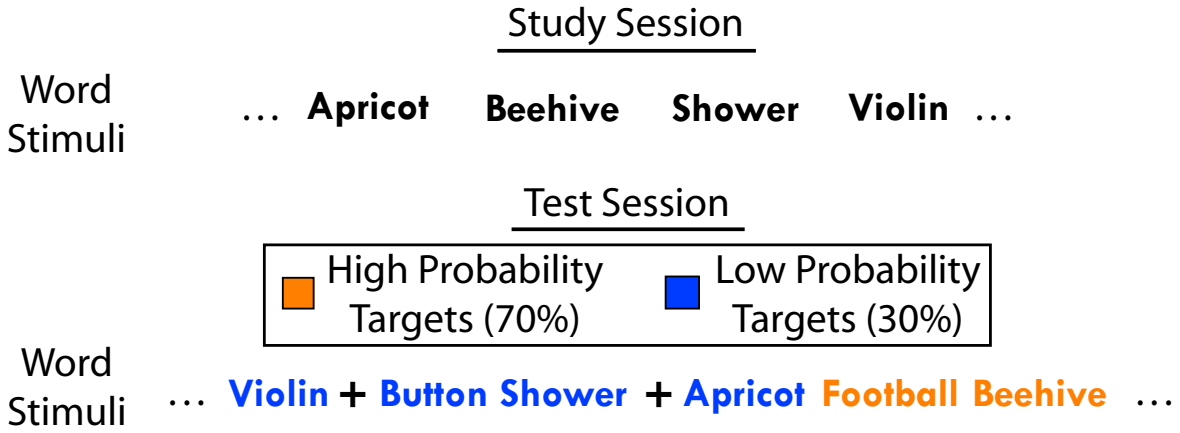

Figure 4.1: **Memory task structure.** A schematic depiction of the word memory task design. During a study session, subjects study a list of common words. They are then scanned during the study session, in which they are shown a new set of words and asked to decide whether each word has been seen before in the study session. Words in the test session are colored to provide probabilistic information about their likelihood of having been seen in the study session.

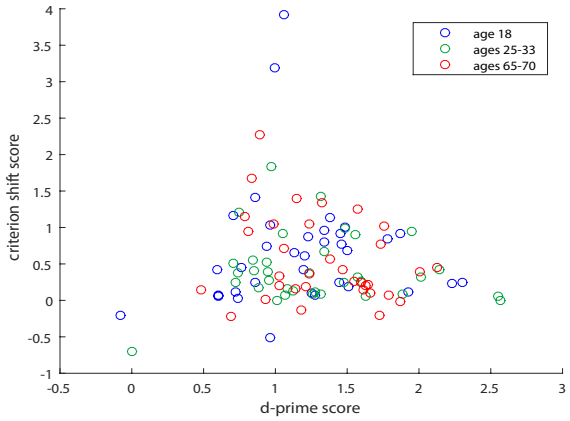

Figure 4.2: **Memory task performance measures.** Distributions of d-prime and criterion shift scores for $N = 104$ subjects. These measures characterize overall accuracy and the extent to which subjects switched strategies between probability contexts, respectively. Subjects are colored by age: blue indicates adolescents (age 18), green young adults (ages 25-33), and red older adults (ages 60-75). There is no apparent correlation between these measures (Pearson's $r = -0.060$, $p > 0.1$).

4.3 Results

In this section, we present the characteristics of dynamic community structure within individuals, and evaluate their correspondence with age and recognition memory performance.

For assessing correlations with age throughout this section, we use the Spearman rank correlation, due to the non-continuity and non-uniformity of the ages in our subject sample. However, we use the Pearson correlation for assessing correlations with all performance measures, which are continuously and approximately normally distributed. We partial out mean relative motion from all subject-wise correlations, as discussed in

Section 4.2.6, and apply a Bonferroni correction for multiple comparisons on the set of overall and system-specific correlations for each pair of measures (e.g. age and flexibility, age and community number, criterion shift score and flexibility, etc.).

4.3.1 Word memory performance

We examine two behavioral measures of performance on the word memory task: d-prime, an indicator of overall accuracy on all memory trials; and criterion shift score, which describes the extent to which subjects change their response strategies in the face of probabilistic information about the correct responses [42]. A more positive criterion shift score indicates that the subject made a comparatively large shift from a liberal to a conservative strategy when responding to high- and low-probability targets, respectively. A more negative criterion shift score indicates the opposite strategy shift (from conservative responses on high-probability targets to liberal responses on low-probability targets). Very few subjects displayed this objectively worse strategy. A criterion shift score of 0 indicates no strategy difference between high- and low-probability targets. We find that among all subjects in this study, the d-prime and criterion shift scores are approximately normally distributed, and they are not significantly correlated with each other (Figure 4.2). In addition, neither measure shows any significant correlation with subject age [127]. Upon exclusion of two apparent outliers in Figure 2, these results and the significance of other task performance correlations reported in this chapter are not affected (see Appendix C for details).

4.3.2 Functional communities in the brain

We focus primarily on the dynamic community structure of individual functional brain networks composed of time windows containing 52 time samples each. The ex-

periment contains a total of 18 such time windows; each window represents functional connectivity within an approximately 80-second period. Community dynamics on time scales between 60 and 200 seconds provide relatively fine time resolution while retaining sufficient time sampling statistics within each window, and have been shown to contain relevant information about brain function in previous studies [14,33,111]. The 80-second windows used here are significantly longer than individual trials (approximately 2 seconds each) or blocks of trials sharing the same probability condition (approximately 12 seconds each). While they cannot resolve functional dynamics related to a specific word or probability condition, these windows are expected to capture the cognitive control and memory processes that are active over the course of several strategy blocks in the task.

For comparison, we also investigate the community dynamics of networks composed of much longer time windows, each containing 316 time samples and corresponding to one functional run of the experiment. The time windows in these networks capture dynamics over a longer time scale, with each window representing approximately 500 seconds or 8.4 minutes of brain activity, and a total of only three windows across the experiment. The results in these 500-second networks are in general qualitatively similar to results from 80-second networks, although the correspondences between demographics or performance measures and community dynamics are often much weaker. This suggests that shorter (80-second) time windows resolve the relevant dynamics better than longer windows. The shorter time windows are also somewhat closer in length to the timescales of cognitive function demanded by the task setup (although still not identical or aligned with blocks of probability context or other specific task features). Therefore, in this chapter we focus on networks with 80-second time windows, except where explicitly noted. Results from 500-second time windows are presented in Appendix C.

Number and size of communities

Figure 4.3 shows the number of static communities identified in each 80-second time window of the functional brain networks during the memory task in each subject. (Here "static communities" refers to the number of distinct, dynamically detected communities present in a single given time window, while "dynamic communities" counts the total number of communities identified by dynamic community detection in all time windows.) For almost all subjects, the number of static communities remains fairly stable across time windows, and the number of dynamic communities is only slightly larger than the number in any one window. This indicates that few communities appear or disappear, and that community number is a measure with more meaningful individual, inter-subject differences than time-dependent intra-subject differences.

Each subject has between 7 and 24 total dynamic communities (mean = 12.7). The correlations with community number in the remainder of this chapter use each subject's dynamics community number as the primary measure, since it corresponds closely with static community number throughout the experiment.

Note that in Figure 4.3, two subjects in particular appear to have much higher static community numbers than the others, and are potential outliers in this regard. Nearly all results pertaining to the correspondence between age, performance, and community structure reported herein remain unchanged when these subjects are removed from the analysis. One discrepancy is noted below and further details are provided in Appendix C.

Figures 4.4A and 4.4C show the distribution of community sizes for each individual, both within individual time windows and across the entire experiment. Community sizes are relatively uniformly distributed, save for an excess of communities of very small size. Results with single-node communities, or "singletons," excluded from the analysis do

not differ substantially from those reported here; see Appendix C for further details on singletons.

Flexibility

Figure 4.5A shows the flexibility of each of the $n = 194$ brain regions, averaged over $N = 104$ subjects, for networks with 80-second time windows. Regions in the occipital lobe, most of which are in visual cortex, tend to show relatively low flexibilities, as do some motor-associated regions in the dorsal anterior frontal and posterior parietal lobes. Most other brain regions have a somewhat higher flexibility. Consistent with previous work [14], we find notably greater variation in flexibility across subjects than across brain regions.

We find that the regions with lower mean flexibility across subjects tend to have a higher cross-subject variance, as shown in Figure 4.5B; in other words, highly flexible nodes are very consistently flexible across subjects, while nodes with lower mean flexibility (such as those in visual and motor cortex) show greater individual differences in dynamics.

This effect differs strikingly from the flexibility patterns seen on longer timescales, in networks composed of 500-second time windows. The identities of the brain regions with the lowest mean flexibility and the variance of those regions are very similar with both 80-second and 500-second time windows. However, with 500-second time windows, the cross-subject variance of the high-flexibility (non-visual, non-motor) regions is much higher than that of the same regions in networks with 80-second time windows, and consistently higher than the variance of low-flexibility regions as well (see Appendix C for further discussion).

Dependence of community structure on overall brain connectivity

In our sample, we find considerable variation in the density of subjects' functional networks, computed as the sum of all functional connectivity weights between brain region pairs. That is, some subjects have higher overall brain connectivity or coherence than others. In addition, we find that this overall connectivity (OC) is significantly correlated with subject age (Spearman's $\rho = -0.50$, $p < 0.001$).

To ensure that the community structures we identify are not primarily driven by OC alone, but instead capture underlying dynamics in functional connectivity, we construct a null model in which we destroy the underlying connectivity structure in each subject by redistributing network edge weights among region pairs uniformly at random, preserving only the symmetry of the edge matrix, the lack of self-edges, and the total sum of edge weights (i.e., overall connectivity) in each time window of each subject's network. Note that this null model also destroys the inherently constrained structure of the coherence matrix, such that the randomly permuted matrices are not necessarily examples of coherence matrices, as the original networks are. This null model thus cannot speak to the inherent effect of coherence structure on community structure, but can only elucidate density effects.

We create 100 randomly permuted null networks for each subject, and analyze the communities identified therein. As shown in Figure 4.6, the null distributions of total community number are relatively uniform across subjects, save the two subjects with the highest OC, who display consistently lower community number across null networks. A nonlinear relationship between OC and the number of communities identified is evident here–in contrast to this behavior at high OC, the number of communities increases with increased OC at the low-OC end of the distribution.

For most subjects, the number of communities identified in Figure 4.6 is smaller

and differs fundamentally from the number identified in randomized networks. There is no significant correlation between a subject's number of communities and the mean number of communities identified in the corresponding networks. However, if the two outlier subjects with high OC are excluded, the number of communities in null models preserving only by OC does significantly correspond to the mean number of communities in subjects' brain networks (Pearson's $r = 0.32$, $p < 0.001$).

The distributions of community sizes for each subject, shown in Figures 4.4A and 4.4C, are in general relatively uniform, save for elevated numbers of communities of small sizes. Null networks (Figures 4.4B and 4.4D) show very different community size distributions, which include a clear peak at intermediate sizes and much smaller maximum community sizes. These stark differences show that while some aspects of the community structure are related to OC, others are driven by characteristics of the underlying connectivity structure that cannot be explained by OC alone.

4.3.3 Relationship of functional community structure to age and performance

Analysis of community number and flexibility distributions reveals that both measures vary across subjects notably more than across time windows or brain regions within individual subjects. We investigate these individual differences in community number and flexibility, and whether they are related to age or performance, by examining the *total community number* in each subject's entire time-dependent functional network, as well as the *whole-brain flexibility* of each subject, or the mean flexibility over all that subject's nodes. Here we summarize the results of these comparisons, which are also presented in Table 4.1.

We find that total community number is significantly positively correlated with sub-

ject age (Spearman's $\rho = 0.29$, $p < 0.05$). This indicates that cohesive functional communities in the brains of older subjects tend to be more fragmented than those in younger subjects. A significant correlation between community number and age (Spearman's $\rho = 0.31$, $p < 0.05$) is also seen on average in null networks that preserve OC but randomize other topological/spatial network structure. Six out of 100 instances of randomized networks have a stronger Spearman correlation between age and total number of randomized communities than the correlation between age and number of communities reported above (Figure 4.7).

We find that whole-brain flexibility is also significantly positively correlated with age, as shown in Figure 4.8 (Spearman's $\rho = 0.53$, $p < 0.001$). This indicates that younger subjects have brain regions that switch between communities significantly less frequently, and thus more stable community partitions over the course of the experiment.

We find no significant correlations between task performance metrics – either d-prime or criterion shift score – and any of the three metrics of community dynamics, including flexibility, number of communities, and recruitment. This holds true for global brain metrics and for those localized to specific functional systems. In addition, we conduct a multivariate regression analysis to test whether task performance is predictive of brain metrics in individual brain regions. This analysis uses the flexibility scores of the 194 nodes in all subjects as outcomes, and the two performance metrics – d-prime and criterion shift score – as predictors. We also include mean relative motion as a predictor, to ensure it is accounted for as in previous analyses. To test the significance of the fit, we use a permutation null model in which we shuffle the d-prime and criterion shift scores uniformly at random (separately for each measure), perform the fit again 1000 times, and compare the original beta value for each node to the null distribution of beta values produced by these fits to permuted data. When uncorrected for multiple comparisons, 20 non-zero d-prime beta values and 20 non-zero criterion shift score beta values are

significantly different from random ($p < 0.05$). After a false discovery rate correction for multiple comparisons, only a single beta, with an original fit value of zero, is significantly different from its null distribution. This suggests that there is no significant correspondence between the performance measures from this task and the brain metrics we investigate here, even on a node-by-node basis.

To test the possibility that interactions between age and task performance are predictive of neural dynamics measures, we performed a multiple regression analysis including the effects of age, d-prime and criterion shift scores, as well as the interaction between age and each of the two performance scores. We also included mean relative motion as a predictor to ensure it was accounted for. However, we found that none of the interaction terms between age and performance – nor, indeed, any other terms save age itself and mean relative motion – had any significant influence on flexibility, number of communities, or recruitment.

Interdependence of community measures

We observe a very strong correlation between the number of communities in a subject's brain network and that subject's whole-brain flexibility (Pearson's $r = 0.65$, $p < 0.001$). The correspondences with age for number of communities and flexibility are likely also related and may in fact be different measures of what is fundamentally the same phenomenon. For example, consider two separate communities in an older subject, which still have mutually coherent activity and recruit from the same set of brain regions. These brain regions may flexibly switch allegiances between the two communities during the experiment if they could be nearly equally well associated with either community. However, in a younger subject with stronger overall functional associations, these regions would be more likely to all be grouped into a single community throughout the experiment, and thus display far lower flexibility, stemming from the smaller community

number.

To understand the dependence of flexibility on community number, we construct a null model in which we shuffle the community assignments of nodes in each subject's brain network uniformly at random. This preserves the number and size of communities in each brain network while destroying other structure that may be contributing to the flexibility. We re-compute flexibility in each of 100 null-model community structures for each subject.

We consider a null model in which community assignments were randomized individually within each time window of each subject's network. This destroys spatial/topological community structure as well as the continuity of communities over time. Figure 4.9 shows the whole-brain flexibility of each subject (ordered) as well as the corresponding null distribution of flexibilities for 100 community-number-preserving randomized community structures for the same subject. Clearly, preserving the mere number of communities separately in each network time window produces much higher flexibilities for all subjects, and does so very consistently, with a low variance among the 100 random instance of the network. Indeed, there are much greater differences in null flexibility across subjects than across random instances within a single subject.

When community structure information other than the number and size of communities is destroyed, we still see significant correlations between subject age and the flexibility computed from the shuffled community structure. However, these null-model correspondences are notably less strong than the correlation between age and flexibility, which contains information on flexibility beyond that explained by mere community number and size distributions (Figure 4.10).

Although the measure of flexibility is not completely explained by community number in this case, there is still a clear and strong correlation between a subject's whole-brain flexibility and the mean flexibility of regions in the corresponding null model (Pearson's

$r = 0.76, p < 0.001$). This indicates that the information contained in community size and number distributions alone does predict relative subject flexibility quite well.

4.3.4 Community Organization and Functional Circuits

Having examined the dynamic community structure of individual functional networks largely on its own, in a data-driven manner, we aim to further understand and quantify how this structure corresponds to known functional systems in the brain. Figure 4.11 shows the locations (A) and flexibilities (B) of the ten functional systems considered: auditory (AU), cingulo-opercular (CO), default mode (DM), dorsal attention (DA), fronto-parietal (FP), other (OT), somatosensory (SM), subcortical (SC), ventral attention (VA), and visual (VS). Consistent with Figure 4.5A, the visual system is the least flexible, followed by the somatosensory. The high inter-subject variance in ventral attention regions likely reflects the relatively small size of that system (4 brain regions).

In addition to total community number and whole-brain flexibility, we examine whether the relationships between community structure and age differ across these specific functional systems, as visualized in Figure 4.11A and described in Section 4.2.6. We find that the number of distinct communities into which regions of each individual functional system are grouped is significantly positively correlated with age, for all ten functional systems. Mean flexibility and age are positively correlated in all ten functional systems, with all correlations significant ($p < 0.05$) except in the visual system. The visual regions have the lowest mean flexibility overall (Figure 4.11B) and the highest variance in flexibility across subjects. However, when mean relative motion is partialed out of the correlation, they also have the weakest relationship between system-wide flexibility and age.

We further examine the correspondence between functional communities and known

functional systems using the *recruitment coefficient*, a measure of how cohesively regions from the same functional system are grouped together. Figure 4.11C shows the recruitment coefficient of each region (Eq. 4.3), and 4.11D shows the self-recruitment of each entire system (Eq. 4.4). Again consistent with our flexibility findings, as well as previous reports of recruitment in the literature [33], the visual and somatosensory systems have the highest self-recruitment, indicating that they are the systems most consistently grouped together in communities across time windows.

Whole-brain recruitment, or the average of region recruitment over all brain regions, is significantly anticorrelated with subject age (Spearman's $\rho = -0.32$, $p < 0.05$). However, we find that system-specific self-recruitment is affected differently by age in different circuits. System recruitment is significantly anticorrelated with age only in cingulo-opercular, somatosensory, subcortical, and ventral attention regions, but no correlation is apparent in other regions, such as the visual system (see Figures 4.13 and 4.14).

Dependence of recruitment on community size distributions.

To ensure that the recruitment values reported here are not driven primarily by the size and number of communities detected, we again use a null model that permutes the community assignments of nodes within each subject's network uniformly at random, but preserves community size and number distributions as well as time continuity. We compute the mean recruitment over all nodes in each subject's brain in each of these random null networks, as well as the recruitment of each functional node system. Figure 4.15 summarizes the results. All subjects have mean whole-brain recruitment values significantly higher than those expected from networks with identical community size distributions but no other structure, indicating that the association of algorithmically identified communities with known functional systems is significantly greater than random. Within individual functional systems, results vary. Some systems, including sensory

and motor cortices (auditory, somatosensory, and visual) and subcortical structures, are consistently associated with identified communities at a rate significantly greater than random. Others, including systems identified with executive control, both focused and bottom-up attention, and the resting state, have recruitment values that could reasonably be explained by chance in several subjects (i.e., similar values were found in randomized community structures that share only community size and number distributions with the corresponding human brain community structures).

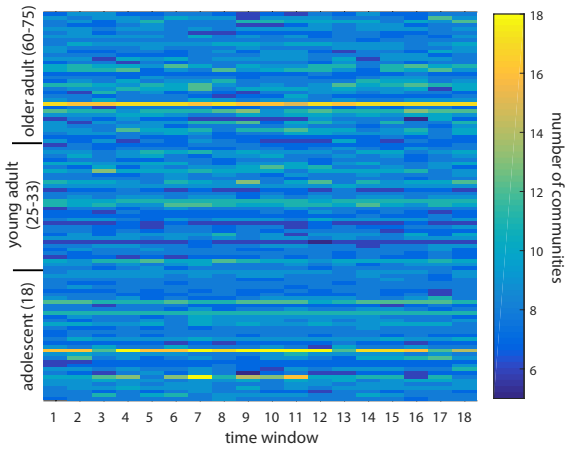

Figure 4.3: **Number of communities.** Color indicates the number of communities detected within each 80-second time window in each subject. Subjects (on the vertical axis) are ordered by age. Notable individual differences exist between subjects, but community number changes comparatively little over the course of the experiment within individual subjects.

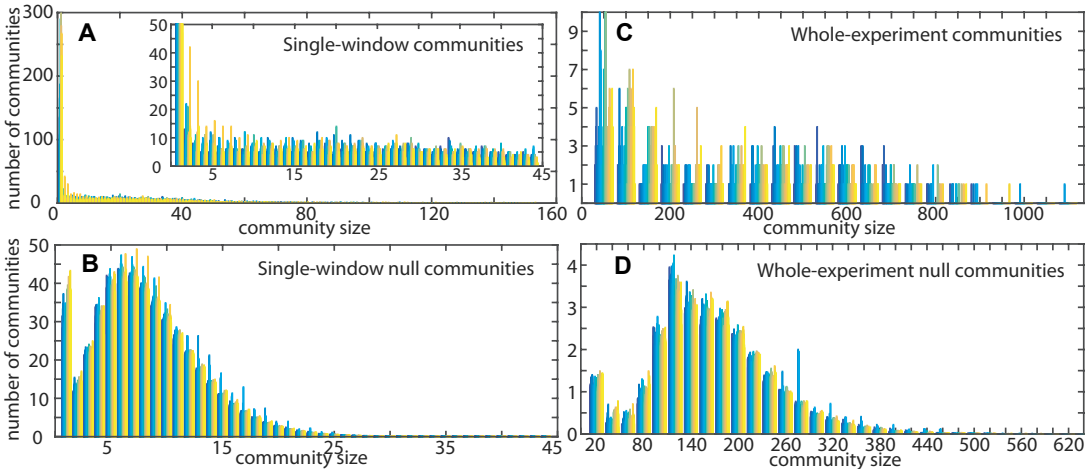

Figure 4.4: **Community size distributions.** A and B show histograms of community sizes within individual time windows, for the observed data and averaged over 100 null networks, respectively. Values are plotted individually for each subject, each represented by one color. The inset shows the data in A restricted to the same axes as B for comparison. C and D show histograms of the sizes of dynamic communities across the whole experiment, also comparing observed data (C) to an average over 100 null networks (D). Community sizes tend to be larger at maximum and to be distributed much more evenly in functional brain networks than in randomized null models.

4.4 Discussion

These findings relating functional community dynamics to age provide important insight into factors affecting the significant individual differences in community dynamics. The community structure appears to act as a signature of individual functional dynamics that is strongly influenced by age, indicating that cognitive organization during such a memory task differs across the lifespan of participants.

Interestingly, despite marked differences in community dynamics, we find no significant correspondence between community structure measures and performance on the memory task, and no age-related differences in memory performance or strategy were detected in this experiment. This is likely related to the choice to study only healthy adults with no measured deficits in cognitive function. It may also be partly explained

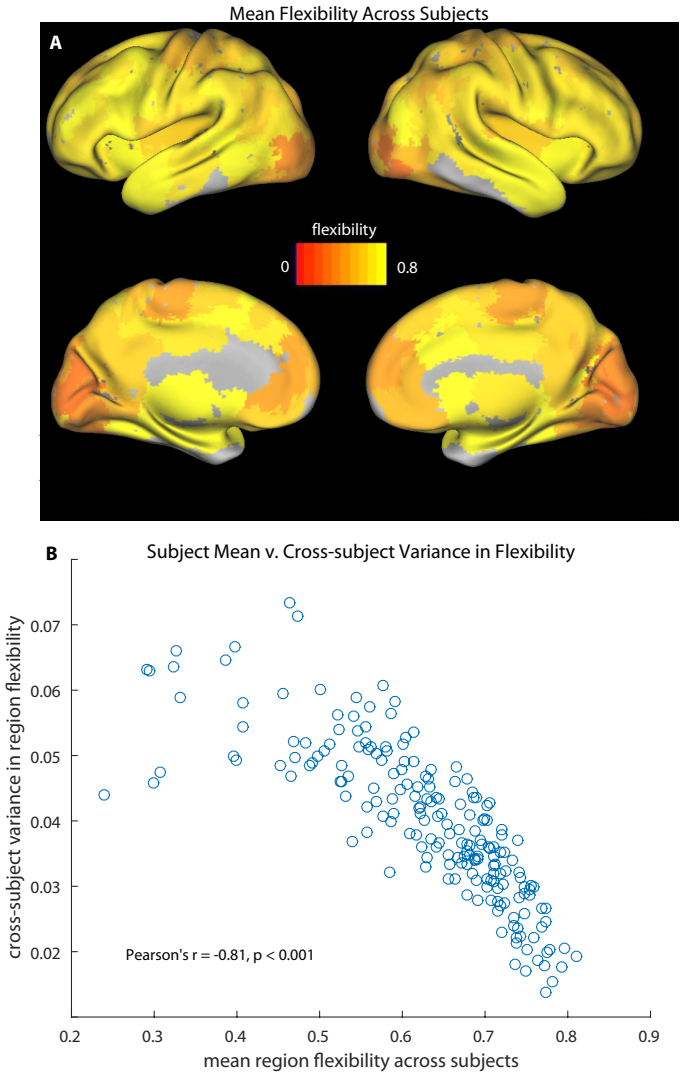

Figure 4.5: **Brain region flexibility.** A: Flexibility of the 194 brain regions used as network nodes (mean over $N = 104$ subjects). Visual cortex and somatosensory regions in particular have exceptionally low mean flexibility. B: Scatter plot of mean region flexibility against variance in region flexibility across subjects. Brain regions that are more flexible on average have a strong tendency to also display lower cross–subject variability in flexibility.

by the timescales which we are able to probe; if criterion shift score and task accuracy are related to changes in brain dynamics primarily at the level of single trials or strategy blocks, these changes may be somewhat obscured in our dynamic networks. However, it is clear that the dynamic community structure delineated by the slower fluctuations –

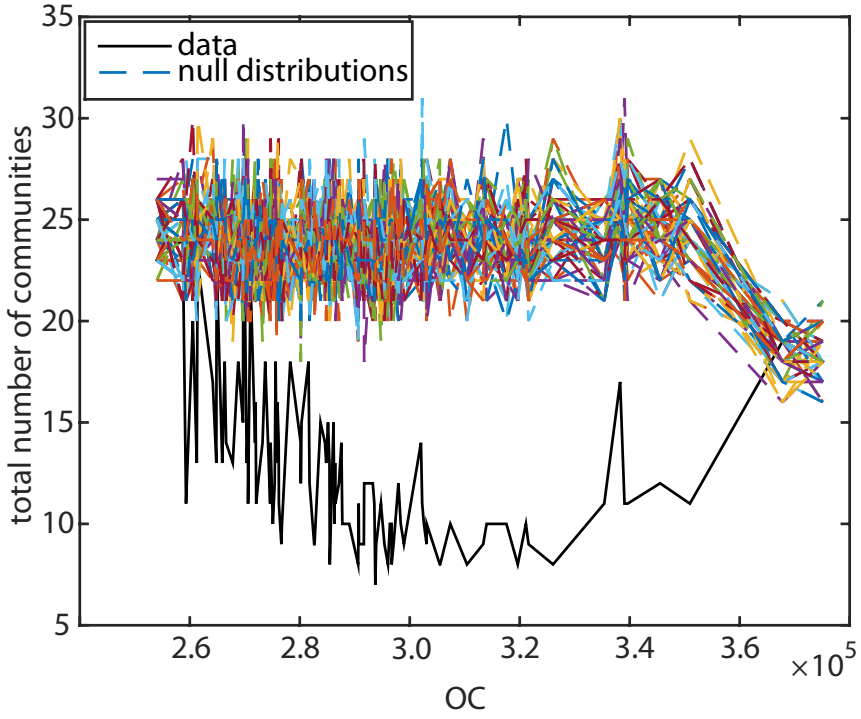

Figure 4.6: **Total community number and null distributions.** Total number of communities identified in the brain network of each subject (solid black line) compared with total number of communities identified in each of 100 overall-connectivity-preserving null model networks for the same subjects (dashed colored lines). Subjects are sorted by overall connectivity (OC). The numbers of communities identified in the data are very different for most subjects from those found in null distributions with identical OC, suggesting that the communities detected are driven largely by characteristics of the underlying connectivity structure that cannot be explained by OC. The relationship between OC and number of communities appears to be nonlinear, with large and small OC tending to lead to numbers of communities that are more strongly driven by the OC value (i.e., more similar to the null model that preserves OC alone).

e.g., in phasic arousal, attention, or strategy – do show significant changes related to demographics, to which our analyses are sensitive. Future studies designed to elicit greater performance differences, either by increasing task difficulty or by including a population of individuals with age-related cognitive impairment, could probe whether individual patterns of community dynamics are associated with these age-related changes in memory

	Community No.		**Flexibility**		**Recruitment**	
	ρ	p-value	ρ	p-value	ρ	p-value
Whole brain	0.28852	0.0031207	0.52984	8.66E-09	-0.32053	9.64E-04
Auditory	0.38723	5.33E-05	0.49118	1.38E-07	*-0.21268*	*3.10E-02*
Cingulo-opercular	0.35580	2.26E-04	0.55497	1.18E-09	-0.31670	1.12E-03
Default Mode	0.40296	2.44E-05	0.44264	2.84E-06	*-0.22947*	*1.97E-02*
Dorsal Attention	0.30945	1.47E-03	0.36514	1.49E-04	*-0.06906*	*4.88E-01*
Fronto-parietal	0.31721	1.10E-03	0.53856	4.41E-09	*-0.22348*	*2.33E-02*
Other	0.35313	2.53E-04	0.44943	1.92E-06	*-0.14008*	*1.58E-01*
Somatosensory	0.39382	3.86E-05	0.46037	9.94E-07	-0.31483	1.20E-03
Subcortical	0.38995	4.67E-05	0.46622	6.93E-07	-0.29679	2.33E-03
Ventral Attention	0.39083	4.47E-05	0.32840	7.07E-04	-0.33822	4.75E-04
Visual	0.37777	8.36E-05	*0.27263*	*5.34E-05*	*-0.15554*	*1.17E-01*

Table 4.1: **Correlations between subject age and community measures.** Spearman rank correlation ρ values and associated p-values for correlations between age and each of three community metrics: community number, flexibility, and age. Mean relative motion has been partialed out of all correlations. Italics indicate correlations that are not significant ($p > 0.05$) after family-wise error rate correction for multiple comparisons within each column; all non-italicized values represent significant correlations ($p < 0.05$).

ability, and determine which dynamics at which timescales correspond to retention or deterioration of performance.

We found that age correlates positively with community number and flexibility. That is, older adults tend to have more fragmented communities with less coherent activity than those in younger adults. Furthermore, brain regions are more likely to switch their community membership in older adults, a result only partially explained by the existence of more communities to switch between. Taken together, these results agree with previous findings from task-free paradigms, insofar as the brains of older adults tend to show a loss of the within-network integrity that might have led to them being grouped in fewer and larger coherent communities, while at the same time losing segregation between communities and seeing more fluidity of community membership over the course of the memory task.

Using null models, we show that the overall connectivity, or density, of a subject's

brain network has some influence on the number of communities detected, and hence also on the flexibility. This may be related to the resolution limit inherent in modularity maximization algorithms for community detection, in which the network density determines an intrinsic scale to the modularity that prevents the detection of communities below a certain size [171]. In multi-slice modularity maximization in general, the addition of links (here of weight ω) to connect communities across network slices affects the resolution limit of the problem, potentially biasing the number of communities as well as the flexibility. This complicated interaction between the time resolution parameter and overall network connectivity makes the mechanism underlying the changes in flexibility and community number more difficult to isolate. Here we have employed null models to probe the extent of the influence of OC on our results; however, future work is needed to fully elucidate this relationship.

We also investigate the correspondence of communities to known functional systems in the brain, and find that this correspondence is modulated by age in several circuits involved in cognitive control, including ventral attention, cingulo-opercular, and subcortical systems. The ventral attention system is involved in bottom-up attention, or response to infrequent or unexpected cues [172]. The cingulo-opercular circuit, composed of anterior cingulate cortex as well as the supramarginal gyrus, rostral middle frontal gyrus, and sections of inferior frontal gyrus, is thought to underlie tonic alertness and the maintenance of available function during a task, and to be important for cognitive control during working memory [31, 173]. The cingulo-opercular functions also include contribution from thalamus, which is categorized as a subcortical region in this scheme. The subcortical regions are less finely divided than the cortical regions in this atlas, so the subcortical nodes have larger volume and are more functionally heterogeneous [44]. Thus, the results involving subcortical regions likely contain less information on meaningful functional correlations than results involving cortex, since the signal from these

regions is averaged over a larger area containing distinct functional responses.

Overall, these results show that age-related differences are evident during the memory task in specific circuits related to attention and cognitive control (as well as the task-related somatosensory network), which is consistent with past findings that cognitive control is modulated by age [149–151]. The relationship between regions identified as theoretically meaningful on the basis of prior GLM-based analyses of BOLD activity, and the sort of dynamic, system-level connectivity of interest here, is not yet well understood. However, this study demonstrates that we can use a data-driven method to discover regions of interest for aging and task function about which it is still very difficult to make *a priori* hypotheses at this scale, based on our previous understanding of the neural processes involved in this task. The results of this and other similar investigations can be used to guide further study with different methodologies, and provide a valuable complementary body of knowledge to that gleaned from traditional, more static methods of analyzing BOLD activity.

The finding that age selectively modulates the cohesive functional grouping of these cognitive control circuits, as well as the task-involved somatosensory cortex, shows that specific cognitive systems differ notably across the lifespan, while others remain relatively unaffected by age. Importantly, although we can identify the extent of each circuit's functional changes across the lifespan, the behavioral effects of differences in these circuits remain unclear. All participants in the experiment were cognitively healthy and none showed memory impairment; furthermore, no age-related differences in performance were evident despite the clear changes we observed in functional organization. The presence of such widespread neural changes, with no manifest change in behavior, strongly suggests that compensatory mechanisms may be playing a role in this cognitive task for older adults, as proposed in previous work [174, 175]. While this study cannot identify which age-related changes are beneficial to memory performance rather than detrimental,

the methods used here provide a framework for quantifying such changes in community structure and dynamics, in future studies where age-related performance differences are evident.

4.5 Conclusion

Overall, this work confirms that the dynamics of functional community structure in the human brain during a memory task vary considerably with age. In particular, both whole-brain flexibility, which measures the tendency of brain regions to switch between communities over time, and the overall number of functional communities show notable individual differences and are strongly correlated with age, with older subjects demonstrating significantly higher flexibility and more fragmented functional communities. Using quantitative methods of comparing the community structure to known functional brain systems, we also examine the tendency of brain systems to be grouped cohesively together in communities during the memory task. We find that this tendency is significantly modulated by age in brain regions associated with cingulo-opercular, somatosensory, ventral attention, and subcortical circuits, but not in other brain areas. These results identify age as an important driver of individual variation in functional community dynamics, and provide insight into how aging differentially impacts the functional organization of different brain systems, even in healthy adults who do not experience declines in performance. Additionally, they demonstrate methods which promise to be useful in quantifying which circuits drive changes in network organization across a broad range of situations, including in task-active networks.

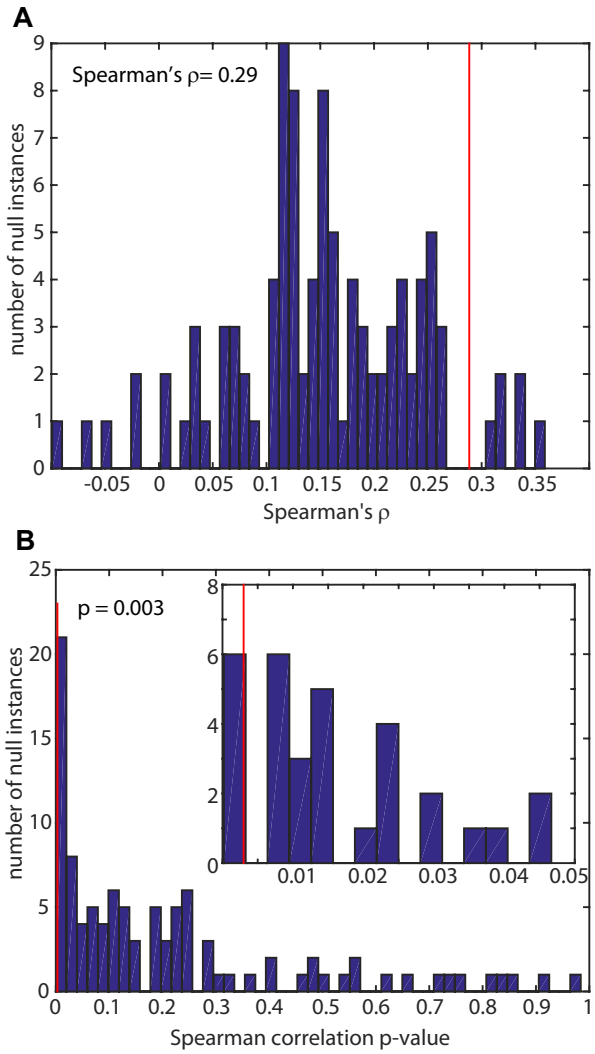

Figure 4.7: **Relationship between age and number of communities in functional brain data and corresponding null models.** A: The solid red line shows the Spearman correlation value between subject age and number of communities; the bars show a histogram of the same correlation values, each computed from one set of 100 OC-preserving null networks. B: The solid red line shows the p-value of the Spearman correlation observed in the data; the bars show the null distribution of p-values corresponding to the null Spearman correlation values in A. For six out of 100 randomized community structures, the correlation between age and number of random communities is stronger than the observed correlation between age and number of functional brain communities. However, for most randomized community structures, no significant correlation is found between age and number of random communities is found.

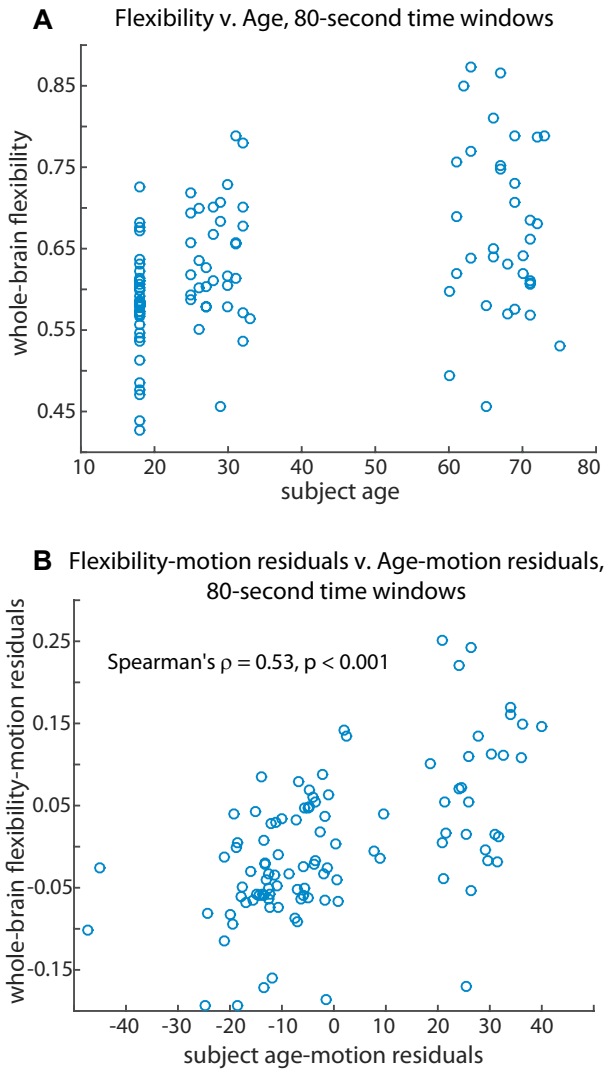

Figure 4.8: **Whole-brain flexibility and age.** A: Scatter plot of the correspondence between subject age and whole-brain flexibility in networks with 80-second time windows. B: Scatter plot showing the significant positive correlation between age and whole-brain flexibility with mean relative motion partialed out (plot shows residuals of separately regressing each measure on mean relative motion).

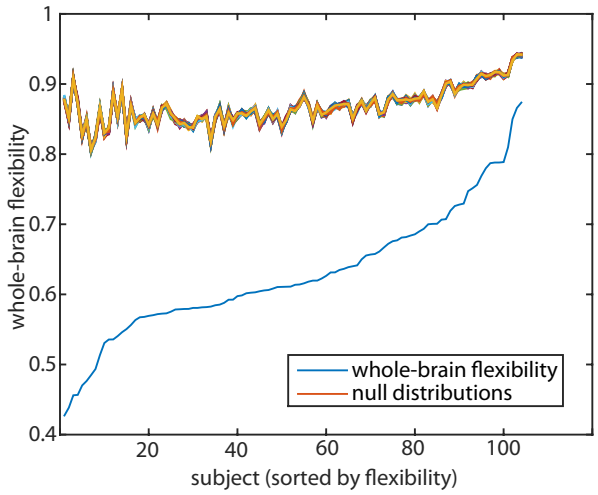

Figure 4.9: **Whole-brain flexibility and null distributions.** The solid black line shows the whole-brain flexibility of each subject (sorted); the colored lines show the distribution of whole-brain flexibilities computed from 100 instances of the corresponding community-number-preserving null model. There are clear differences between the whole-brain flexibility and the null distributions, but the flexibility values computed from null models, based only upon the number and size of communities, remain strong predictors of a subject's whole-brain flexibility.

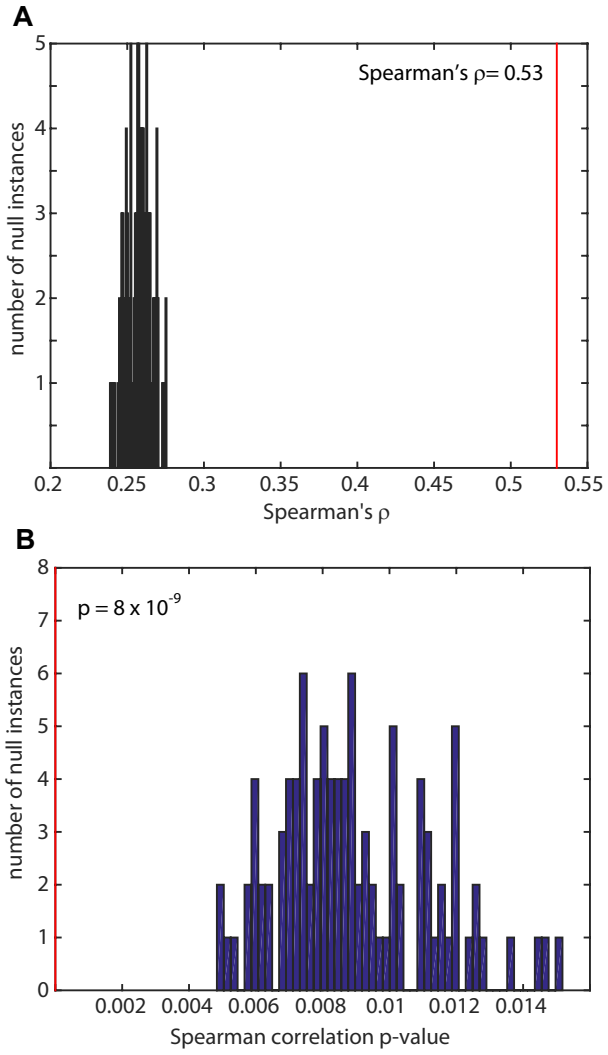

Figure 4.10: **Age-flexibility relationship in functional brain networks and null models.** A: The solid red line shows the Spearman correlation value between whole-brain flexibility and age; the bars show a histogram of the same Spearman correlation values, each computed from one set of 100 community-number-preserving null models. B: The solid red line shows the p-value of the Spearman correlation observed in the data; the bars show the null distribution of p-values corresponding to the null Spearman correlation values in A. While all null correlations between age and flexibility in shuffled community structures are fairly strong and statistically significant, all are quite distant from and weaker than the correlation between age and whole-brain flexibility, indicating that the number of communities alone cannot explain the entire correspondence between flexibility and age.

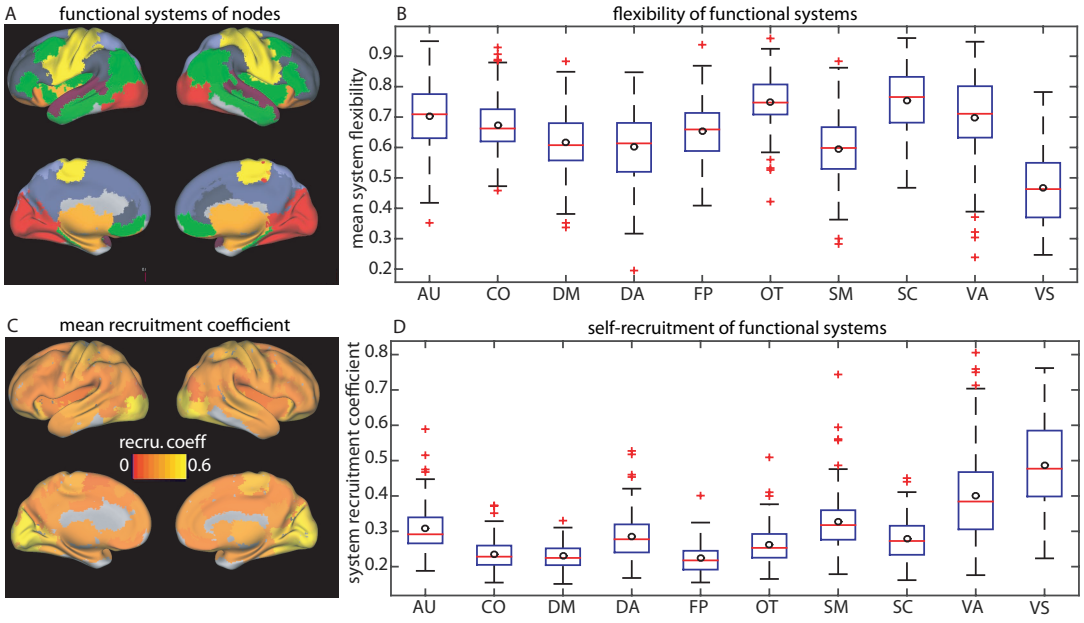

Figure 4.11: **Community organization of functional systems.** A: The functional system partition of brain regions, with systems indicated by color. Systems identified (in color order from purple to red) are auditory (AU), cingulo-opercular (CO), default mode (DM), dorsal attention (DA), fronto-parietal (FP), other (OT), somatosensory (SM), subcortical (SC), ventral attention (VA), and visual (VS). B: Box plot showing the mean flexibility of brain regions in each functional system, and the distribution of this mean flexibility over subjects. C: Recruitment coefficients of each brain region (network node); visual cortex and somatosensory regions in particular have exceptionally high recruitment. D: Box plot showing the self-recruitment of each functional system, and its distribution over subjects.

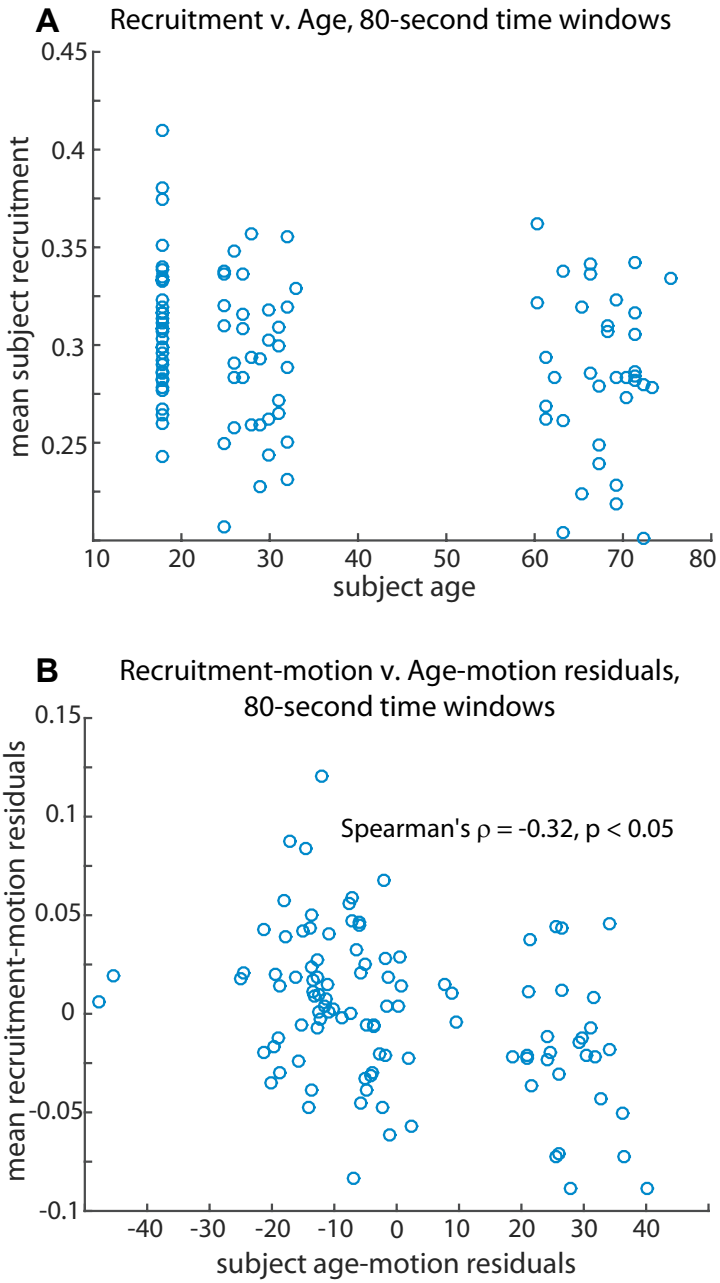

Figure 4.12: **Brain region recruitment and age.** A: Scatter plot of the correspondence between subject age and average recruitment across all brain regions in networks with 80-second time windows. B: Scatter plot showing the significant negative correlation between these measures with mean relative motion partialed out.

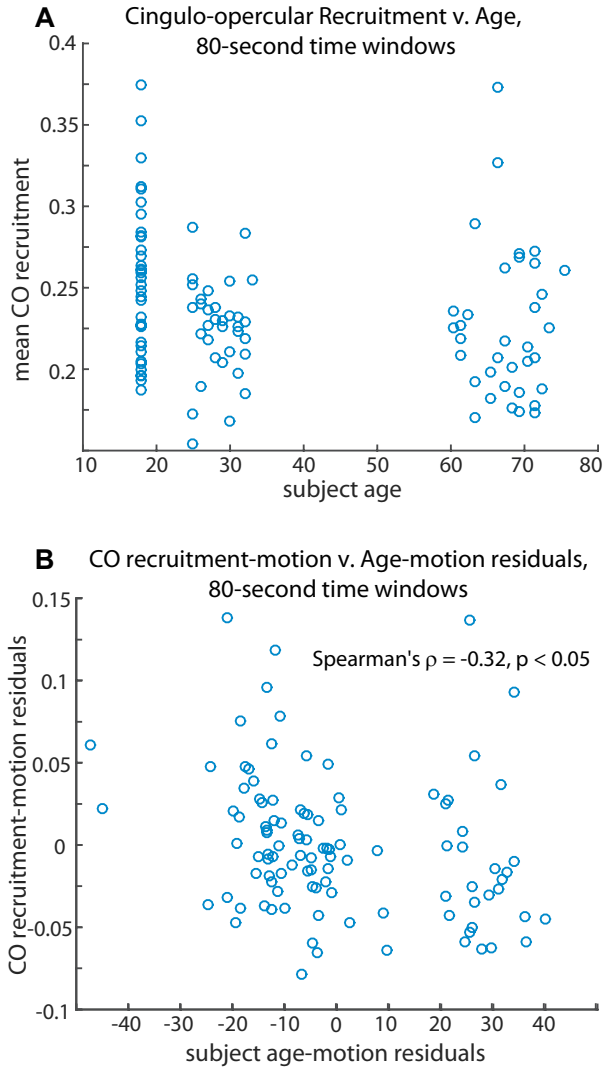

Figure 4.13: **System self-recruitment and age.** A: Scatter plot of the correspondence between cingulo-opercular system self-recruitment and subject age in networks with 80-second time windows. B: Scatter plot showing a significant anticorrelation between these measures with mean relative motion partialed out (plot shows residuals of separately regressing each measure on mean relative motion). C: Scatter plot of the correspondence between visual system self-recruitment and subject age in networks with 80-second time windows. D: Scatter plot of the correspondence between visual self-recruitment and age with mean relative motion partialed out; there is no apparent correlation.

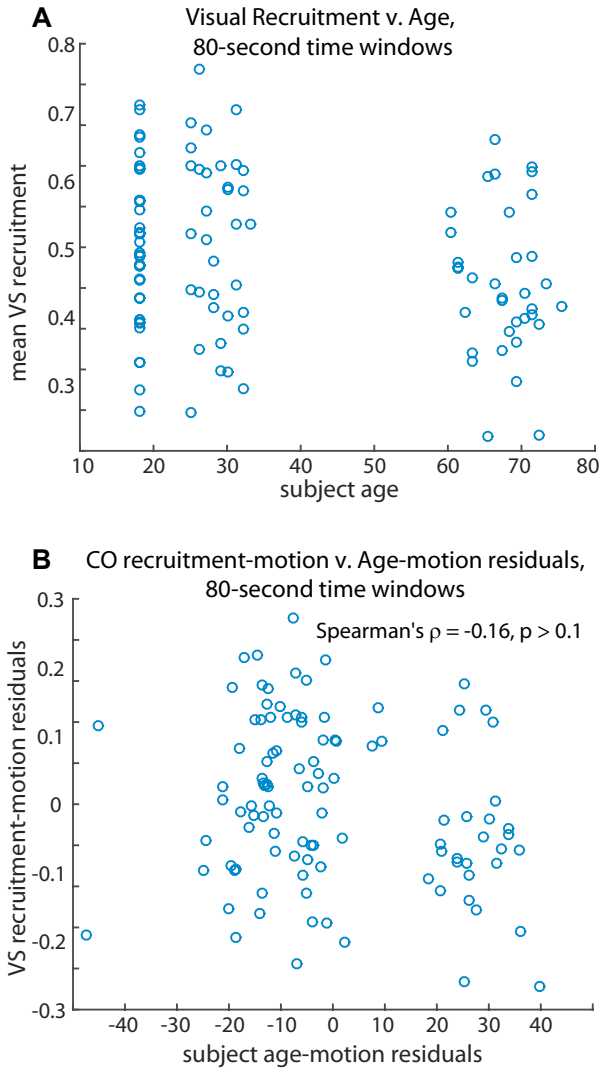

Figure 4.14: **System self-recruitment and age.** A: Scatter plot of the correspondence between visual system self-recruitment and subject age in networks with 80-second time windows. B: Scatter plot of the correspondence between visual self--recruitment and age with mean relative motion partialed out; there is no apparent correlation.

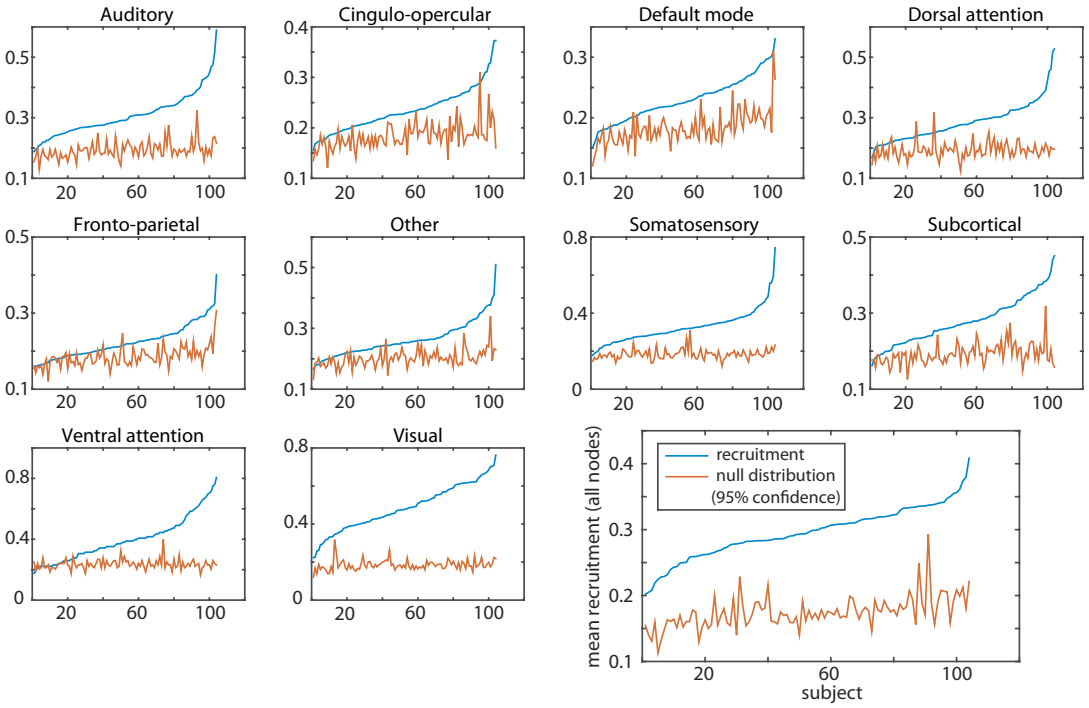

Figure 4.15: **Dependence of recruitment on community size distributions.** Bottom right: Recruitment values for each subject (blue line), averaged over all brain regions, alongside 95% confidence line (red line) from null distribution of 100 recruitment values computed in community-size-preserving null networks. All subjects have mean whole-brain recruitment values significantly higher than those expected in null networks, indicating that the association of algorithmically identified communities with known functional systems is statistically significant. Smaller panels: Recruitment values for each subject (blue lines), averaged over brain regions in known functional systems, alongside 95% confidence lines (red lines). Some individual systems, including sensory and motor cortices (auditory, somatosensory, and visual) and subcortical structures, are consistently associated with identified communities at a statistically significant rate. Others–including systems identified with executive control, both focused and bottom-up attention, and the resting state–have recruitment values that could reasonably be explained by chance in several subjects.

Chapter 5

Targeted Node Removal for Improved Resolution of Dynamic Brain Communities

5.1 Introduction

In the study of complex networks, dynamic community detection is a method for identifying highly intraconnected clusters of nodes within a network and quantifying how these clusters change over time. In many cases, the identified clusters, or communities, correspond to modules that perform an identifiable functional or structural role, thus giving insight into the composition and organization of a network [176]. Detecting temporally changing clusters enables an analysis of how the roles of these modules evolve, and how the network reorganizes itself on various timescales [15]. In network neuroscience, communities highlight the organization of interacting neurons or brain regions [8]. The application of these methods to large-scale MRI-based structural and functional brain networks has identified broad organizational similarities shared between

distinct brains [10, 26, 169], and quantified changes in brain dynamics across cognitive states, demographic measures, and time [14, 113, 137].

Although dynamic community detection has contributed to a more complete understanding of the brain, challenges remain in applying it to large-scale functional brain networks for predictive and diagnostic purposes [8]. Functional modules exist at a variety of sizes and temporal ranges, and community methods rely on parameter choices to resolve clusters at scales relevant to a specific question or investigation [8, 14, 177]. In the absence of clear "ground truth" knowledge against which to evaluate methods, strategies for choosing these parameters have varied widely and are often based upon the statistical robustness of results [8, 35]. Thus, groups of brain regions that are especially strongly correlated may dominate the identified community structure, obscuring the resolution of other functional modules or dynamic properties. For example, during tasks with a visual component, brain regions in the visual cortex form a highly coherent community that strongly affects the selection of resolution parameters. This may prevent a community detection algorithm from resolving modules that perform other cognitive functions during the task. An incomplete understanding of the effects of these parameter choices also complicates comparison across analyses with different parameters.

Here, we introduce an approach for targeted removal of network nodes to improve resolution in dynamic community detection. We demonstrate the approach in a synthetic network of oscillators, in which we precisely quantify detection performance by comparing to well-defined "ground truth" communities. We show that the presence of multi-scale organization inhibits community detection in these oscillator networks. We further demonstrate that removal of targeted subsets of nodes during community detection improves the resolution of communities among the remaining nodes.

We demonstrate the utility of targeted node removal in neuroscience applications by applying this method to dynamic functional brain networks from two distinct fMRI exper-

iments. One features repetition of a single cognitive task and the other encompasses the performance of multiple different tasks. Commonly used community detection methods fail to resolve the substantial dynamic differences between these single- and multi-task data sets. However, with targeted removal of the regions in visual cortex, which contains the most functionally coherent brain regions, community detection reveals clear differences in the dynamic network properties of the two data sets on a population level. In addition, removal of visual regions improves the ability of these methods to spatially resolve groups of brain regions known to be functionally similar, especially in the multi-task data set. These results show that targeted node removal can both improve resolution of community dynamics in a single data set, and also enable comparison of community structures across data sets.

Background and Motivation

Community detection algorithms aim to use the connectivity information of a network to identify a network partition, or a division of the network nodes into clusters, such that each cluster is composed of nodes strongly connected within the cluster and weakly connected to other clusters. Unsupervised community detection methods often uncover useful or intuitive groupings: they can extract official affiliations based on interactions in human social networks [178]; identify similar regions in a field of view to aid image processing and compression [176,179,180]; and classify biochemical species based on their dynamics within metabolic networks [181].

Developing practical applications of community detection for specific networked systems requires choosing various context-dependent model parameters. In network neuroscience, this choice is especially challenging. Very little is known about fundamental principles that underlie the dynamic organization of large-scale brain regions, and a "ground truth" benchmark for functional communities is not well-understood in a net-

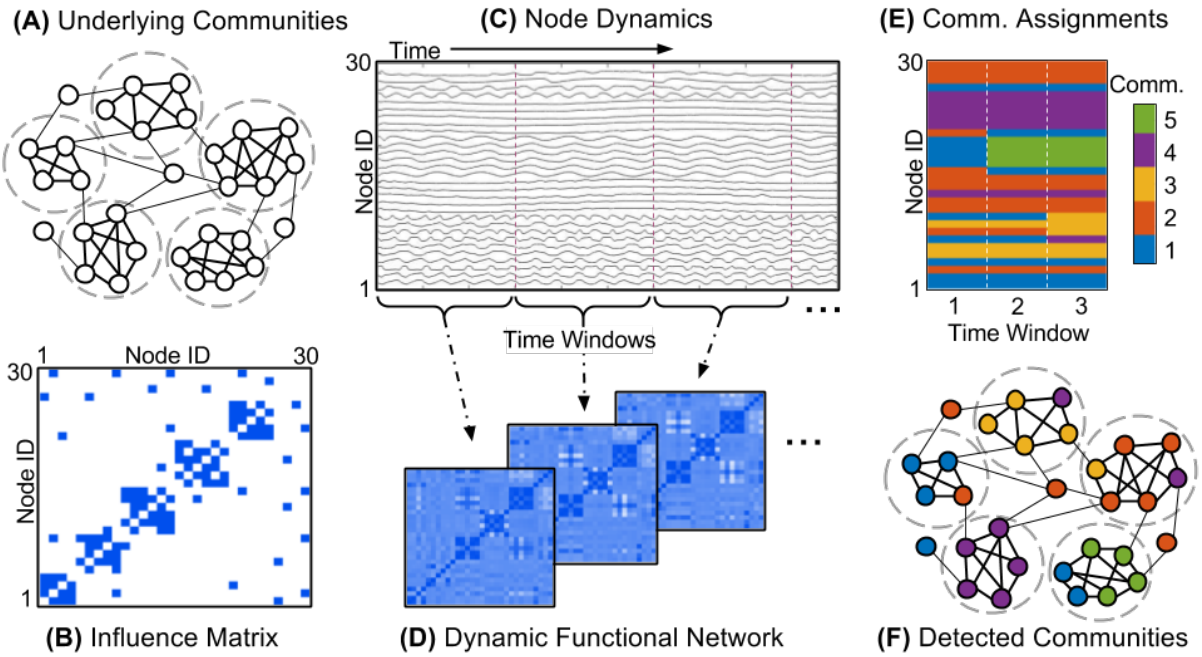

Figure 5.1: **Schematic: Community detection on dynamic networks.** A: Representation of the nodes and edges of a modular network, with dashed circles indicating the underlying communities. B: Binary *influence matrix* corresponding to this network, in which blue entries indicate a direct influence between node pairs and white entries indicate no direct influence. C: An example of node dynamics produced by this network, where each node is modeled as an oscillator with an intrinsic frequency, and nodes are influenced by their neighbors according to Equation 5.5. The evolving phases of successive nodes are stacked along the y-axis. D: Synchronization matrices representing the dynamic functional network derived from the time series in C. Each sequential matrix shows the synchronization between node pairs, averaged across the corresponding time window. E: Dynamic communities detected in the dynamic functional network in D. Each node is assigned to a single community (denoted by color) in each time window, and may switch community assignments between time windows. F: Comparison of dynamic community assignments from the third time window in E (denoted by color) to the underlying communities of influence from A (denoted by dashed circles).

work context. In current research, evaluation of community "correctness" is largely based on correspondence with anatomy, pre-existing knowledge of functional roles, or statistical analysis of community robustness [8, 169]. There is therefore little consensus about the best methods for imposing constraints on community detection algorithms and for evaluating results.

Additionally, the human brain has inherently multi-scale organization. This includes hierarchical modularity (communities within communities), interlocking communities of varying sizes, and communities that dynamically reconfigure over time [10, 14, 15, 182, 183]. Furthermore, brain dynamics related to a given experimental condition cannot be easily isolated. During an MRI scanning session, many brain circuits of multiple sizes and strengths, both related and unrelated to the phenomenon under study, are simultaneously active. Brain regions with particularly strong or consistently coherent activity, such as those in sensory or motor cortices during recruitment of those functions, may dominate the community structure detected in a data-driven analysis, masking the dynamic properties of other brain areas more relevant to the study.

The noisy, multi-scale, and dynamic nature of brain dynamics is not necessarily suited to common community detection algorithms. The "modularity maximization" method for community detection is widely used, computationally efficient, easily implemented, and natural to extend to weighted, signed, and dynamic networks [8,167]. However, it has also been shown to have an inherent resolution limit [171] and requires the choice of parameters defining spatial and temporal scales [8, 15, 35]. Several heuristics have been developed to ensure that the detected communities represent organization at an informative scale. These typically involve either exploring many possible scales, or choosing communities at a scale that gives the most consistent partitions [8, 35, 177]. Such methods have enabled reasonable statistical confidence in results from a single individual or data set. However, approaches for comparing results of different resolutions between different individuals or data sets remain elusive.

Strategy

This chapter demonstrates that targeted removal of network nodes during community detection can be leveraged to improve the resolution of communities on multiple

scales, and to aid in the principled comparison of community structure across data sets. Specifically, we study the effect of removing subsets of nodes with particularly strongly connected or coherent dynamics, as determined by observation or through knowledge of their functional role. For example, sensory or motor regions in large-scale functional brain networks are often highly coherent during tasks that recruit these functions, which may hinder the resolution of the dynamics of other network nodes in community detection.

We first illustrate the node removal approach in synthetic oscillator networks, in which results are evaluated based on a clear underlying "ground truth" network. A schematic of the approach to community detection in these networks is shown in Figs 5.1 and 5.2. We designate an underlying adjacency matrix (Fig 5.1B) for each synthetic network (Fig 5.1A), organized into modules that we define. This matrix, which we refer to as the "influence matrix," represents the causal influences between oscillators that drive the synchronization dynamics of the network. This influence matrix serves as a rough analogy to the set of underlying anatomical connections and/or functional influences that produce the dynamics of brain activity measured with fMRI. It is not meant to represent the physical structure or architecture of the brain, but rather the basic functional organization that underlies the observed neural activity, and which we elucidate with community detection.

Once an influence matrix is defined for a synthetic oscillator network, we simulate the resulting network dynamics (Fig 5.1C), which we then "measure" by computing the synchronization between node pairs across a set of time windows (Fig 5.1D). Node pairs within the same underlying community tend to synchronize more closely than pairs in different communities. We produce a community partition (Fig 5.1E) from the dynamic network formed by the synchronization measurements, by applying a modularity-maximization community detection algorithm to this synchronization network [15, 167]. We use true and false positive rates to quantify the extent to which this partition matches

the original underlying communities (Fig 5.1F). Finally, we apply community detection again, this time using information from only a subset of nodes in the synchronization matrices, and compare the results to community detection on the full network (Fig 5.2). Note that when removing nodes, the dynamics are still simulated for the full underlying network. Nodes are removed only from the dynamic synchronization networks (shown in Fig 5.1D and Fig 5.2A) before the community detection algorithm is applied.

We find that in the precisely controlled oscillator networks, removal of certain subsets of nodes during community detection can improve the resolution of communities among the remaining nodes. We then apply our methods to functional human brain networks using two data sets, corresponding to two separate functional MRI experiments with different temporal task structures. In these brain networks, removal of the strongly correlated visual cortex from the dynamic adjacency matrix allows for better resolution of the differences between community structure during different cognitive tasks.

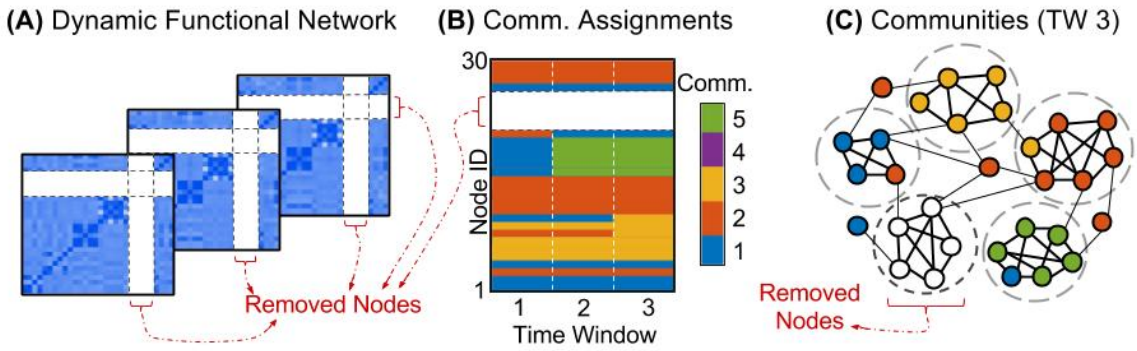

Figure 5.2: **Schematic: Targeted removal of nodes for community detection.** A: From the measured dynamics of the modular network in Figure 5.1, new dynamic functional network matrices are computed, with the functionally cohesive community removed (nodes 22-26). B: This truncated functional network is then used to detect dynamic communities, producing assignments for all remaining nodes. C: The resulting community structure from the third time window, denoted by color, provides a clearer identification of the underlying communities than the community structure detected with all nodes taken into account (Figure 5.1F).

5.2 Methods

We first describe the method of modularity maximization used to detect dynamic communities throughout this chapter, including the algorithms used to perform the maximization and find a stable community structure. We focus on this technique due to its widespread use for community detection in network neuroscience [8], which stems from its clear conceptual definition, the ease of adapting it to weighted and dynamic networks [184], and the existence of multiple computationally efficient implementations [15, 167]. We also define the metrics used to quantify community properties and assess detection performance. Finally, we introduce the two types of networks to which we apply these community detection methods: synthetic networks of nonlinear Kuramoto oscillators, with dynamics simulated *in silico*; and dynamic networks of human brain function, derived from fMRI measurements of brain activity while participants performed different sets of cognitive tasks.

5.2.1 Community Detection Methods

We consider a network of N nodes connected by weighted, unsigned edges. Each edge may take on a different positive, real-valued weight in each of T time windows; A_{ijt} denotes the weight of the edge between node i and node j in time window t. In order to identify the optimal partition of nodes into modular communities, we seek the partition that maximizes the multislice modularity,

$$Q = \frac{1}{2\mu} \sum_{ijtr} \left\{ \left(A_{ijt} - \gamma P_{ijr} \right) \delta_{tr} + \omega \delta_{ij} \right\} \delta \left(g_{it}, g_{jr} \right), \tag{5.1}$$

which indicates the quality of the modular structure of a partition in comparison to a randomized "null" network, P_{ijt} [15].

This quantity considers all node pairs i, j and all time window pairs t, r in which the community assignment of node i in window t (g_{it}) is the same as the community assignment of node j in window r (g_{jr}). For each node pair assigned to the same community in the same time window, the first term in the brackets provides a positive contribution to Q if the actual edge weight between the pair compares favorably to that in the null model. A spatial resolution factor γ determines the relative weight given to the null model. For each node j and each pair of time windows, the second term provides a positive contribution of ω to Q when j is assigned to the same community in both time windows [15]. Thus, maximizing Q favors network partitions in which the weights between nodes in the same community are greater than those expected in the null model, as well as those that group more nodes in the same community as themselves across multiple time windows.

In this chapter we use the Newman-Girvan null model, which treats edge weights as randomly distributed within each time window while preserving the node degree distribution [178]. We maximize Q over network partitions with a Louvain-like locally greedy algorithm implemented in MATLAB [15, 167]. Due to the stochasticity of the algorithm and the expected high degeneracy of solutions near the maximum value of Q, we use a community consensus procedure to distill a statistically representative partition from an ensemble of 100 solutions for each network (for more details, see [35] and [113]).

Resolution parameters

The value of the multislice modularity Q depends upon the values of γ, a spatial resolution parameter, and ω, a temporal resolution parameter. These parameters control the relative weight given to the null model in calculating Q, and thus alter the spatial and temporal scales at which communities will be found. A higher value of γ gives more weight to the null model, requiring much stronger connections to be present within a subset of nodes before they are counted as sufficiently different from the null model to

constitute a community. A lower value of γ requires less connectivity within a potential community before it is deemed to be significant. Thus, lower values encourage few and large communities ($\gamma = 0$ will always return a single community containing every connected network node), while higher γ values tend to produce more, smaller communities.

In terms of temporal resolution, higher ω values place a greater value on maintaining the community assignment of a node across time windows of the network, and tend to produce partitions in which the node assignments are more similar to each other across time windows. Lower ω values give temporally consistent node assignments less weight, and in the case where $\omega = 0$, the dynamic community detection task becomes equivalent to performing static community detection on each time window separately.

Metrics for Community Structure and Detection Performance

Community number. For a given community partition, the *community number* is defined as the total number of distinct community assignments given to the network nodes. The community number is always between 1 (all nodes in the same community) and N (each node in a different community) for modularity maximization methods.

Flexibility. The *flexibility* of a node in a dynamic network is defined as the number of times that node switches communities between adjacent time windows, normalized by the total possible number of switches:

$$f(i) = \frac{1}{T-1} \sum_{t=1}^{T-1} \left[1 - \delta(g_{it}, g_{i(t+1)})\right]. \tag{5.2}$$

Here, T is the total number of time windows; $\delta(g_{it}, g_{it'})$ equals 1 if node i is assigned to the same community in slice t and slice t', and 0 otherwise. A node with high flexibility changes communities in every or nearly every time window and has a flexibility at or near 1, while a node with low flexibility may remain in the same community in all windows

and have a flexibility of 0. For example, in the schematic representation of community assignments for a dynamic network in Fig 5.1E, nodes 1 through 6 remain in the same community across all three time windows and have a flexibility of 0. However, node 7 switches from community 1 to community 4 between the last two time windows, and thus has a higher flexibility.

The *whole-network flexibility* is defined as the mean flexibility over all network nodes, and can be used to compare network dynamics under different conditions.

Community detection performance. In cases with available "ground truth" communities, we use *true* and *false positive rates* to quantify community detection performance. We define these in terms of node pair co-assignments, comparing whether each node pair is truly in the same community (a "true" co-assignment) or not, and whether the node pair is "identified" as belonging to the same community by the algorithm or not. These quantities are defined as follows:

$$\text{true positive rate (TPR)} = \frac{tp}{tp + fn}, \tag{5.3}$$

$$\text{false positive rate (FPR)} = \frac{fp}{fp + tn}. \tag{5.4}$$

Here, tp indicates the number of true positives – i.e., node pair co-assignments both existent in the underlying influence matrix ("true") and identified by the algorithm ("identified"). Similarly, fp indicates the number of false positives ("identified" co-assignments which are not "true"); fn the number of false negatives ("non-identified" co-assignments that are "true" in the influence matrix); and tn the number of true negatives ("non-identified" co-assignments which are also not "true"). Here, we compute the TPR and FPR separately for each network instance in an ensemble.

For each network, we summarize the performance of the community detection algorithm over an entire ensemble of networks with a *detection probability matrix* D. Each entry D_{ij} gives the fraction of network instances in which nodes i and j are identified as belonging to the same community. Perfect detection performance over all instances would result in $D_{ij} = 1$ for all i, j pairs in the same underlying community, and $D_{ij} = 0$ for all other pairs. $D_{ij} < 1$ (for i and j in the same underlying community) or $D_{ij} > 0$ (for i and j in different underlying communities) indicate that there are network instances in which the node pair co-assignment was incorrectly identified. When compared to the underlying influence matrix, D specifies which particular node pairs were correctly or incorrectly identified, contributing to the overall true and false positive rates.

We choose TPR, FPR, and the detection probability matrix to evaluate the community detection performance in order to shed light on the differences between different types of detection failures. Separating true positive rate and false positive rate – as opposed to using a measure such as Rand index which counts all correctly identified pair associations (both true positives and true negatives) together – allows easy visualization of the difference between community partitions that tend to combine underlying communities together (high false positive rate) and those that tend to split them apart (low true positive rate). Plotting TPR against FPR, as in Figs. 5.3,5.5,and5.4, provides a visual interpretation of the performance of the detection algorithm compared to its expected random-chance performance (where $TPR = FPR$), and perfect delineation of communities (the upper left-hand corner, where $TPR = 1$ and $FPR = 0$).

5.2.2 Synthetic Networks of Kuramoto Oscillators

We apply these community detection methods and techniques for improving resolution of community detection to synthetic networks of Kuramoto oscillators, in which the

underlying influence matrix can be precisely controlled.

Following [16], we define each of the N nodes of a Kuramoto network as an oscillator indexed by i ($i = 1, 2, ..., N$), whose time-dependent internal state is given by the angle $\theta_i(t) \in [0, 2\pi)$. The state of each oscillator i evolves according to

$$\frac{d\theta_i}{dt} = \omega_i + \sum_j \kappa C_{ij} \sin(\theta_j - \theta_i), \tag{5.5}$$

where ω_i is the intrinsic frequency and the second term describes the influence of other oscillators in the network. Interactions between oscillators are governed by a scaling factor κ and a time-independent, N x N binary matrix C, where $C_{ij} = 1$ denotes a *direct* influence between oscillators i and j, such that their dynamics will tend to synchronize over time. $C_{ij} = 0$ denotes no direct influence between i and j, although there may still be synchronization between them as a result of indirect influence or by chance. We refer to C as the *influence matrix* of the network.

We design C to have a modular structure, consisting of a number of communities of influence. Any pair of nodes in the same community have a directly influential relationship with probability P_{in}, while a pair of nodes in two different communities is directly related with probability P_{out}.

We examine two distinct network types determined by their underlying influence structure: single-scale and multi-scale. *Single-scale* networks are composed of communities that are all the same size. We consider single-scale communities of either 20 nodes or 8 nodes each. *Multi-scale* networks contain a collection of communities of different sizes. We focus on a multi-scale network of $N = 100$ nodes, consisting of eight total communities, with three 20-node communities and five 8-node communities. For all networks, we use $P_{in} = 0.9$ for 20-node communities, $P_{in} = 0.7$ for 8-node communities, and $P_{out} = 0.01$ for all out-of-community connections.

After determining an influence matrix C, we initiate each network with normally distributed intrinsic oscillator frequencies ($\sigma = 1$) and solve the network dynamics numerically with $\kappa = 0.2$. The dynamics of such a network begin in a random state, and the oscillators approach a steady state of partial synchronization after a short transient period. We quantify the observed functional dynamics of the network by computing the time-dependent *synchronization* [16]

$$\phi_{ij}(t) = |\cos(\theta_i(t) - \theta_j(t))|. \tag{5.6}$$

We then apply the community detection method to these synchronization dynamics, represented in the adjacency matrix Φ. The entries Φ_{ijr} are determined by averaging the synchronization $\phi_{ij}(t)$ over each of a chosen set of time windows (indexed by r), as depicted schematically in Fig 5.1D. Here, we use eight time windows of 50 time steps each. These time-dependent synchronization matrices are used by the community detection algorithm to identify dynamic communities, as in Fig 5.1E. We compare the detected community structure to the underlying "ground truth" of the influence matrix (Fig 5.1F) to quantify the performance of the community detection algorithm under various conditions.

For each arrangement of underlying communities, we generate an ensemble of 20 influence matrices, all using the same intra- and inter-community connection probabilities (P_{in} and P_{out}). We perform the dynamic simulation and community detection procedure separately on each network instance in the ensemble, resulting in a distribution of performance metrics. This distribution provides an estimate of the variance in community detection performance across networks with the same underlying structure, but noise or other variation affecting the existence or the measurement of some connections.

5.2.3 Functional Brain Networks from fMRI Data

To demonstrate their utility in large-scale human brain networks, we apply these community detection and node removal techniques to two distinct functional brain network data sets, derived from two separate fMRI experiments on different groups of participants. We refer to these two functional data sets as the "single-task" and the "multi-task" data sets. Both experiments require participants to perform cognitive tasks during successive fMRI runs, but they provide a contrast between a set of behaviorally similar functional runs ("single-task" experiment) and a set of runs designed to elicit distinct cognitive functions ("multi-task" experiment).

Informed written consent was obtained from each participant prior to experimental sessions, and all procedures were approved by the University of California, Santa Barbara Human Participants Committee.

Experimental Procedure

Single-task experiment. 126 healthy adult participants were scanned while performing a recognition memory task with lexical stimuli. During each of three identically designed functional runs, both previously examined and novel words were shown, and participants were required to distinguish between them with the aid of probabilistic cues. Each run was approximately 8.5 minutes long. Due to various sources of attrition and technical issues, data from 22 participants was excluded, leading to a final analysis of 104 participants. For additional experimental details, see [113] and [42].

Multi-task experiment. Functional MRI data were collected from 116 healthy adult participants during a set of distinct cognitive states. Participants were scanned at rest (task-free) and while engaging in three functional tasks: an attention-demanding task, a memory task with lexical stimuli similar to that used in the single-task experiment,

and another similar memory task with face stimuli. Due to various sources of attrition, only 77 participants are included in the final analysis. For more experimental details, see [111, 185], and [12].

Functional Brain Networks

A dynamic network is constructed separately from each participant's measured functional activity. Each network contains N nodes, corresponding to the $N = 194$ brain regions of a "hybrid" anatomical atlas, an adaptation of the multi-resolution Lausanne2008 atlas minimizing variability in region size [44,111]. This atlas was registered to MNI space for each participant. The same set of brain regions is used for all participants, and the brain regions do not change over time. Region-specific time series from each functional run were generated for each node by averaging the BOLD signal time series across all voxels within the brain region [111].

Each network has $E = N(N-1)/2$ edges, each with a real-valued, non-negative connection weight that may change over time, taking a new value in each of T sequential time windows spanning the experiment. The weight of an edge between nodes i and j in a given time window t, denoted A_{ijt}, is defined as the mean low-frequency (0.06-0.125 Hz) wavelet coherence between the BOLD time series of i and j within that time window [11, 33, 56, 113, 166].

Community Detection

For each participant's dynamic functional network, the community detection method described above is applied to find the partition that maximizes the multislice modularity Q. In these brain networks, the spatial resolution parameter γ is chosen with the analysis described in [113]. The most informative spatial scale is expected to be the one giving the most consistent community partitions across randomly seeded stochastic runs of the

locally greedy modularity maximization algorithm. Following this reasoning, community detection is performed independently across a range of γ values, and the mean z-score of the Rand index between each pair of partitions generated by 100 algorithm runs at each γ value is computed, providing a quantitative measure of similarity across the partitions [186]. The Rand z-score is chosen because it inherently provides a comparison to a null distribution that takes the number and size of communities in each partition into account; it can be calculated analytically; and its behavior in the context of modularity maximization on functional brain networks has been previously characterized [35, 186]. The optimal γ value is that giving the highest average Rand z-score across pairs of algorithm runs and across participants, indicating the most consistent community partitions. When there is no clear choice (i.e., when the γ landscape is relatively flat), a near-optimal value is chosen based upon the expected number of functional communities [113]. After choosing a spatial scale, a temporal scale (ω value) is determined by choosing the value that maximizes the variance in flexibility across network nodes, where a node's flexibility measures the number of times it switches community assignments between adjacent time windows (see Eq 5.2). This ensures that the algorithm will resolve high-flexibility nodes from those that remain within the same community throughout the experiment. See Section 5.4 for further treatment of resolution parameters.

Comparison to Known Functional Systems

In the absence of a clear "ground truth" benchmark for human functional brain networks, precise evaluation of community detection performance is a challenge. We use a basic partition of functional systems within the brain to assess whether these methods resolve brain systems with different functional roles from each other [169]. In this partition, each of the 194 brain regions is assigned to one of ten systems: auditory, cingulo-opercular, default mode, dorsal attention, fronto-parietal, somatosensory, subcortical,

ventral attention, visual, and other. These systems were identified with a network-based clustering approach [169], and have been used to describe and quantify system-specific functional brain interactions [24, 26, 33, 113, 170, 187]. The assignment of regions to systems used here is based upon the primary functional roles of different anatomical brain areas, as detailed in [187] and [33].

We use the *recruitment* to quantify the relationship between known functional systems and the communities detected in the data. The *node-specific recruitment* of a brain region is a measure of the consistency with which that region is assigned to the same community as other nodes in its own functional system. It is given by

$$R(i) = \frac{1}{n(s_i) - 1} \sum_{j \neq i} \delta(c_i, c_j)\delta(s_i, s_j). \tag{5.7}$$

Here, s_i denotes the functional system of brain region i, $n(s_i)$ gives the total number of regions in system s_i, and c_i denotes the community assignment given to i by the community detection algorithm. The Kronecker deltas $\delta(s_i, s_j)$ and $\delta(c_i, c_j)$ count the region pairs (i, j) that belong to both the same known functional system and the same data-driven community. Thus, a brain region will have a high recruitment coefficient if the data-driven community to which it is assigned also contains a high fraction of functionally similar nodes.

We define the *system-specific* or *system recruitment* Ψ of a given system S as the average node-based recruitment of all nodes in the system, given by

$$\Psi(S) = \frac{1}{n(S)(n(S) - 1)} \sum_{ij, i \neq j} \delta(c_i, c_j)\delta(s_i, S)\delta(s_j, S). \tag{5.8}$$

This measures the extent to which nodes in system S are cohesively grouped together in the same community.

These recruitment metrics are inspired by similar quantities used in Refs. [33] and [188], but have been adapted to allow comparison of the metrics across scanning runs and subjects, and to avoid self-comparisons between nodes.

This basic partition into functional systems enables a quantitative assessment of the overlap between detected communities and the broad functional organization of the human brain. However, it does not fulfill the function of the "ground truth" benchmarks used in the synthetic oscillator networks. In particular, any "true" underlying functional modules driving the measured brain dynamics in these fMRI experiments would be expected to change dynamically as different brain systems are recruited for different cognitive tasks. This partition also does not capture the likely individual, temporal, and situational variation in the organization of brain function within a single task.

5.3 Results

5.3.1 Community Detection in Synthetic Networks

Here, we demonstrate the performance of the modularity maximization community detection technique on a set of synthetic Kuramoto oscillator networks, detecting communities both before and after the targeted removal of nodes, and assessing the effect on performance.

As described in Section 5.2.2, the basic system studied is a network of N oscillators, each with its own intrinsic frequency ω_i, as well as influences from the other oscillators which are described by a time-independent binary influence matrix C. We simulate the network dynamics resulting from this underlying pattern of influence, and track the observed synchronization dynamics that result. These synchronization networks serve as a simplified analogy to functional brain networks, reflecting both direct and indirect

influence among network nodes through a measurement of observed coherence [16]. By performing community detection on the observed synchronization dynamics of Kuramoto networks, we determine the accuracy with which these techniques can uncover the true underlying network communities in the presence of inherently multi-scale dynamics.

Comparing Single-scale and Multi-scale Networks

Single-scale networks. We begin by investigating networks with communities of a single size. The first of these is a network of $N = 60$ nodes, containing three underlying communities of 20 nodes each. This modular pattern of underlying influence between oscillators is captured in the influence matrix C, which is generated randomly for each of an ensemble of 20 instances of the network, according to the relevant P_{in} and P_{out} probabilities. The mean of C over this ensemble is depicted for this network in Fig 5.3A, in which the three underlying communities of influence are clearly visible in a block-diagonal arrangement.

The dynamics of the Kuramoto network obeying these influences is then numerically simulated by solving Eq 5.5, and the average synchronization is calculated across each of eight separate time windows of 50 time steps each. Fig 5.3B shows an example synchronization matrix from a single time window in a single network instance. Here, although node pairs from the same underlying community are more synchronized than those from different communities on average, the delineation between communities appears more ambiguous than in the underlying influence matrix from panel A.

The community detection algorithm is then applied to the observed synchronization, for each of the 20 network instances in the ensemble. The performance is summarized in Fig 5.3C, the detection probability matrix D, in which each entry D_{ij} gives the percentage of instances in the ensemble in which nodes i and j are identified as belonging to the same community. (This matrix is averaged over the eight 50-step time windows for

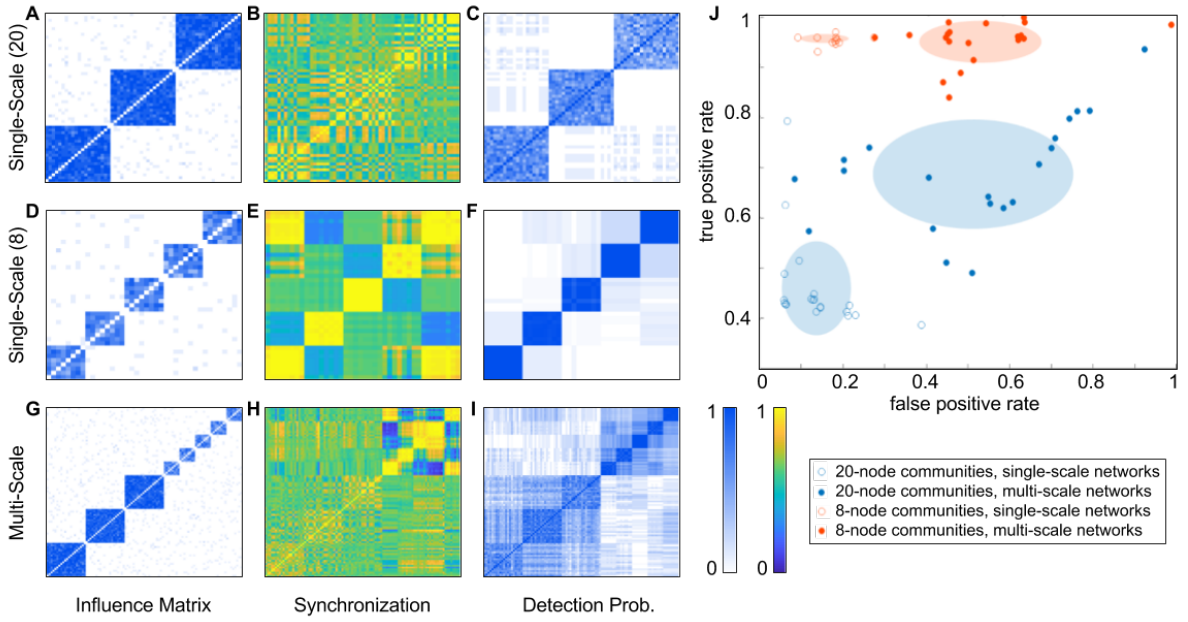

Figure 5.3: **Community detection in single-scale networks outperforms multi-scale networks.** Binary underlying influence matrices, examples of synchronization dynamics, and detection probability matrices for single-scale networks of 20-node communities (A-C), single-scale networks of 8-node communities (D-F), and multi--scale networks with both community sizes (G-I). Each influence matrix is averaged over an ensemble of 20 matrices generated with the same connection probabilities, and each detection probability matrix is averaged over the results from the same ensemble. Panel J summarizes the results, with each dot indicating the true and false positive rates for one network instance from an ensemble. Shaded ellipses highlight the ensemble shapes, with the ellipse centers at the ensemble means and the ellipse axes corresponding to the standard deviations. Unfilled dots represent results for single-scale networks with one community size only, with 20-node communities in blue and 8-node communities in orange. Filled dots represent results for 20-node communities (blue) and 8-node communities (orange) embedded in the multi-scale network of panels G-I. The community detection algorithm shows low false positive rates on single-scale networks, but substantially higher false positive rates on the same communities within multi-scale networks. This indicates that communities of the same size are more difficult to resolve when embedded in multi-scale networks.

each simulation, since the results are very similar across time windows.) The shapes of the underlying communities from the influence matrix are evident here, indicating that the algorithm can detect the underlying communities on average over the whole ensemble. However, the detection probability for individual node pairs is usually less

than 1, meaning that in individual network instances, the algorithm often misses the correct underlying influences.

We compare this performance to that of the same algorithm on a network of ($N = 40$ nodes), consisting of five underlying communities of 8 nodes each. Fig 5.3D, E, and F describe this network in an manner analogous to the previous panels. The average of the binary influence matrix across all networks in the ensemble (Fig 5.3D) again shows a clear delineation of the underlying communities. An example of the synchronization dynamics of the network (Fig 5.3E) shows that the dynamics between node pairs within these smaller communities tend to be more strongly synchronized than within the larger communities from Fig 5.3B. The detection probability matrix in Fig 5.3F indicates that the algorithm detects almost all of the node pairs that share the same underlying community, but also falsely classifies some node pairs as having an underlying influence on each other.

Multi-scale networks. We compare the single-scale networks to the multi-scale network described above, consisting of three 20-node communities and five 8-node communities on underlying influence. Fig 5.3G, H, and I show the binary influence matrix, the synchronization dynamics, and the detection probability matrix for this network, respectively. While the influence matrix and synchronization dynamics look relatively similar in the multi-scale network to the corresponding single-scale networks, the detection probability matrix is notably different. Many false underlying influences are detected, with the algorithm displaying particularly poor performance in distinguishing different communities of the same size from each other.

Fig 5.3J summarizes the changes in community detection performance between single- and multi-scale networks, by plotting the true positive rate for detection of node co-assignments against the false positive rate. For each ensemble of networks, the results from the 20 instances are shown as individual dots, highlighted by an ellipse with its

center at the mean and its semimajor and semiminor axes illustrating the standard deviation of the ensemble results. An instance with perfect detection would fall in the upper left corner ($TPR = 1$ and $FPR = 0$), finding 100% of existing connections and no false positives.

The unfilled dots show performance on single-scale networks, with each representing one instance of a single-scale network with 20-node communities (blue) or 8-node communities (orange). While the community detection algorithm shows low false positive rates on both of these networks, on the 20-node network it displays significantly lower true positive rate, meaning it tends to split up the "true" communities and identify them as smaller subsets.

The filled dots show performance on multi-scale networks. This figure only shows the classification performance on node pairs where both nodes belong to communities of 20 nodes (blue) or where both belong to communities of 8 nodes (orange). The results measure how well the algorithm distinguishes communities *of the same size* from each other, without considering its ability to distinguish between communities of different sizes. This allows direct comparison of results between single-and multi-scale networks for each community size. The algorithm shows a slightly better true positive rate on the 20-node communities within the full multi-scale network, compared to in a single-scale network. However, this comes at a cost of a substantially higher false positive rate. In addition, the detection of 8-node communities deteriorates in a multi-scale network. The algorithm tends to combine multiple 8-node communities into one, producing more false positives than in the single-scale network.

Node Removal

We now demonstrate community detection performance under removal of *a priori*-identified groups of nodes, as schematically represented in Fig 5.2. From the observed

network dynamics, the synchronization matrix is computed while ignoring the removed nodes, creating a new network that includes no direct information about the removed nodes or their connections to the remaining nodes. The community detection method is then applied to these smaller functional networks to produce community assignments for all the remaining nodes.

Removing larger communities. First, we remove the larger communities and test the effect on the ability of the algorithm to identify the smaller communities. Note that in terms of dynamics, the larger communities are also consistently less synchronized compared to the smaller ones. Fig 5.4 shows average influence matrices (A-D) and detection probability matrices (E-H) for the full multi-scale network, compared to the network with one, two, and three of the 20-node communities removed. Fig 5.4I summarizes the changing community detection performance (computed for node pairs in 8-node communities only) as progressively more large communities are removed. Dots indicate performance on individual network instances, with colors distinguishing network ensembles with different numbers of nodes removed, and shaded ellipses highlighting the means and standard deviations of the different ensembles. With one (yellow) or two (red) large communities removed, there is no significant difference in true or false positive rates from the full multi-scale network, shown in orange (one-sample t-test, $p > 0.1$). However, with all three large communities removed (purple), the false-positive rate distribution is significantly different from all three other ensembles, and the true-positive rate distribution is significantly different from the full network ensemble and the ensemble with two large communities removed (t-test, $p < 0.05$, Bonferroni corrected for multiple comparisons).

Removing smaller communities. Next, we remove the smaller, more synchronized communities, and measure the effect on the ability of the algorithm ability to identify

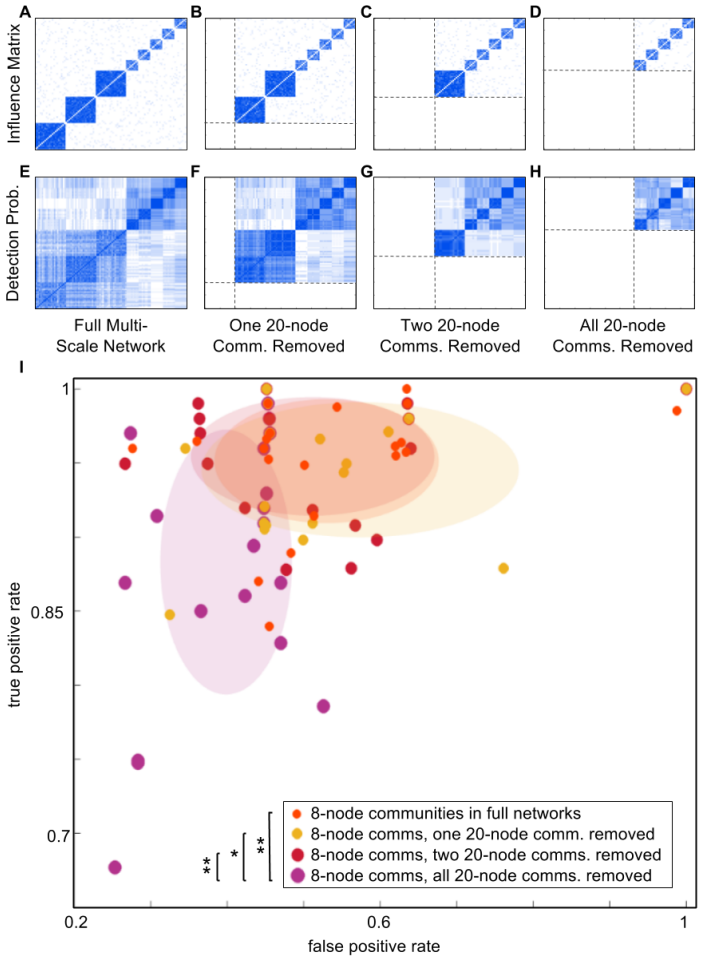

Figure 5.4: **Removing larger communities slightly affects detection of smaller communities.** Removal of larger, less coherent communities during community detection slightly affects detection of smaller, more coherent communities. The figure shows average influence matrices (A-D) and detection probability matrices (E-H) for the full multi-scale networks, as well as the same networks with one, two, and three 20-node communities removed. Panel I summarizes the results, plotting the true versus false positive rates for the detection of co-assignments of pairs of nodes in 8-node communities. Each dot represents results for one instance of an ensemble of 20 networks. A shaded ellipse highlights each ensemble, with the center at the ensemble mean and the axes corresponding to the standard deviations. Successive removal of larger communities (yellow and red distributions) does not make a significant difference in the detection of smaller communities, compared to in the full multi-scale network (orange). Only with all 20-node communities removed (purple) does the distribution of results change. Single stars by the figure legend indicate pairs of distributions that differ significantly in false positives only; double stars indicate a significant difference both true and false positives.

the larger, less synchronized communities. Here, the successive removal of the smaller but more coherently connected 8-node communities substantially improves the ability of the community detection algorithm to distinguish the 20-node communities from each other, as shown in Fig 5.5. Each successive removal decreases false positive rates and increases true positive rates on average. All three distributions differ significantly in false positive rate, and the distribution with all smaller communities removed differs in true positive rate from the full multi-scale network (t-test, $p < 0.05$, Bonferroni corrected for multiple comparisons).

Note that although smaller, the 8-node communities are more cohesively connected overall, tend to be more synchronized, and are substantially easier for the algorithm to identify than the larger communities, both in single-scale networks and in the full multi-scale network. This may reflect that the resolution parameters used, $\gamma = 1$ and $\omega = 1$, are more well-suited to detecting communities at the 8-node scale than the 20-node scale, given the dynamics in these networks.

Summary. For the smaller, more synchronized 8-node communities, the choice of resolution parameter in the context of the other network parameters and dynamics allows relatively clean detection in both the original single-scale network and the full multi-scale network, although the multi-scale network causes an increase in false positive identifications. When the 20-node communities are removed from the multi-scale network, the false positive rates for the 8-node communities drop again, but the performance does not match that in the single-scale network (Fig 5.3J).

The larger and less synchronized 20-node communities are detected with relatively low true positive and false positive rates when in single-scale networks, as the algorithm tends to "split" the larger communities and identify them as smaller subsets. When

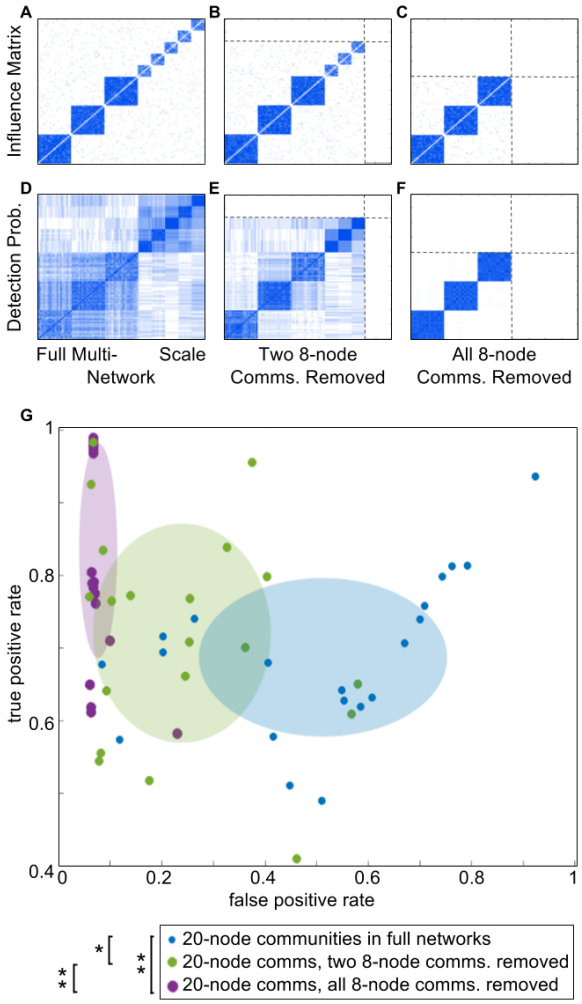

Figure 5.5: **Removing smaller, more coherent communities improves detection of larger, less coherent communities.** Average influence matrices (A-C) and detection probability matrices (D-F) for the full multi-scale network, as well as the same network with two and five 8-node communities removed. Panel G summarizes the results, plotting the true versus false positive rates for the detection of co-assignments of pairs of nodes in 20-node communities. Each dot represents results for one instance of an ensemble of 20 networks. A shaded ellipse highlights each ensemble, with the center at the ensemble mean and the axes corresponding to the standard deviations. Full multi-scale networks are shown in blue; networks with two small communities removed in green; and networks with all five small communities removed in purple. Each successive removal of more 8-node communities substantially improves detection performance, enabling the algorithm to distinguish the 20-node communities from each other. Single stars by the figure legend indicate pairs of distributions that differ significantly in false positives only; double stars indicate a significant difference both true and false positives.

embedded in multi-scale networks with smaller and more synchronized communities, the larger communities are detected with a much higher false positive rate (and only a mild improvement in true positive rate), as the algorithm often groups the larger communities together into an even larger super-community. However, when the smaller 8-node communities are removed from the multi-scale network, the detection of the 20-node communities improves markedly in true positives, while false positives are eliminated almost entirely, and the three underlying communities are cleanly detected in most instances.

Taken as a whole, these results demonstrate that in a relatively simple dynamic system, with highly interconnected communities of influence and low noise, embedding communities within multi-scale networks makes them more difficult to identify than they are in single-scale networks. However, the removal of subsets of network nodes during the community detection process can also dramatically improve detection of the remaining network communities, particularly when removing nodes that belong to highly synchronized communities.

5.3.2 Community Detection and Node Removal in Brain Networks

We now apply these methods to human functional brain networks extracted from two distinct fMRI data sets. As described in Section 5.2.3, the "single-task" data set involves the performance of a single cognitive memory task for three sequential scanning runs, while the "multi-task" data set records the activity during a set of four sequential cognitive tasks, including rest and three different memory- and attention-recruiting tasks. Thus, we expect these data sets to display fundamentally different dynamics.

Spatial and Temporal Community Structure

In this section, we summarize basic characteristics of identified dynamic community structure in the two experiments, with all brain regions included in the analysis. Subsequently, we will consider how these spatial and temporal characteristics are altered by the targeted removal of nodes before community detection.

In the single-task experiment, each participant has a *community number* between 7 and 24, with an average of 12.4. The community structure of one example participant in this experiment is visualized in the top panel of Fig 5.6A. In the multi-task experiment, each participant has between 5 and 12 dynamic communities, with an average of 6.9. The top panel of Fig 5.6B shows the community structure of a single example participant in this experiment.

We observe that for both experiments, brain regions in visual cortex tend to form the most consistently cohesive communities across across time windows, which represent separate functional runs within the experiments. For example, visual cortex regions form the clearly consistent orange community in the top panel of Fig 5.6A, and the corresponding teal community in time windows 2-4 of the top panel of Fig 5.6B. (Time window 1 in the multi-task experiment corresponds to the resting state scan, in which visual functions are not explicitly required.)

We use the *flexibility* (Eq 5.2), which can range from 0 to 1, to quantitatively assess these observations on the consistency of community assignments. The average flexibility over all brain regions and participants is 0.31 for the single-task data, and 0.27 for the multi-task data. These are remarkably similar values given the distinctly different dynamic structure of the experiments, and the presence of task switching in the multi-task experiment only. Fig 5.7 shows the flexibility of each separate brain region, averaged across participants. As reported in [113], and consistent with observations in the example

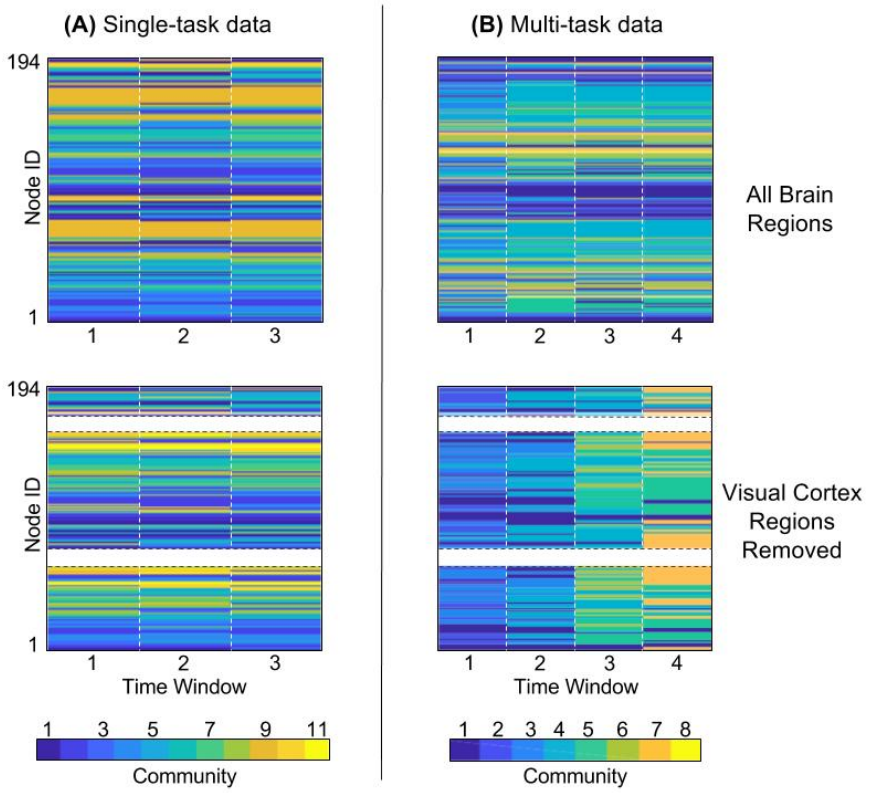

Figure 5.6: **Node removal reveals dynamic changes in community structure in task-based functional brain networks.** A: Visualization of community structure of a single participant during the single-task experiment. Each time window represents a single functional run during which the same lexical memory task is performed. The top panel shows community structure with all brain regions considered; visual cortex, which typically contains the least flexible nodes across functional runs, is visibly grouped into the temporally stable orange community (9). The bottom panel shows community structure with regions in the visual cortex removed from the functional network; these regions are now shown in white and have no community assignment. With the removal of vision nodes, the single-task experiment still shows largely temporally consistent communities. B: Community structure of a single participant in the multi-task experiment, in which each time window represents a single functional run containing a different cognitive task or a resting-state scan. Visual regions, which are explicitly required to perform tasks in time windows 2-4, again form the most stable community with the lowest flexibility (community 4) in the top panel. In the bottom panel, with visual regions removed, distinct differences are seen between the community structure in different tasks, including new communities arising as brain dynamics reconfigure for each new task.

in Fig 5.6, the least flexible regions are largely found in visual and motor cortices. These areas are consistently recruited throughout all three functional runs of the single-task

experiment, and runs 2-4 of the multi-task experiment, which require viewing a lexical stimulus and pressing a button to respond.

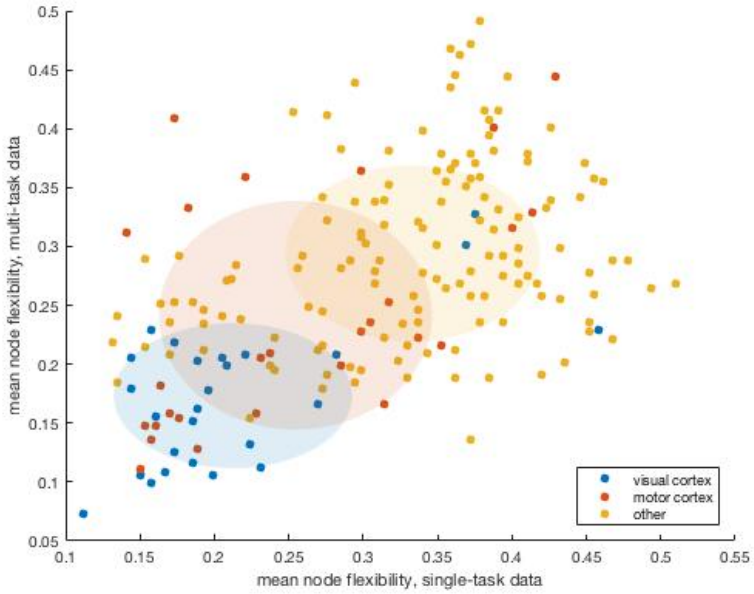

Figure 5.7: **Region-specific flexibility.** Average flexibilities of the 194 network nodes corresponding to brain regions. Each dot represents a single region/node, plotted according to its flexibility in switching between communities across runs of the single-task experiment (x-axis) and the multi-task experiment (y-axis). Flexibility is averaged over all participants in each study. There is a correlation between the flexibility of a brain region in the two different experiments. In particular, visual regions (blue circles) tend to be the least flexible.

This similarity in flexibility distributions, despite the dynamic differences in the experiments, may result from the choice of a temporal resolution parameter that maximizes variation in flexibility across nodes. This effectively broadens the distribution in this direction and increases its likelihood of being centered. A Pearson correlation of $r = 0.5$ ($p < 0.01$) exists between the average flexibility of brain regions in the single-task experiment and the average flexibility of the same regions in the multi-task experiment.

Node Removal from Visual Cortex

Both the single- and multi-task experiments strongly recruit visual brain regions, which consequently form strongly cohesive dynamic communities. However, the dynamic differences between the two experiments are not resolved by the community detection approach, and it is possible that the strength of the visual cortex masks the existence of other functional communities at different spatial scales, as in the synthetic oscillator networks. To determine whether better community resolution is possible in the non-visual brain regions, we repeat the community detection procedure after removing selected subsets of nodes, as shown schematically in Fig 5.2.

We remove the brain regions associated with vision, as classified according to [169]. The bottom panels of Figs 5.6A and B show the community detection results when the visual cortex regions are left out of the network, for a single example participant. The dynamic community structure in single-task data does not appear largely different. However, the multi-task dynamic community structure drastically shifts in this example; without the vision regions to provide a coherent community that persists across functional runs at this scale, rearrangements of brain regions in different task conditions are identified as entirely new communities. Since the resolution parameters in the algorithm are not re-scaled with the removal of vision nodes, these results provide a glimpse into the change in the relevant scales of brain network resolution across functional task conditions.

Resolution of Known Functional Systems

To move beyond qualitative observation of a single example participant, we use the *recruitment coefficient* to quantify the extent to which targeted removal of visual nodes improves the resolution of other functional brain systems.

For each participant, we quantify the extent to which the community detection re-

solves known functional systems by computing the recruitment coefficient for each brain region (Eq 5.7), and the mean *system-specific recruitment* (Eq 5.8) for each of the ten functional systems. Fig 5.8A depicts the locations of the ten functional systems in the brain. Fig 5.8B plots the system-specific recruitment for each of the ten systems, during the first functional run of the single-task experiment. Each bar denotes the mean of the system-specific recruitment over all participants, and the error bars indicate the standard deviation over participants. Blue bars show the recruitment for each functional system when community detection is performed on the entire brain network, while yellow bars show the recruitment when visual regions are removed from the community detection analysis. In the single-task data, node removal does not make a significant difference in the recruitment of most systems, but it does significantly raise recruitment in the dorsal attention system (see statistics in Fig 5.8E). This indicates that brain regions associated primarily with top-down attention are more consistently grouped together into the same community when vision is removed.

Fig 5.8C and D show system-specific recruitment results for two functional runs of the multi-task experiment: the resting state scan and the attention task scan, respectively. (Results from all time windows and tasks in both experiments are included in Appendix D.) Here, the removal of visual cortex regions significantly increases recruitment in almost every remaining functional system. The only exceptions are the ventral attention system in both cognitive states, and the sensory-motor system in the resting state (see details in Fig 5.8F). This indicates that in the multi-task experiment, targeted removal of the highly coherent visual brain regions enables the community detection algorithm to better resolve the involvement of the remaining brain in several other broad functional roles.

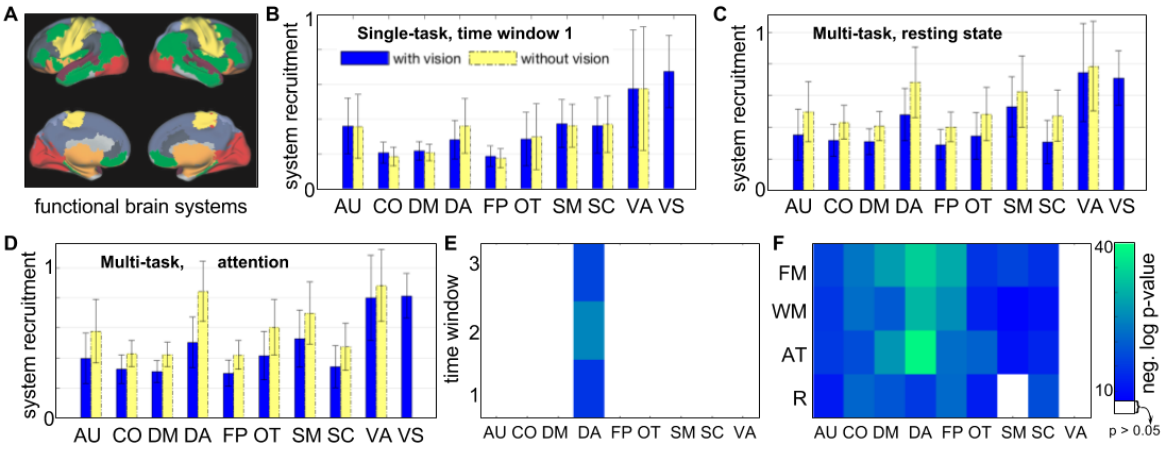

Figure 5.8: **Effect of targeted node removal on resolution of known functional systems.** A: Location in the brain of the ten functional systems: auditory (AU), cingulo-opercular (CO), default mode (DM), dorsal attention (DA), fronto-parietal (FP), somatosensory (SM), subcortical (SC), ventral attention (VA), visual (VS), and other (OT). B: System-specific recruitment coefficients with (blue) and without (yellow) the targeted removal of visual cortex regions, for the first functional run of the single-task experiment. Colored bars show the mean and black error bars the standard deviation over participants in each experiment. C: System-specific recruitment coefficients, analogous to panel B, for the resting state portion of the multi-task experiment. D: System-specific recruitment coefficients for the attention-demanding portion of the multi-task experiment. (Results from all functional runs are displayed in Appendix D.) E: Depiction of systems and time windows in which targeted removal of visual cortex regions leads to significant increase in system-specific recruitment in the single-task experiment (one-sided paired t-test, $p < 0.05$, Bonferroni corrected for multiple comparisons). F: Depiction of systems and tasks/time windows in the multi-task experiment with significant increase of system-specific recruitment after node removal. Tasks include resting state (R), attention (AT), word memory (WM), and face memory (FM). In E and F, colored entries indicate a significant increase, with the color corresponding to the level of significance (negative logarithm of corrected p-value). In the multi-task experiment, targeted node removal significantly increases recruitment in most functional systems. In the single-task experiment, this only occurs in the dorsal attention system.

Resolution of Dynamic Task Structure

Having demonstrated that targeted removal of brain regions in the visual cortex during community detection improves spatial resolution of other functional systems, we use measures of flexibility to assess the effect of this targeted removal on the detection

of task-switching dynamics.

For each participant, the *whole-brain flexibility* is computed by averaging the flexibility of all $N = 194$ brain regions to give a single measure. The distribution of whole-brain flexibility over participants is plotted in Fig 5.9, for community detection performed with all nodes (blue bars) and with the visual cortex nodes removed from the network (yellow bars). Single-task results are shown in the top panel, and multi-task data in the bottom panel. The flexibility distributions derived from all brain regions are similar between the two data sets, which is unsurprising given that the temporal resolution parameters were chosen independently to maximize variation in flexibility for each dataset. However, when visual regions are removed from the network, the differing properties of the two data sets become strongly evident. While the mean of the flexibility distribution across participants in the single-task experiment remains statistically indistinguishable before and after the removal of visual nodes (paired t-test, $p > 0.1$), the mean of the multi-task distribution shifts significantly (paired t-test, $p < 0.001$).

In this case, the removal of visual nodes distinguishes the community structure of these data sets beyond what can be observed through looking at the entire network. The upward shift in flexibility seen when visual nodes are removed from the multi-task dataset suggests that the strong connectivity and coherence within the visual cortex was the key source of temporal consistency in the originally detected communities. Without the consistently low-flexibility community of visual regions, the multi-task nature of this experiment becomes evident, as the community structure becomes better described by large differences between communities across tasks (as in the example in Fig 5.6B).

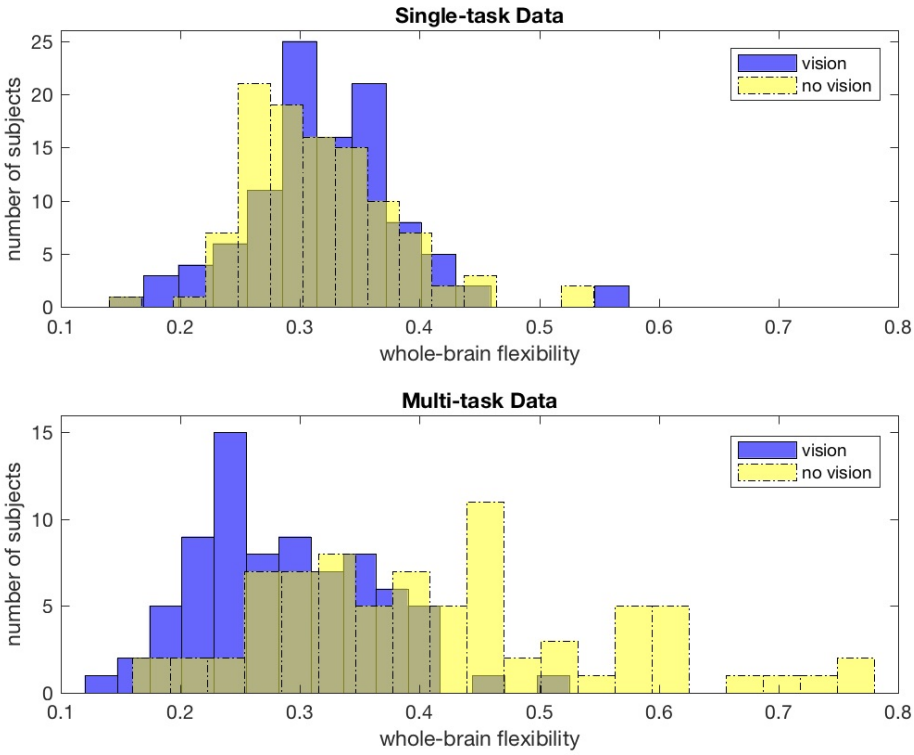

Figure 5.9: **Effect of targeted node removal on resolution of task-switching dynamics.** Whole-brain flexibility distributions as a result of dynamic community detection both with (dotted lines) and without (solid lines) brain regions from visual cortex, with single-task results in blue and multi-task results in red. Single-task distributions do not significantly change with removal of visual regions (paired t-test, $p > 0.1$), but multi-task distributions show a significant shift (paired t-test, $p < 0.001$).

5.4 Discussion

We introduce a technique of targeted node removal during dynamic community detection in complex networks, which can improve the resolution of community structure and dynamics. Using synthetic networks of Kuramoto oscillators, in which the underlying influences between nodes are well-defined, we quantify the performance of a common modularity-maximization community detection algorithm. We show that this algorithm can fail to resolve communities that occur at multiple spatial scales within the same network, compared to its detection of similar communities in single-scale networks. We

also demonstrate that targeted removal of subsets of nodes, especially those that form the most functionally cohesive communities, can improve the resolution of communities among the remaining nodes.

It is important to note that the clusters of underlying influence in this synthetic network serve as a rough proxy for the basic functional organization that underlies the observed brain activity, but should not be interpreted as directly analogous to structural connectivity networks measured in the brain with diffusion MRI. The brains structural connectivity networks are sparse and heavily constrained by their spatial embedding [67], and are not necessarily expected to coincide with the results of community detection on dense functional connectivity networks measured with fMRI.

Instead, we hypothesize that community detection on functional connectivity matrices will provide information on a set of underlying functionally related areas which at least partially drive the observed activity patterns. In the synthetic oscillator networks, we interpret the influence matrix as a representation of this set of underlying functional groupings. In the brain, we use the known functional systems from Ref. [169], representing brain areas with shared functional roles, as a rough proxy for these underlying groupings.

In multi-scale networks like the human brain, it is unlikely that a single resolution parameter choice will elucidate the full community structure of interest, especially if the communities involved span widely disparate scales [177]. In many networks, there are conditions under which no parameter choice can fully distinguish the underlying communities, even when the communities of interest are of similar sizes [189]. In such situations, the choice of a single parameter may cause communities to be obscured. However, the targeted removal of node subsets that are strongly coherent can allow for clearer resolution of the communities formed by the remaining nodes.

To illustrate targeted node removal in a network neuroscience context, we apply it to functional networks derived from fMRI measurements of large-scale human brain activity.

We use data from two distinct fMRI experiments: one involving repeated performance of a single task, and the other consisting of a succession of tasks designed to elicit differing cognitive functions. We expect the identified community structure in the single-task experiment to display consistency across functional runs as the same task is repeated, and that in the multi-task experiment to show rearrangement of communities over time in conjunction with task switching. However, when comparing the two experiments, the established algorithm alone finds no significant difference in the flexibility of brain regions between communities. It does, however, identify a strongly cohesive and inflexible community of nodes in the visual cortex. Upon removing the regions in the visual cortex during dynamic community detection, the ability of the algorithm to resolve certain other known functional systems within the brain significantly improves.

In the single-task experiment, removal of visual nodes significantly improves the recruitment coefficient of the dorsal attention system, composed of brain regions associated with the top-down directing of attention [172]. None of the other examined brain systems show significant improvement in the single-task experiment, possibly indicating that the presence of a strong visual community at this particular scale does not mask other functional dynamics in this experiment.

In the multi-task experiment, removal of visual nodes significantly improves the system-specific recruitment coefficient of almost every other functional system. This suggests that the coherence of the visual regions, especially as one of the few communities that is not expected to experience major shifts across different tasks, had significantly masked the functional coherence of several other functional modules in this experiment. The dorsal attention system again shows the most significant shift in recruitment coefficient, especially in the attention-related task, and to a lesser degree in the memory-related tasks, which also require directed attention. The ventral attention system, which is associated with responses to infrequent or unexpected cues [172], does not show a significant

increase in recruitment; this may be related to the relatively small size of this brain system (4 regions).

Furthermore, upon removal of visual cortex regions, the dynamic differences between these experiments become evident. The average flexibility of brain regions in the multi-task experiment shifts sharply up, reflecting task switching, while the flexibility in the repeated single-task experiment stays the same. These results show that targeted node removal can not only improve the ability of the algorithm to resolve important dynamic changes in community structure, but also allow for meaningful comparisons between data sets through observation of the changes in identified community structure when specific nodes are removed.

The inability of the algorithm to identify the dynamic differences between experiments likely stems in part from the method of selecting the resolution parameters, especially the temporal resolution parameter ω in Eq 5.1. Having chosen a value for the spatial resolution γ, ω is chosen to maximize the variance in flexibility among nodes, as averaged across all participants in the experiment [8, 35]. This favors a broad distribution of flexibilities between 0 and 1, with an average likely to be nearer the center (0.5) than the extremes. Thus, a separate choice of γ and ω for each data set serves to effectively push all whole-network flexibilities toward a similar mean. This minimized differences in whole-network flexibility distributions between data sets, even if their nodes have relatively different flexibility patterns.

However, when the same nodes are removed from multiple networks, the algorithm responds differently if those nodes played different roles in the organization of the networks. As we see in the single-task and multi-task data sets, when highly coherent visual cortex regions are removed from the brain network during community detection, the algorithm is able to successfully resolve differences in community dynamics between the two experiments, recognizing the increased flexibility of brain regions in participants when

switching between cognitive tasks.

Comparing community structures between networks from different participant demographics and experimental conditions is critical for understanding how these conditions shape brain organization, and how the brains organization affects them in turn. Our results demonstrate that even with commonly used heuristics for selecting resolution parameters, there is no guarantee that the resulting community structure will capture information that is comparable between data sets, especially when those data sets are based upon different experimental designs, participants, or imaging protocols. Without "ground truth" knowledge or an understanding of precisely how parameter choices affect the results, there is no clear way to establish an equivalence between parameters – and, by extension, community structures – for different data sets. In the neuroscience data studied here, it is only by performing a *relative* comparison of these two data sets, with and without a subset of regions removed, that the clear dynamic difference between the experiments comes into focus. This method provides a promising approach for performing such relative comparisons among other data sets, in order to resolve the dynamic roles of brain regions within functional networks.

5.4.1 Methodological Considerations

The Kuramoto oscillator simulations in this chapter include a "ground truth" community structure, specified by the influence matrix, which allows us to precisely evaluate the ability of the community detection algorithm to uncover the underlying clusters that fundamentally drive activity patterns. However, these synthetic networks serve as a relatively abstract and simplified model of brain network dynamics, and are limited in their ability to account for more complex features of brain function that are likely to affect community structure. For example, the synthetic influence matrix contains purely topo-

logical communities, in contrast to functional brain networks derived from fMRI data, whose connection weights are substantially influenced by the spatial proximity of brain regions. Although the effects of targeted node removal in the synthetic networks cannot be naïvely assumed to hold true in functional brain networks, our complementary analysis of networks derived from brain data provides evidence that the technique can be practically employed to improve resolution of known functional systems in the brain. It will be an important avenue for future research to develop and investigate synthetic networks that more realistically approximate the properties of functional brain networks, but remain straightforward to simulate and evaluate.

There are various possible strategies for targeted node removal, several of which have been previously proposed. We have focused on removing subsets of nodes defined by previous knowledge of their role in the system under study. For example, the fact that our experiments have visual stimuli provides good reason to expect the brain regions in visual cortex to be especially coherent, although they are not of particular interest to questions about memory and attention. Many fMRI experiments may have similar functional regions worth targeting for removal, or even multiple sets of regions that can be removed hierarchically.

Other strategies for choosing nodes to remove include data-driven removal, in which nodes are removed based on statistics of their network role and connectivity [190]; and model-guided removal, in which nodes are removed in order to optimize a model of underlying communities in the presence of noise [191]. Future work will quantify the performance of these methods, and assess their potential for identifying useful communities in human brain networks.

We have also focused our synthetic network experiments on multi-scale networks in which each node belongs to a single underlying community. This framework can be used in future studies to test community detection performance in more complex network

configurations that are also likely to be relevant in network neuroscience applications. Examples include network nodes that belong to multiple communities, nodes that do not belong to any communities, and underlying communities of influence that change their organization over time.

This chapter is in part an attempt to better understand the behavior of community detection methods that have an inherent scale (often controllable by a resolution parameter), and to address resolution issues that may arise when such methods are applied to networks with varying community sizes, as is common in network neuroscience. It should be noted that since the modularity maximization method used here is subject to an inherent scale and a resolution limit that depends on the overall network size [171,192], our proposed technique of removing highly coherent underlying modules from the network has the potential to reveal new communities among the remaining nodes that are not viable or meaningful communities in the context of the entire network.

We have demonstrated here that our technique works in practice to improve resolution of true underlying modules in synthetic networks, as well as known functional systems in brain networks, on which this modularity maximization method is widely used. We have also shown that our technique can provide important information about dynamics in temporal brain networks. However, recent literature suggests an alternative approach: the use of resolution-limit-free clustering methods, in which the optimal partitioning results for subgraphs are guaranteed to remain the same as in the full graph [192–194]. In will be important in future work to further investigate the behavior of these methods on temporal brain networks, including their ability to reveal underlying information about network dynamics on different time scales.

Finally, we note that we have not applied a threshold to our functional networks in advance of community detection. Weaker associations in functional human brain networks have been shown to contain functionally relevant and predictive information, and can

feasibly occur even in strongly interconnected communities [55, 91, 92]. Our choice not to threshold retains this information, and avoids the issue of choosing a suitable threshold value with no ground truth information. However, retaining low-strength functional associations also increases the influence of noise in the network, which can impact the ability of clustering algorithms to identify communities cleanly. In future work, it will be important to evaluate the effect on community detection performance of combining thresholding with the techniques proposed here.

Appendix A

In this Appendix, we include the following material to support the work described in Chapter 2.

1. An examination of hyperedge properties across analyses that vary either the length or number of samples in time windows used for dynamic network construction.
2. Figure A.1: Comparison of hyperedge size distributions in analyses with differing time window selection procedures.
3. A discussion of hyperedge node degree and its relation to brain region size.
4. Figure A.1: Relation of hyperedge node degree to brain region size.

Examination of time window selection across different TRs

fMRI data was sampled during rest and attention tasks with a TR of 2 seconds, and during memory tasks with a TR of 2.5 seconds. As a result, when choosing time windows for our dynamic network, we could not hold both the time window length and the number of samples per window constant over the entire experiment. To ensure that

a discontinuity in either window length or window samples does not drive our results, we performed two separate analyses: one with a constant time window length of 60 seconds but a change in the number of samples per window, and the other with a constant number of samples in all windows but a change in the window length. The hyperedge size distributions from both analyses are shown in Fig. 1.

The distribution with the number of samples per time window held constant (Fig. 1B) has slightly smaller hyperedges in both the small and large regimes than the distribution with constant window lengths (Fig. 1A), likely driven by the longer time windows (75 rather than 60 seconds) in the memory tasks. Despite these small differences between the distributions, our results are quite robust to variation of the window length selection strategy, showing a very similar distribution shape and overall number of hyperedges in both. We choose to use results from the analysis with constant time window lengths in Chapter 2, so that each time layer in our dynamic network covers the same duration of time.

Correlation of hyperedge node degree with brain region size

The hyperedge node degree is a measure of the number of hyperedges in which a node participates. We expect it to contain information about a given nodes likelihood of co-evolving with other nodes in the network. To check whether a nodes hyperedge degree is influenced by the size of that node, we investigated the correlation between physical node size and hyperedge node degree (presented in Fig. 2). We find that the correlation between size and hyperedge node degree is not extremely strong (Spearman's $\rho = 0.401$, Pearson's $r = 0.324$). However, the p-values for these relationships are highly

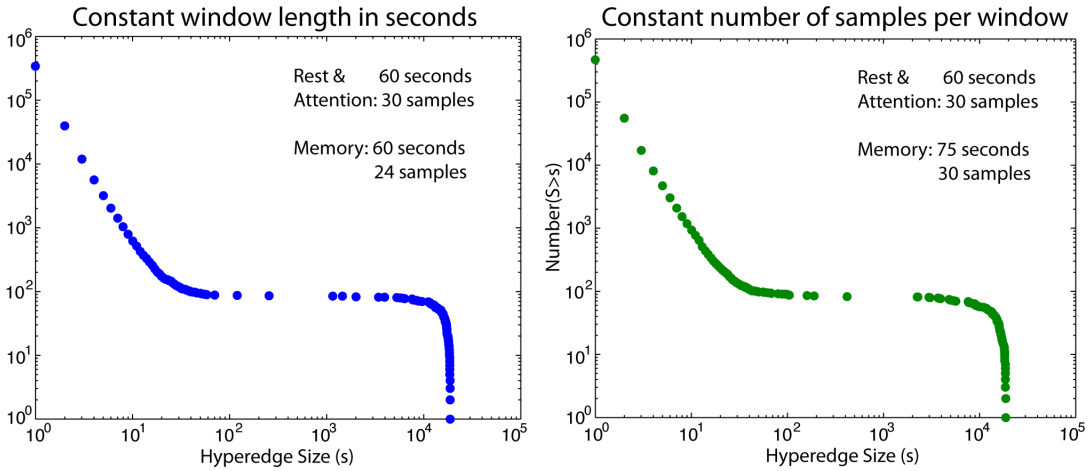

Figure A.1: Cumulative Hyperedge Size Distributions: Left: Cumulative hyperedge size distribution as presented in Chapter 2, from an analysis in which time windows were held at a constant length of 60 seconds across tasks with different TRs. Right: Cumulative hyperedge size distribution from an analysis in which the number of data points used to compute the correlation in each time window was held constant at 30 samples across tasks. In this case, time windows in memory tasks covered 75 seconds. These two analyses produce very similar results, with slightly fewer large hyperedges and more small hyperedges in the right-hand distribution.

statistically significant (Spearman p-value $= 7.01 \times 10^{-9}$, Pearson p-value $= 4.12 \times 10^{-6}$).

While the hyperedge node degree is somewhat influenced by node size, the relatively weak correspondence shown in Fig. 2 indicates that the hyperedge degree is not simply a reflection of node size alone; we also expect it to include information on the extent of the nodes co-evolution with other brain regions. We have explicitly constructed a *hybrid atlas*, as described in detail in Chapter 2, in order to minimize the variation in brain region size across both subjects and regions, and control the effect of variations in node size on our results as much as possible.

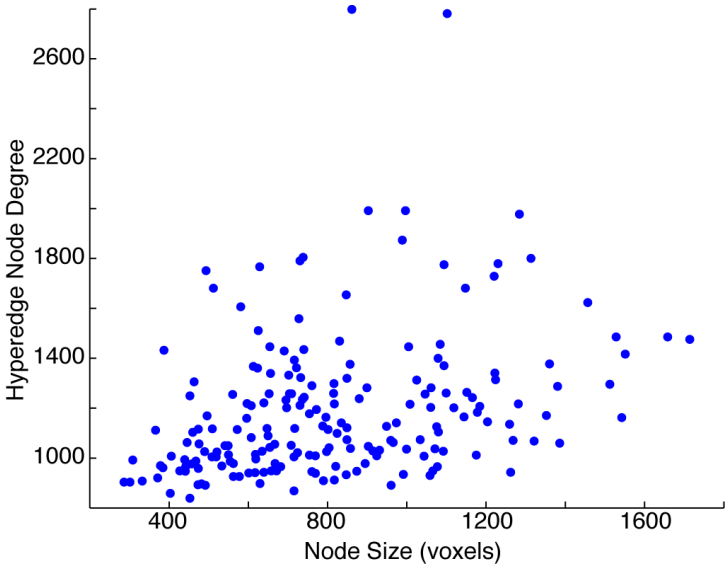

Figure A.2: Relation of hyperedge node degree to brain region size: A scatter plot of the size of each node (brain region) in voxels plotted against its hyperedge node degree. Each voxel is a cubic volume with sides of $2mm$. The two are not especially strongly correlated (Spearman's $\rho = 0.401$, Pearson's $r = 0.324$), but the p-values for these relationships are highly significant (Spearman p-value $= 7.01 \times 10^{-9}$, Pearson p-value $= 4.12 \times 10^{-6}$).

Appendix B

The following information is included in this Appendix to support the results presented Chapter 3.

1. Figure B.1 and Figure B.2: Cumulative size distributions for several methods for minimizing the effect of concatenation.

2. Tables B.1, B.2, and B.3: Tables of individual difference measures grouped by category for the full analysis, multi-task data, and age-memory data.

3. Figure B.3: R^2 changes for the task-specific hypergraph cardinality regression analysis.

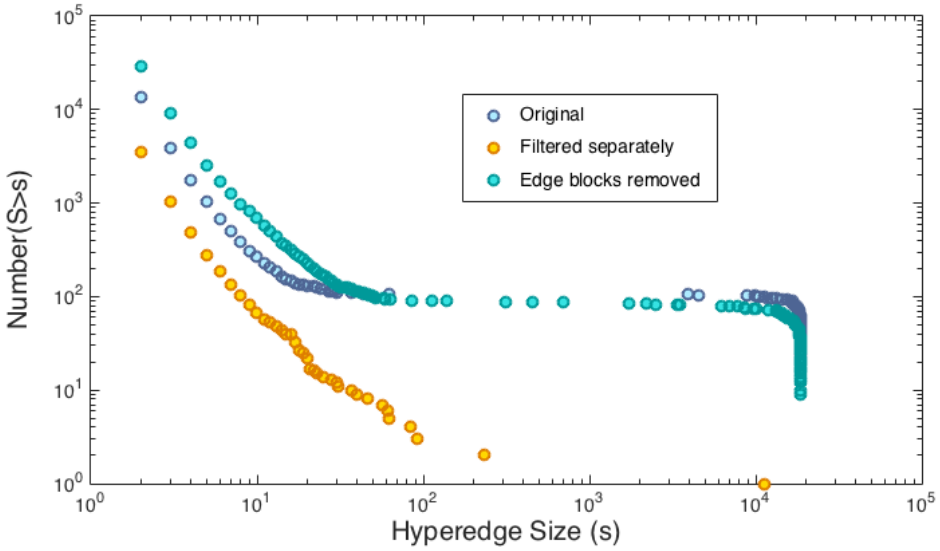

Figure B.1: **Edge compensation comparison:** Cumulative size distributions for the original age-memory data set (with no changes to remove effects of the edges) and two methods for removing potential effects from the edges. The "edge blocks removed" method is used in all analyses in Chapter 3.

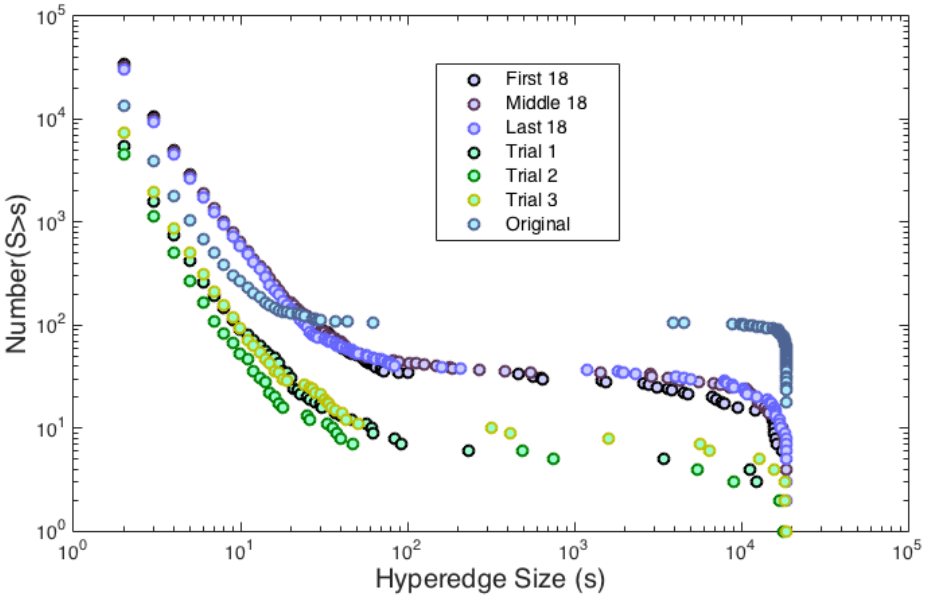

Figure B.2: **Trial separation comparison:** Cumulative size distributions for two different methods for separating edge effects. In the trial-by-trial method, hypergraphs are constructed separately for each trial, while in the 18-split analysis, hypergraphs are constructed from the first, middle, or last 18 edge time series data points.

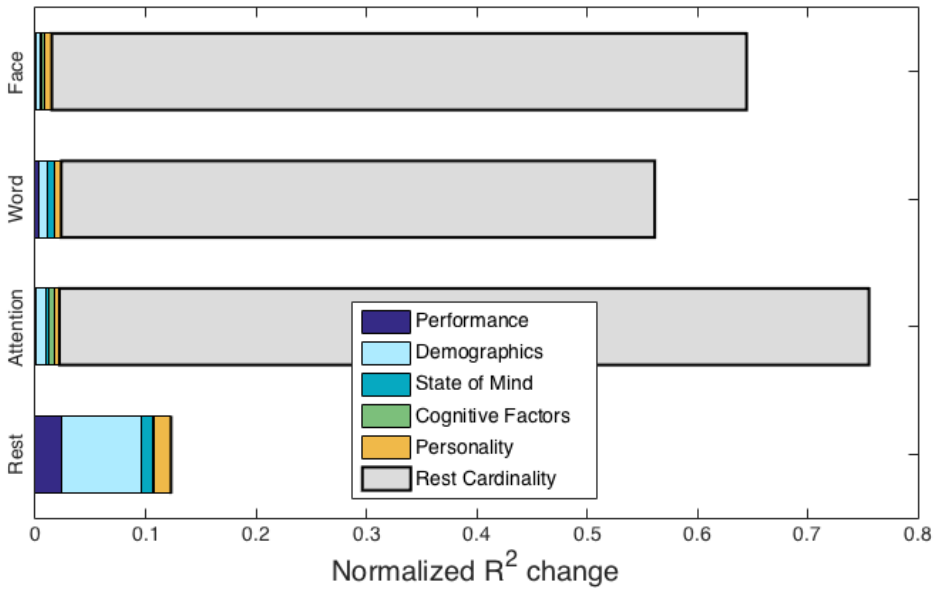

Figure B.3: **Task-specific multi-task R^2 changes:** Normalized R^2 changes with respect to task-specific hypergraph cardinality for each of the four task-specific hypergraphs. Rest-specific hypergraph cardinality is included as an independent variable for the other three tasks and is the only significant predictor, which is denoted with a bold outline.

Performance (Word)	Demographics	Personality	Cognitive Factors	State of Mind
Criterion shift score	Age	PANAS (6)	OSIQ-S/O	Arrival time
Liberal Dprime	Sex	Big 5 (5)	VVQ-W/P	Meal (hours since)
Conservative Dprime	Educ. (yrs)	BIS/BAS (4)	Need for cognition	Hours of sleep
Overall Dprime	Dom. hand		SBCSQ visual	Physical/mental comfort
			SBCSQ verbal	Beck Depression Inventory
			Paper folding	Alcohol (Y/N)
			Card rotation	Exercise (Y/N)
				Smoking (Y/N)
				Caffeine (Y/N)

Table B.1: **Common behavioral measures in both data sets:** Categories containing measures of interest (42). For the state of mind measures, (Y/N) indicates measures where participants were asked whether they had performed the activity in the past 24 hours.

Performance	Demographics	Personality	Cognitive Factors	State of Mind
Attention CS	Military rank	EPQ-R (4)	Working memory	MSW/MSF
Face memory CS			Vocabulary test	PTSD Score
Attention Dprime				PTSD (Y/N)
Face memory Dprime				Concussion score
				Concussion 5 inventory

Table B.2: **Additional behavioral measures in multi-task data:** Categories containing measures of interest. For the state of mind measures, (Y/N) indicates measures where participants were asked whether they had performed the activity in the past 24 hours.

Performance	Demographics	Personality	State of Mind
Hit rates	Height	Distracted	Stressed (Y/N)
Failure rates	Weight	Motivated	Days since period
Reaction time	Contraceptive use		Usual hours of sleep
	Children (Y/N)		Drugs past 48h (Y/N)
	Number of children		MMSE (dementia)

Table B.3: **Additional behavioral and brain measures in age-memory data:** Categories containing measures of interest. For the state of mind activity measures, (Y/N) indicates measures where participants were asked whether they had performed the activity in the past 24 hours. Questions about daily, weekly, and monthly amounts of activity, including whether activity in the past 24 hours were more or less than usual were also recorded for all (Y/N) state of mind activities in the age-memory study.

Appendix C

In this Appendix, we present the following information to support the results presented in Chapter 4.

1. dynamic community structure results in dynamic functional networks with 500-second time windows;

2.1. details on the choices of spatial and temporal resolution parameters for the community detection algorithm;

2.2. details on categorical versus time-ordered definitions of flexibility;

2.3. an analysis of single-node communities, or "singletons," and results when they are excluded from the analysis;

2.4. discussion of behavioral and brain-measures outliers and their effect on the results;

2.5. details and discussion of the statistical correction for mean relative motion.

Results from networks with 500-second time windows

Here we present further results from networks with 500-second time windows, which capture dynamics associated with a much longer time scale than the 80-second time

windows that are the focus of Chapter 4. Overall, these results are mostly qualitatively similar to those found for 80-second time windows, but the correlations between demographics or performance measures and brain community structure measures are weaker.

Brain region flexibility

As discussed in Chapter 4, we find that in networks with 80-second time windows, highly flexible brain regions are very consistently flexible across subjects, while those with lower mean flexibility show greater inter-subject variance in flexibility. On longer timescales, in networks with 500-second time windows, the identities of the brain regions with the lowest mean flexibility – i.e., regions in visual and motor cortex – are largely the same as those found with 80-second time windows (Figure C.1A), and their cross-subject variance is quantitatively similar as well. However, the regions with higher mean flexibility – i.e., non-visual and non-motor regions – are much more variable in flexibility across subjects in networks composed of 500-second time windows. This leads to a strong correlation between mean flexibility and cross-subject variance with 500-second time windows, shown in Figure C.1B, whereas 80-second time windows lead to a similarly strong anticorrelation (as shown in Figure 4.5 in Chapter 4).

The cohesive dynamics of visual and motor systems thus show similar flexibility patterns across subjects even on very different timescales, while individual differences in the dynamics of other brain regions are more strongly impacted by the choice of time resolution. This may be because visual and motor components of the memory task do not differ across trials, leading to strong functional similarity throughout the entire experiment.

Whole-brain flexibility, number of communities, and recruitment

The strong correspondence between age and whole-brain flexibility seen with 80-second time windows, and presented in Chapter 4, is not statistically significant in networks with 500-second time windows (Figure C.2). This indicates that the community dynamics modulated by age here are relevant on shorter timescales of only a few minutes, which correspond more closely to the timescales of cognitive function demanded by the task setup.

With 500-second time windows, there is also no significant correspondence between dynamic community number and age, and the number of distinct communities in individual functional systems is significantly positively correlated with age only in four systems: auditory, somatosensory, subcortical, and ventral attention. The weakening of both the age-flexibility correspondence and the age-community number correspondence on this longer timescale is consistent with the possibility that flexibility and community number are related.

The correspondence between subject age and the mean recruitment coefficient over all brain regions is shown in Figure C.3. There is a highly significant anticorrelation between these measures (Spearman's $\rho = -0.42, p < 0.001$) with both 500-second and 80-second time windows.

System-specific recruitment

In networks with 500-second time windows, system-specific recruitment is significantly anticorrelated with age only in cingulo-opercular (Figure C.4), subcortical, ventral attention, and auditory systems, but not in other systems (such as the visual system, shown in Figure C.5). This effect is consistent across timescales for cingulo-opercular, subcortical, and ventral attention regions.

Methodological details and considerations

Resolution parameters

The spatial resolution parameter γ determines the relative weight given to the randomized null model as compared to the data in each time window when finding a partition. Varying γ changes the number and size of communities found in the partition – higher values of γ favor many small communities, while lower values favor fewer, larger communities. In order to choose a spatial resolution that will give meaningful results about brain organization on the scale of our chosen atlas, we prefer γ values at which the stochastic algorithm tends to produce less variation in partitions across algorithm runs. We measure variation among partitions with the z-score of the Rand coefficient, which measures the extent to which two partitions are similar compared to the expected similarity of randomized partitions [186], averaged over all pairs of partitions produced by the algorithm. It has been shown in simulated networks of oscillators that the lowest cross-subject variance in Rand z-score occurs at the value of γ that produces communities corresponding to the size and number of "ground-truth" communities in the network [35]. However, since human brain functional networks have meaningful activity at various scales, we see no clear maximum in Rand z-score corresponding to a minimum in Rand z-score variance at any single value of γ. In networks with 80-second time windows, we choose $\gamma = 1.2$ – a value between $\gamma = 1$, which often gives just two or three large communities, and $\gamma = 1.4$, which in many subjects gives as many at 100 communities (more than half the total number of nodes) – in order to obtain communities that are on average similar to the size of the functional systems we are interested in. For 500-second time windows, we choose $\gamma = 1.15$ for the same reason. Both of these choices lie in a range of values over which Rand z-score and its variance are relatively uniform, indicating that the consistency of the communities detected does not depend sensitively upon

this parameter. In addition, the Rand z-score is high for all choices, indicating that the community partitions detected are significantly more consistent across these parameter values than would be expected of community partitions with the same community size distributions selected at random.

The time resolution parameter ω determines the relative weight given to intra-window (non-temporal) and inter-window (temporal) considerations when finding a partition. Here, in order to most clearly resolve the differences in the flexibilities of different brain regions, we choose the value of ω that maximizes the variance in flexibility across nodes. This value is $\omega = 0.05$ for 80-second time windows, and $\omega = 0.001$ for 500-second time windows.

Categorical versus time-ordered flexibility

Equation 2 in Chapter 4, reproduced here, defines the metric of flexibility for each brain region:

$$f(i) = \frac{1}{T(T-1)} \sum_{t \neq t'} \left[1 - \delta(g_{it}, g_{it'})\right].$$

This method of calculating flexibility, known as "categorical" flexibility, compares the community assignments of nodes between all possible pairs of time windows, not just time-adjacent windows. Typical uses of categorical flexibility compare community assignments between categories or tasks without considering temporal changes. In this work, we use categorical flexibility to emphasize the consistency of nodes across long time windows. We choose to compare between all time windows equally (without imposing time order) since each window is long compared to the differing elements of the task on a trial or probability block level, and all windows are statistically identical with respect to task design (for 500-second windows) or nearly so (for 80-second windows). In this way we avoid assuming that changes in brain dynamics happen progressively over

the course of the task, but instead focus on assessing stability of community structure over the entire task at once.

All results reported in Chapter 4 use categorical flexibility. For comparison, we repeat our analysis using time-ordered flexibility:

$$f_{to}(i) = \frac{1}{T-1} \sum_{t=1}^{T-1} \left[1 - \delta(g_{it}, g_{i(t+1)})\right].$$

We find that the values of node flexibility and subject-wise whole brain flexibility are extremely closely correlated, as shown in Figure C.6. In addition, all correlations with categorical flexibility reported in Chapter 4 are essentially unchanged when computed with time-ordered flexibility. This suggests that progressive changes in brain dynamics over the course of the task are less important at this time scale than overall consistency or variability of community assignments throughout the experiment, as we might expect with time windows representing multiple statistically similar portions of the same task.

Analysis of single-node communities

As seen in Figure 4.4, the community detection algorithm identifies communities of size 1 in the brain networks of many subjects. As described in Chapter 4, we identify both "dynamic singletons," or single-node communities that contain only one brain region across all time windows, and "static singletons," or communities with a single brain region in one time window, regardless of whether that community also extends across multiple time windows.

We find that across all subjects, there is only a single dynamic singleton identified in our data. Static singletons are more numerous, but still remain sparse. Figure C.7 shows the number of communities identified in each time window for each subject with single-node communities excluded from each time window (cf. Figure 4.3 in Chapter 4).

These statistics appear qualitatively similar for almost all subjects.

A closer look at the cross-subject and cross-region distributions of static singletons is given in Figure C.8. Panel A shows the total number of static singletons (summed over 80-second time windows) in each subject and each brain region. Most static singletons to not tend to persist across time windows, either in particular subjects or in particular regions. However, one subject (subject 35) does have a handful of regions which are consistently singletons in 16 out of 18 time windows; this is very unusual and only occurs once in one other subject (subject 28). In panel B of Figure C.8 – which depicts the number of singletons by subject, with each color representing one brain region – the large contributions from these consistently single regions visibly boost the total singleton count for these two subjects, making them appear as outliers. (More details on outliers are given below.) Panels C and D both show the number of singletons for each brain region. In C, the colors represent the contributions from individual subjects, while in D, the colors represent contributions from the three age groups.

To ensure that singletons do not drive results, we repeat our analyses with these communities excluded from consideration. The correlation between age and number of communities, both overall and in specific functional systems, is nearly unchanged, as shown in Table C.1. Although found in most subjects, static singletons are not significantly correlated with age and do not substantially affect age-related changes in community dynamics.

Analysis of outliers in task performance and brain measures

As noted in Chapter 4, two subjects appear to be bivariate behavioral outliers (see Figure 4.2). In order to ensure that these anomalous performance values do not affect the behavioral correlations, we repeated our analysis with these two subjects removed. We

had originally found no significant correlations, either between the d-prime and criterion shift performance measures, or between either of these measures and the brain measures of interest. With the outliers removed, we similarly find that all Pearson correlations between behavior measures and brain measures, as well as the correlation between d-prime and criterion shift, remain insignificant.

Similarly, two subjects in Figure 4.3 in Chapter 4 appear to have notably higher numbers of static communities than the rest. One of these subjects, subject 35, also has a notably higher number of static singletons, as seen in Figure C.8B, along with another subject who is not an outlier in number of non-singleton static communities. To ensure that these outliers are not driving results, we also repeat our analysis while excluding these three brain-measures outlier subjects. We find that the significance or non-significance of all correlations between brain measures (flexibility, number of communities, recruitment) and age or performance remain the same, both overall and in individual functional systems, with a single exception. That exception is the system-specific recruitment of the subcortical nodes, which is significantly correlated with age with the outliers included (Spearman's $\rho = -0.30$, $p = 2.33 \times 10^{-3}$, as reported in Table 4.1), but not once the outliers were removed (Spearman's $\rho = -0.26$, $p = 9.69 \times 10^{-3}$, which is not significant after correction for multiple comparisons).

Statistical correction for mean relative motion

As discussed in Chapter 4, since subject age is correlated with mean relative motion in these data, we expect motion to substantially affect the correspondence measures of community dynamics and age, and potentially other performance and demographic measures as well, due to the broad and non-uniform distribution of ages in our sample. Thus, all subject-wise correlations in this study are performed with mean relative subject

motion partialed out – i.e., each correlation variable was first regressed separately on mean relative motion, and we assessed the correlation between the residuals of these regressions, to ascertain the extent of their relationship that could not be explained by motion. Some of the observed results are indeed affected by motion, showing a different level of correlation and significance with and without the motion correction. Here we report the differences we observe.

- The correlation between age and whole-brain flexibility is consistently highly significant both with ($r = 0.53,\ p < 0.001$) and without ($r = 0.40,\ p < 0.001$) motion correction in networks with 80-second time windows. Indeed, the correlation is stronger when motion is accounted for. However, in networks with 500-second time windows, a significant correlation ($r = 0.30,\ p < 0.05$) is observed only when not correcting for motion. When motion is accounted for, the correlation is weaker and does not pass the significance test.

- The correlation between age and number of communities evident in networks with 80-second time windows ($r = 0.29,\ p < 0.05$) is not significant without accounting for motion ($p > 0.1$). With 500-second time windows, there is no evident correlation between age and number of communities, and motion does not impact this result.

- The anticorrelation between age and average recruitment observed in networks with 80-second time windows ($r = -0.32,\ p < 0.05$) is not significant without accounting for motion ($p > 0.1$). Similarly, the anticorrelation between age and average recruitment with 500-second time windows is only significant when motion is accounted for.

- Mean relative motion affects the correspondence between system-specific flexibility and age in several systems. Flexibility over 80-second time windows in the dorsal

attention, subcortical, and ventral attention systems shows no significant correlation with age when motion is not partialed out, but does correlate with age when motion is accounted for. On the other hand, flexibility over 80-second time windows in the visual system does not correlate with age when motion is accounted for, but correlates only when motion is not partialed out. The correlation between age and system-specific flexibility over 500-second time windows is not affected by this motion correction in any specific systems.

- Motion also affects system-specific recruitment and its correlation with age. When motion is not accounted for in 80-second time window networks, none of the functional systems have self-recruitments that significantly correlate with age. When motion is not accounted for in 500-second time window networks, three systems show a significant anticorrelation between self-recruitment and age: cingulo-opercular and subcortical, which show the same results with motion partialed out, and fronto-parietal, which does not.

- Overall, we find that mean relative motion is most likely to affect recruitment in small systems (i.e., those composed of fewer brain regions). This is depicted in Figure C.9, which shows a significant correlation between system size and the strength of the correlation between motion and system self-recruitment. However, we do not see a similar relationship between system size and the effect of mean relative motion on system flexibility or number of communities in the system.

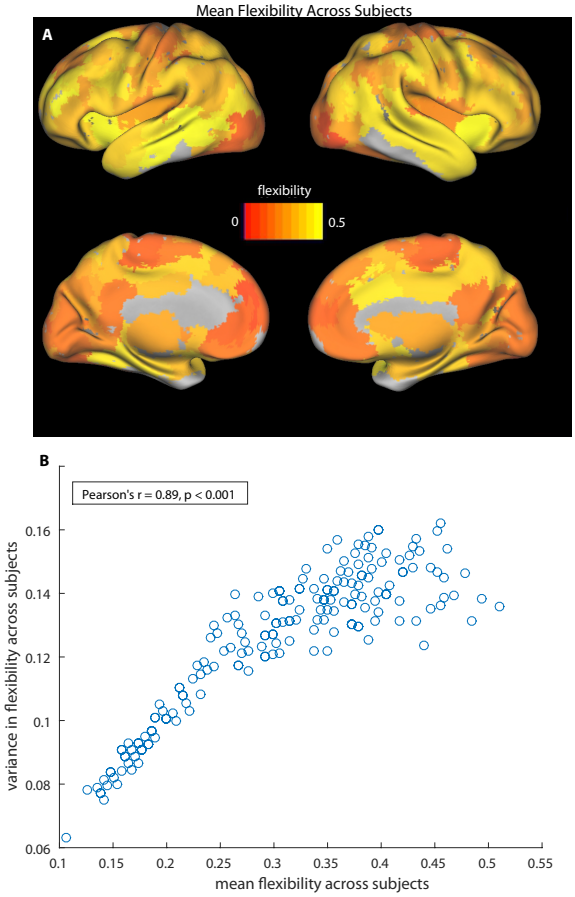

Figure C.1: A: Flexibility of the 194 brain regions used as network nodes in networks with 500-second time windows. Color indicates mean flexibility over $N = 104$ subjects. Visual cortex and somatosensory regions in particular have exceptionally low mean flexibility, as also observed with 80-second time windows. B: Scatter plot of mean region flexibility against variance in region flexibility across subjects. Brain regions that are more flexible on average have a strong tendency to also display higher cross-subject variance in flexibility. This is the opposite effect from that seen in networks with 80-second time windows (Figure 4.5 in Chapter 4), due largely to the much higher cross-subject variability seen in high-flexibility (non-motor, non-visual) regions with 500-second time windows.

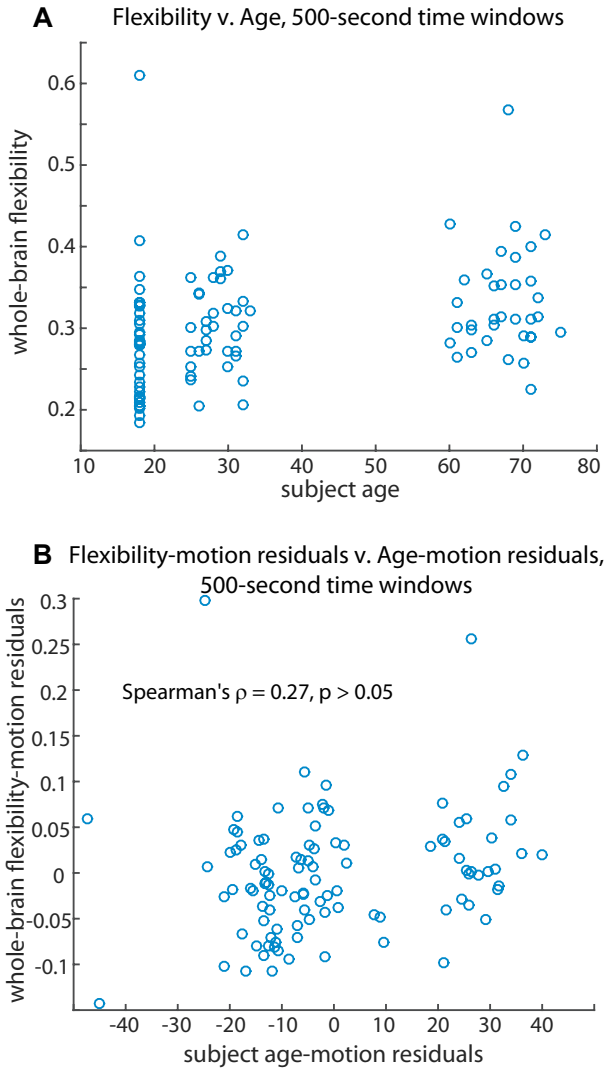

Figure C.2: A: Scatter plot of the uncorrected correspondence between subject age and whole-brain flexibility in networks with 500-second time windows. B: Scatter plot showing residuals of separately regressing each measure on mean relative motion. The correlation between these residuals is not significant, indicating that there is no significant relationship between age and flexibility in these networks that cannot be explained by mean relative motion.

Figure C.3: A: Scatter plot of the uncorrected correspondence between subject age and average recruitment across all brain regions in networks with 500-second time windows. B: Scatter plot showing the significant negative correlation between these measures with mean relative motion partialed out. Older subjects have significantly lower recruitment on average over brain regions than younger subjects on both timescales investigated.

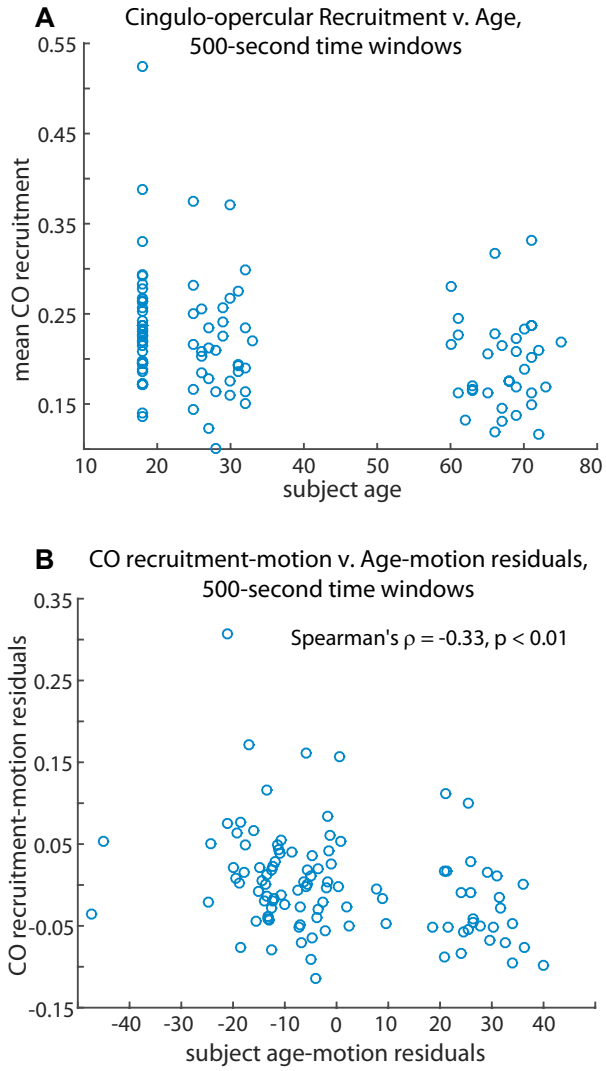

Figure C.4: A: Scatter plot of the uncorrected correspondence between cingulo-opercular system self-recruitment and subject age in networks with 500-second time windows. B: Scatter plot showing a significant anticorrelation between these measures with mean relative motion partialed out. Older subjects have significantly lower cingulo-opercular recruitment coefficients on both timescales investigated; this correspondence is also consistent across timescales in the subcortical and ventral attention systems. However, there are changes in the significance of this correspondence across timescales in some systems. In 500-second time window networks (unlike in 80-second time window networks), the recruitment-age correspondence is not significant in the somatosensory system, and it is significant in the auditory and default mode systems.

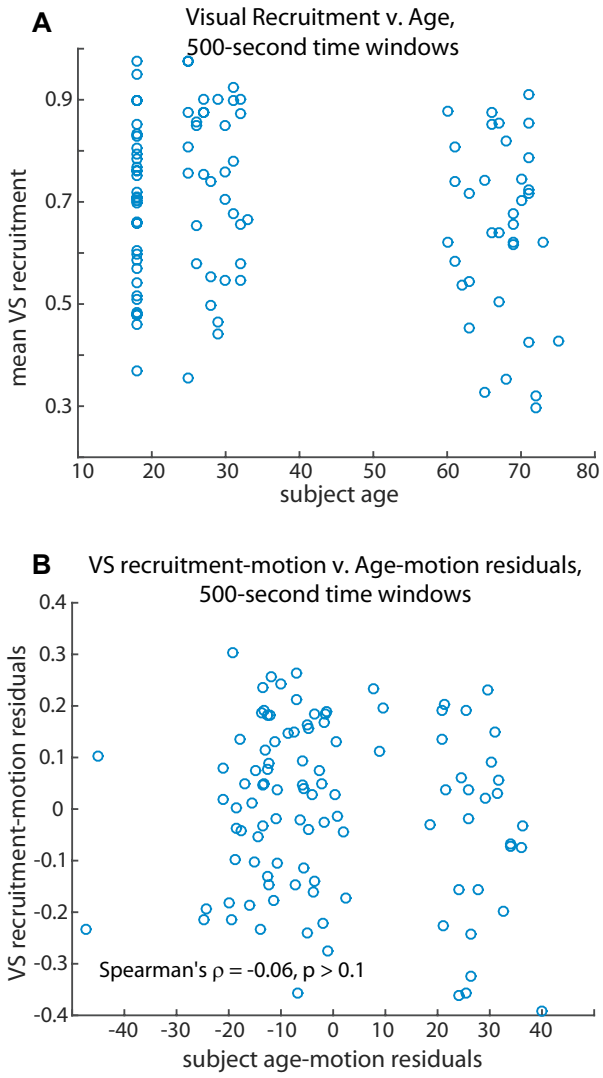

Figure C.5: A: Scatter plot of the correspondence between visual system self-recruitment and subject age in networks with 500-second time windows. B: Scatter plot of the correspondence between visual self-recruitment and age with mean relative motion partialed out; there is no apparent correlation on this coarser timescale, consistent with the result in 80-second time window networks.

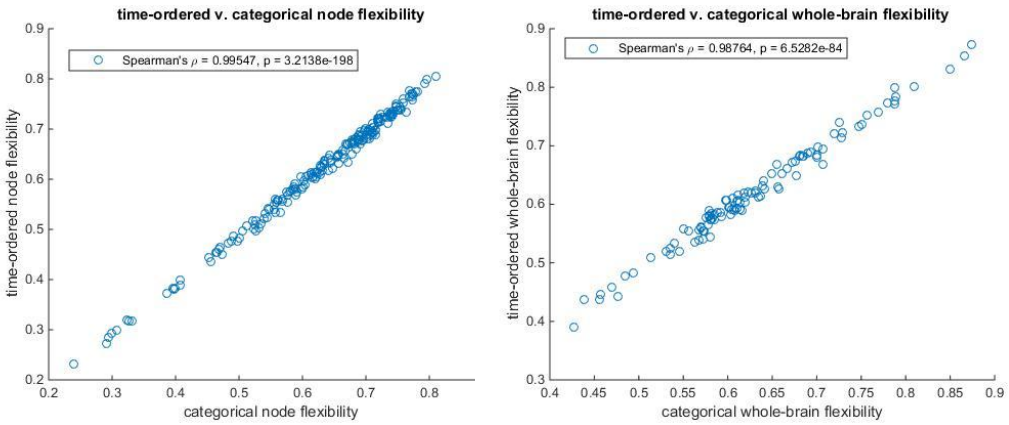

Figure C.6: Correspondence between categorical and time-ordered flexibility. A: Flexibility of each brain region (averaged over subjects). B: Whole-brain flexibility of each subject. Both measures show near-perfect correlation between categorical and time-ordered flexibility (Spearman's $\rho = 0.99$, $p \approx 0$)

	Age v. Community Number (single-node communities excluded)	
	Spearman's ρ	p-value
Whole brain	0.28852	0.0031207
Auditory	0.38723	5.33E-05
Cingulo-opercular	0.36970	1.21E-04
Default Mode	0.40874	1.82E-05
Dorsal Attention	0.34042	4.34E-04
Fronto-parietal	0.33858	4.68E-04
Other	0.35313	2.53E-04
Somatosensory	0.40325	2.41E-05
Subcortical	0.38999	4.66E-05
Ventral Attention	0.39083	4.47E-05
Visual	0.37777	8.36E-05

Table C.1: **Correlations between subject age and number of non-singleton communities.** Spearman rank correlation ρ values and associated p-values for correlations between age and community number, with single-node communities excluded. Mean relative motion has been partialed out of all correlations. All correlations are statistically significant ($p < 0.05$) after family-wise error rate correction for multiple comparisons, and values are essentially unchanged from corresponding values with single-node communities included.

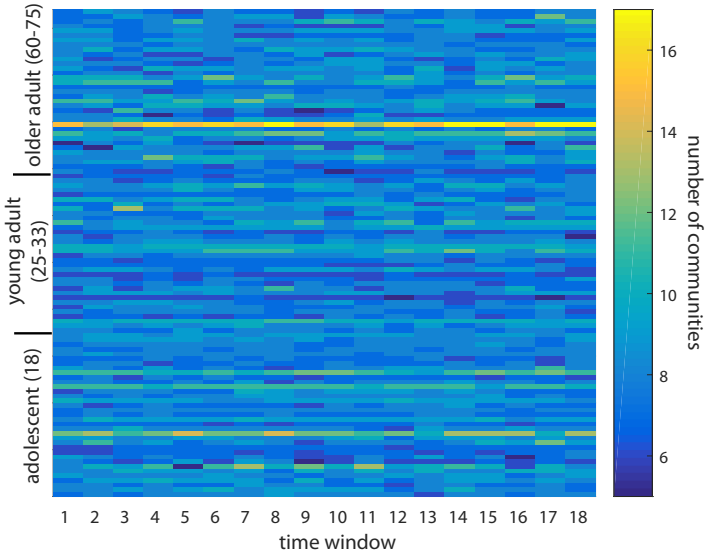

Figure C.7: **Number of non-single-node communities.** Color indicates the number of communities detected within each 80-second time window in each subject, excluding communities composed of only a single brain region within that time window. Subjects (on the vertical axis) are ordered by age. These results are qualitatively similar to the numbers of communities found in each subject with single-node communities included (see Figure 4.3).

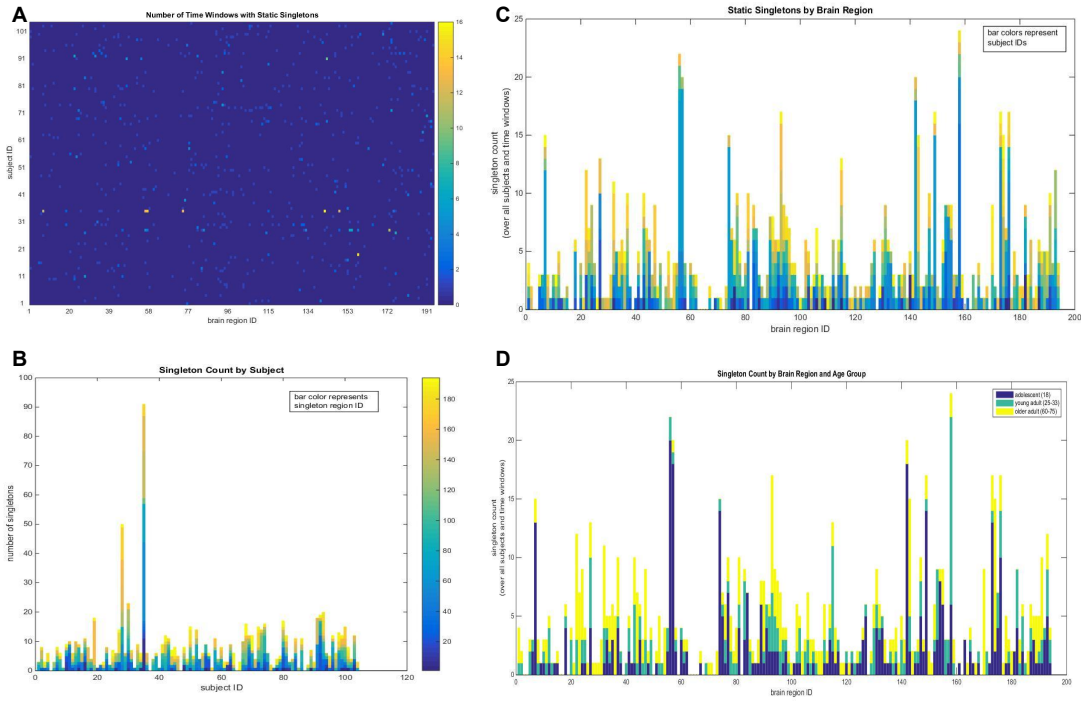

Figure C.8: A: Number of time windows in which each brain region is a static singleton in each subject. Static singletons are relatively sparse, and most regions are not consistently singletons across subjects or time windows. However, two subjects (subjects 28 and 35) have regions which are singletons in most of the 18 80-second time windows. B: Distribution of static singletons over subjects. Colors represent contributions from individual brain regions. Due largely to contributions from just one or two brain regions, subjects 28 and 35 have many more singletons than the others. C: Distribution of static singletons over brain regions. Colors here represent individual subjects. D: Distribution of static singletons over brain regions, as in C. Here, colors represent contributions from one of the three age groups.

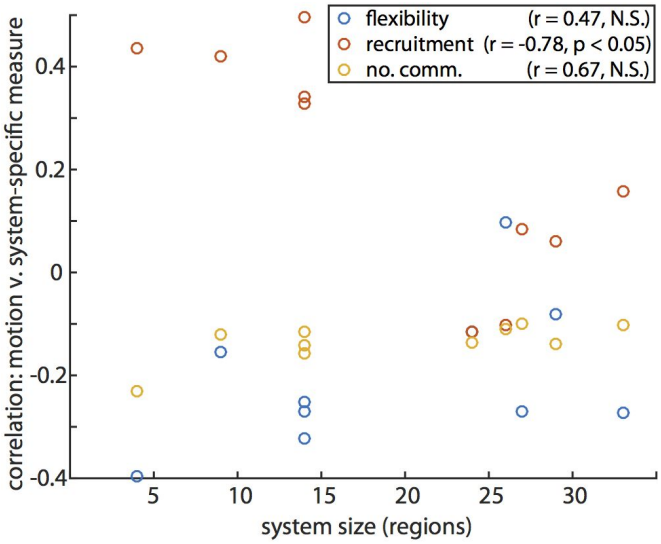

Figure C.9: Scatter plot of the correspondence between functional system size and the effect of motion on system-specific community structure diagnostics. System size reliably predicts the strength with which motion will correlate with system self-recruitment, but this effect is not observed for other diagnostics, such as system flexibility or number of communities in a system.

Appendix D

Effect of visual cortex brain region removal on system-specific recruitment.

As described in Section 5.2.3, we perform community detection on individual functional brain networks, both with and without the targeted removal of brain regions in visual cortex. We use system-specific recruitment (Section 5.2.3) to quantify the overlap of detected communities with known functional regions of the brain [113,169]. Figure 5.8 includes examples of recruitment during different cognitive tasks and experiments, for each functional system in the brain.

For completeness, we present in Figure D.1 the same data for all three time windows of the single-task experiment (with the same word memory task performed in each), as well as for all four cognitive states tested in the multi-task experiment. Colored bars represent the mean recruitment over experiment participants in the corresponding functional system, both for community detection that includes visual cortex (blue bars) and for community detection with targeted removal of visual cortex regions (yellow). Black error bars show the standard deviation over participants. The functional systems include auditory (AU), cingulo-opercular (CO), default mode (DM), dorsal attention (DA),

fronto-parietal (FP), somatosensory (SM), subcortical (SC), ventral attention (VA), visual (VS), and other (OT). Note that panels A, D, and E are exact reproductions of subfigures in Chapter 5.

Panels H and I summarize that statistics of these recruitment comparisons, and are also reproductions of subfigures in Chapter 5. White entries denote functional systems in which a one-sided paired t-test found no significant increase in recruitment in the corresponding time window or cognitive task. Here, significance is defined as $p < 0.05$, after a Bonferroni correction for multiple comparisons. Colored entries denote significant increases in recruitment after targeted removal of visual cortex regions, and the color represents the negative logarithm of the p-value. In the multi-task experiment, targeted node removal significantly increases recruitment in most functional systems. In the single-task experiment, this only occurs in the dorsal attention system.

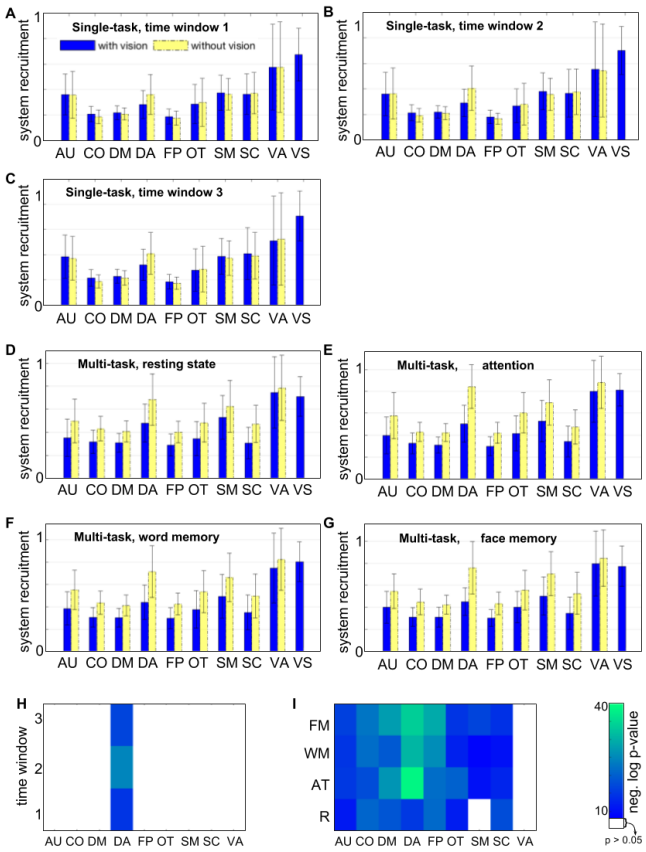

Figure D.1: **Effect of targeted node removal on resolution of known functional systems.** System-specific recruitment coefficients with (blue) and without (yellow) the targeted removal of visual cortex regions, for the ten functional systems. Colored bars show the mean and black error bars the standard deviation over participants in each experiment. Panels A, B, and C show the three functional runs of the single-task experiment. All three runs consist of the same recognition memory task with lexical stimuli, and the runs are treated as three time windows in the dynamic functional brain networks. Panels D-G show the four time windows of the multi–task experiment, with each window encompassing a different task or cognitive state. These include resting state (D), an attention-demanding task (E), a recognition memory task with lexical stimuli (F), and a recognition memory task with face stimuli (G). H: Depiction of systems and time windows in which targeted removal of visual cortex regions leads to significant increase in system-specific recruitment in the single-task experiment. I: Depiction of systems and tasks (resting state (R), attention (AT), word memory (WM), and face memory (FM)) in the multi-task experiment with significant increase of system-specific recruitment after node removal. In H and I, colored entries indicate a significant increase, with the color corresponding to the level of significance (negative logarithm of corrected p-value).

Bibliography

[1] H. Damasio, T. Grabowski, R. Frank, A. M. Galaburda, A. R. Damasio, *et. al.*, *The return of phineas gage: clues about the brain from the skull of a famous patient*, *Science* **264** (1994), no. 5162 1102–1105.

[2] B. J. Baars and N. M. Gage, *Cognition, brain, and consciousness: Introduction to cognitive neuroscience*. Academic Press, 2010.

[3] W. B. Scoville and B. Milner, *Loss of recent memory after bilateral hippocampal lesions*, *Journal of neurology, neurosurgery, and psychiatry* **20** (1957), no. 1 11.

[4] M. S. Gazzaniga, J. E. Bogen, and R. W. Sperry, *Observations on visual perception after disconnexion of the cerebral hemispheres in man*, *Brain* **88** (1965), no. 2 221–236.

[5] M. S. Gazzaniga and R. W. Sperry, *Language after section of the cerebral commissures*, *Brain* **90** (1967), no. 1 131–148.

[6] R. B. Buxton, *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge university press, 2009.

[7] E. Bullmore and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems*, *Nature Reviews Neuroscience* **10** (2009), no. 3 186–198.

[8] O. Sporns and R. F. Betzel, *Modular brain networks*, *Annual review of psychology* **67** (2016) 613–640.

[9] D. S. Bassett and E. Bullmore, *Small-world brain networks*, *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry* **12** (Dec., 2006) 512–523.

[10] D. Meunier, R. Lambiotte, and E. T. Bullmore, *Modular and hierarchically modular organization of brain networks*, *Frontiers in neuroscience* **4** (2010) 200.

[11] D. S. Bassett, N. F. Wymbs, M. P. Rombach, M. A. Porter, P. J. Mucha, and S. T. Grafton, *Task-based core-periphery organization of human brain dynamics*, *PLoS Computational Biology* **9** (2013), no. 9 e1003171.

[12] A. M. Hermundstad, D. S. Bassett, K. S. Brown, E. M. Aminoff, D. Clewett, S. Freeman, A. Frithsen, A. Johnson, C. M. Tipper, M. B. Miller, S. T. Grafton, and J. M. Carlson, *Structural foundations of resting-state and task-based functional connectivity in the human brain*, *Proceedings of the National Academy of Sciences* **110** (2013), no. 15 6169–6174.

[13] A. M. Hermundstad, K. S. Brown, D. S. Bassett, E. M. Aminoff, A. Frithsen, A. Johnson, C. M. Tipper, M. B. Miller, S. T. Grafton, and J. M. Carlson, *Structurally-constrained relationships between cognitive states in the human brain*, *PLoS Computational Biology* **10** (2014), no. 5 e1003591.

[14] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, *Dynamic reconfiguration of human brain networks during learning*, *Proceedings of the National Academy of Sciences* **108** (2011), no. 18 7641–7646.

[15] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, *Community structure in time-dependent, multiscale, and multiplex networks*, *Science* **328** (2010), no. 5980 876–878.

[16] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, and S. T. Grafton, *Cross-linked structure of network evolution*, *Chaos* **24** (Mar., 2014) 013112.

[17] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, *A default mode of brain function*, *Proceedings of the National Academy of Sciences* **98** (2001), no. 2 676–682.

[18] J. Damoiseaux, S. Rombouts, F. Barkhof, P. Scheltens, C. Stam, S. M. Smith, and C. Beckmann, *Consistent resting-state networks across healthy subjects*, *Proceedings of the national academy of sciences* **103** (2006), no. 37 13848–13853.

[19] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, *The human brain is intrinsically organized into dynamic, anticorrelated functional networks*, *Proceedings of the National Academy of Sciences* **102** (2005), no. 27 9673–9678.

[20] M. Hampson, N. Driesen, J. K. Roth, J. C. Gore, and R. T. Constable, *Functional connectivity between task-positive and task-negative brain areas and its relation to working memory performance*, *Magnetic Resonance Imaging* **28** (2010), no. 8 1051–1057.

[21] K. J. Friston, *Functional and effective connectivity: a review*, *Brain Connectivity* **1** (2011), no. 1 13–36.

[22] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, and E. Bullmore, *Adaptive reconfiguration of fractal small-world human brain functional networks*, *Proceedings of the National Academy of Sciences* **103** (2006) 19518–19523.

[23] M. Mennes, C. Kelly, X. N. Zuo, A. Di Martino, B. B. Biswal, F. X. Castellanos, and M. P. Milham, *Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity*, *NeuroImage* **50** (2010), no. 4 1690–1701.

[24] M. W. Cole, J. R. Reynolds, J. D. Power, G. Repovs, A. Anticevic, and T. S. Braver, *Multi-task connectivity reveals flexible hubs for adaptive task control*, *Nature Neuroscience* **16** (2013), no. 9 1348–1355.

[25] M. Mennes, C. Kelly, S. Colcombe, F. X. Castellanos, and M. P. Milham, *The extrinsic and intrinsic functional architectures of the human brain are not equivalent*, *Cerebral Cortex* **23** (2013), no. 1 223–229.

[26] M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, and S. E. Petersen, *Intrinsic and task-evoked network architectures of the human brain*, *Neuron* **83** (July, 2014) 238–251.

[27] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, *Dynamic functional connectivity: promise, issues, and interpretations*, *NeuroImage* **80** (Oct., 2013) 360–378.

[28] M. Ekman, J. Derrfuss, M. Tittgemeyer, and C. J. Fiebach, *Predicting errors from reconfiguration patterns in human brain networks*, *Proceedings of the National Academy of Sciences* **109** (2012), no. 41 16714–16719.

[29] K. W. Doron, D. S. Bassett, and M. S. Gazzaniga, *Dynamic network structure of interhemispheric coordination*, *Proceedings of the National Academy of Sciences* **109** (2012), no. 46 18661–18668.

[30] F. Siebenhuhner and D. S. Bassett, *Multiscale Analysis and Nonlinear Dynamics. From Genes to the Brain*, vol. In Press, ch. Multiscale Network Organization in the Human Brain. Wiley & Sons, 2013.

[31] J. R. Cohen, C. L. Gallen, E. G. Jacobs, T. G. Lee, and M. D'Esposito, *Quantifying the reconfiguration of intrinsic networks during working memory*, *PloS ONE* **9** (2014), no. 9 e106636.

[32] R. P. Monti, P. Hellyer, D. Sharp, R. Leech, C. Anagnostopoulos, and G. Montana, *Estimating time-varying brain connectivity networks from functional MRI time series*, *NeuroImage* **103** (2014) 427–443.

[33] D. S. Bassett, M. Yang, N. F. Wymbs, and S. T. Grafton, *Learning-induced autonomy of sensorimotor systems*, *Nature Neuroscience* **18** (2015), no. 5 744–751.

[34] A. V. Mantzaris, D. S. Bassett, N. F. Wymbs, E. Estrada, M. A. Porter, P. J. Mucha, S. T. Grafton, and D. J. Higham, *Dynamic network centrality summarizes learning in the human brain*, *Journal of Complex Networks* **1** (2013), no. 1 83–92.

[35] D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, and P. J. Mucha, *Robust detection of dynamic community structure in networks*, *Chaos* **23** (Mar., 2013) 013142.

[36] F. Samaria and A. Harter, *Parameterisation of a stochastic model for human face identification*, 1994. 2nd IEEE Workshop on Applications of Computer Vision. Sarasota (Florida).

[37] A. Martinez and R. Benavente, *The AR face database*, 1998. CVC Technical Report no. 24.

[38] P. Peer, "Computer Vision Laboratory Face Database, University of Ljubljana, Slovenia."

[39] F. Solina, P. Peer, B. Batagelj, S. Juvan, and J. Kova, *Color-based face detection in the '15 seconds of fame' art installation*, in *Mirage 2003: Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, Image Analysis and Graphical Special Effects*, pp. 38–47, March, 2003.

[40] M. Minear and D. Park, *A lifespan database of adult facial stimuli*, *Behaviour Research Methodology Instrumentation Computer* **36** (2004) 630–633.

[41] B. Weyrauch, J. Huang, B. Heisele, and V. Blanz, *Component-based face recognition with 3D morphable models*, 2004. First IEEE Workshop on Face Processing in Video, Washington, D.C.

[42] E. M. Aminoff, D. Clewett, S. Freeman, A. Frithsen, C. Tipper, A. Johnson, S. T. Grafton, and M. B. Miller, *Individual differences in shifting decision criterion: a recognition memory study*, *Memory & Cognition* **40** (Oct., 2012) 1016–1030.

[43] A. Daducci, S. Gerhard, A. Griffa, A. Lemkaddem, L. Cammoun, X. Gigandet, R. Meuli, P. Hagmann, and J.-P. Thiran, *The connectome mapper: an open-source processing pipeline to map connectomes with MRI*, *PLoS ONE* **7** (2012), no. 12 e48121.

[44] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, *Mapping the structural core of human cerebral cortex*, *PLoS Biology* **6** (2008), no. 7 e159.

[45] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, *et. al.*, *Advances in functional and structural MR image analysis and implementation as FSL*, *NeuroImage* **23** (2004) S208–S219.

[46] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith, *Bayesian analysis of neuroimaging data in FSL*, *NeuroImage* **45** (2009), no. 1 S173–S186.

[47] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, *FSL*, *NeuroImage* **62** (Aug., 2012) 782–790.

[48] R. W. Cox, *AFNI: software for analysis and visualization of functional magnetic resonance neuroimages*, *Computers and Biomedical Research* **29** (June, 1996) 162–173.

[49] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, *Improved optimization for the robust and accurate linear registration and motion correction of brain images*, *NeuroImage* **17** (Oct., 2002) 825–841.

[50] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, *Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion*, *NeuroImage* **59** (2012), no. 3 2142–2154.

[51] H. J. Jo, S. J. Gotts, R. C. Reynolds, P. A. Bandettini, A. Martin, R. W. Cox, and Z. S. Saad, *Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state fMRI*, *Journal of Applied Mathematics* (2013).

[52] D. N. Greve and B. Fischl, *Accurate and robust brain image alignment using boundary-based registration*, *NeuroImage* **48** (Oct., 2009) 63–72.

[53] A. Zalesky, A. Fornito, I. H. Harding, L. Cocchi, M. Yücel, C. Pantelis, and E. T. Bullmore, *Whole-brain anatomical networks: does the choice of nodes matter?*, *Neuroimage* **50** (2010), no. 3 970–983.

[54] M.-E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore, *Functional connectivity and brain networks in schizophrenia*, *The Journal of Neuroscience* **30** (2010), no. 28 9477–9487.

[55] D. S. Bassett, B. G. Nelson, B. A. Mueller, J. Camchong, and K. O. Lim, *Altered resting state complexity in schizophrenia*, *NeuroImage* **59** (2012), no. 3 2196–2207.

[56] F. T. Sun, L. M. Miller, and M. D'Esposito, *Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data*, *NeuroImage* **21** (2004), no. 2 647–658.

[57] J. A. Cadzow, *Discrete-time systems: An introduction with interdisciplinary applications*, .

[58] A. Fornito, A. Zalesky, and E. T. Bullmore, *Network scaling effects in graph analytic studies of human resting-state fMRI data*, *Frontiers in Systems Neuroscience* **4** (2010) 22.

[59] C. R. Genovese, N. A. Lazar, and T. Nichols, *Thresholding of statistical maps in functional neuroimaging using the false discovery rate*, *NeuroImage* **15** (Apr., 2002) 870–878.

[60] T. Evans and R. Lambiotte, *Line graphs, link partitions, and overlapping communities*, *Physical Review E* **80** (2009), no. 1 016105.

[61] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, *Link communities reveal multiscale complexity in networks*, *Nature* **466** (2010), no. 7307 761–764.

[62] M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, and R. Lambiotte, *Memory in network flows and its effects on spreading dynamics and community detection*, *arXiv preprint arXiv:1305.4807* (2013).

[63] N. Eagle, A. S. Pentland, and D. Lazer, *Inferring friendship network structure by using mobile phone data*, *Proceedings of the national academy of sciences* **106** (2009), no. 36 15274–15278.

[64] Y. Hochberg, *A sharper Bonferroni procedure for multiple tests of significance*, *Biometrika* **75** (Dec., 1988) 800–802.

[65] O. Sporns and J. D. Zwi, *The small world of the cerebral cortex*, *Neuroinformatics* **2** (June, 2004) 145–162.

[66] M. P. v. d. Heuvel, C. J. Stam, R. S. Kahn, and H. E. H. Pol, *Efficiency of functional brain networks and intellectual performance*, *The Journal of Neuroscience* **29** (June, 2009) 7619–7624.

[67] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. W. Moore, and E. T. Bullmore, *Efficient physical embedding of topologically complex information processing networks in brains and computer circuits*, *PLoS computational biology* **6** (2010), no. 4 e1000748.

[68] X.-Y. Long, X.-N. Zuo, V. Kiviniemi, Y. Yang, Q.-H. Zou, C.-Z. Zhu, T.-Z. Jiang, H. Yang, Q.-Y. Gong, L. Wang, K.-C. Li, S. Xie, and Y.-F. Zang, *Default mode network as revealed with multiple methods for resting-state functional MRI analysis*, *Journal of Neuroscience Methods* **171** (June, 2008) 349–355. PMID: 18486233.

[69] N. B. Albert, E. M. Robertson, P. Mehta, and R. C. Miall, *Resting state networks and memory consolidation*, *Communicative & Integrative Biology* **2** (2009), no. 6 530–532. PMID: 20195459 PMCID: PMC2829828.

[70] M. Dastjerdi, B. L. Foster, S. Nasrullah, A. M. Rauschecker, R. F. Dougherty, J. D. Townsend, C. Chang, M. D. Greicius, V. Menon, D. P. Kennedy, and J. Parvizi, *Differential electrophysiological response during rest, self-referential, and non-self-referential tasks in human posteromedial cortex*, *Proceedings of the National Academy of Sciences* **108** (Feb., 2011) 3023–3028. PMID: 21282630.

[71] P. Fransson and G. Marrelec, *The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis*, *NeuroImage* **42** (Sept., 2008) 1178–1184.

[72] F. Cauda, G. Geminiani, F. D'Agata, K. Sacco, S. Duca, A. P. Bagshaw, and A. E. Cavanna, *Functional connectivity of the posteromedial cortex*, *PLoS ONE* **5** (Sept., 2010) e13107.

[73] M. Corbetta and G. L. Shulman, *Control of goal-directed and stimulus-driven attention in the brain*, *Nature Reviews Neuroscience* **3** (Mar., 2002) 201–215. PMID: 11994752.

[74] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, *Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems*, *Proceedings of the National Academy of Sciences* **103** (June, 2006) 10046–10051. PMID: 16788060.

[75] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius, *Dissociable intrinsic connectivity networks for salience processing and executive control*, *The Journal of Neuroscience* **27** (Feb., 2007) 2349–2356. PMID: 17329432 PMCID: PMC2680293.

[76] A. C. Nobre, G. N. Sebestyen, D. R. Gitelman, M. M. Mesulam, R. S. Frackowiak, and C. D. Frith, *Functional localization of the system for visuospatial attention using positron emission tomography.*, *Brain* **120** (Mar., 1997) 515–533. PMID: 9126062.

[77] M. Corbetta, F. M. Miezin, G. L. Shulman, and S. E. Petersen, *A PET study of visuospatial attention*, *The Journal of Neuroscience* **13** (Mar., 1993) 1202–1226. PMID: 8441008.

[78] J. B. Hopfinger, M. H. Buonocore, and G. R. Mangun, *The neural mechanisms of top-down attentional control*, *Nature Neuroscience* **3** (Mar., 2000) 284–291.

[79] S. Shomstein and M. Behrmann, *Goal-directed attentional orienting in patients with dorsal parietal lesions*, *Journal of Vision* **5** (Sept., 2005) 690–690.

[80] I. Levy, U. Hasson, G. Avidan, T. Hendler, and R. Malach, *Center-periphery organization of human object areas*, *Nature Neuroscience* **4** (May, 2001) 533–539.

[81] A. Nestor, M. Behrmann, and D. C. Plaut, *The neural basis of visual word form processing: a multivariate investigation*, *Cerebral Cortex* **23** (July, 2013) 1673–1684. PMID: 22693338.

[82] P. E. Turkeltaub, G. F. Eden, K. M. Jones, and T. A. Zeffiro, *Meta-analysis of the functional neuroanatomy of single-word reading: method and validation*, *NeuroImage* **16** (July, 2002) 765–780. PMID: 12169260.

[83] T. A. Polk and M. J. Farah, *Functional MRI evidence for an abstract, not perceptual, word-form area*, *Journal of Experimental Psychology* **131** (Mar., 2002) 65–72. PMID: 11900104.

[84] A. M. Rauschecker, R. F. Bowen, L. M. Perry, A. M. Kevan, R. F. Dougherty, and B. A. Wandell, *Visual feature-tolerance in the reading network*, *Neuron* **71** (Sept., 2011) 941–953.

[85] A. C. Vogel, F. M. Miezin, S. E. Petersen, and B. L. Schlaggar, *The putative visual word form area is functionally connected to the dorsal attention network*, *Cerebral Cortex* (June, 2011) bhr100. PMID: 21690259.

[86] J. V. Haxby, E. A. Hoffman, and M. I. Gobbini, *Human neural systems for face recognition and social communication*, *Biological Psychiatry* **51** (Jan., 2002) 59–67.

[87] N. Kanwisher, J. McDermott, and M. M. Chun, *The fusiform face area: a module in human extrastriate cortex specialized for face perception*, *The Journal of Neuroscience* **17** (June, 1997) 4302–4311.

[88] I. Gauthier, M. J. Tarr, J. Moylan, P. Skudlarski, J. C. Gore, and A. W. Anderson, *The fusiform "face area" is part of a network that processes faces at the individual level*, *Journal of Cognitive Neuroscience* **12** (May, 2000) 495–504. PMID: 10931774.

[89] G. McCarthy, A. Puce, J. C. Gore, and T. Allison, *Face-specific processing in the human fusiform gyrus*, *Journal of Cognitive Neuroscience* **9** (Oct., 1997) 605–610.

[90] Z. M. Saygin, D. E. Osher, K. Koldewyn, G. Reynolds, J. D. E. Gabrieli, and R. R. Saxe, *Anatomical connectivity patterns predict face selectivity in the fusiform gyrus*, *Nature Neuroscience* **15** (Feb., 2012) 321–327. PMID: 22197830.

[91] M. W. Cole, T. Yarkoni, G. Repovš, A. Anticevic, and T. S. Braver, *Global connectivity of prefrontal cortex predicts cognitive control and intelligence*, *Journal of Neuroscience* **32** (2012), no. 26 8988–8999.

[92] E. Schneidman, M. J. Berry, R. S. II, and W. Bialek, *Weak pairwise correlations imply strongly correlated network states in a neural population*, *Nature* **440** (2006), no. 7087 1007.

[93] A. Zalesky, A. Fornito, and E. Bullmore, *On the use of correlation as a measure of network connectivity*, *NeuroImage* **60** (2012), no. 4 2096–2106.

[94] I. Tavor, O. P. Jones, R. Mars, S. Smith, T. Behrens, and S. Jbabdi, *Task-free mri predicts individual differences in brain activity during task performance*, *Science* **352** (2016), no. 6282 216–220.

[95] D. Zhang and M. E. Raichle, *Disease and the brain's dark energy*, *Nature Reviews Neurology* **6** (2010), no. 1 15–28.

[96] M. Greicius, *Resting-state functional connectivity in neuropsychiatric disorders*, *Current Opinion in Neurology* **21** (2008), no. 4 424–430.

[97] E. L. Dennis and P. M. Thompson, *Functional brain connectivity using fMRI in aging and Alzheimer's disease*, *Neuropsychology Review* **24** (2014), no. 1 49–62.

[98] J. A. Contreras, J. Goñi, S. L. Risacher, O. Sporns, and A. J. Saykin, *The structural and functional connectome and prediction of risk for cognitive impairment in older adults*, *Current Behavioral Neuroscience Reports* **2** (2015), no. 4 234–245.

[99] R. Sala-Llonch, D. Bartrés-Faz, and C. Junqué, *Reorganization of brain networks in aging: a review of functional connectivity studies*, *Frontiers in Psychology* **6** (2015).

[100] K. Onoda, M. Ishihara, and S. Yamaguchi, *Decreased functional connectivity by aging is associated with cognitive decline*, *Journal of Cognitive Neuroscience* **24** (Nov., 2012) 2186–2198.

[101] D. Tomasi and N. D. Volkow, *Aging and functional brain networks*, *Molecular Psychiatry* **17** (2012), no. 5 549–558.

[102] L. Wang, L. Su, H. Shen, and D. Hu, *Decoding lifespan changes of the human brain using resting-state functional connectivity MRI*, *PloS ONE* **7** (2012), no. 8 e44530.

[103] J. Song, R. M. Birn, M. Boly, T. B. Meier, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, *Age-related reorganizational changes in modularity and functional connectivity of human brain networks*, *Brain Connectivity* **4** (2014), no. 9 662–676.

[104] L. K. Ferreira, A. C. B. Regina, N. Kovacevic, M. d. G. M. Martin, P. P. Santos, C. de Godoi Carneiro, D. S. Kerr, E. Amaro, A. R. McIntosh, and G. F. Busatto, *Aging effects on whole-brain functional connectivity in adults free of cognitive and psychiatric disorders*, *Cerebral Cortex* (2015) 190.

[105] L. Geerligs, R. J. Renken, E. Saliasi, N. M. Maurits, and M. M. Lorist, *A Brain-Wide Study of Age-Related Changes in Functional Connectivity*, *Cerebral Cortex* **25** (July, 2015) 1987–1999.

[106] K. K. Ng, J. C. Lo, J. K. W. Lim, M. W. L. Chee, and J. Zhou, *Reduced functional segregation between the default mode network and the executive control network in healthy older adults: a longitudinal study*, *NeuroImage* **133** (June, 2016) 321–330.

[107] M. Y. Chan, D. C. Park, N. K. Savalia, S. E. Petersen, and G. S. Wig, *Decreased segregation of brain systems across the healthy adult lifespan*, *Proceedings of the National Academy of Sciences* **111** (2014), no. 46 E4997–E5006.

[108] P. Holme and J. Saramäki, *Temporal networks*, *Physics reports* **519** (2012), no. 3 97–125.

[109] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, *The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery*, *Neuron* **84** (2014), no. 2 262–274.

[110] J. Gonzalez-Castillo, C. W. Hoy, D. A. Handwerker, M. E. Robinson, L. C. Buchanan, Z. S. Saad, and P. A. Bandettini, *Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns*, *Proceedings of the National Academy of Sciences* **112** (July, 2015) 8762–8767.

[111] E. N. Davison, K. J. Schlesinger, D. S. Bassett, M.-E. Lynall, M. B. Miller, S. T. Grafton, and J. M. Carlson, *Brain network adaptability across task states*, *PLoS Computational Biology* **11** (2015), no. 1 e1004029.

[112] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear, *Time-resolved resting-state brain networks*, *Proceedings of the National Academy of Sciences* **111** (2014), no. 28 10341–10346.

[113] K. J. Schlesinger, B. O. Turner, B. A. Lopez, M. B. Miller, and J. M. Carlson, *Age-dependent changes in task-based modular organization of the human brain*, *NeuroImage* **146** (2017) 741–762.

[114] J. Qin, S.-G. Chen, D. Hu, L.-L. Zeng, Y.-M. Fan, X.-P. Chen, and H. Shen, *Predicting individual brain maturity using dynamic functional connectivity*, *Frontiers in Human Neuroscience* **9** (2015).

[115] A. T. Beck, N. Epstein, G. Brown, and R. A. Steer, *An inventory for measuring clinical anxiety: psychometric properties.*, *Journal of Consulting and Clinical Psychology* **56** (1988), no. 6 893.

[116] L. J. Massa and R. E. Mayer, *Testing the ATI hypothesis: should multimedia instruction accommodate verbalizer-visualizer cognitive style?*, *Learning and Individual Differences* **16** (2006), no. 4 321–335.

[117] O. Blajenkova, M. Kozhevnikov, and M. A. Motes, *Object-spatial imagery: a new self-report imagery questionnaire*, *Applied Cognitive Psychology* **20** (2006), no. 2 239–263.

[118] J. T. Cacioppo and R. E. Petty, *The need for cognition.*, *Journal of Personality and Social Psychology* **42** (1982), no. 1 116.

[119] A. Richardson, *Verbalizer-visualizer: a cognitive style dimension*, *Journal of Mental Imagery* **1** (1977) 109–126.

[120] R. B. Ekstrom, J. W. French, and H. H. Harman, *Manual for kit of factor referenced cognitive tests*. Educational Testing Service, Princeton, NJ, 1976.

[121] O. P. John and S. Srivastava, *The Big Five trait taxonomy: history, measurement, and theoretical perspectives*, *Handbook of Personality: Theory and Research* **2** (1999), no. 1999 102–138.

[122] C. S. Carver and T. L. White, *Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales.*, *Journal of Personality and Social Psychology* **67** (1994), no. 2 319.

[123] D. Watson, L. A. Clark, and A. Tellegen, *Development and validation of brief measures of positive and negative affect: the PANAS scales.*, *Journal of Personality and Social Psychology* **54** (1988), no. 6 1063.

[124] MATLAB, *version 8.5.0 (R2015a)*. The MathWorks Inc., Natick, Massachusetts, 2015.

[125] S. W. Raudenbush and A. S. Bryk, *Hierarchical linear models: applications and data analysis methods*, vol. 1. Sage, 2002.

[126] M. B. Miller, C.-L. Donovan, C. M. Bennett, E. M. Aminoff, and R. E. Mayer, *Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals*, *NeuroImage* **59** (2012), no. 1 83–93.

[127] B. O. Turner, B. Lopez, T. Santander, and M. B. Miller, *One dataset, many conclusions: Bold variabilitys complicated relationships with age and motion artifacts*, *Brain Imaging and Behavior* **9** (2015), no. 1 115–127.

[128] S. M. Smith, *Fast robust automated brain extraction*, *Human Brain Mapping* **17** (2002), no. 3 143–155.

[129] M. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. Smith., *Bayesian analysis of neuroimaging data in FSL*, *NeuroImage* **45** (2009), no. 2 73–86.

[130] M. Jenkinson and S. Smith, *A global optimisation method for robust affine registration of brain images*, *Medical Image Analysis* **5** (2001), no. 2 143–156.

[131] M. Jenkinson, P. Bannister, J. Brady, and S. Smith, *Improved optimisation for the robust and accurate linear registration and motion correction of brain images*, *NeuroImage* **17** (2002), no. 2 825–841.

[132] R. M. Birn, E. K. Molloy, R. Patriat, T. Parker, T. B. Meier, G. R. Kirk, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, *The effect of scan length on the reliability of resting-state fMRI connectivity estimates*, *NeuroImage* **83** (2013) 550–558.

[133] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty, *Resting-state functional connectivity reflects structural connectivity in the default mode network*, *Cerebral Cortex* **19** (2009), no. 1 72–78.

[134] D. S. Bassett, E. Bullmore, B. A. Verchinski, V. S. Mattay, D. R. Weinberger, and A. Meyer-Lindenberg, *Hierarchical organization of human cortical networks in health and schizophrenia*, *The Journal of Neuroscience* **28** (2008), no. 37 9239–9248.

[135] L. Geerligs, N. M. Maurits, R. J. Renken, and M. M. Lorist, *Reduced specificity of functional connectivity in the aging brain during task performance*, *Human Brain Mapping* **35** (Jan., 2014) 319–330.

[136] D. J. Madden, M. C. Costello, N. A. Dennis, S. W. Davis, A. M. Shepler, J. Spaniol, B. Bucur, and R. Cabeza, *Adult age differences in functional connectivity during executive control*, *NeuroImage* **52** (Aug., 2010) 643–657.

[137] D. Meunier, S. Achard, A. Morcom, and E. Bullmore, *Age-related changes in modular organization of human brain functional networks*, *NeuroImage* **44** (Feb., 2009) 715–723.

[138] R. F. Betzel, L. Byrge, Y. He, J. Goñi, X.-N. Zuo, and O. Sporns, *Changes in structural and functional connectivity among resting-state networks across the human lifespan*, *NeuroImage* **102** (Nov., 2014) 345–357.

[139] L. K. Ferreira and G. F. Busatto, *Resting-state functional connectivity in normal brain aging*, *Neuroscience & Biobehavioral Reviews* **37** (Mar., 2013) 384–400.

[140] J. D. Medaglia, M.-E. Lynall, and D. S. Bassett, *Cognitive network neuroscience*, *Journal of Cognitive Neuroscience* **27** (Aug., 2015) 1471–1491.

[141] R. Sala-Llonch, C. Junqué, E. M. Arenaza-Urquijo, D. Vidal-Piñeiro, C. Valls-Pedret, E. M. Palacios, S. Domènech, A. Salvà, N. Bargalló, and D. Bartrés-Faz, *Changes in whole-brain functional networks and memory performance in aging*, *Neurobiology of Aging* **35** (2014), no. 10 2193–2202.

[142] R. F. Betzel, B. Mišić, Y. He, J. Rumschlag, X.-N. Zuo, and O. Sporns, *Functional brain modules reconfigure at multiple scales across the human lifespan.* arXiv preprint arXiv:1510.08045, 2015.

[143] A. Irimia and J. D. Van Horn, *Scale-dependent variability and quantitative regimes in graph-theoretic representations of human cortical networks*, *Brain Connectivity* **6** (2016), no. 2 152–163.

[144] K. Wei, M. Cieslak, C. Greene, S. T. Grafton, and J. M. Carlson, *Sensitivity analysis of human brain structural network construction.* Submitted, 2016.

[145] J. Wang, L. Wang, Y. Zang, H. Yang, H. Tang, Q. Gong, Z. Chen, C. Zhu, and Y. He, *Parcellation-dependent small-world brain functional networks: a resting-state fMRI study*, *Human Brain Mapping* **30** (2009), no. 5 1511–1523.

[146] R. Magalhães, P. Marques, J. Soares, V. Alves, and N. Sousa, *The impact of normalization and segmentation on resting-state brain networks*, *Brain Connectivity* **5** (2015), no. 3 166–176.

[147] D. A. Balota, P. O. Dolan, and J. M. Duchek, *Memory changes in healthy older adults*, *The Oxford handbook of memory* (2000) 395–409.

[148] C. L. Grady and F. I. Craik, *Changes in memory processing with age*, *Current opinion in neurobiology* **10** (2000), no. 2 224–231.

[149] N. J. Cepeda, A. F. Kramer, and J. Gonzalez de Sather, *Changes in executive control across the life span: examination of task-switching performance.*, *Developmental psychology* **37** (2001), no. 5 715.

[150] R. West, K. J. Murphy, M. L. Armilio, F. I. Craik, and D. T. Stuss, *Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control*, *Brain and cognition* **49** (2002), no. 3 402–419.

[151] F. H. Treitz, K. Heyder, and I. Daum, *Differential course of executive control changes during normal aging*, *Aging, Neuropsychology, and Cognition* **14** (2007), no. 4 370–393.

[152] C. Grady, S. Sarraf, C. Saverino, and K. Campbell, *Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks*, *Neurobiology of Aging* **41** (2016) 159–172.

[153] H.-Y. Zhang, W.-X. Chen, Y. Jiao, Y. Xu, X.-R. Zhang, and J.-T. Wu, *Selective vulnerability related to aging in large-scale resting brain networks*, *PLoS ONE* **9** (2014), no. 10 e108807.

[154] R. Sala-Llonch, C. Junqué, E. M. Arenaza-Urquijo, D. Vidal-Piñeiro, C. Valls-Pedret, E. M. Palacios, S. Domènech, A. Salvà, N. Bargalló, and D. Bartrés-Faz, *Changes in whole-brain functional networks and memory performance in aging*, *Neurobiology of Aging* **35** (2014), no. 10 2193–2202.

[155] J. Gomez-Ramirez, Y. Li, Q. Wu, and J. Wu, *A quantitative study of network robustness in resting-state fMRI in young and elder adults*, *Frontiers in Aging Neuroscience* **7** (2015).

[156] J. Gonzalez-Castillo, D. A. Handwerker, M. E. Robinson, C. W. Hoy, L. C. Buchanan, Z. S. Saad, and P. A. Bandettini, *The spatial structure of resting state connectivity stability on the scale of minutes*, *Frontiers in Neuroscience* **8** (2014) 138.

[157] E. C. Hansen, D. Battaglia, A. Spiegler, G. Deco, and V. K. Jirsa, *Functional connectivity dynamics: modeling the switching behavior of the resting state*, *NeuroImage* **105** (2015) 525–535.

[158] X. Di and B. B. Biswal, *Dynamic brain functional connectivity modulated by resting-state networks*, *Brain Structure and Function* **220** (2015), no. 1 37–46.

[159] J. Dubois, *Brain age: A state-of-mind? On the stability of functional connectivity across behavioral states*, *The Journal of Neuroscience* **36** (2016), no. 8 2325–2328.

[160] L. L. Jacoby, A. J. Bishara, S. Hessels, and J. P. Toth, *Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults*, *Journal of Experimental Psychology: General* **134** (2005), no. 2 131.

[161] E. Aminoff, S. Freeman, D. Clewett, C. Tipper, A. Frithsen, A. Johnson, S. Grafton, and M. Miller, *Maintaining a cautious state of mind during a recognition test: A large-scale fmri study*, *Neuropsychologia* **67** (2015) 132–147.

[162] F. Siebenhühner, S. A. Weiss, R. Coppola, D. R. Weinberger, and D. S. Bassett, *Intra-and inter-frequency brain network structure in health and schizophrenia*, *PLoS ONE* **8** (2013), no. 8 e72351.

[163] E. Bullmore and O. Sporns, *The economy of brain network organization*, *Nature Reviews Neuroscience* **13** (2012), no. 5 336–349.

[164] D. S. Bassett, E. T. Bullmore, A. Meyer-Lindenberg, J. A. Apud, D. R. Weinberger, and R. Coppola, *Cognitive fitness of cost-efficient brain functional networks*, *Proceedings of the National Academy of Sciences* **106** (2009), no. 28 11747–11752.

[165] M. F. Folstein, S. E. Folstein, and P. R. McHugh, *"Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician*, *Journal of Psychiatric Research* **12** (1975), no. 3 189–198.

[166] A. Grinsted, J. C. Moore, and S. Jevrejeva, *Application of the cross wavelet transform and wavelet coherence to geophysical time series*, *Nonlinear Processes in Geophysics* **11** (2004), no. 5/6 561–566.

[167] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, *Fast unfolding of communities in large networks*, *Journal of statistical mechanics: theory and experiment* **2008** (2008), no. 10 P10008.

[168] S. Gu, F. Pasqualetti, M. Cieslak, S. T. Grafton, and D. S. Bassett, *Controllability of brain networks*, *arXiv preprint arXiv:1406.5197* (2014).

[169] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, *et. al.*, *Functional network organization of the human brain*, *Neuron* **72** (2011), no. 4 665–678.

[170] J. D. Power, B. L. Schlaggar, C. N. Lessov-Schlaggar, and S. E. Petersen, *Evidence for hubs in human functional brain networks*, *Neuron* **79** (2013), no. 4 798–813.

[171] S. Fortunato and M. Barthelemy, *Resolution limit in community detection*, *Proceedings of the National Academy of Sciences* **104** (2007), no. 1 36–41.

[172] S. Vossel, J. J. Geng, and G. R. Fink, *Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles*, *Neuroscientist* **20** (2014) 150–159.

[173] S. Sadaghlani and M. D'Esposito, *Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness*, *Cerebral Cortex* **25** (2015) 2763–73.

[174] R. Cabeza, N. D. Anderson, J. K. Locantore, and A. R. McIntosh, *Aging gracefully: Compensatory brain activity in high-performing older adults*, *NeuroImage* **17** (2002), no. 3 1394–1402.

[175] C. L. Grady, *Cognitive neuroscience of aging*, *Annals of the New York Academy of Sciences* **1124** (2008), no. 1 127–144.

[176] S. Fortunato, *Community detection in graphs*, *Physics reports* **486** (2010), no. 3 75–174.

[177] C. Lohse, D. S. Bassett, K. O. Lim, and J. M. Carlson, *Resolving anatomical and functional structure in human brain organization: identifying mesoscale organization in weighted network representations*, *PLoS Comput Biol* **10** (2014), no. 10 e1003712.

[178] M. Girvan and M. E. Newman, *Community structure in social and biological networks*, *Proceedings of the national academy of sciences* **99** (2002), no. 12 7821–7826.

[179] A. Browet, P.-A. Absil, and P. Van Dooren, *Community detection for hierarchical image segmentation*, in *International Workshop on Combinatorial Image Analysis*, pp. 358–371, Springer, 2011.

[180] J. Shi and J. Malik, *Normalized cuts and image segmentation*, *IEEE Transactions on pattern analysis and machine intelligence* **22** (2000), no. 8 888–905.

[181] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, *Hierarchical organization of modularity in metabolic networks*, *science* **297** (2002), no. 5586 1551–1555.

[182] R. Lambiotte, *Multi-scale modularity in complex networks*, in *Modeling and optimization in mobile, ad hoc and wireless networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*, pp. 546–553, IEEE, 2010.

[183] R. Lambiotte, J.-C. Delvenne, and M. Barahona, *Laplacian dynamics and multiscale modular structure in networks*, *arXiv preprint arXiv:0812.1770* (2008).

[184] M. Rubinov and O. Sporns, *Weight-conserving characterization of complex functional brain networks*, *Neuroimage* **56** (2011), no. 4 2068–2079.

[185] E. N. Davison, B. O. Turner, K. J. Schlesinger, M. B. Miller, S. T. Grafton, D. S. Bassett, and J. M. Carlson, *Individual differences in dynamic functional brain connectivity across the human lifespan*, *PLOS Computational Biology* **12** (2016), no. 11 e1005178.

[186] A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, *Comparing community structure to characteristics in online collegiate social networks*, *SIAM Review* **53** (2011), no. 3 526–543.

[187] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, B. Y. Alfred, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, *et. al.*, *Controllability of structural brain networks*, *Nature communications* **6** (2015).

[188] M. G. Mattar, M. W. Cole, S. L. Thompson-Schill, and D. S. Bassett, *A functional cartography of cognitive systems*, *PLoS computational biology* **11** (2015), no. 12 e1004533.

[189] A. Lancichinetti and S. Fortunato, *Limits of modularity maximization in community detection*, *Physical review E* **84** (2011), no. 6 066122.

[190] H. Wen, E. Leicht, and R. M. DSouza, *Improving community detection in networks by targeted node removal*, *Physical Review E* **83** (2011), no. 1 016114.

[191] P.-Y. Chen and A. O. Hero, *Deep community detection*, *IEEE Transactions on Signal Processing* **63** (2015), no. 21 5706–5719.

[192] V. A. Traag, P. Van Dooren, and Y. Nesterov, *Narrow scope for resolution-limit-free community detection*, *Physical Review E* **84** (2011), no. 1 016114.

[193] C. Nicolini and A. Bifone, *Modular structure of brain functional networks: breaking the resolution limit by surprise*, *Scientific reports* **6** (2016).

[194] C. Nicolini, C. Bordier, and A. Bifone, *Community detection in weighted brain connectivity networks beyond the resolution limit*, *Neuroimage* **146** (2017) 28–39.