Title
Large-Scale Brain Network Mechanics

Author
Bertolero, Maxwell Arthur

Publication Date
2017

Peer reviewed|Thesis/dissertation

Large-Scale Brain Network Mechanics

by

Maxwell Arthur Bertolero

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

and the Designated Emphasis

in

Behavioral and Systems Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Chair Professor Mark D'Esposito,
Professor Robert Knight,
Professor Richard Ivry,
Professor Christos Papadimitriou

Spring 2017

Abstract

Large-Scale Brain Network Mechanics

By

Maxwell Arthur Bertolero

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Mark D'Esposito, Chair

The brain of any species can be modeled as a network of regions and connections between those regions. Here, I analyze the brain's large-scale network mechanics. I show that the brain can be divided into subnetworks, each with a discrete function. I provide evidence that the each subnetwork's processing is mostly modular; however, certain regions exist that perform integrative functions. I characterize this integrative set of regions in depth and discover that this set of regions exists across various species' brains and even in man-made networks. I also identify an optimal network structure for cognitive processing. Finally, I demonstrate one mechanism for why the brain's network structure was selected by evolution.

Dedications

Sarah Bertolero
Arthur Bertolero
Rachel Zagarino
Ruth Russell
Hubert Russell
Rena Bertolero
Arthur Bertolero Sr.
Linda Zagarino
Jeff Miller
Mark D'Esposito
Dapha Shohamy
Ed Smith
Achille Varzi
Katja Vogt
Patricia Kitcher
Thomas Griffiths
Robert Knight
Richard Ivry
Christos Papadimitriou

Table of Contents

Introduction

While Plato and Hippocrates were the first to claim that the mind is what the brain does, Galen of Pergamon was perhaps the first to assign specific functions to specific anatomical locations in the body (1, 2). Galen proposed that each organ had a specific function, even mental functions. He claimed that there is no distinction between the mental and the physical world. While he accurately attributed the rational aspect of Plato's tripartite soul–what we would call the mind–to the brain, he fallaciously attributed the spiritual and the appetitive aspects to the heart and the liver, respectively. 1400 years and a dark age later, the more empirically driven Andreas Vesalius countered that claim with the idea that the brain and the nervous system are responsible for all mental processes(3). While Galen made the mistake of attributing mental processes to the heart and the liver, he made the interesting claim that interactions among those organs give rise to mental processes. Vesalius too, believed that proper analysis of the components of a system could only be understood in relation to the whole(4).

A focus on the localization of mental functions in the brain, not interactions among different aspects of the brain, is what persisted. And not without good reason. While the phrenologist Franz Joseph Gall's claims regarding the shape of the skull and what mental functions existed (e.g., the literary function) were extremely misguided, his focus on localization was founded in the correct belief that the brain is not a homogenous unit (5) The brain is a collection of components, each with a different function. The first real localization of a cognitive function involved language. Why language? Not being able to speak is a very acute and obvious deficit that brings one to the neurologist rather quickly. Different aspects of language are well-localized in Broca's and Wernicke's areas. Ludwig Lichtheim did claim that connections between brain regions were critical for language(6), which has been verified(7). However, one-to-one mappings between a localized brain lesion and a cognitive deficit dominated the field, likely due to how accurate these models are despite their simplicity. Moreover, it is difficult to understand how the components of a system function together before one understands what the components of the system are.

The advent of magnetic resonance imaging (MRI) opened the floodgates for localization. While Broca and Wernicke had to wait until their patients died to examine their brains, MRI allows neurologists to see, immediately, where damage had occurred and correlate this damage with the patient's cognitive deficit. Furthermore, using functional MRI (fMRI), one can have a healthy subject do a particular cognitive task, and locate an area or set of areas for which neural activity increased. Thus, so it seems, whatever cognitive functions that are isolated by that task are localized to that region or set of regions.

We are still trying to carve nature at its joints. Cognitive neuroscience is still largely steeped in localization. If you ask someone in cognitive neuroscience what they study, many will respond that they seek to understand how the brain executes a particular cognitive function. However, many of their answers predominately state where in the

brain a particular cognitive function seems to occurs, or where a particular type of information (e.g., sensory stimuli, semantic information, or the value of a reward) is represented, not how the brain executes the function. This is true even of very distinguished, resplendent, and highly technical work(8). But to what explanatory end? Imagine obtaining a complete dictionary of every cognitive function, and where in the brain it is localized. Do you now understand how the brain works? Even if you think you do, you do not. This dictionary, a priori, cannot be accurate. First, the terms. We cannot be certain we have carved cognitive functions at their joints. Is vision, like water, a natural kind(9)? If our terms are fallacious, which many of them certainly are, then the definitions are meaningless. In 50 years, we might look back at our current ontology of cognitive functions the same way we look at Gall's. Second, the relationship between the mind (the terms) and brain (the definitions) is, to put it mildly, ambiguous. There is little reason to believe that any cognitive function–say, vision–maps in any clean and simple way to a location in the brain. More technically, mental kinds like vision are not identical with neurophysiological kinds. Vision is not simply occipital lobe activity like water is h_20 (10). However, the history of neurology and the relative success of localization with fMRI does suggest that specialized and localized functioning occurs in the brain(11). An accurate dictionary of which functions exists and where each function is located in the brain is necessary, not sufficient, for understanding the brain.

Cognitive neuroscience, particularly in the past 10 years, has begun to move away from the simple localization paradigm towards neural mechanisms. While the term 'neural mechanism' is often abused and underdefined, it actually has a very clear and precise meaning: a neural mechanism describes how, not merely where. For example, take the neural mechanics of the hippocampus. The hippocampus contains "place cells" that fire when an animal enters a particular location in space, and this network of neurons represents the animal's current environment(12, 13). This likely aides in spatial awareness and navigation. This goes beyond spatial awareness and navigation being localized to the hippocampus to the neural mechanism of representing space via a particular neural pattern. Other research has shown that the formation of memories in the hippocampus is related to the particular covariance structure of distributed neural activity during the experience of the event being maintained after the event(14). To remember an event, you have to neurally replay that event after it happens. That is a mechanism. The neural mechanics of other cognitive functions have also been elucidated. The dorso-lateral prefrontal cortex potentially maintains information (i.e., working memory) by directing attention to internal representations of sensory stimuli and motor plans that are stored in other regions of the cortex(15). Working memory has actually defied localization and is dependent on activity in sensory, parietal, and prefrontal brain regions(16). Finally, asking how the brain represents and calculates temporal information, instead of finding "the time area", has led to elegant systems level explanations of different ways in which the brain represents time(17) and the underlying neural mechanisms that contribute to each type of representation, like which ones require the cerebellum(18).

Critically, many scientists have begun to elucidate how brain regions interact. For example, coupling between brain rhythms coordinates activity in distributed cortical areas, and different cognitive tasks evoke distinct patterns of coupling across the cortex (19-21). Moreover, mathematical tools from statistical physics have been co-opted to analyze the entirety of interactions between brain regions, treating the brain as a large scale network(22). I want to know what the general principles and mechanics of these interactions are. I want to know why the brain works that way, and not some other way. Functional MRI, despite its historical connection to localization, actually presents a unique opportunity–one can observe, albeit with low resolution, the neural activity of the entire human brain. More recently, whole brain imaging of smaller animals has been executed with far greater temporal and spatial resolution(23). This allows us to ask very different questions than ever before, going both beyond localization and beyond studying the neural mechanics of individual structures or even a small set of structures. We can ask what principles explain how the brain's constituent parts function together as a complex network.

The answer, I believe, will enhance all the work done in cognitive neuroscience. Understanding the brain's network mechanics will further elucidate precisely how the constituent parts function. For example, how does spatial information in the hippocampus get routed to other regions of the brain that must utilize that information? How does the dorso-lateral prefrontal cortex interact with the rest of cortex? Semantic information is coherently tiled across the cortex(8); how does language processing access all of that information? How is any of this done in a computationally tractable manner? Can we find network structures and mechanisms that explain all of these processes? Answering these network questions will undoubtedly also give us more insight into how those individual structures and processes function. Defining vision in the same way we define water as h_20 requires understanding occipital lobe activity in the context of the brain as a large scale network. Moreover, general artificial intelligence (i.e., artificial intelligence that can solve multiple novel problems(24)) requires this network level understanding of the brain.

More pointedly, my research has focused on the brain's functional network architecture and mechanics. How modular or "informationally encapsulated" is the brain's processing(25). Can we really think of the brain as a collection of components, each with a discrete function? How can we best identify these components and their functions? How do these components interact? Are there areas that integrate across these functions? How can we best identify these areas? Are there network features that exists across species' brains, and even in man-made networks? What is the least and most optimal network architecture for cognitive processing? Why did the brain evolve to have the particular network architecture that it does? These are the questions I address here.

Modularity

Is Holism A Problem For Inductive Inference? A Computational Analysis

Maxwell A. Bertolero & Thomas L. Griffiths

Fodor (26) claims that, while processes like vision can function adaptively almost purely based on sensory input from the eyes, central cognitive processes like decision-making only function adaptively if they integrate every bit of relevant information (i.e., if *every* variable is connected). Fodor argues that this holism—that everything we experience is potentially relevant to every conclusion we might want to draw—makes inductive inference intractable. This argument revisits Hume's (27) observation that there are no rules of induction—one cannot know *a priori* what variables to evaluate for evidence of a belief, as every variable is potentially relevant to the belief in question. This seems to make inductive inference intractable, as one must do an exhaustive analysis of every connected variable during inductive inference.

If the intractability of inductive inference was not sufficiently concerning by itself, Fodor(25) raises the stakes by arguing that holism may also make cognitive science a futile effort. If central cognitive processes involve inductive inference, then they are potentially holistic. Science carves nature at its joints, dividing and conquering. However, holistic systems have no joints, so Fodor argues that scientists should give up studying central cognitive processes, like decision making, where holism is an issue. He recommends focusing on non-holistic processes like vision.

Concerns about the impact of holism are based on the idea that all human concepts are connected; more precisely, that any one proposition involving particular concepts could bear on the confirmation of any other proposition, even ones involving different concepts. However, research in statistical machine learning has shown that variables being connected is not intrinsically a problem for inductive inference. Work on Bayesian networks shows that the complexity of probabilistic inference increases as a (potentially exponential) function of the size of the largest "clique" (i.e., fully connected set of nodes) in the graph of dependencies between variables(28). Thus, the connectivity structure of our concepts might not render inductive inference intractable if the size of the largest clique only increases slowly as the network grows.

This presents an interesting framework. We examine whether the connectivity structure of human concepts supports efficient probabilistic inference, providing a potential response to concerns about holism. We analyze a large network of concepts intended to capture human commonsense knowledge by using tools from graph theory and results concerning the computational complexity of inference in Bayesian networks. By examining the rate at which the largest clique grows—and the factors that contribute to the growth—we determine whether probabilistic inference could be conducted efficiently over human concepts. In essence, does computational complexity increase logarithmically as the size of the network increases? Similar functions has been observed for the small-world property of networks. In many networks, from social

networks to the world wide web, the average path length between two nodes only increases logarithmically with the size of the network (29). Thus, even as a network becomes very large, there are short path lengths between all the nodes. However, this presents a problem; the number of nodes within k edges from a node increases exponentially as the network grows, as k only increasing logarithmically with the number of nodes in the graph. This means that the number of potentially relevant concepts increased exponentially as the network grows. Thus, how does the computational complexity of a network scale with the size of the network?

Bayesian Networks and Holism

A graph is a mathematical model that represents a network. This model represents the network's entities as nodes, and connections between entities as edges that connect the nodes (see Figure 1). A Bayesian network is probabilistic graphical model that represents a set of variables as nodes and the dependencies between those variables as edges in a graph(30). The graph structure can be used to support probabilistic inference, calculating the probability distribution over the variables given the observed values of a subset of those variables.

Concerns about holism are based on the fact that all human concepts seem potentially related, and, as stated above, many networks contain a small-world structure in which every node is a small number of edges from every other node. If we imagine these concepts being arranged as the nodes in a graph, with edges indicating relationships, then we might be concerned about holism if the graph is connected—if we can find a sequence of edges that links any pair of nodes.

The fundamental concern is about the tractability of inductive inference—whether it is possible to make inductive inferences under these circumstances. In computer science, problems are considered intractable if the amount of time taken to solve the problem increases exponentially in the size of the problem. This establishes a basic criterion for determining whether holism is a challenge for inductive inference: determining whether the time taken to perform probabilistic inference in a Bayesian network that captures the dependency structure of human concepts increases exponentially as the network grows.

Probabilistic inference in Bayesian networks is #*P*-hard, and the (worst-case) computational complexity of known exact (e.g.(28)) and approximate algorithms (e.g.,(31, 32)) for solving this problem is exponential with the size of the largest clique in the moralized and triangulated version of the graph. In other words, the size of the largest clique determines the computational complexity of inference across the entire network. Maximum clique size is simply the size of the largest set of fully connected nodes, where every node has an edge to every other node in the set. When we moralize the graph, we connect any two nodes that have edges that point to a common node, and then we drop the directionality of edges. When we triangulate the graph, we make it "chordal," in that there is always an edge between nodes in loops greater than

three. Thus, there are no chordless loops (loops without edges across nodes) greater than three in the graph. Figure 1 shows the progression from a directed graph to a moralized graph to a triangulated graph.
Using these results, we can formally define the criterion for holism to present an obstacle to inductive inference. Measuring clique size as the graph grows makes it possible to extrapolate about the computational complexity of probabilistic inference for graphs approaching the scale of human knowledge: if the largest clique grows near linearly as nodes are added to the graph, then holism is a real concern; if it grows sublinearly, then connections between concepts might not pose a real challenge for inductive inference, as the time required to calculate relevant probabilities would increase subexponentially. In the remainder of the paper we explore which of these regimes characterizes the properties of a graph based on human concepts.

Methods

Constructing a Graph of Human Concepts

To execute our analyses, we needed a network that represents human conceptual knowledge. We chose the Concept-Net database, which contains 280,000 facts, such as “eating is motivated by survival,” contributed by a large online community (33). This database was made to support common sense reasoning in artificial intelligence, but it can also potentially provide clues about aspects of cognition and the connectivity structure of our concepts.

We constructed a graph from the ConceptNet database. We took concepts to be nodes, and relationships between concepts to be edges. While a node is clearly a different notion of a concept than is assumed by Fodor, it is a working assumption that allows us to begin a quantitative exploration of these issues. There are 27 different types of relationships in ConceptNet; we treated all relationships as equal. For example, “eating” and “survival” are nodes, and “is motivated by” is an edge from “eating” to “survival”. ConceptNet relations are judged by an online community, which allowed us to only use the highly rated relationships between concepts. This results in a directed and unweighted graph with 25,561 nodes and 33,768 edges that represents the network of human concepts.

We reduced the graph in this way for two reasons: accuracy and computational limits. ConceptNet contains many relationships that simply are not valid, so we wanted to make use of the ratings that the online community gave the relationships between concepts. Moreover, some measurements, like triangulation and clique finding, become highly computationally expensive on networks much larger than 10,000 nodes.

Growth Models

We used two different methods to simulate the growth of knowledge, allowing us to examine the properties of the resulting graph as it increased in size. One method randomly

added nodes, the other method randomly added edges. We used these two different growth methods as they reflect different assumptions about learning. Both methods keep the graph connected during growth.

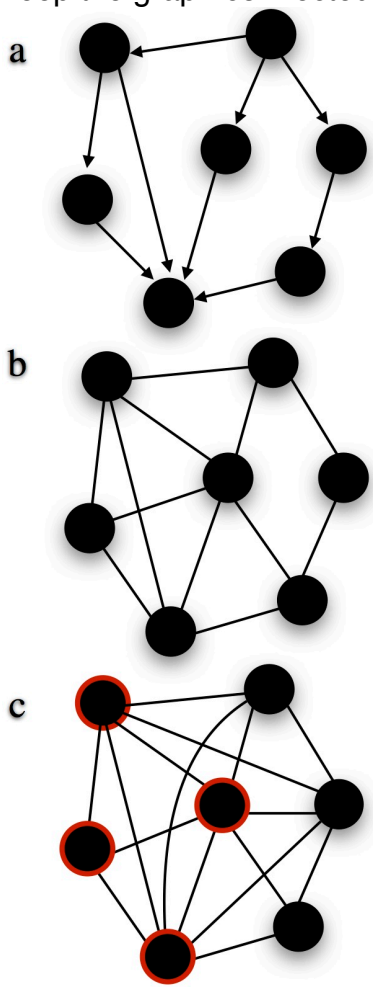


Figure 1: Graphs. (a) A directed graph, with nodes indicating concepts and edges indicating relationships. (b) The moralized graph. (c) The triangulated graph, largest clique nodes outlined in red.

Node Sampling. In the *node sampling* growth method, we randomly select a node to initialize the graph. Next, we randomly select one edge that connects to this first initial node, and add the node at the other end of the edge to the graph. The graph is now initialized. After this, we randomly select a node that already exists in the graph, and then randomly select one of that node's edges to find a new node that is not in the graph yet. We add the new node that was found at the end of the randomly selected node's edge, along with any edges the new node has to nodes that exist in the graph. Thus, at each step, one node is added, and at least one edge is added. This growth method reflects a learning model in which we learn a concept, and we acquire all its relations to concepts we already know.

Edge Sampling. In the *edge sampling* growth method, we randomly select an edge to initialize the graph. Both nodes get added to the graph. The graph is now initialized. After this, we randomly sample from all the edges that connect to nodes that exist in the graph. If an edge is picked that connects two nodes that already exist in the graph, only the edge is added. If an edge is picked that is between a node that exists in the graph and a node that has not been added to the graph yet, both the new edge and the node that does not exist in the graph yet are added. Thus, at each step, one edge is added, and either zero or one node is added. This growth method reflects a learning model in which we sequentially learn relationships between concepts, acquiring concepts when we learn a relation from a known concept to an unknown concept.

Graph-Theoretic Analyses

Clique Size. We measured the largest clique size in a moralized and triangulated version of the graph. As described above, moralizing the graph first creates edges between nodes with edges that direct to the same target node, and then drops the directionality of edges, resulting in an undirected graph. Triangulating the graph ensures that there are no loops greater than three. Triangulation can be performed by selecting an ordering of the nodes, then proceeding through these nodes in turn, adding connections between all nodes to which they are connected. The ordering influences the size of the resulting cliques, and finding the ordering that minimizes the size of the largest clique is *NP*-hard. Consequently, we used the heuristic of starting with the nodes with the fewest edges, which is standard in the Bayesian network literature (28).

Degree. The degree of a node is the number of edges that connect to it. We were primarily interested in degree measurements of the graph as it grows, but we also measured the degree distribution of the entire graph. Moreover, we wanted to know when higher degree nodes are added to the graph. Thus, when a node is added, we recorded its degree in the full ConceptNet graph.

Modularity. The graph-theoretic notion of modularity measures how decomposable a network is. Many different networks, from metabolic networks(34) to the world wide web (35), exhibit a modular structure, which means that the network's nodes can be grouped into smaller communities such that within-community edges are denser than between-community edges. Modularity quantifies the ability of the network to decompose into separable communities, and is an essential property found in many complex networks that allows the network to easily evolve, develop, and engage in flexible and dynamic behaviors (36). We measured modularity using Newman's Q (37), which is the fraction of the edges in the network that connect nodes from within the same (automatically-identified, as described below) communities minus the expected value of the same quantity in a network with the same community divisions but random connections between the nodes. If the fraction of within-community edges to between-community edges is no better than random, we will get a Q value of 0. The Q value increases as that fraction becomes greater than random. Q values approaching 1, which is the maximum, indicate a very modular structure. Modular networks identified in the previous research mentioned above exhibit Q values from 0.3 to 0.7.

Community Detection. For many networks, including our conceptual network, the partition of the network into communities is not known. Thus, we extract the communities of the graph by using Louvain heuristics (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) until modularity (*Q*) is maximized. First, we assign a different community to each node of the network, so the initial partition contains as many communities as nodes. We then iterate through two steps. Step 1: for each node *x*, we consider the neighbors *y* of *x*, and we evaluate the modularity score gain of removing *x* from its community and by placing it in the community of *y*. The node *x* is then placed in the community that maximizes the gain in modularity, but only if this gain is positive. If no positive gain is possible, *x* stays in its original community. This process is applied until modularity is maximized, such that no nodes can be moved to increase modularity. Step 2: each community is treated as a node, and Step 1 is repeated. The algorithm stops once invoking Step 2 does not increase modularity. The output is a modularity value of the network and a hierarchical partition of the network into communities. This process has been validated by applying it to networks that have a known community structure, with the percentage of correctly grouped nodes at 0.98(38).

Results

Our primary focus was on the growth of the largest clique, but we used the additional graph theory analyses as a source of insight into the patterns of clique growth. We will thus present these analyses in turn. Results were very similar across the two growth models, so they are presented together.

Clique Size Growth

In both growth models, we found that the largest clique initially grew roughly linearly, but the rate of growth decreased significantly as the number of nodes in the graph increased (see Figure 2). The increase in clique size is thus sub-linear, and the increase in the computational cost of probabilistic inference is sub-exponential. In the remainder of this section we consider the factors that contributed to this growth pattern.

Degree

Our analysis of the degree of nodes included evaluating the overall degree distribution in ConceptNet, and looking at the role that degree played in the growth of the graphs. Overall Degree Distribution Taking the full ConceptNet graph, we found that the overall degree distribution is heavy-tailed (consistent with a power-law or log-normal distribution), as shown in Figure 3. This is consistent with the degree distribution observed in other graphs based on human semantic knowledge (39), and is a common observation in other complex networks (40). It indicates that while many nodes have low degree, a few nodes have very large degree, forming connections with a large number of other nodes.

Degree of Nodes added to Graph. We observed that, despite the random sampling process behind both of our growth methods, high degree nodes were added to the graph early (see Figure 4). If nodes are sampled randomly, the nodes with the most

edges are most likely to be found and added to the graph early. If edges are sampled randomly, the nodes with the highest degree are more likely to be added to the graph, as they have the most edges. This result is a priori predictable, assuming a heavy tailed degree distribution, which we observed. As high-degree nodes are added early, the average degree of the network is higher while the graph is small. As the graph grows, and lower degree nodes are added, the average degree begins to shrink. As soon as the average degree begins to shrink, clique size growth slows. The observation that high-degree nodes are added early provides some validation for the assumptions behind our growth models. Previous models (39) have made predictions about the time course of semantic acquisition. These models show that the order in which concepts are acquired is crucial in determining their connectivity—concepts that enter the network early are expected to show higher degree. This has been verified against age of acquisition norms (41), which show that high degree concepts are learned early. In this growing network model, the degree of a node decreases as a function of the time since it was first connected to the network. We observed this same phenomenon—as the graph grew, the degree of nodes added to the graph decreased.

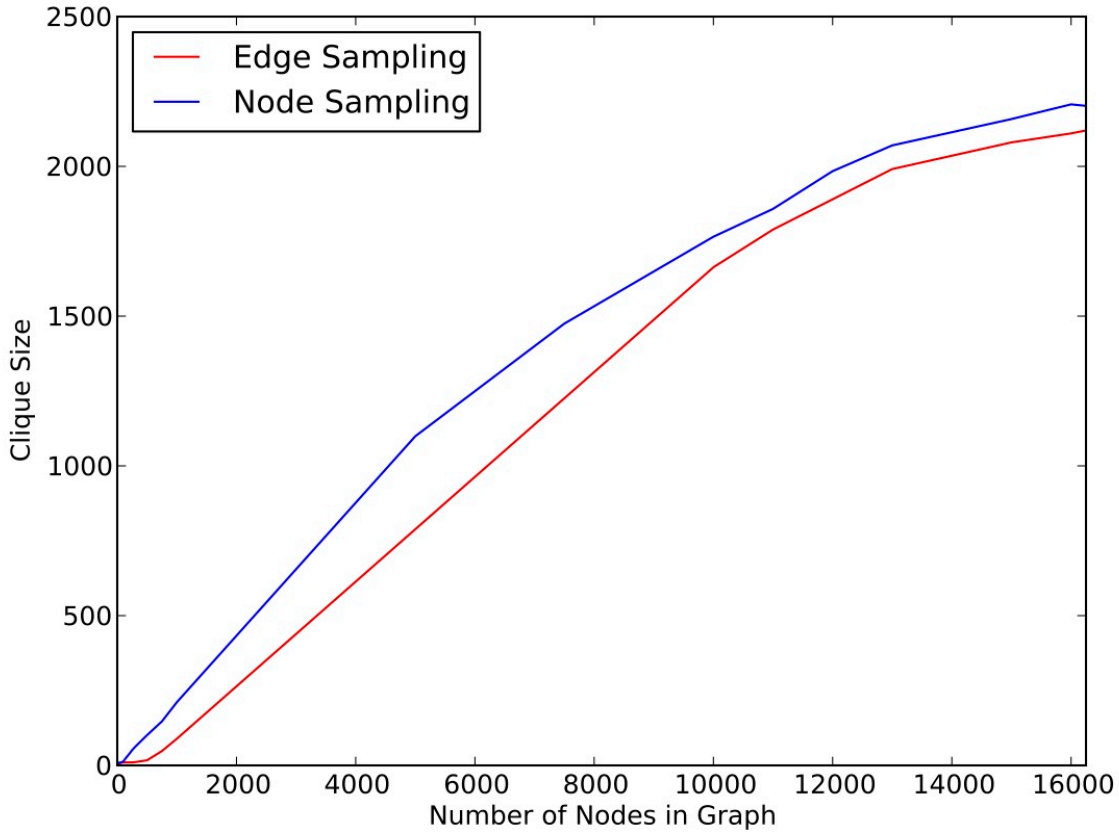


Figure 2: Size of the largest clique as a function of the number of nodes in the graph, for both growth methods.

Communities

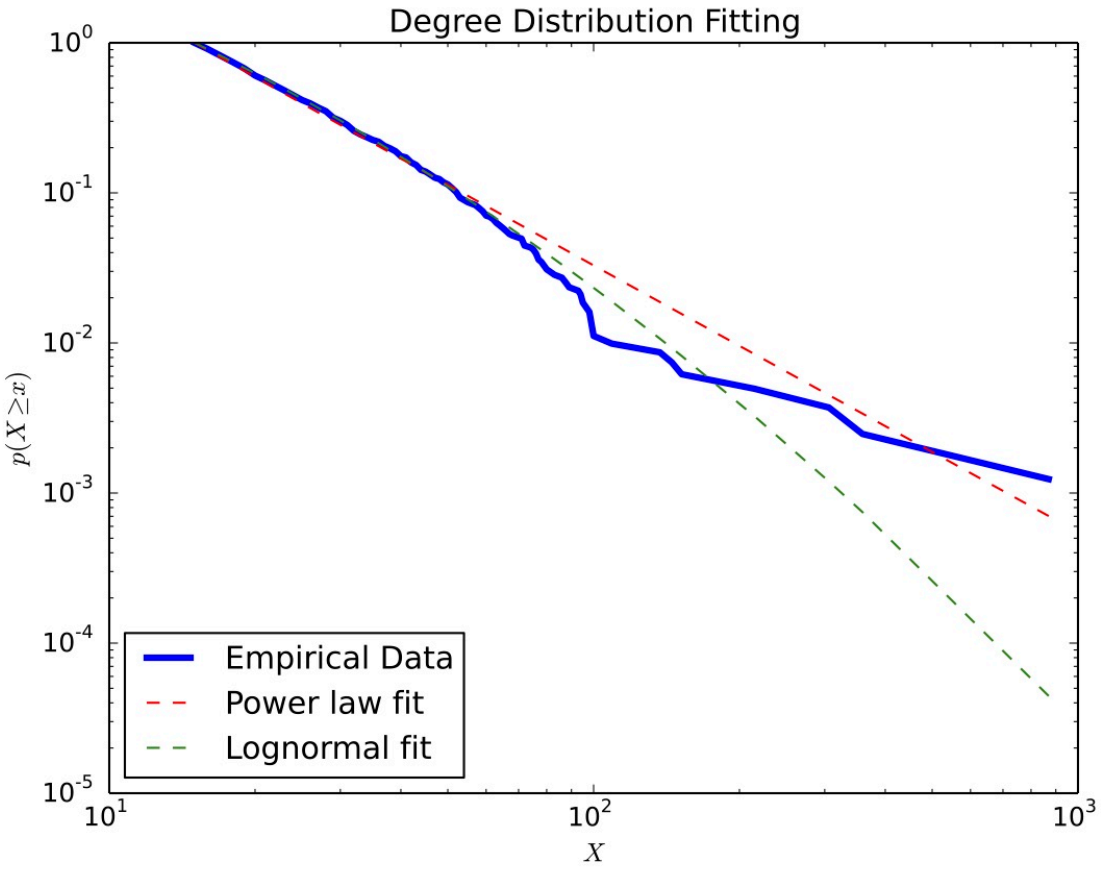


Figure 3: Overall degree distribution for ConceptNet, plot- ted on log-log axes. The near-linear relationship is consistent with a heavy-tailed distribution.

The community-finding analysis identified clear partitions of the graph into subsets that had many internal connections but few external connections (see Figure 5). We also observed a linear growth trend in the number of communities in the graph, suggesting the dynamic ability of the network of concepts to form new, more specific communities as more concepts are added, which ensures that modularity stays stable or increases. More specific communities allows for more efficient searches of what connected variables are critical for an inference, which is required as the graph grows. For example, an early, small version of the graph might contain the following nodes in one community: “horse”, “President Obama”, “Richard Feynman”, “redwood”, “dog”, “Carl Sagan”, “cactus”. All of these concepts can coherently exist in one community, as they are all living organisms. However, as more concepts are added, three new communities might be created, each only containing plants, animals, or humans, or even a further division of humans into politicians and scientists. Thus, the search space for important connected variables remains constant, even though the number of variables in the graph continues to grow.

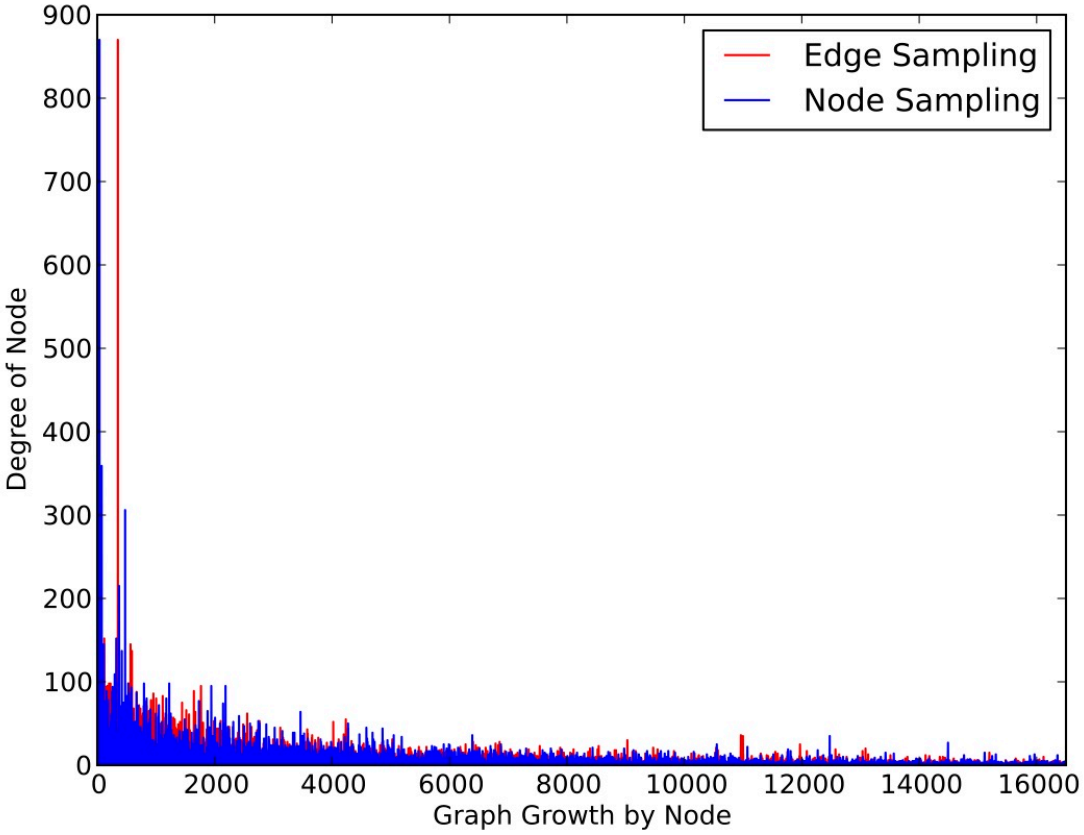


Figure 4: Degree of nodes during growth, shown in order of addition to the graph under both growth models. High-degree nodes tend to be added early.

Modularity

The modularity of the full ConceptNet graph is very high ($Q = 0.75$). During the initial stages of growth, we observed that modularity decreased as the clique size grew. After this initial growth, modularity stabilized, and clique size growth slowed (see Figures 2 & 6). This suggests a relationship between modularity and clique size growth. A modular network will have smaller cliques than a non-modular network, as edges between modules are limited in a modular network, which results in many small cliques, assuming the sizes of the communities stay stable during growth. In a non-modular network, there are no decomposable entities, resulting in one large clique. We speculate that, as long as network growth does not lead to decreases in modularity, clique size will continue to increase slowly as nodes are added beyond the range that we considered in our analysis.

Moreover, the observed increase of modularity of the network as the size of the network grows is precisely in line with previous computational studies of the efficiency of networks. Modular networks are less efficient than non-modular networks when networks are small, but modular networks become more efficient as the size of the

network increases. Thus, if our growth-model is capturing the important aspects of concept learning, modularity should increase as the network becomes larger, as a more modular network is required to keep the network efficient.

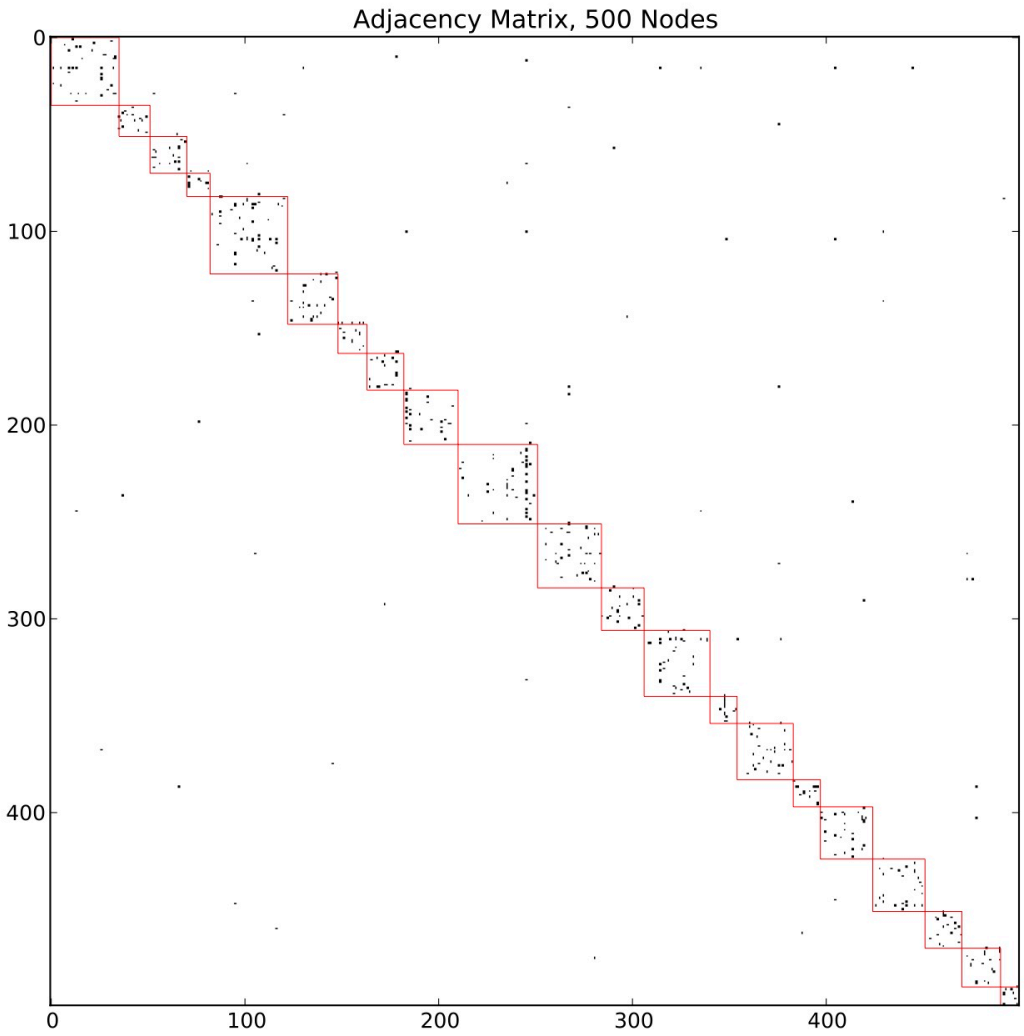


Figure 5: Adjacency matrix during growth under the node sampling growth model, with 500 Nodes. Black indicates where edges exist, white where edges do not exist. The red boxes represent communities. The graph is highly modular, with few connections outside communities.

To more rigorously test the relationship between modularity and clique size growth, we ran two additional analyses; first, we constructed 1000 graphs with a heavy-tailed degree distribution, and only varied the modularity of the graph. The correlation between the modularity of the graph and the size of the largest clique was r = 0.95,

p=1e-216. This suggests a very strong relationship between the modularity of a graph and the computational complexity of the graph. Next, we built 1000 graphs and varied how heavy-tailed the degree distribution is. Here, there was no relationship between how heavy tailed the degree distribution was and the size of the largest clique. We also modified the original ConceptNet graph by looping through edges in the graph, and, if that edge can be moved to decrease modularity while keeping the in-degree, out-degree, and total degree distribution perfectly equal, then it is moved. When this is done, we were able to decrease the modularity of the graph to 0.70; despite only a 0.05 decrease in modularity, the size of the largest clique increased nearly 10 percent; however, the growth is still logarithmic, further suggesting that a heavy tailed degree distrubtion allows for logarithimic growth. Finally, we built a synthetic graph with a heavier tailed degree distribution than the original ConceptNet graph (it is a scale free network), but with even lower modularity than the previous two graphs, at 0.65. Despite 0.65 still being very high modularity score, the 0.1 dip in modularity increased the size of the largest clique almost 50 percent, even though growth was logarithmic. These growth patterns are shown in Figure 7). Together, these results suggest that, while a heavy-tailed degree distribution might contribute to logarithmic growth of the largest clique in the network, modularity is what keeps the largest clique size small.

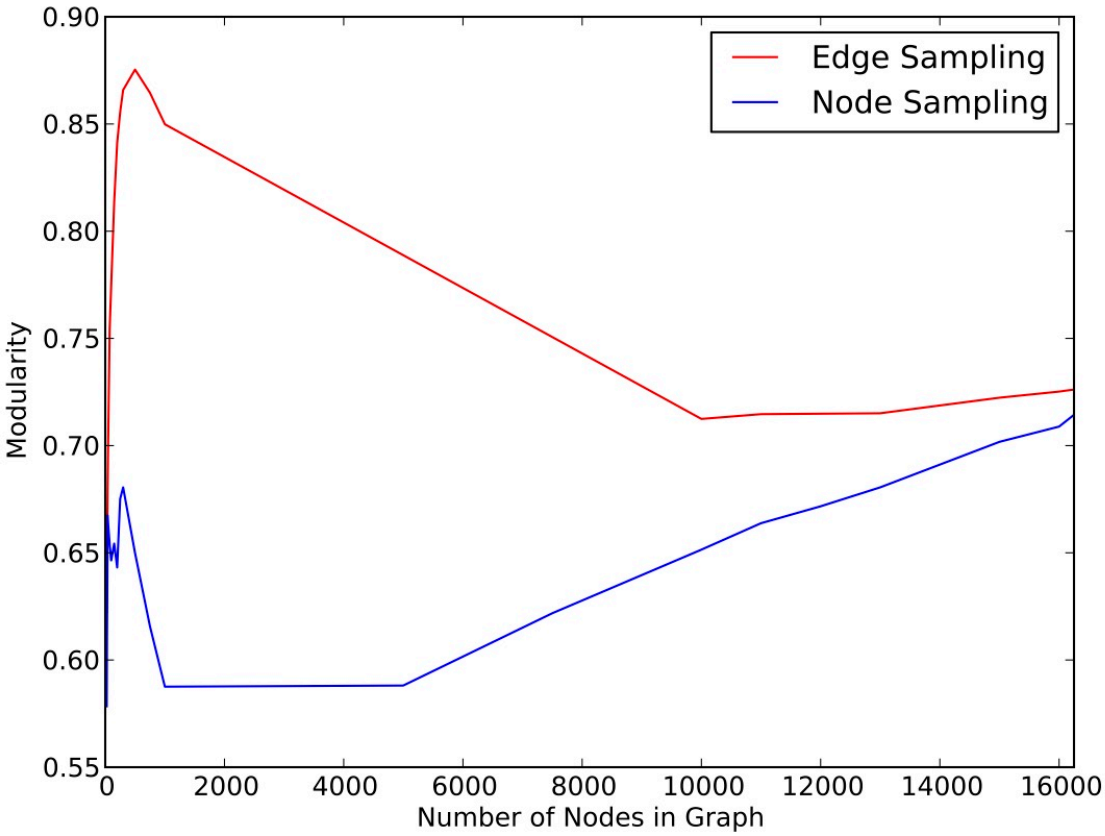


Figure 6: Modularity as a function of the number of nodes in the graph, for both growth methods.

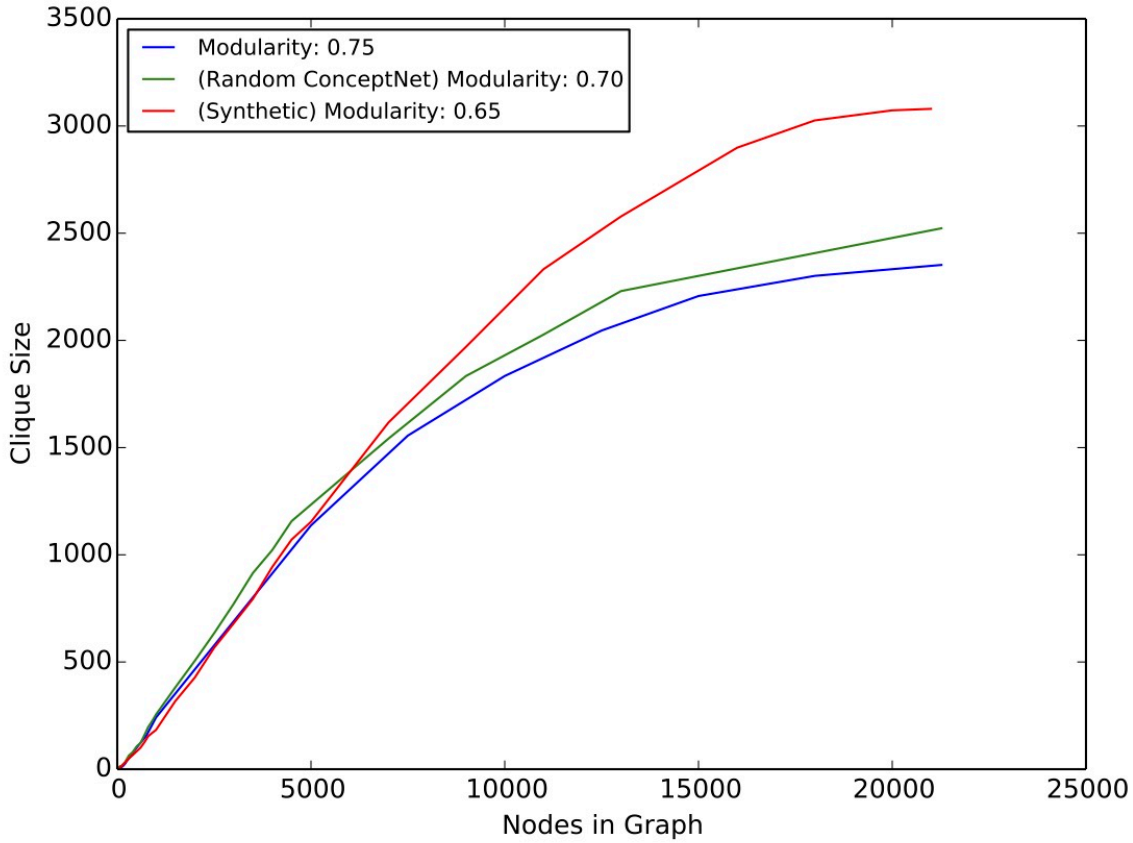


Figure 7: Blue, original ConceptNet; Green, less modular version of ConceptNet; Red, less modular synthetic graph. Final clique size, but not growth, is modulated by modularity.

Discussion

Hume's solution to the problem of induction was that Nature builds in the proper connections between concepts, claiming that "Nature, by an absolute and uncontrollable necessity has determined us to judge as well as to breathe and feel". Hume was on the right track. Holism does not present a problem for inductive inference, as the nature of the acquisition of our concepts and the modular nature of the connections between them, as well as the heavy tailed degree distribution, lead to sub-linear growth in clique size, meaning that efficient probabilistic inference is potentially possible.

We observed that, because our conceptual network has a heavy tailed degree distribution, high-degree nodes are added early. While these nodes are added, modularity decreases, and clique size growth is linear. However, when lower-degree nodes are added to the graph, modularity stabilizes and slowly increases (in part by increasing the number of communities), and clique size growth slows. This suggests that the time-course of semantic acquisition and the degree distribution and modular structure of our conceptual network makes inductive inference tractable. In other words, the connectivity of a variable (i.e., a concept) constrains what variables one must

observe to form a belief regarding that variable, as the connectivity structure of our conceptual network is modular, which limits the size of cliques. Thus, it is not the case that one must do an exhaustive search of every connected variable.

Fodor claimed that, because our central processing capacities, like reasoning and decision making, potentially require the truth of any proposition to provide confirmation to any other, inductive inference is intractable. Perhaps worse, the holistic nature of our conceptual network makes the study of central processing capacities a futile effort, as one can not tease apart holistic systems. However, we found evidence that the assumptions of Fodor's claims are false. We found that, while our conceptual network is holistic, in that it is connected, this does not present a theoretical problem for inductive inference or the study of central processes, as this network is modular in the mathematical sense. While it might not be informationally encapsulated—the criterion that Fodor believes makes analysis of certain psychological faculties possible—this does not mean that it is not decomposable. Finally, Fodor differentiates two aspects of holism: confirmation is isotropic, in that any proposition can in principle bear on any other, and also "Quinean", in that the confirmation of a proposition could depend on properties determined by the entire belief system, such as its simplicity. While our analysis has focused on this isotropic aspect of holism, Quinean concerns are dependent on the network being isotropic; thus, we suspect that the modular structure of human knowledge may also address the Quinean aspect of holism, as global properties like simplicity might not be as global as Fodor assumes.

Holism is a significant challenge, and the present work just represents a first step towards evaluating its seriousness. One concern for proponents of inductive inference is that, while the size of the largest clique grows sub-linearly, the largest clique remains large enough to pose a significant computational challenge. Understanding how this continues to scale will require analyzing larger graphs. Future analyses should also include null growth models, as well as models that vary modularity, community growth, and degree distribution. This will allow us to make stronger claims as to what drives sub-linear growth in clique sizes. For example, a heavy tailed degree distribution and modularity might actually be independent but necessary attributes of sub-linear growth, or the heavy tailed degree distribution might determine the modular structure, which in turn limits clique size growth. Moreover, future analyses should use different knowledge networks and inference algorithms, as the theoretical significance of the current analysis depends on how well Bayesian networks capture inference, and on how well ConceptNet serves as an approximation of knowledge.

More broadly, most graph theory analyses focus on describing properties of graphs, rather than considering algorithmic processes that occur in graphs. The approach taken in this paper demonstrates the utility of going beyond investigating the "descriptive statistics" of complex networks, and asking questions about network structure motivated by algorithmic properties. We asked, giving the structure of this network, what algorithms are computationally possible? While we used Bayesian inference, there might be other, similar questions about algorithms and representations that scientists can ask that go beyond the traditional "complex network" approach. In particular, a

similar algorithmically-motivated analysis could be applied to network analysis of the human brain.

Classification: Biological Sciences; Neuroscience, Psychological and Cognitive Sciences

Title: The Modular and Integrative Functional Architecture of the Human Brain

Authors: Maxwell Bertolero[a,1], B.T. Thomas Yeo[b], Mark D'Esposito[a]

[a]Helen Wills Neuroscience Institute and Department of Psychology, University of California Berkeley, California, United States of America
[b]Electrical and Computer Engineering, Clinical Imaging Research Centre, Singapore Institute for Neurotechnology & Memory Networks Programme, National University of Singapore, Singapore

[1]To whom correspondence may be addressed. Email: bertolero@berkeley.edu.

Keywords: modularity, hubs, network, graph theory, cognition

Abstract

Network based analyses of brain imaging data consistently reveal distinct modules and connector nodes with diverse global connectivity across the modules. How discrete the functions of modules are, how dependent the computational load of each module is to the other modules' processing, and what the precise role of connector nodes is for between module communication remains underspecified. Here, we employ a network model of the brain derived from resting-state functional MRI (rs-fMRI) data, and investigate the modular functional architecture of the human brain by analyzing activity at different types of nodes in the network across 9,208 experiments of 77 cognitive tasks in the BrainMap database. Using an author-topic model of cognitive functions, we find a strong spatial correspondence between the cognitive functions and the network's modules, suggesting that each module performs a discrete cognitive function. Crucially, activity at local nodes within the modules does not increase in tasks that require more cognitive functions, demonstrating the autonomy of modules' functions. However, connector nodes do exhibit increased activity when more cognitive functions are engaged in a task. Moreover, connector nodes are located where brain activity is associated with many different cognitive functions. Connector nodes potentially play a role in between module communication that maintains the modular function of the brain. Together, these findings provide a network account of the brain's modular yet integrated implementation of cognitive functions.

Significance Statement

Many complex networks are modular—they are composed of mostly autonomous and discrete units or *modules* that form an interconnected network. We sought to elucidate the nature of the brain's modular function by testing the autonomy of the brain's

modules and the potential mechanisms underlying their interactions. By studying the brain as a large-scale complex network and measuring activity across the network during 77 cognitive tasks, we demonstrate that, despite connectivity between modules, each module appears to execute a discrete cognitive function relatively autonomously from the other modules. Moreover, brain regions with diverse connectivity across the modules appear to play a role in enabling modules to interact while remaining mostly autonomous. This generates the counterintuitive idea that regions with diverse connectivity across modules might be necessary for modular biological networks.

The principle of modularity, in which a system or process is mostly decomposable into distinct units or *modules*, explains the architecture of many complex systems. Biological systems, including the human brain, are particularly well explained by the principle of modularity (36, 42-47). For example, gene expression in the brain is modular; the transcriptomes of human brain regions are robustly organized into modules of coexpressed genes that reflect the underlying cellular composition of brain tissue (48) and the spatial topography of cortex is also strongly reflected in its genetic topography—the closer two cortical regions, the more similar their transcriptomes (49). Moreover, these genetic divisions corresponded largely to meaningful structural and functional divisions (50, 51). This suggests a modular evolution of the brain, as functions that are adjusted modularly require modular genetic bases (44). Brain regions also exhibit cytoarchitectonic differences (52). A modular structural wiring network (i.e., white matter tracts), with dense connectivity within modules and weak connectivity between modules, has also been consistently found in human brain imaging data (46, 47, 53, 54). Finally, the brain's functional architecture—how the brain's modules interact to produce cognition—appears modular (36, 55-57).

One of the most powerful mathematical frameworks for studying the functional architecture of the human brain is the network science (i.e., graph theory) approach (22, 46, 58-60). Network models represent the brain as a graph with a set of nodes, usually around 10^1 to 10^5 areas of the brain, with edges (i.e., connections) between nodes. Here, edges represent the strength of the correlation between the nodes' time-series of spontaneous neural activity (usually derived from "resting-state" fMRI (rs-fMRI) data where the subject is not administered a cognitive task). These edges represent "functional connections", which are not a direct proxy for anatomical connectivity but are largely constrained by anatomical connectivity (53, 61-64). A functional connection between two nodes represents the phase-locking of the two nodes' low frequency oscillations, which is associated with the transfer of information between two nodes or the modulation of activity in one node by another node (65, 66). While not every node and functional connection in the brain is identical, modeling all of the functional connections in the brain as edges between nodes in a network allows for the study of global properties of brain, such as modularity. This level of analysis complements analyses that measure activity magnitudes within particular brain regions, not connectivity between all regions. These two approaches are not redundant; a brain region can decrease in activity, but increase in connectivity with other brain regions

(67). Thus, in this study, we employ both types of analyses.

Nodes can be divided into *modules* by grouping the nodes that are most tightly connected to each other into a single module. When applied to spontaneous neural activity in humans measured via rs-fMRI, this procedure reveals a modular network architecture, in that a large fraction of the edges fall within the modules compared to the expected fraction if edges were distributed at random (36, 56). These modules reflect the underlying structural connectivity architecture of the brain, in that a large number of the functional connections in a module reflect direct anatomical connections (61, 68, 69). Moreover, the spatial organization of these modules also corresponds to regions that have more highly correlated gene expression than expected by chance (51). Importantly, the brain's network architecture during task performance is shaped primarily by the network architecture present during resting-state (i.e., spontaneous neural activity), as spontaneous neural activity is likely a prior or constraint on task activity (70, 71). This has been demonstrated in humans using fMRI (21, 72-75), in monkeys using multi-electrode recordings (76), and in the zebrafish using two-photon Ca2+ imaging (77). Thus, predictions regarding the brain's network structure—and potentially nodes' activity magnitudes—during tasks can be made based on the brain's network structure during spontaneous neural activity.

Given a division of nodes into modules, each node's topological role in the network can be characterized by graph theory metrics that measure the node's connectivity within its own module and to other modules. A metric called the participation coefficient measures how evenly distributed a node's edges are across modules, dividing nodes into two classes—connector nodes with many global edges across modules, and local nodes with mostly within-module edges. Connector nodes are thought to have access to information among different modules and thus potentially integrate information across or coordinate connectivity between the modules, while local nodes support the specialized function of individual modules (20, 78). A second metric, called the within-module-degree z-score, measures how well connected a node is to nodes in its own module. Thus, connector nodes are further subdivided into "connector hubs" and "satellite connectors" that both have high participation coefficients, but only connector hubs have high within-module-degree z-scores. Local nodes are further subdivided into "provincial hubs" and "peripheral nodes" that both have low participation coefficients, but only provincial hubs have high within-module-degree z-scores (79-81). These two metrics capture the topological roles of nodes in many large-scale complex networks, such as the mouse brain (82), the cat brain, the macaque brain (40), the Internet, air transportation networks, A. thaliana, the protein interactome of C. elegans, and metabolic networks (81). Moreover, in the macaque brain, a brain region's participation coefficient is positively correlated with its dendritic tree size, spine count, spine density, and layer III pyramidal soma size, suggesting a relationship between a brain region's topological position in the macroscale brain network and the region's cytoarchitecture (83). In humans, regions with high participation coefficients are implicated in a diverse range of tasks (20, 84, 85) and located where many modules are

within close physical proximity (78).

The brain's modular, yet integrated, functional architecture could potentially involve each module executing a discrete cognitive function mostly autonomously or informationally encapsulated from the other modules (25, 44, 86-88), where the computational load in one module is not heavily influenced by processing in the other modules. How discrete the functions of modules are, how dependent the computational load of each module is to the other modules' processing, and what the precise role of connector nodes is for between module communication is underspecified. Here, to specify the brain's modular functional architecture and the role of connector nodes, we measure each node's *connectivity* in the network during spontaneous neural activity (measured with rs-fMRI), and then make predications about those nodes' *activity magnitude probabilities* across 9,208 experiments of 77 tasks in the BrainMap database (89). Moreover, we employ a highly principled model of cognitive functions—an author-topic (i.e., hierarchical Bayesian) model of the BrainMap database (85)—that provides an ontology of cognitive functions, which are each represented by their probability of engagement in each task in the BrainMap database and the probabilistic spatial distribution of the cognitive function's activity across the brain. Thus, the model allows for a precise estimate of the number of cognitive functions engaged in each BrainMap task. We refer to the cognitive functions estimated via the author-topic model as cognitive components. In the context of this approach and the large array of cognitive tasks in the BrainMap database, a modular functional architecture with connector nodes makes very specific testable predictions (illustrated in Figure 1). First, if each module is dedicated to a discrete cognitive function, the cognitive components and modules should exhibit similar spatial distributions in the brain and show similar engagement in each BrainMap task. Second, when more cognitive functions (in our analyses, cognitive components or potentially modules) are engaged in a particular task, more discrete information is generated across the entire brain network and transferred between modules. If the modules' processing is relatively autonomous (i.e., modular), local nodes (i.e., nodes with mostly within-module connections) will not be required to process more information when more cognitive functions are engaged. Thus, activity at local nodes in each module should not increase during the performance of tasks that engage more cognitive functions. Third, during tasks in which more cognitive functions are engaged, connector nodes will increase activity to maintain modularity while additional information is generated and transferred across the brain, perhaps by integrating information across modules or coordinating the connections between those modules. Finally, connector nodes, in order to execute the function suggested by the third prediction, should be located where brain activity is associated with many different cognitive components.

Results

Overview. To test the predictions of the brain's modular functional architecture with connector nodes, we built a network model of the brain by measuring spontaneous

neural activity with rs-fMRI and correlated the activity probabilities during each BrainMap task (i.e., how often a region's BOLD activation magnitude was high enough to be reported active across BrainMap experiments for the task) at the four types of nodes in the network with the number of modules or cognitive components engaged in that task. We also compared the spatial distribution of the cognitive components (derived from BrainMap) to the modules (derived from rs-fMRI of spontaneous neural activity), as well as their engagement across BrainMap tasks.

Network Model of the Brain. Our network model of the brain was built from 24 healthy human subjects. We recorded six separate ten minute blocks of whole-brain spontaneous neural activity with rs-fMRI in each subject (2,610 time points per subject). As there is no agreement regarding the optimal brain atlas to parcellate the brain into nodes in the network, we used four different brain atlases, which we refer to by the name of the first author of the publication. We also validated our results with a publicly available spontaneous neural activity (measured via rs-fMRI) correlation matrix that utilized a fifth brain atlas (74). Thus, we built five different networks—four with our data using four different atlases to define the nodes in the network (which we refer to as the Shen (90), Power (56), Gordon (91), and Craddock (92) networks), and one with the publicly available data and a fifth brain atlas, which we refer to as the Crossley network. For ease of presentation, only results from the Shen network are presented in Figure 2 and 3. We chose the Shen network because it has extensive brain coverage, and the number of nodes is common for graph theory analyses of rs-fMRI in humans (56, 91).

In each network, two nodes are connected by a weighted edge, with the weight being the Fisher transformed Pearson correlation value (z) between the time series of activity in the two nodes, if z survives cost thresholding, where a cost of 0.15 retains the strongest 15 percent of possible edges and their edge weights (i.e., z values) in the network. This cost is reported in the main results. Figure 4, Figure 5, and Supplemental Figure 2 show results from analyses of costs from 0.05 to 0.2 in 0.01 steps, and Supplemental Tables 1, 3, and 4 show results based on the average of each nodes' participation coefficient and within-module-degree z-score across those costs (93). Modules were identified via consensus clustering (94) based on individual subjects' networks and the InfoMap algorithm, which decomposes the network into modules based on the probabilistic flow of information through the network. We leveraged individual subjects' module organization, as group-level averaging often does not accurately represent the topological properties of individuals' networks (95). For each subject, a consensus matrix (a value of 1 where the two nodes are in the same module and a value of 0 elsewhere) was created by implementing the InfoMap algorithm on the subject's network. The average of these matrices was used to form a weighted network, to which the InfoMap algorithm was applied. Because individual data is not available for the Crossley network, we ran the InfoMap algorithm on the group-level network. Using this procedure, eleven modules were identified in the Shen network (Figure 2a, Supplemental Figure 1). This division of nodes into modules was then applied to a group level network, where the edge weights were the average of subjects' Fisher

transformed correlation matrices, only keeping edges (and their weight values) that survive cost thresholding. Participation coefficients and within-module-degree z-scores were then computed for each node. These two metrics were then used to identify the four types of nodes previously defined. Other community detection approaches, the Louvain algorithm, and using motion scrubbed (96) time-series yielded similar results (see Supplemental Table 1).

Author-Topic Model of Cognitive Functions. To derive an ontology of cognitive functions in the human brain and estimate the number of cognitive functions engaged in each BrainMap task, we used an author-topic (i.e., hierarchal Bayesian) model of the BrainMap database originally built and then validated in an independent dataset (85). Unlike previous attempts to derive an ontology of cognitive functions by applying Independent Component Analysis to the BrainMap database and then mapping the components to tasks post-hoc (72, 73), the approach we implemented *jointly* models the association between brain activation and tasks under the appropriate premise that each task requires a unique number and set of discrete cognitive functions, with each cognitive function being executed by a distinct, but possibly partially overlapping, set of brain regions. In the current model, cognitive functions (referred to as cognitive components) were modeled as latent variables, explaining the relationship between 77 BrainMap tasks and corresponding brain activity. A cognitive component is a cognitive function that generates discrete information (e.g., visual representations of objects, motor movements, attentional biases) required to complete a task. We use the term "cognitive" in this broad sense. Figure 2b and Supplemental Figure 1 show that different cognitive components are executed by different sets of brain regions, and that these cognitive components have a spatial distribution that is qualitatively similar to the modules identified from a network analysis of spontaneous neural activity (Figure 2b).

The author-topic model of cognitive components provides information crucial for our subsequent analyses that previous models of BrainMap data (72-74) were unable to provide. Each BrainMap task has a precise probability of each cognitive component being engaged in that task. This distribution allows for the quantification of the number of cognitive components (i.e., cognitive functions) engaged in each BrainMap task (see Methods). This allows for an examination across BrainMap tasks of how task activity probabilities at different types of nodes in the network model of the brain are modulated by the number of cognitive functions engaged in the task. Moreover, each brain region has a probability value for each cognitive component, which allows us to compare the spatial distribution of the cognitive components to the modules derived from network analysis of spontaneous neural activity (Figure 2, Supplemental Figure 1). Also, we could identify brain regions where the probability values are high (> 1e-5) for multiple cognitive components, which means that activity at that region is associated with multiple cognitive components. We reasoned that these regions, similar to connector nodes, likely interact with and have access to information from many cognitive functions. Thus, their spatial distribution was compared to brain areas identified as connector nodes.

Spatial Distribution and Engagement of Cognitive Components and Network Modules. We first compared the spontaneous neural activity based network model to the BrainMap based author-topic model. If each network module executes a discrete cognitive function, then the spatial distribution of the modules derived via network analysis of rs-fMRI of spontaneous neural activity should be similar to the author-topic cognitive components model derived from BrainMap task data. Also, there should be a linear relationship between the number of modules and cognitive components engaged in each BrainMap task, in that, as more cognitive components are engaged in a task, more modules are also engaged.

We quantified the similarity of the spatial distribution of modules and cognitive components with Normalized Mutual Information (NMI, (97)). To compute this, we assigned each node to the cognitive component that had the highest average probability at the voxels in the node. Thus, we have two divisions of the nodes, one based on the network analysis of spontaneous neural activity, and one based on the author-topic model of BrainMap. If the two divisions are identical, NMI is 1, whereas NMI is 0 if they are completely dissimilar. If NMI is low or there is not a linear relationship between the number of modules and cognitive components engaged in each task, this suggests that modules are not performing discrete cognitive functions. In all networks, we found a high level of NMI between the modules and the cognitive components, suggesting that they share a similar spatial distribution (Shen Network NMI=0.523; Power Network NMI: 0.603; Gordon Network NMI=0.542; Craddock Network NMI=.410; Crossley Network NMI = 0.410; $p<1e\text{-}323$ in every network, see Supplemental Information for p-value calculation). Figure 2 and Supplemental Figure 1 illustrate this finding. For a relative comparison of NMI values for brain networks, the mean NMI between individual subject partitions and the group partition was 0.41, which is equal to or less than the NMI between the modules and cognitive components in every network. We also verified these results with the z-score of the Rand coefficient (Supplemental Table 2), which can be interpreted statistically.

We then quantified the number of modules engaged by a BrainMap task. Given the significant amount of smoothing in the activation images, we considered a module engaged if the activity probability at any voxel in the module was greater than the mean of activity probabilities in all active brain nodes engaged by the task. There was a strong correlation between the number of cognitive components and modules engaged in each task in every network (Figure 2c; Shen: $r=0.722$, $p=1.3e\text{-}13$; Power: $r=0.577$, $p=3.8e\text{-}8$; Gordon: $r=0.807$, $p=8.2e\text{-}19$; Craddock: $r=0.693$, $p=2.9e\text{-}12$; Crossley: $r=0.598$, $p=9.5e\text{-}9$; $df=75$). Together, these results suggest that each module executes a discrete cognitive function.

The Relationship of Nodal Activity Probabilities to Engagement of Cognitive Components and Modules. The author-topic model allows us to precisely quantify the number of cognitive components engaged in each BrainMap task. With each activation

map for each BrainMap task, we quantified the number of modules engaged in each task, as well as activity probabilities at the four different types of nodes. This allows us to test our predictions regarding the modules' autonomy (Figure 1) by examining how different types of nodes' activity probabilities in a task are modulated by the number of cognitive functions (i.e., cognitive components or modules) engaged in the task. For each BrainMap task, we quantified a type of node's activity probability during the task as the mean of activity probabilities at the active nodes of that type divided by the mean of activity probabilities throughout all active brain nodes during the BrainMap task. Then, across all tasks, we correlated activity probabilities at each type of node in each task with the number of cognitive components or modules engaged in the task.

Consistent with predictions from a modular functional brain architecture, we did not observe a positive correlation in any network between provincial hub activity probabilities and the number of cognitive components or modules engaged in a task (Figure 3a, cognitive components: Shen Network: r=-0.395, p=0.0004; Power Network: r=-0.570, p=6.6e-08; Gordon Network: r=-0.530, p=7.1e-07; Craddock r=-0.407, p = 0.0002; Crossley Network: r=-0.571, p=5.8e-08; df=75; Figure 3b, modules: Shen Network: r=-0.384, p=0.0006; Power Network: r=-0.489, p=6.2e-6; Gordon Network: r=-0.644, p=2.6e-10; Craddock Network: r=-0.415, p=0.0001; Crossley Network: r=-0.374, p=0.0007; df=75). We also did not observe a positive correlation in any network between peripheral node activity probabilities and the number of cognitive components or modules engaged in a task (Figure 3a, cognitive components: Shen Network: r=-0.451, 3.9e-5; Power Network: r=-0.461, p=2.4e-5; Gordon Network: r=-0.356, p=0.002; Craddock Network: r=-0.567, p=7.7e-8; Crossley Network: r=-0.423, p=0.0001; df=75; Figure 3b, modules: Shen Network: r=-0.354, p=0.002; Power Network: r=-0.379, p=0.0007; Gordon Network: r=-0.343, p=0.002; Craddock Network: r=-0.328,p=0.003; Crossley Network: r=-0.116, p=0.316; df=75).

In contrast, but consistent with the modular functional brain architecture predictions, in all networks, we observed a positive correlation between connector hub activity probabilities and the number of cognitive components or modules engaged in a task (Figure 3c, cognitive components: Shen Network: r=0.600, p=8.3e-9; Power Network: r=0.536, p=5.1e-7; Gordon Network: r=0.569, p=6.8e-8; Craddock Network: r=0.512, p=1.9e-6; Crossley Network: r=0.641, p=3.2e-10; df=75; Figure 3d, modules: Shen, r=0.528, p=7.9e-7; Power Network: r=0.409, p=0.0002; Gordon Network: r=0.616, p=2.5e-9; Craddock Network: r=0.398, p=0.0003; Crossley Network: r=0.337, p=0.002; df=75). Significant positive correlations between satellite connector activity probabilities and the number of cognitive components or modules engaged in a task were also observed (Figure 3c, cognitive components: Shen Network: r=0.309, p=0.006; Power Network: r=0.514, p=1.7e-6; Gordon Network: r=0.488, p=6.6e-6; Craddock Network: r=0.593, p=2.8e-12; Crossley Network: r=0.615, p=2.7e-09; df=75; Figure 3d, modules: Shen Network, r=0.364, p=0.001; Power Network: r=0.393, p=0.0004; Gordon Network: r=0.627, p=1.1e-9; Craddock Network: r=0.527, p=8.1e-7; Crossley Network: r=0.447, p=4.5e-05; df=75)

While a cost threshold of 0.15 was used in these analyses, to test for the effect of cost thresholds, we ran the same analyses at costs from 0.2 to 0.05. All results were similar and significant across costs (see Supplemental Figure 2). We also averaged participation coefficients and within-module-degree z-scores for nodes across these costs, and results were similar and significant (Supplemental Table 1). Moreover, nodal participation coefficient values are publicly available for the Power Network from another cohort of subjects (56, 78). With this data, we found a negative correlation between activity probabilities at local nodes (nodes with low participation coefficients) and the number of cognitive components (r=-0.572, p=5e-08; df=75) or modules (r=-0.436, p=7e-05; df=75) engaged in a task. We also found a positive correlation between activity probabilities at connector nodes (nodes with high participation coefficients) and the number of cognitive components (r=0.651, p=1e-10; df=75) or modules (r=0.489, p=6e-06; df=75) engaged in a task. For calculating the number of modules engaged in each task in these correlations, we used the division of nodes into modules originally published with the Power atlas and used to derive these participation coefficient values, which are often referred to as the "Power Networks" (56). These modules also have a high NMI with the cognitive components (NMI=0.520).

Individual Subject Analyses. While the modular functional brain architecture predictions were confirmed in the group level networks, we sought to ensure that the network structure of individual subjects' networks also showed the above pattern of correlations, such that only connector nodes increase activity probabilities in tasks that engage more cognitive functions. Thus, for each subject, for each network (except for the Crossley Network, because individual subject data is not available) and cost threshold (0.05 to 0.2), we used the group level module division (for example, for the Shen network, the modules in Figure 2a and Supplemental Figure 1), but the edges between nodes in the network were defined by the subject's Fisher transformed correlation matrix. We then calculated activity probabilities at the four types of nodes in the network in the BrainMap tasks. Thus, each subject had 60 data points for each type of node; 15 data points for each of the four networks, representing the 15 different cost thresholds. Across subjects, on average, we observed a negative correlation between activity probabilities at local nodes and the number of cognitive components or modules engaged in a task. Across subjects, on average, we also observed a positive correlation between activity probabilities at connector nodes and the number of cognitive components or modules engaged in a task. In Figure 4, these results are plotted as kernel density estimations with the median and the 25th and 75th percentile of Pearson *r* values across subjects and costs shown for the particular type of node in that particular network.

Alternative Analyses of Nodal Activity. While we did observe a seemingly counter-intuitive decrease of activity at local nodes in tasks where more cognitive functions are engaged, follow-up analyses suggest that activity is not decreasing at these nodes, but is only increasing at connector nodes. We measured "hub weighted activity" by

calculating, for each task, the sum of [each voxel's activity probability multiplied by that voxel's participation coefficient or within-module-degree z-score score], divided by the sum of all voxels' activity probabilities. These calculations only consistently find a strong correlation between participation coefficient weighted activity probabilities and the number of components or modules engaged in the task (cognitive components: Shen: r=0.571, p=6e-8; Power: r=0.330, p=0.003; Gordon: r=0.503, p=3e-6; Craddock: r=0.498, p=3.9e-6; Crossley: r=0.595, p=1e-8; df=75; modules: Shen: r=0.659, p=7.3e-11; Power: r=0.290, p=0.01; Gordon: r=0.675, p=1.6e-11; Craddock: r=0.457, p=2.8e-5; Crossley: r=0.444, p=5.2e-5; df=75). The relationship between within-module-degree z-score weighted activity probabilities and the number of cognitive components or modules engaged in the task were, overall, very weak (cognitive components: Shen: r=-0.093, p=0.419; Power: r=-0.163, p=0.156; Gordon: r=-0.196, p=0.088; Craddock: r=-0.190, p=0.09; Crossley: r=-0.275, p=0.015; df=75; modules: Shen: r= -0.226, p=0.05; Power: r=-0.224, p=0.05; Gordon: r=-0.317, p=0.005; Craddock: r= -0.417, p=0.0001; Crossley: r=-0.457, p=2.8e-5; df=75).

We also measured activity by computing a Spearman *r* between task-active voxels' activity probabilities and the active voxels' participation coefficients or within-module-degree z-scores, and then correlated these *r* values with the number of cognitive components or modules engaged in the task. This calculation led to similar results. We observed positive correlations between *r* values of participation coefficients and task activity probabilities and the number of cognitive components or modules engaged in the task (cognitive components: Shen: r=0.524, p=1e-6; Power: r=0.581, p=3e-8; Gordon: r=0.555, p=1e-7; Craddock: r=0.591, p=1.5e-8; Crossley: r=0.592, p=1e-8; df=75; modules: Shen: r=0.480, p=9.9e-06; Power: r=0.385, p=0.0006; Gordon: r=0.660, p=6e-11; Craddock: r=0.439, p= 6.5e-5; Crossley: r= 0.402, p=0.0003; df=75). The correlations between *r* values of within-module-degree z-score and task activity probabilities and the number of cognitive components or modules engaged in the task were, overall, very weak (cognitive components: Shen: r=-0.114, p=0.323; Power: r=0.184, p=0.109; Gordon: r=-0.190, p=0.1; Craddock: r=-0.067, p=0.559; Crossley: r=-0.025, p=0.827; df=75; modules: Shen: r=-0.113, p=0.328; Power, r= 0.071, p=0.540; Gordon: r=-0.253, p=0.027; Craddock: r=-0.165, p=0.150; Crossley: r=-0.347, p=0.002; df=75).

These secondary analyses demonstrate that activity probabilities are not systematically shifting from local nodes to connector nodes, but that, in tasks that require more cognitive components or modules, there is greater activity probabilities at connector nodes. The decrease in activity probabilities at local nodes in tasks that require more cognitive components or modules observed in our main calculation is likely due to an increase in activity probabilities at connector nodes, which increases the mean of activity probabilities in the task relative to local nodes. This suggests that the brain does not obey the law of conservation of energy in network terms, as the total energy of the system does not appear constant.

Location of Connector Nodes and Convergence of Cognitive Components. If connector nodes are integrating information and coordinating connectivity across modules, and modules are each executing a discrete cognitive function (as evidenced by the above results), then connector nodes should reside in brain areas where activity is associated with many different cognitive components. This location would allow for them to simultaneously interact with brain regions dedicated to multiple cognitive functions with minimal additional wiring costs. Qualitatively, we found that connector nodes reside in brain areas where activity is associated with many different cognitive components (Figure 5 and Supplemental Figure 3). Quantitatively, the number of cognitive components that activity is associated with at a given brain area at greater than a probability of 1e-5 is greater at connector nodes than local nodes (Shen, $t=29.121$, $p=7.1e-186$, $df=150,520$; Power, $t=60.059$, $p\approx0.0$, $df=21,383$; Gordon: $t=43.4$, $p\approx0.0$, $df=56251$; Craddock: $t=71.317$, $p\approx0.0$, $df=142,999$; Crossley: $t=44.006$, $p\approx0.0$, $df=160,159$). While we previously showed that the modules and cognitive components share a similar spatial distribution, this finding demonstrates that the spatial distribution of areas that interact with many modules or cognitive components is also very similar.

Discussion

The Modular Brain. A detailed argument for the modular function of the brain was first presented over 30 years ago (25), with many versions of the concept presented since then (22, 56, 58, 98-109). A specific account of modularity with empirical support is *massive modularity* (56, 101, 103). Under this view, the brain consists of many modules that each executes a discrete cognitive function relatively autonomously from the other modules. Proponents of massive modularity (e.g.,(103)) argue that cognitive adaptations to the environment that have evolved are just as likely to have been specific solutions as any other physical adaptation. That is, a discrete module for visual processing is just as likely to have been naturally selected as an opposable thumb is for grabbing. Moreover, proponents argue that a biological system that is composed of mostly autonomous modules with discrete functions will perform more effectively, more efficiently, and adapt faster than a system with a few general functions; thus, a modular functional network architecture is most probable from an evolutionary perspective. Several lines of evidence support this evolutionary argument and the discrete and autonomous nature of modules.

As mentioned in the Introduction, the genetic basis of each module is significantly specialized, a necessary condition for each module to have a discrete function that was shaped modularly by natural selection (44). Moreover, there is a strong spatial correspondence between modules derived from spontaneous neural activity obtained with rs-fMRI and the task-evoked activity in the BrainMap database, and these modules can be linked to broad behavioral domains or to specific groups of related tasks (72-74), suggesting that these modules might be dedicated to discrete cognitive functions. However, these analyses did not empirically derive discrete cognitive functions from the data in BrainMap that could be linked to specific regions in the brain. Here, we observed

a strong correspondence between the spatial distribution of the cognitive components derived from the data in Brainmap and modules derived from rs-fMRI of spontaneous neural activity. Given that the cognitive components represent a highly principled model of cognitive functions, our results support the hypothesis that each module executes a discrete cognitive function.

Computational studies have demonstrated that selecting networks based on performance, but not wiring costs (i.e., the number of connections between nodes in the network) produces non-modular networks that are slow to adapt to new environments. However, adding a selective pressure to minimize wiring costs, or varying the goals of the network, leads to the evolution of modular networks that quickly adapt to new environments (110, 111). Moreover, modular neural networks can be trained to solve problems with less connections than non-modular networks (112). Thus, when the costs of adding connections is significant in the overall energy budget of an organism, as is likely for the brain, modular networks with weaker connectivity between modules than within the modules are superior to non-modular networks, as they function as well and adapt faster without consuming as much energy. Moreover, a double dissociation was found between two modules, such that, for both modules, damage to a node in one module only caused dysfunction (i.e., a decrease in functional connectivity) in the damaged module, suggesting their autonomy (113). Finally, empirical functional connectivity studies have shown that modules are weakly functionally connected to each other, likely because their computations are predominantly distinct, suggesting modular function (21, 36, 47, 56, 114). However, connections do occur between modules that transfer information between the modules and influence local activity in the modules, potentially increasing the modules' computational loads (21, 36, 47, 56, 66, 67, 115-118). Thus, to validate the autonomous nature of modules, it is necessary to demonstrate that, as more modules are engaged simultaneously and more information is generated across the network and transferred between the modules, the modules' computational loads do not increase (25, 86, 87) (Figure 1). Here, we used activity magnitude probabilities at the local nodes in each module to measure the computational load on each module—when more information needs to be computed, activity probabilities increase. Given that activity can dissociate from connectivity (67) and no previous analysis has quantified connectivity changes in relation to the number of modules engaged in a task, this type of evidence is not present in previous analyses of connectivity data. If the additional information that is generated and transferred between the modules when more modules are simultaneously engaged does not increase the modules' computational load, the modules' local nodes' activity will not increase. This would suggest that each module, regardless of the information present across the network or transferred to it from other modules, executes a distinct function without an increased computational load (i.e., it is relatively autonomous). On the contrary, if activity does increase, this would suggest important computational dependencies between the modules, and thereby that brain function is potentially not modular (25, 86). We observed the former, demonstrating a decisive characteristic (25, 86, 87) of the human brain's modular nature.

The Role of Connector Nodes in a Modular Network. Brain function is often conceptualized as a balance between modular and integrative processing (22, 54, 86, 99, 106-109, 114, 119, 120). The previous results generate an important question regarding this balance—how does the modules' functioning remain autonomous despite connectivity between the modules and the need for information to be integrated across the modules? One possible mechanism is to have brain regions that integrate across the modules and coordinate connectivity between the modules, keeping the modules' function mostly autonomous. For example, understanding syntax is likely a discrete cognitive *function*, supported by a dedicated brain module, whereas getting the gist of a sentence is likely a distributed cognitive *process*, requiring a combination and integration of various discrete cognitive functions (e.g. vision, attention, semantics, syntax) and processing in multiple modules. Undoubtedly, even a seemingly simple cognitive process likely arises from the simultaneous engagement of multiple modules, which could be orchestrated by connector nodes.

In support of this notion, we found that connector nodes are located where activity is associated with many different cognitive components, and, crucially, activity increases at connector nodes as the number of cognitive functions engaged in a task increases. Thus, unlike local nodes, connector nodes' increase in activity is proportional to the additional computational load required when many modules are engaged—when more information is generated across the brain that must be integrated and connectivity between many modules must be coordinated without sacrificing the autonomy of the modules' function (i.e., modularity is maintained). Thus, connector nodes are potentially integrating information across the modules and coordinating connectivity between the modules (i.e., modulating direct connections between modules that are not routed through connector nodes) to ensure modular function, such that information generated in one module does not increase the computational load in the other modules, as we found. An alternative proposal regarding connector nodes is that they are flexible regions that are reused in various ways depending on the current task (84, 121). The connectivity of connector nodes is flexible, in that their connectivity changes based on the particular task demands and modules engaged in the task (20). However, if their function is not relatively fixed, it would be unlikely to observe a strong relationship between their computational load and the number of modules engaged in a task, as the interaction between computational load and the number of modules engaged in a task would be different based on the function they are executing in a given task. Thus, our evidence suggests that connector nodes have a relatively invariable function that becomes more demanding as more modules are engaged in a task.

A wide range of evidence supports this conclusion. It has been proposed that integration across modules occurs via brain regions with strong connectivity across many modules (78, 105, 106). Connector node regions are engaged in a diverse range of tasks (84, 85), because many different tasks require multiple modules, and thus connector nodes' function. Also, similar to our finding regarding cognitive components,

connector nodes are located where many modules are within close physical distance (78); this requires the least amount of physical wiring for maximal or simultaneous interactions across multiple modules. Along with being physically close to many modules, connector nodes appear to *change* membership to modules often, and this is associated with higher performance and modularity. For example, an fMRI study found that, when switching from a single task to a dual-task, nodes changing their module membership was related to both increases and deceases in performance. However, the membership changes of connector node regions (i.e., brain regions identified as connectors in our analysis) was only associated with an increase in performance. Moreover, the high performing subjects (i.e., subjects with connector nodes that changed module membership) had less connectivity between single-task modules (i.e., more modular) (118). Another fMRI study found that increased module membership changes of 11 regions, 7 of which are connector node regions in our analysis, predicted increased learning rates (122). Moreover, an fMRI study of learning across time showed that, while connectivity between motor and visual modules is broadly evident in all nodes within those modules (not just at connector nodes), motor and visual modules become less connected (i.e., more modular) to each other during learning, and this process appears to be driven by the temporo-parietal junction, and entorhinal cortex, and a fronto-cingulate network, all of which contain connector nodes in our analysis (67). Finally, an fMRI study of a visual attention task found that connectivity changes in visual cortex that led to smaller local modules were associated with stronger top-down directional influence from, and stronger connectivity between, the dorsal frontal eye field, the intraparietal sulcus, and the superior parietal lobule, which are all connector nodes in our analysis (71). Together these studies suggest that the connectivity of connector nodes allows for effective integration across modules and the coordination of between-module connectivity to maintain modularity, which improves performance. This generates the counterintuitive idea that nodes with diverse connectivity across modules might be necessary for modularity in biological networks.

Connectivity changes (measured via fMRI) from spontaneous to task activity also support integrative and coordinative functions for connector nodes. In an analysis of changes in connectivity from spontaneous activity to a range of task evoked activity, most connections were maintained; however, connector node regions' connectivity patterns varied as a function of the task (75). In a similar analysis of the connectivity changes from spontaneous activity to task-evoked activity across 64 tasks (21), the network architecture was modular, but there was decreased within-module connectivity during task performance (as compared to resting-state), with increased between-module connectivity. It is possible that connector nodes are involved in coordinating some of these changes in connectivity. In another analysis of this data, the fronto-parietal network had the highest mean participation coefficient of all of brain networks, exhibited the greatest changes in connectivity with the rest of the brain across tasks out of any network, and the whole-brain connectivity of the fronto-parietal network could predict which task the subject was engaged in. This finding was interpreted as suggesting that the fronto-parietal network can coordinate multiple modules that are

engaged in the task (20).

Connector nodes integrating across or coordinating connectivity between task-relevant modules to maintain modular function is also consistent with observations of patients with brain pathology. For example, in patients with focal brain lesions, damage to connector nodes, but not local nodes, decreases the modularity of the global network structure (123). Moreover, TMS to connector regions in two different modules (fronto-parietal and cingulate-opercular) increased global connectivity (likely decreasing modularity) (124). These results are precisely what one would expect if connector nodes' function plays a role in global connectivity that maintains modularity. Moreover, patients with lesions damaging connector nodes exhibit widespread cognitive deficits, whereas damage to other brain regions (i.e., local nodes) causes specific cognitive deficits (125). Finally, connector node regions have been found to be metabolically demanding (126). Perhaps due to their crucial roles of integration and coordination and high biological cost (extensive wiring and metabolism), abnormal connector node function is associated with nine brain disorders, including Schizophrenia and Alzheimer's disease (127). Thus, the effects of brain damage and dysfunction can be interpreted in the context of our findings derived from brain imaging data in healthy individuals. That is, damage to or dysfunction of local nodes, due to their role in discrete cognitive functions, causes specific impairments without a degradation of the brain's modular structure, whereas damage to or dysfunction of connector nodes, likely due to their role in the integration and coordination across many modules, which maintains modularity, causes widespread cognitive deficits and a degradation of the brain's modular structure.

Connector nodes are a single facet of a large network of many complex interactions, and their function must be interpreted in this context. It is unlikely that all cognitive tasks require the engagement of connector nodes. For example, it has been shown that they are not recruited when a task is well learned (57, 67). Crucially, connector nodes are also not involved in all between-module connectivity. Non-connector node regions play a role as well (67, 118, 122, 128), and, as noted above, increased between-module connectivity that is not routed through connector nodes has been found during many different tasks, and differentiates task connectivity from resting-state connectivity (21). While the current and previous findings suggest that connector node regions coordinate some of these between module connections, it is likely that other mechanisms for between-module communication exist that are executed by non-connector nodes. In line with this, connector nodes are not the only nodes that change module membership during tasks (122, 128). For example, an fMRI study of a 2-back task showed that, while many connector node regions changed module membership and this predicted higher performance, the regions with the most module membership changes were in ventral-prefrontal cortex, where no connector nodes were located in our analysis (128). Finally, while connector node regions all exhibit the graph theoretic property of a high participation coefficient, these regions do not all have identical cytoarchitecture and their connectivity profiles are to different modules. Thus, each connector node's function

likely has unique properties that are not captured by our graph theory approach. Certainly, the precise type or types of computations that occur at connector nodes, how this varies based on the particular nodes involved in the task, the role of connector nodes' computations in the global functioning of the brain, and the role of distributed and self-organizing processes that are independent of connector nodes (57) needs to be elucidated by future work.

Methods

At the time of analysis, the BrainMap database contained findings from 2194 journal articles, which contained 83 tasks and 10,449 experimental contrasts and their activation images. BrainMap data was processed identically to the original model's application (85). For each BrainMap task, we use the average of binary activation reported. For example, if there were 100 experiments for a particular task, and 88 of the experiments reported activity at a voxel, that voxel would have an activity score of 0.88 in our analyses. Thus, activity in our analyses is a probability of activation in a task. For our analyses, we only used tasks with more than 10 experiments, which reduced the number of tasks to 77. The author-topic model has the best cross-validation with 10-14 cognitive components (the cognitive components are estimated based on a random 95% subset of the BrainMap data, and the resulting generalization power is computed on the remaining 5% of the data). Thus, for our analyses, we used the 14 component model, as it has the finest-grained resolution while still obtaining strong cross-validation; however, models with other numbers of cognitive components led to similar results (See Supplemental Table 3 for full results from all calculations with the Shen Network).

We calculated the number of cognitive components engaged in each BrainMap task as recommended by the original analysis (85). For each task, the probabilities of cognitive components engaged sums to 1. This probability distribution should be interpreted as the probability of a task recruiting a cognitive component for any activated voxel and not the probability of a task recruiting a cognitive component for the entire task. Thus, the probability (P) for a given cognitive component being engaged in a task is equal to 1 – (1 – X)^Z, where X is the cognitive component's probability for that task (using the probability from the distribution that sums to 1) and Z is the sum of voxel activity in the task (See (85) for a detailed discussion). We considered a cognitive component engaged in a task if it had at least a P of 0.9. We also used the original X values (which sum to 1) to quantify the number of cognitive components engaged in a task in three ways. Results were similar to those reported with these alternative calculations (See SI Text and SI Table 4 for full results from all calculations with the Shen Network).

Spontaneous Neural Activity was measured in healthy human subjects with BOLD fMRI (see Supporting Information). Voxel time series were averaged within each node in the atlas. We note that, for the Craddock atlas, we used the 950 node version that, for each node, maximized the similarity of the each voxel's whole-brain spatial correlation map and used the two-level averaging scheme in which the data of each participant were

clustered separately and the results were combined for a group level parcellation (92). The Shen, Power, Crossley, and Gordon networks had 278, 264, 638, and 333 nodes, respectively. Functional connectivity was assessed in each participant by computing time series Pearson correlations between all pairs of nodes, resulting in a correlation matrix for each participant, which was then Fisher transformed. This correlation matrix was then thresholded, which then served as the basis for defining a weighted, undirected graph. A graph is defined as a set of nodes that are connected by edges, which, in our analyses, represent Fisher z transformed correlation values between the nodes that survive the cost threshold. We use the term "network" to refer to a graph.

To average individual networks into a group level network, we applied the InfoMap algorithm to each subject's network. We chose to use the InfoMap algorithm, because, when compared to other algorithms on networks for which the true division of nodes into modules is known, it achieves the most accurate division of nodes into modules (129, 130). Moreover, while most algorithms maximize modularity, InfoMap is based on the probabilistic flow of information through a network, making it principled for use in networks for which information is hypothesized to flow. For each subject, a consensus matrix (a value of 1 where the two nodes are in the same module and a value of 0 elsewhere) was formed on the subject's network, thresholded at 0.035 in an attempt to achieve a similar number of modules and cognitive components. The average of these matrices was thresholded at the higher cost of 0.07, as all edges in this matrix represented that those two nodes were grouped into the same module, allowing for a less stringent threshold than was applied to the individual subjects' networks. InfoMap was run on this final matrix, forming a group level division of nodes into modules. In every iteration, InfoMap was run 5000 times, with the optimal division (i.e., minimal length of the InfoMap equation) chosen. For the Crossley Network, InfoMap was run on the full matrix, which had been previously thresholded based on multiple comparison correction, resulting in a cost of 0.0916 (74). Other clustering methods and algorithms were used (see Supplemental Information) and produced very similar results (See Supplemental Table 1).

Edges in the group level network were based on the average of Fisher transformed Pearson correlations across subjects. Various cost thresholds were used for this network, and all led to similar results (see Figure 4, Supplemental Table 1, and Supplemental Figure 2). Results from a cost threshold of 0.15 are presented in the main text for our data, and 0.0916 for the Crossley data. Higher cost thresholds were used for the final networks than for the module detection, as module detection at higher cost thresholds often leads to only three or four large modules (56), and we aimed to compare the modules at the same level of resolution as the cognitive components (i.e., a similar number of modules and cognitive components).

Each node's topological role in the graph was characterized by measuring the node's participation coefficient and within-module-degree z-score (79). Edge weights were used in the calculation instead of a binary count of edges. While the mean of

participation coefficients was 0.52 (for the Shen Network at a cost of 0.15), other thresholds for the cutoff of a high participation coefficient (i.e., dividing nodes into connector nodes or local nodes) were tested as well. In the original paper (79), a participation coefficient of 0.62 was proposed—thresholds up to 0.80 also led to similar results (See Supplemental Information). Supplemental Table 1 shows the mean participation coefficients values for each atlas and clustering method. Results were also robust to variations in the cut off for the within-module-degree z-score (See Supplemental Information).

Given the large variety of experimental and data processing procedures utilized across all of the studies in the BrainMap database, a direct comparison of activity scores across tasks is not justifiable. Thus, we calculated the activity in a particular type of node (e.g., connector hubs or provincial hubs) by calculating the mean activity of all voxels within the particular type of node divided by the mean of activity across all nodes. Only active voxels (voxels with non-zero activity scores) were used in this calculation. Thus, our main activity measure reflects the amount of activity at a type of node relative to whole brain activity for the task. We also used two secondary measures (described in the main text) that reflect how well activity aligns with participation coefficients and within-module-degree z-scores.

Figures

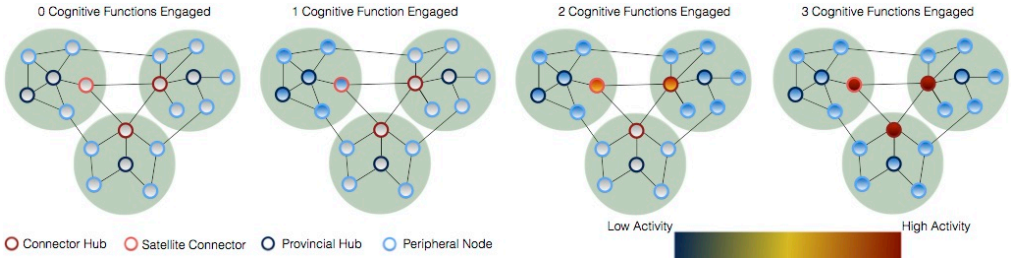


Figure 1 | Empirical Predictions of a Modular Functional Architecture with Connector Nodes. Each light green circle represents a set of nodes comprising a unique module. The four types of nodes are represented in different colors. Activity is shown from low activity in blue to high activity in dark red. From left to right, as the number of cognitive functions (measured in our analysis by modules and cognitive components) engaged in a task varies from 0 to 3, a modular functional architecture predicts that activity will increase at connector hubs and satellite connectors (i.e., connector nodes with high participation coefficients), but not at provincial hubs or peripheral nodes (i.e., local nodes with low participation coefficients).

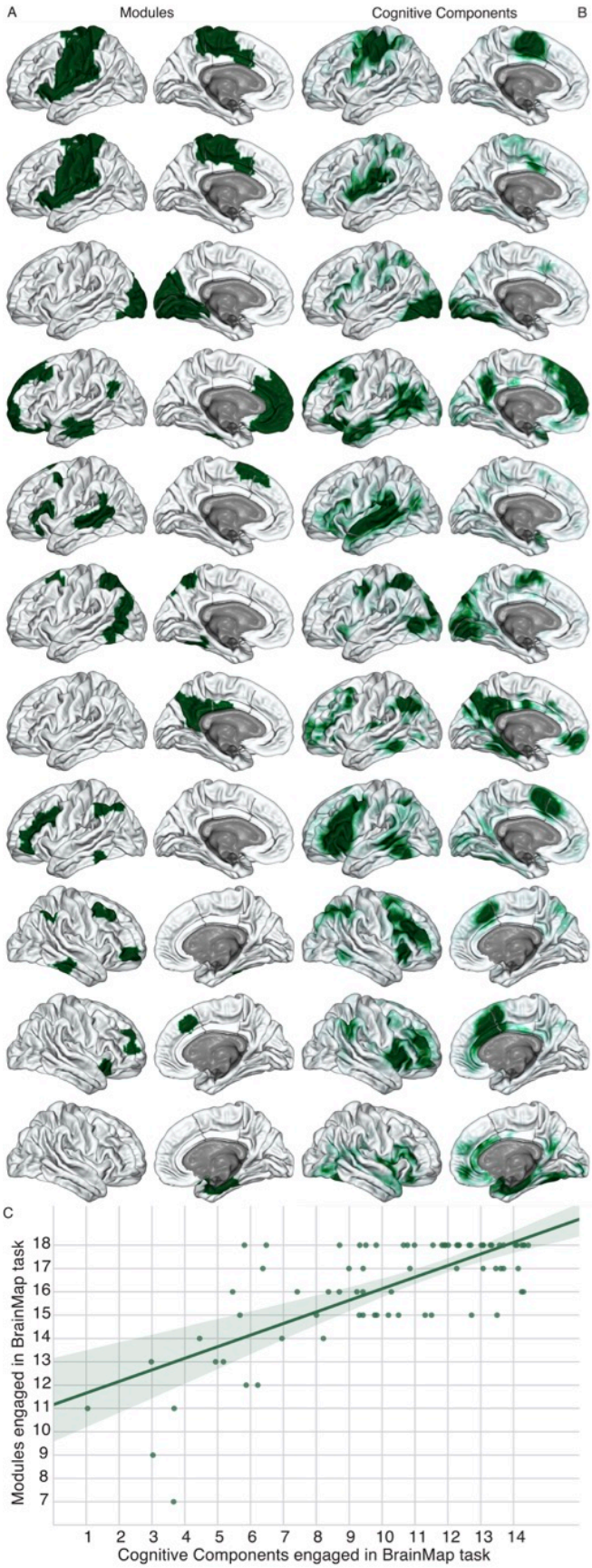

A
Modules
Cognitive Components
B
C
Modules engaged in BrainMap task
18
17
16
15
14
13
12
11
10
9
8
7
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Cognitive Components engaged in BrainMap task

Figure 2 | Spatial Distribution of Network Modules and Author-Topic Cognitive Components. A, left column, modules derived from spontaneous neural activity measured via rs-fMRI with graph theory. B, right column, cognitive components derived from the BrainMap database with an author-topic model. Each row comprises a unique module and the cognitive component with the highest probability of activity at voxels in the module. However, the module on the first two rows is a duplicate of the same module, as two different cognitive components had a high probability of activating voxels in it. The spatial distribution of network modules and cognitive components in subcortical regions are presented in Supplemental Figure 1. All modules with more than 5 nodes are shown (10 cortical, 1 subcortical), with the cognitive component that has the highest probability of activating voxels in the module on the same row, to the right. As a default, the left hemisphere is plotted, unless the module or cognitive component was predominantly located in the right hemisphere. This figure illustrates the high level of shared spatial distribution between modules and cognitive components. Sub-Cortical results are shown in Supplemental Figure 1. C, lower, the correlation between cognitive components engaged and modules engaged in each BrainMap task. Each dot represents a BrainMap task. The number of modules and cognitive components engaged in each BrainMap task is strongly correlated.

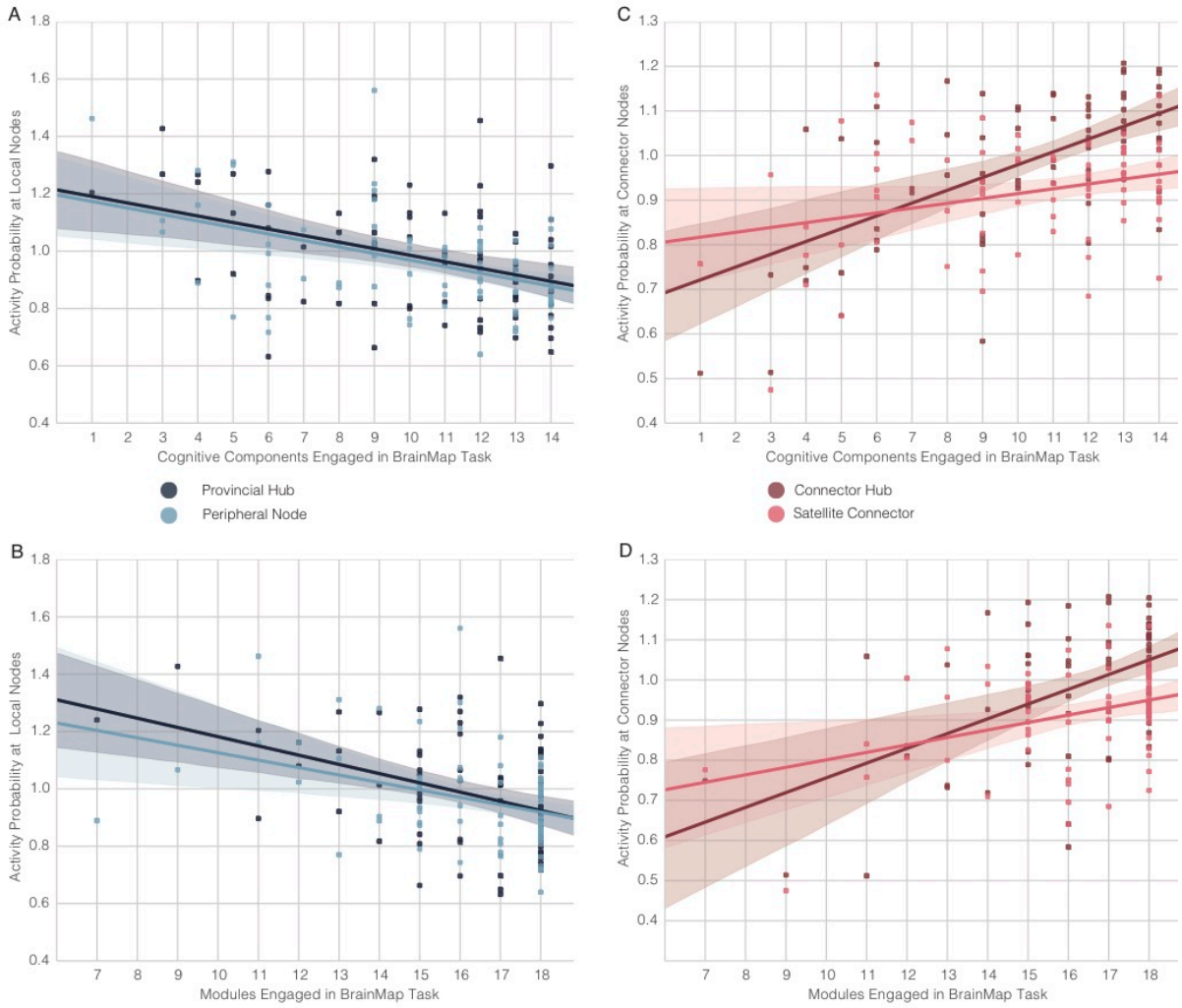


Figure 3 | Activity Probabilities at Different Types of Network Nodes Across BrainMap Tasks Based on the Number of Modules or Cognitive Components Engaged. Each dot is the mean of activity probabilities in a BrainMap task at the active nodes of interest divided by the mean of activity probabilities across all active brain nodes in the task. In line with predictions from Figure 1, there is no increase in activity probability at provincial hubs and peripheral nodes (i.e., local nodes) in tasks that engage more components (A) or modules (B), but there is an increase at connector hubs and satellite connectors in tasks that engaged more cognitive components (C) or modules (D).

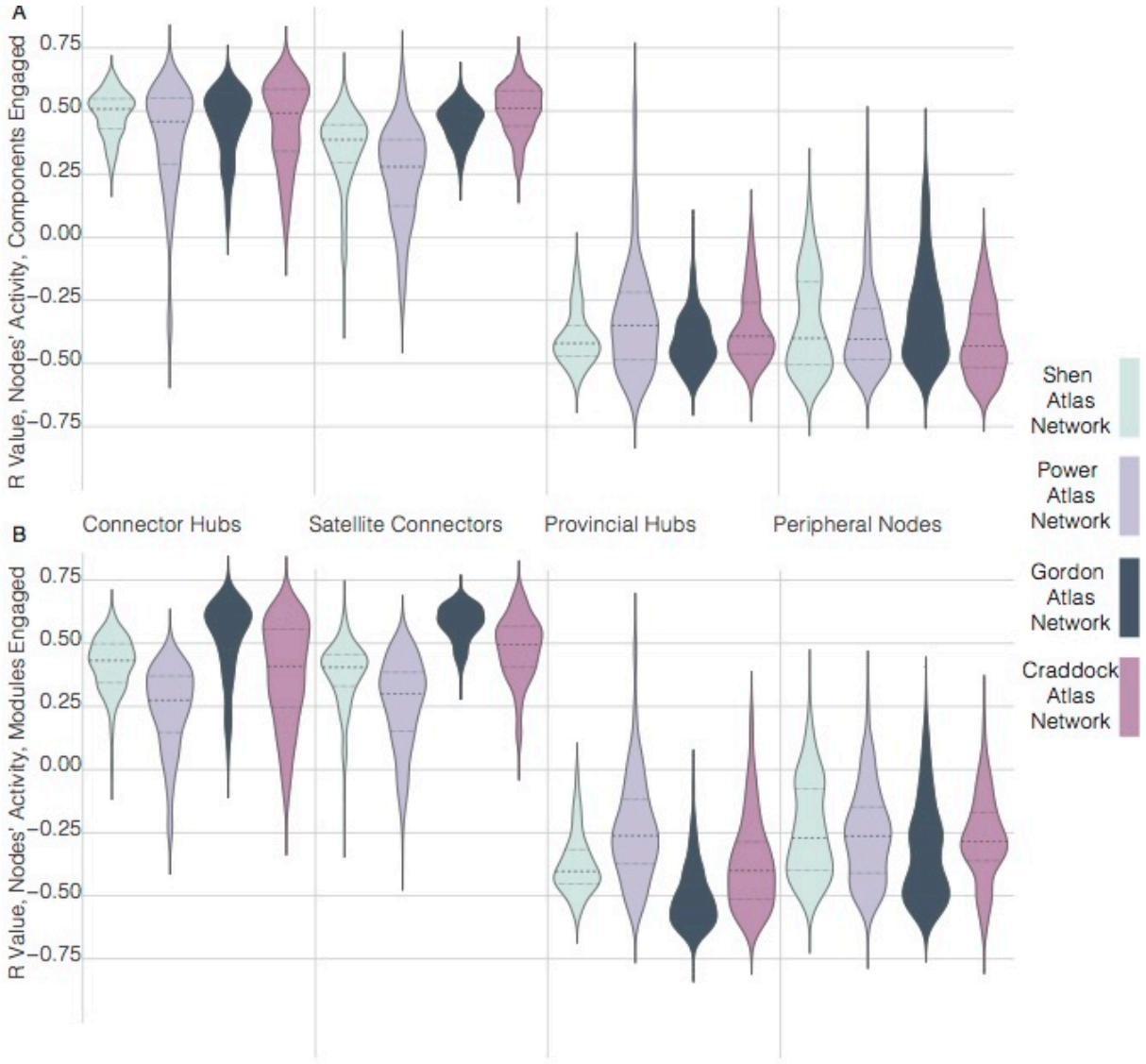


Figure 4 | Correlations Between Different Types of Network Nodes' Activity Probability and the Number of Cognitive Components (A) or Modules (B) Engaged in the Task Across Atlases and Costs in Individual Subjects. A kernel density plot for each of the four brain networks and four types of nodes is presented. Each kernel represents all subjects at all cost thresholds in the given network for that type of node. The median and the 25th and 75th percentile of Pearson *r* values are shown in each kernel, with the median as the dashed line, and percentiles as a dotted line.

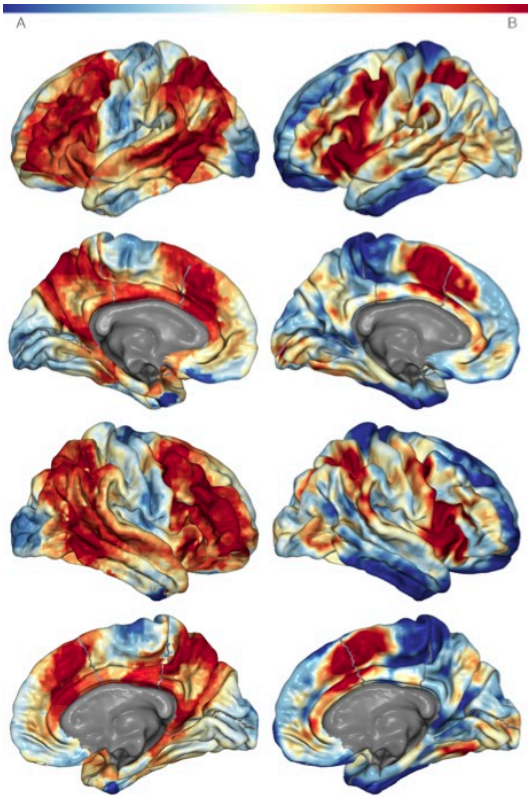


Figure 5 | Spatial Distribution of Connector Nodes and Areas Where Many Cognitive Components Have High Probability of Activity. A, left column, the mean of each voxel's participation coefficient across costs for all networks (i.e., atlases) that had coverage at that voxel (i.e., the calculation ignores 0 values if the network did not have a node that covered the voxel). This data is derived from the network model of spontaneous neural activity measured with rs-fMRI. B, right column, the number of cognitive components that activity at the voxel is associated with above a probability greater than 1e-5. This data is derived from the author-topic model of BrainMap. The color bar is shown at the top. Values have been normalized to sum to 1 for an accurate comparison across models. Thus, darker red areas represent higher values for both metrics, with lower values in blue. The results from this analysis for subcortical regions is presented in Supplemental Figure 3.

Acknowledgements: This work was supported by NIH Grant NS79698. National Science Foundation Graduate Student Fellowship, NUS Tier 1, Singapore MOE Tier 2 (MOE2014-T2-2-016), NUS Strategic Research (DPRT/944/09/14), NUS School of Medicine Aspiration Fund (R185000271720) and Singapore NMRC CBRG14nov007. BrainMap® is a registered trademark of the University of Texas. The Brain-Map database is a copyrighted electronic compilation. Comprehensive access to the BrainMap database was authorized by a collaborative-use license agreement (http://www.brainmap.org/collaborations.html, last accessed September 12, 2014). BrainMap is supported by National Institutes of Health, National Institute of Mental Health Award R01 MH074457.

Supplemental Information

Spontaneous Neural Activity Measured with BOLD fMRI

24 healthy participants (age range = 18–37 years, mean age = 24 years, 12 male) were studied. All healthy participants were prescreened to exclude individuals with a history of neurologic or psychiatric conditions. Informed consent was obtained from participants in accordance with procedures approved by the Committees for Protection of Human Subjects at the University of California, Berkeley. Brain images were collected on a 3-T Siemens MAGNETOM Trio MRI scanner using a 12-channel head coil. Structural images were acquired using a T1-weighted MPRAGE (Magnetization Prepared Rapid Acquisition Gradient Echo; TR = 2,300 msec, TE = 2.98 msec, 9° flip angle, 1 × 1 × 1 mm voxels). For each subject, 6 blocks of 10 minutes each of T2*-weighted BOLD sensitive gradient echo echo-planar imaging sequence data (EPI) were analyzed (2,610 time points for each subject, TR = 1370 msec, TE = 26 msec, twenty-four 3.85mm thick axial slices, alt+z2 order (interleaved ascending, beginning at second slice), 2.344 x 2.344 x 3.850 mm voxels, 62° flip angle, FoV =1125 x 1125, Percent Phase Field of View = 100, matrix = 96 x 96). All participants were instructed to simply stay awake with their eyes open. No other instructions were given.

Construction of the Graph Theory (Network) Model of the Brain

Image preprocessing was carried out in CPAC (Configurable Pipeline for the Analysis of Connectomes). Advanced Normalization Tools (ANTS) was used to register the images to MNI152 (Montreal Neurological Institute, Montreal, QC, Canada). FSL/FAST were used to automatically segment brain images into white matter (pr=0.96), gray matter (pr=0.7), and cerebral spinal fluid (pr=0.96). Boundary-Based Registration was used to register the EPI values to the anatomical image. This uses the anatomical segmentation outputs to improve the co-registration of EPI images to the anatomical image. Slice timing was used to adjust the time course of voxels in each slice to account for the difference in time between the acquisition of the first and last slice. Volume realignment used the Friston 24-Parameter Model, which is the 6 motion parameters of the current volume and the preceding volume, plus each of these values squared. The motion parameters estimated in this set are then included in the General Linear Model to regress out motion-related artifacts (i.e. regression of motion parameters). The mean white matter and cerebral spinal fluid time series are calculated by averaging signal over all voxels within the white matter or cerebral spinal fluid masks for each time point. Mean white matter and cerebral spinal fluid time series are then used as temporal covariates and, along with linear and quadratic signals, are removed from the time series through linear regression. The time series is also bandpass filtered from 0.009–0.08 Hz to remove physiological noise such as cardiac and respiratory artifacts. A 5.5mm gaussian filter (i.e., smoothing kernal) is then applied to the image. While our main analyses do not include the usage of motion scrubbing, in order to ensure that none of our results are impacted by subject motion, analyses were performed after

motion scrubbing and average frame-wise displacement for each subject was not correlated with any of our measures of interest (see Supplemental Information and Supplemental Table 1).

An Ontology of Cognitive Functions

One common goal of cognitive neuroscience has been to localize cognitive functions. This has been a difficult goal to achieve for numerous reasons (131-138), foremost being the significant challenge of building an ontology of cognitive functions (139). As one scientist eloquently put the challenge: “for example, while we think that ‘‘working memory’’ is a unique function implemented in the brain, it may be the case that there is no such function implemented by the brain and that what we call working memory is in reality a combination of some other functions” (134). Deriving an ontology of the fundamental cognitive functions requires a principled analysis of brain activity across many different cognitive tasks under the appropriate assumptions that each cognitive function is recruited across a variety of cognitive tasks and multiple cognitive functions are recruited in each cognitive task. The author-topic model is the first mathematical model that formally takes into consideration those two assumptions, and it does not require any assumptions concerning the existence of any particular cognitive function or functions (e.g., “working memory”). The author-topic model was used to build an ontology of cognitive functions by jointly modeling the association between brain activity and 77 cognitive tasks, based on the two assumptions above, across over 9,000 experiments (85). With this approach, the only parameter that must be set is the number of cognitive components; however, from 6-16 cognitive components, a nested ontology was found. For example, when changing the model from 11 to 12 components, one cognitive component divides into 2 cognitive components, while the other cognitive components remain unchanged. This suggests that each cognitive component is likely comprised of more specific, micro-level cognitive functions. While every micro-level cognitive function is certainly not differentiated by this model, the analyses in this study only depended on an accurate macro-level description of cognitive functions that are super sets of the micro-level cognitive functions (which is suggested by the observed nested ontology) and a quantification of how many of these macro-level cognitive functions were engaged in each BrainMap task. Thus, our results apply to a macro-level description of cognitive functions that are super sets of possible micro-level cognitive functions. Moreover, there is recent evidence that the principle of modularity governs micro-level functioning as well. For example, Aplysia locomotion has been found to be executed by 12 modules that each map onto physically discrete brain regions as well as a particular function during locomotion (55). Future work will bridge the gap between our macro and micro-level understanding of the modular functional architecture of the brain.

Replications

We chose to replicate our findings in a number of ways. We analyzed three different

and independent datasets of spontaneous neural activity measured with rs-fMRI. First, we used four different brain atlases to analyze our own dataset. Moreover, we tested our hypotheses at both the individual and group levels. Second, we analyzed a spontaneous neural activity (measured via rs-fMRI) correlation matrix from another group (74) using a fifth brain atlas. Third, we performed our analyses with published graph metrics for nodes, including a module division and participation coefficients from a dataset that averaged across two cohorts (56, 78). This variety of datasets, analyses, and processing pipelines assured that our results are robust and can be fully replicated by other researchers.

Availability of Data and Code

All efforts were made to make our experiment completely reproducible. Analysis, visualization, and plotting code was written in Python using igraph, numpy, scipy, scikit-learn, igraph, nibabel, pycortex, and seaborn. The z-score of the Rand Coefficient code was executed by MATLAB Engine for Python and was download from NetWiki. Graph theory code and analysis code is available on GitHub upon request, and was reviewed in a collaborative code meeting at UC Berkeley. The cognitive component model is available via Freesurfer. The original unprocessed rs-fMRI data is available upon request. The rs-fMRI processing pipeline is fully detailed in the Methods, and can be used to exactly (or not exactly, if one wishes) recreate the data we used for the analysis via CPAC, which is also freely available. Thus, this entire analysis can be recreated by anyone with no software purchases or development required.

Alternative Community Detection Methods and Algorithms

To ensure that our results were not dependent on any particular cost, we executed two additional methods that make use of a wide range of costs. First, community detection was ran on each subject across a range of costs, from 0.01 to 0.1, in 0.001 steps. Each cost resulted in a consensus style matrix with 1's between nodes in the same modules and 0's elsewhere. These matrices are averaged (or, for the Crossley Network, we used the group average correlation matrix here), and then community detection is run from 0.1(0.0916 for the Crossley Network) to 0.01 in 0.001 steps, storing a consensus style matrix for each run. The average of these matrices was left un-thresholded, and community detection was run 100 times. A consensus style matrix is stored for each run. If all 100 partitions are identical, the procedure ends. If any partitions are different, the 100 consensus style matrices are averaged and community detection is run again. This is the procedure described in (59). However, a second iteration was never required, because there was enough consensus from the previous community detection techniques to result in a matrix that always leads to the same solution. Results were dramatically similar to our main method. We refer to this as community detection across costs (see Supplemental Table 1; "InfoMap Across Costs").

Second, in order to further test our community detection procedure, for each subject, we

ran community detection at 0.1. A consensus style matrix is formed. The cost is then decreased by 0.001, and community detection is run again. The consensus style matrix is then updated for the new partition, except for rows and columns for which the node has no edges in the current version of the graph or the node is not in a community with at least 5 nodes. This procedure continues until the cost is equal to 0.01. Thus, for each subject, the consensus style matrix that is formed represents the community assignments for each pair of nodes at their sparsest level possible (i.e., before they become disconnected from the graph). This method is very similar to previous methods (16, 56). The average of these matrices was then clustered according to the method described above, the only difference being the original subject matrices were formed with this technique instead of averaging across costs. Results were dramatically similar to our main method. We refer to this as the recursive method (see Supplemental Table 1).

To make sure that our results were not dependent on the InfoMap algorithm we chose, we used the Louvain community detection in the community detection across costs method. Note, however, that the results form the Louvain community detection were very variable across subjects a lower costs (< 0.05), and led to a very large number of modules (> 30 in many cases) so we only included costs from .1 to 0.05. Again, results were dramatically similar to our main method (see Supplemental Table 1; “Louvain Across Costs”).

Note that, in all of the Supplemental Tables, we report findings based on the average of each node’s participation coefficient and within-module-degree across the original range of costs (.2 to 0.05, in 0.01 steps).

Comparison of Modules and Cognitive Components

To calculate a p value, nodes were randomly reassigned to cognitive components and modules, and Normalized Mutual Information (NMI) was recalculated. In this calculation, the number of nodes in each module and cognitive component were kept the same. For example, if module 1 contained 12 nodes, and cognitive component 1 contained 14 nodes, we replaced module 1 with 12 random nodes and cognitive component 1 with 14 random nodes. This was performed 1e8 times, calculating NMI at each iteration. A one sample t-test between the random partition normalized mutual information scores and the real score in every atlas is $p<1e-323$.

We also note that the average NMI between individual subjects’ partitions and the group average partition was 0.41. This suggests that the group level partition is usually more similar to the cognitive component model than individual subjects are to the group level partition.

Finally, we also used the z-score of the Rand coefficient to compare the modules to the cognitive component (see Supplemental Table 2). This offers a clear statistical

interpretation of the values.

Sub-Cortical Views of Modules, Cognitive Components, Participation Coefficients and Areas Where Activity Is Associated with Multiple Cognitive Components

Figure 2 and Figure 5 are surface rendered views of the cortex (i.e., values from the cortical surface and below the cortical surface are projected onto the cortical surface), which do not allow for a view of sub-cortical brain structures. Thus, Supplemental Figure 1 shows the sub-cortical module (A) and the corresponding cognitive component (B). Supplemental Figure 3 shows axial views of the participation coefficients (A) and the areas where activity is associated with multiple cognitive components (B).

Effects of Subject Head Motion

While our main analyses do not include the usage of motion scrubbing, in order to ensure that none of our results are impacted by subject motion, we removed any frames with frame-wise displacement greater than .05, as well as the frame before and after. For every atlas, NMI between the scrubbed time-series partition and intact time-series partition was greater than 0.90. Participation Coefficients and Within-Module-Degree Z-Scores in the scrubbed time-series partition and intact time-series partition all correlated at greater than $r = 0.70$. Modularity (i.e., Newman's Q), average Participation Coefficients, and maximum Participation Coefficients and Within-Module-Degree Z-Scores were not correlated with subject's average frame-wise displacement ($p > 0.25$ in all cases). Moreover, we ran our main analyses on the scrubbed data using the InfoMap community detection across costs, and the results were very consistent with the un-scrubbed data (see Supplemental Table 1; "Scrubbed").

Effects of Cost Thresholds

Given that there is there is no ''correct'' cost threshold, activity was calculated for the four different types of nodes in the four networks at cost thresholds from 0.05 to 0.20 in 0.01 steps (0.01 to 0.0916 for the Crossley Network). For every network we analyzed, across costs, there were positive correlations between activity at connector nodes and the number of cognitive components or modules engaged in a task, but a non-significant or negative correlation between activity at local nodes and the number of cognitive components or modules engaged in a task. Supplemental Figure 2 shows these results plotted as kernel density estimations with the median and the 25th and 75th percentile of Pearson r values across cost thresholds shown for the particular type of node in that particular network. This ensures that our findings are not sensitive to the specific definition of the network's nodes or the densities of the networks constructed based on those nodes. Moreover, we averaged each node's participation coefficients and within-module-degree z-scores across costs, and results were very similar (see Supplemental Table 1).

Effects of Cutoffs for Node Type

The values we used were not arbitrary; as the within-module-degree is z-scored, the mean is 0.0. However, in the rare case that a node is not assigned to a module (mathematically, it is calculated as belonging to it's own module, alone) within-module-degree z-score of 0.0. Thus, to avoid classifying these nodes as provincial hubs or connector hubs, we set the threshold at 1e-5 instead of the mean of the values, as we did with the participation coefficients. However, results were not greatly impacted by changes. For example, with the threshold set to 1 for the Shen Atlas (Connector Hubs, Cognitive Components (r= 0.518, p=1.3e-06), Modules (r=0.494, p= 4.9e-06); Satellite Connectors, Cognitive Components (r=0.576, p= 4e-08), Modules (r= 0.530, p= 7e-07); Provincial Hubs, Cognitive Components (r= -0.151, p= 0.188), Modules (r= -0.201, p= 0.079); Peripheral Nodes, Cognitive Components (r= -0.502, p= 3.2e-06), Modules (r= -0.401, p= 0.0003)). Higher threshold resulted in less than 5 percent of nodes that were provincial or connector hubs. The same pattern was observed in the opposite direction; results were significant at -1 (Connector Hubs, Cognitive Components (r= 0.561, p= 1.1e-07), Modules (r= 0.5033, p= 3e-06); Satellite Connectors, Cognitive Components (r= 0.499, p= 3.9e-06), Modules (r= 0.523, p= 1e-06); Provincial Hubs, Cognitive Components (r= -0.522, p= 1.1e-06), Modules (r= -0.455, p= 3.1e-05); Peripheral Nodes, Cognitive Components (r= 0.035, p= 0.761), Modules (r= 0.0084, p= 0.942)). Lower thresholds resulted in less than 5 percent of nodes that were satellite connectors or periphery nodes.

Results were significant up a participation coefficient of 0.80 (Connector Hubs, Cognitive Components (r=0.592, p=1.4e-08), Modules (r=0.364, p=0.001); Satellite Connectors, Cognitive Components, (r =0.351, p=0.001), Modules (r=0.295, p=0.009); Provincial Hubs, Cognitive Components (r=-0.281 p=0.01), Modules (r=-0.22, p=0.04); Peripheral Nodes, Cognitive Components (r=-0.41, p=0.0002), Modules (r=0.22, p=0.05)). At higher cutoffs, less than 5 percent of nodes were characterized as connectors.

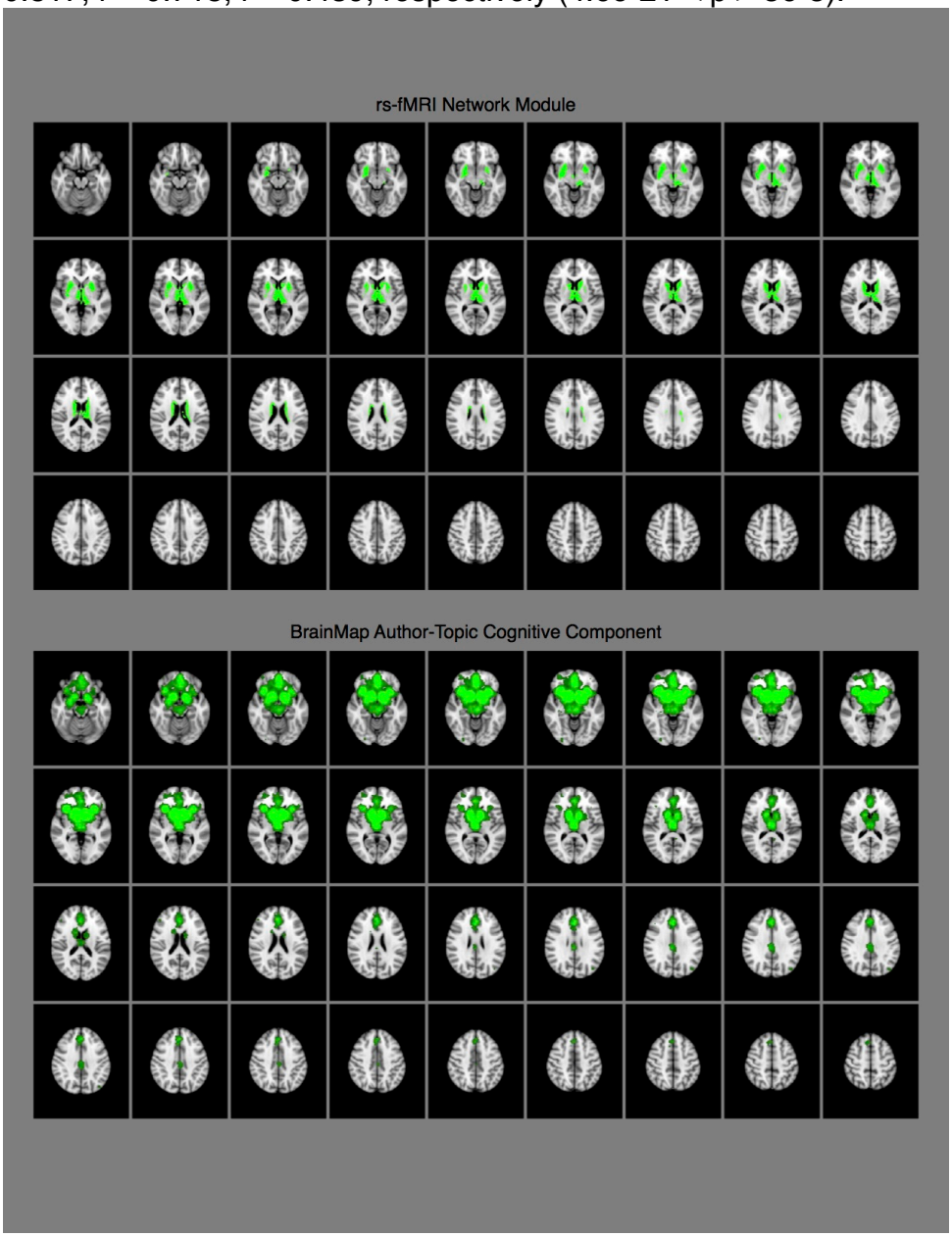
Author-Topic Cognitive Component Model with Different Numbers of Cognitive Components

The Author-Topic model can be applied to estimate a particular number of Cognitive Components. We insured that our results are replicable across Author-Topic models with different numbers of Cognitive Components by running our main analyses across the model with 10-13 cognitive components. These results are shown in Supplemental Table 3.

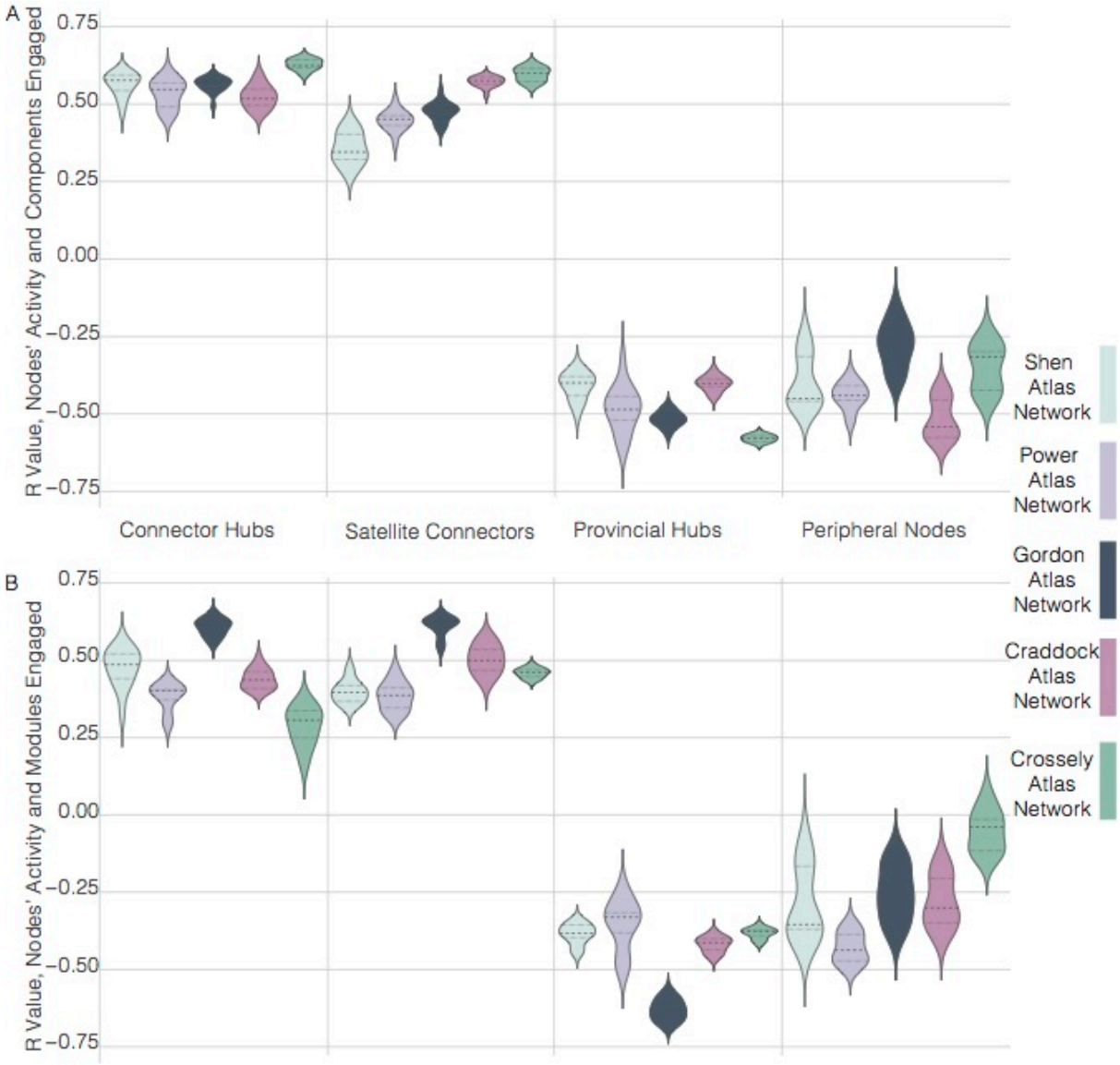
Different Methods for Calculating the Number of Cognitive Components Engaged

In addition to the method implemented in the main text, we also used the original probability (X) values (which sum to 1) to quantify the number of cognitive components

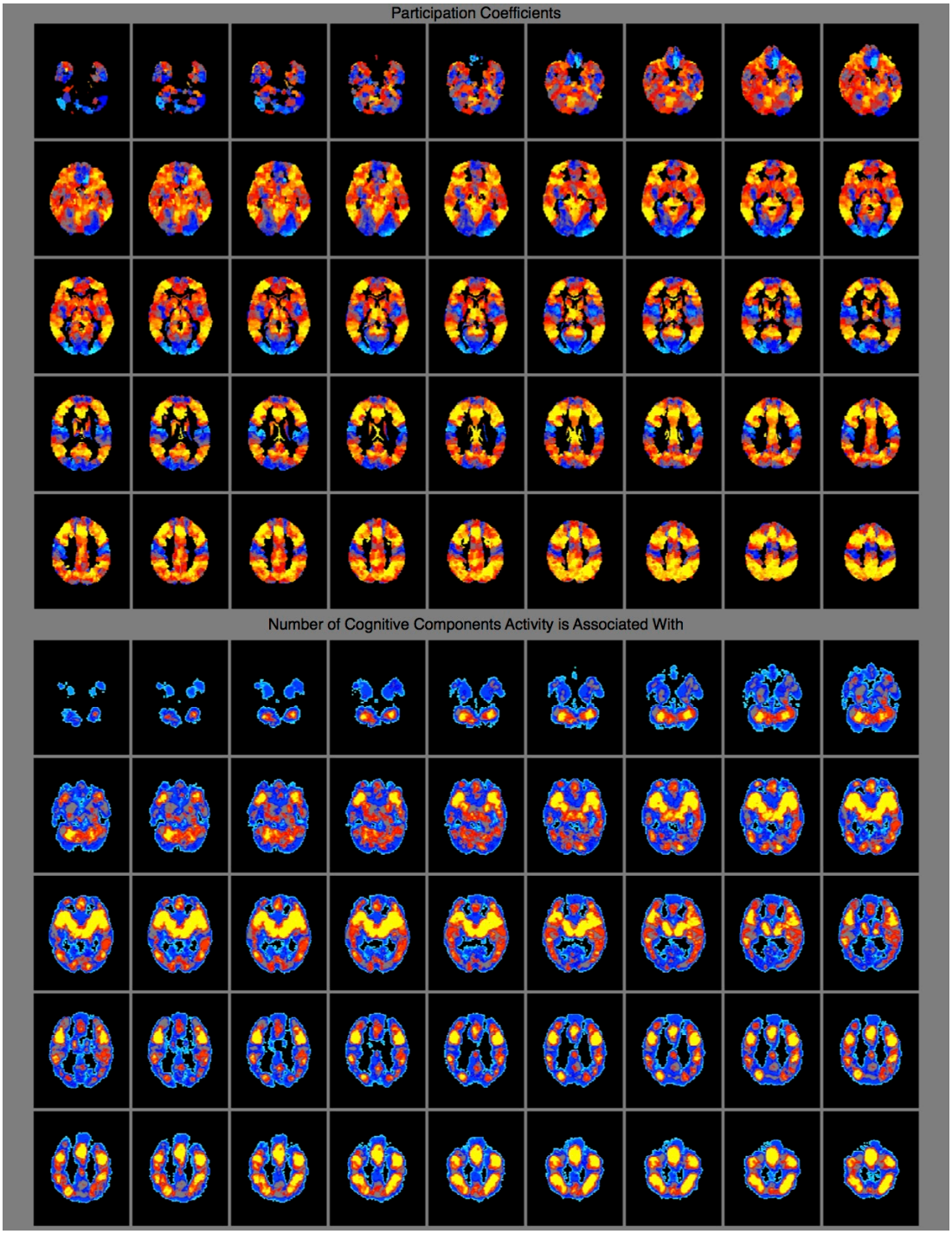
engaged in a task in three ways. In Supplemental Table 4, we present the results where we quantified the number of cognitive components engaged by the entropy of the distribution of X values (Column 1), 1 minus the variance of the distribution of X values (Column 2), and the number of cognitive components with an X value greater than 1/14 (Column 3). These calculations were significantly correlated with our main method at r = 0.817, r = 0.718, r = 0.439, respectively (4.6e-21 < p > 5e-5).



Supplemental Figure 1 | Subcortical rs-fMRI Network Module (top) and BrainMap Author-Topic Cognitive Component (bottom).



Supplemental Figure 2 | Correlations Between Different Types of Network Nodes' Activity and the Number of Cognitive Components (A) or Modules (B) Engaged in the Task Across Atlases and Datasets. A kernel density plot for each of the five brain networks and four types of nodes is presented. Each kernel represents results across cost thresholds, from 0.05 to 0.2 for the Shen, Power, and Gordon atlas networks and 0.01 to 0.1 for the Craddock atlas network. The median and the 25th and 75th percentile of Pearson *r* values are shown in each kernel, with the median as the dashed line, and percentiles as a dotted line.



Supplemental Figure 3 | Subcortical Spatial Distribution of Participation Coefficients (Connector Nodes) (top) and Areas Where Many Cognitive Components Have High Probability of Activity (bottom).

Supplemental Tables

The Diverse Club: The Integrative Core of Complex Networks
M.A. Bertolero, B.T.T. Yeo, & M. D'Esposito

A complex system can be represented and analyzed as a network, where nodes represent the units of the network and edges represent connections between those units. For example, a brain network represents neurons as nodes and axons between neurons as edges. In many networks, some nodes have a disproportionately high number of edges. These nodes also have many edges between each other, and are referred to as the rich club. In many different networks, the nodes of this club are assumed to support global network integration. However, another set of nodes potentially exhibits a connectivity structure that is more advantageous to global network integration. Here, in a myriad of different biological and man-made networks, we discover the diverse club—a set of nodes that have edges diversely distributed across the network. The diverse club exhibits, to a greater extent than the rich club, properties consistent with an integrative network function—these nodes are more highly interconnected and their edges are more critical for efficient global integration. Moreover, we present a generative evolutionary network model that produces networks with a diverse club but not a rich club, thus demonstrating that these two clubs potentially evolved via distinct selection pressures. Given the variety of different networks that we analyzed—the c. elegans, the macaque brain, the human brain, the United States power grid, and global air traffic—the diverse club appears to be ubiquitous in complex networks. These results warrant the distinction and analysis of two critical clubs of nodes in all complex systems.

Many complex systems—neural, the power grid, and air traffic—can be analyzed as a network with graph theory, where units (e.g., neurons or airports) and connections (e.g., axons or flight routes) are treated as nodes and edges in a graph, respectively. These systems all exhibit a *community* structure—nodes cluster into communities such that nodes are more strongly connected to other members of their community than to members of other communities(43, 56, 140, 141). Each node within one of these communities can play a distinct role in the overall network topology. In many different systems, from brains to air traffic, applying two nodal role metrics—degree and participation coefficient—to the graph representing the system identifies nodes that have been proposed to be hubs that perform integrative or coordinative functions(40, 54, 78-82, 127, 142-147).

Degree is a nodal metric of how many edges the node has. While a node's degree captures its magnitude of connectivity, it does not capture the diversity of the node's connectivity across communities in the network. The participation coefficient is a nodal metric of the *diversity* of each node's connections across the network's communities (79, 81). A node's participation coefficient is maximal if it has an equal number of edges to each community in the network. Mathematically, a node's participation coefficient is

independent of the node's degree, as it only measures the diversity of a node's connections across communities. Empirically, across a wide range of networks, the participation coefficient is not correlated with degree, but nodes can be high in both degree and participation coefficient(78).

Across various networks, nodes with a high degree are connected to each other at a rate greater than would be expected in a randomly organized graph(148). This subset of highly interconnected nodes is referred to as the "rich club". The rich club is thought to be critical for global communication given that these nodes have a high betweenness centrality, in that, if the shortest paths between all pairs of nodes is found, many of these shortest paths involve rich club members(142, 149). Evidencing the rich club's criticality, these brain regions are more likely to exhibit brain pathology than other brain regions in many neurological and psychiatric disorders(127). In line with this empirical finding, in silico "attacks" on networks demonstrate that, when edges between nodes in the rich club are removed, global efficiency is decreased (i.e., the sum of shortest paths between all nodes increases)(142). Given these characteristics, the rich club, which has been investigated in over 200 published reports to date, has been proposed to be an integrative and stable core of brain regions that coordinates the transmission of information across the network.

However, as opposed to the high magnitude of connectivity that high degree nodes exhibit, nodes with a high participation coefficient exhibit diverse connectivity. This connectivity pattern places these nodes at the topological center of the network(106), ideal for integration and coordination. In the human brain, these nodes are also located where many communities are within close physical proximity(78), and appear to control or coordinate which regions are "functionally" connected during cognition, in that activity in these nodes predicts changes in the connectivity of other nodes(67, 71), particularly the connectivity between nodes in different communities during cognitive tasks(150). These nodes have also been implicated in a diverse range of tasks(20, 151). Moreover, damage to these brain regions causes a decrease in the modular architecture of the human brain network(152) and widespread cognitive deficits(125). Finally, a recent analysis showed that only these brain regions exhibit increased activity if more communities are engaged in a cognitive task, which suggests that they are involved in processes that are more demanding as more communities are engaged(140). A parsimonious explanation of these empirical findings is that nodes with high participation coefficients integrate information and coordinate connectivity between communities, which allows for modular local processing.

Thus, nodes with both a high participation coefficient and a high degree have been proposed to perform integrative and coordinative functions. Here, we make a comprehensive distinction between high degree and high participation coefficient nodes. We demonstrate, with data from multiple systems, that networks contain a *diverse club*—a set of high participation coefficient nodes that are more highly interconnected than the rich club. We find that the two clubs are largely comprised of different nodes.

Moreover, we analyze the anatomical locations of these clubs in the human brain, the connectivity patterns of these clubs, the functional responses of both clubs in the human brain during cognitive tasks, and how damage to nodes in each club impacts the network's efficiency. Finally, we present a generative evolutionary network model that generates graphs with a diverse club but not a rich club. From these analyses, we demonstrate that the diverse club exhibits, to a greater extent than the rich club, properties that are consistent with an integrative and coordinative function. Together, these results support the distinct roles of the diverse club and rich club in network communication.

Clubness

We analyzed structural and functional networks from multiple species—the c. elegans' structural and functional networks, the macaque structural network, the human functional network, the United States power grid network, and the global air traffic network (see Methods for network construction details). We consider both structural and functional networks, as degree can be artificially inflated in functional (i.e., correlational) networks(78). Graph theory allows for the comparison of network organization among very different systems. While the c. elegans networks, the macaque network, and the human brain networks are clearly different networks, they are all biological neural networks that were shaped by evolution. Thus, we also investigated man-made networks to determine if they exhibit properties similar to the biological networks.

The rich club is the set of high degree nodes in a graph; it is referred to as a club because these nodes are highly interconnected to each other. We refer to how interconnected a club is as "clubness". We measure clubness with the normalized club coefficient, which is the number of intra-club edges the club has relative to the mean of that value in a large set (here, 1000) of random graphs. These random graphs are generated based on the original graph; all nodes maintain their degree, but the edges are randomly placed and the edge weights (in the c elegans functional networks, human functional networks, and air traffic network; other networks contain only binary edges) are shuffled between nodes with the same degree, which accounts for the contribution of both edge placement and edge weight to the normalized club coefficient(148). Results were very similar without shuffling edge weights, which only accounts for the contribution of edge placement (Extended Data Figures 1, 2).

For every network, we defined the rich club and its clubness across different ranks. A rank defines the cutoff for which nodes are in the rich club. For example, in a network with 100 nodes, a rank of 85 contains nodes with a degree greater than or equal to the value of the node with the 15^{th} highest degree. Here, weighted degree (sum of edge weights) was used if the network was weighted. In addition, for every network, we defined the diverse club—the club of high participation coefficient nodes—and its clubness at each rank. We then used a myriad of analyses to characterize, and make

distinctions between, the rich club (i.e. high degree nodes) and the diverse club (i.e. high participation coefficient nodes) in each network.

Clubness of the rich club and the diverse club

We sought to measure if the diverse club or the rich club is more interconnected than the other. For each network, we applied community detection using InfoMap(129) (e.g., Extended Data Figure 3). Next, we calculated the degree and participation coefficient of each node in every network. Using these values, for each network, we calculated the clubness for both clubs at every possible rank. For both clubs, for every network, as the rank increased, clubs with a clubness greater than 1 (i.e., 1 means equal to random) were detected (Figure 1, Extended Data Figures 1, 2, and 4). However, in every network, as the rank increased to only include those nodes with the highest degree or participation coefficient, the clubness of the diverse club was higher than that of the rich club. These results demonstrate that, across a range of networks, the club of high participation coefficient nodes (the diverse club) is more strongly interconnected than the club of high degree nodes (the rich club).

Network topology of the rich club and the diverse club

We further analyzed clubs at the rank that corresponds to the 80^{th} percentile, as this is the rank, across networks, where the normalized club coefficient increased dramatically. For example, in the human brain networks, which contain 264 nodes, the clubs contained 53 nodes each. To visualize the spatial topology of the derived network communities, we used the ForceAtlas2(153) algorithm, which simulates a physical system in which nodes repel each other like charged particles and edges attract their nodes like springs, which results in nodes in the same community pulling together, and different communities pulling apart from one another. We labeled each node in the graph by their community affiliation and their membership in a rich or diverse club (Figure 2 shows the communities in c. elegans and human resting-state; Extended Data Figure 3 shows a non-biological network, e.g. air traffic).

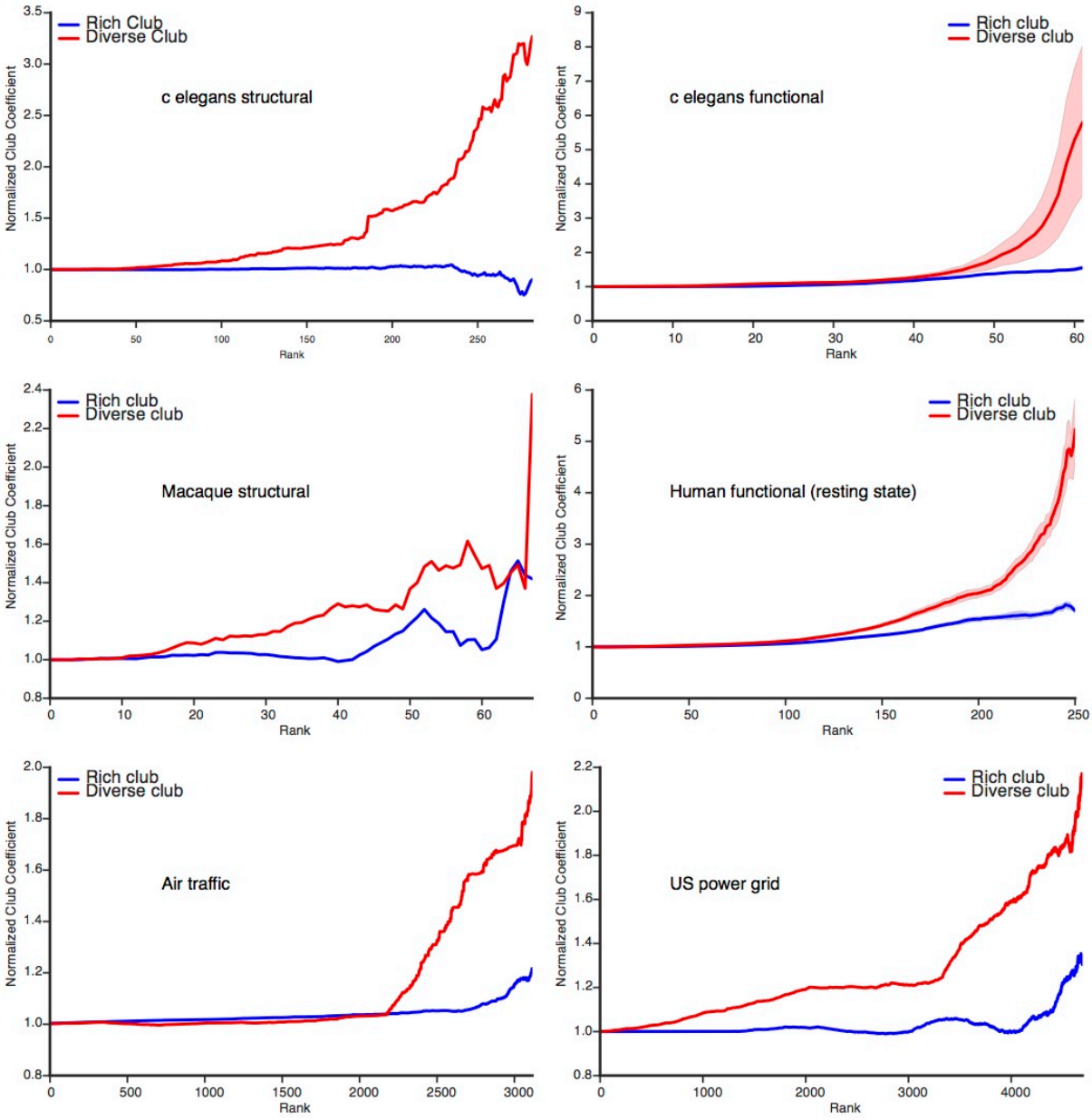


Figure 1 | Clubness of the diverse and rich clubs. Normalized club coefficients (clubness) for the rich and diverse clubs in every network. For functional networks, the mean across network costs for the clubness is plotted, with 95 percent confidence intervals shaded. In every network, as the rank increased and only nodes with a high participation coefficient (red) or degree (blue) are included in the club, the diverse club is consistently higher in clubness than that of the rich club.

Visual inspection of the c. elegans and human functional networks (Figure 2a,b) suggests that rich club nodes exist on periphery of the graph, whereas diverse club nodes are in the center. There are few nodes that are members of both clubs. Anatomically in the human brain resting-state network, the rich club and diverse club are differentially represented in different communities (Figure 2c,d). These analyses

demonstrate that the clubs exist at different anatomical locations in the human brain as well as different topological locations in the graph.

We then quantified how similar the clubs are in each network, measuring the percentage of possible overlap. Zero percent represents that no nodes were members of both clubs, and 100 percent represents that the clubs are identical. In the human networks, across both resting-state and 6 task states, no more than 11 percent of nodes were in both clubs (Figure 3a). In the functional c. elegans networks, the overlap ranged from 11 to 35 percent (Extended Data Figure 5). The structural networks from c. elegans and macaque, and the air traffic and US power grid networks, showed the highest overlap, ranging from 37 to 50 percent (Extended Data Figure 5). These analyses demonstrate that the diverse and rich clubs are predominately comprised of different nodes.

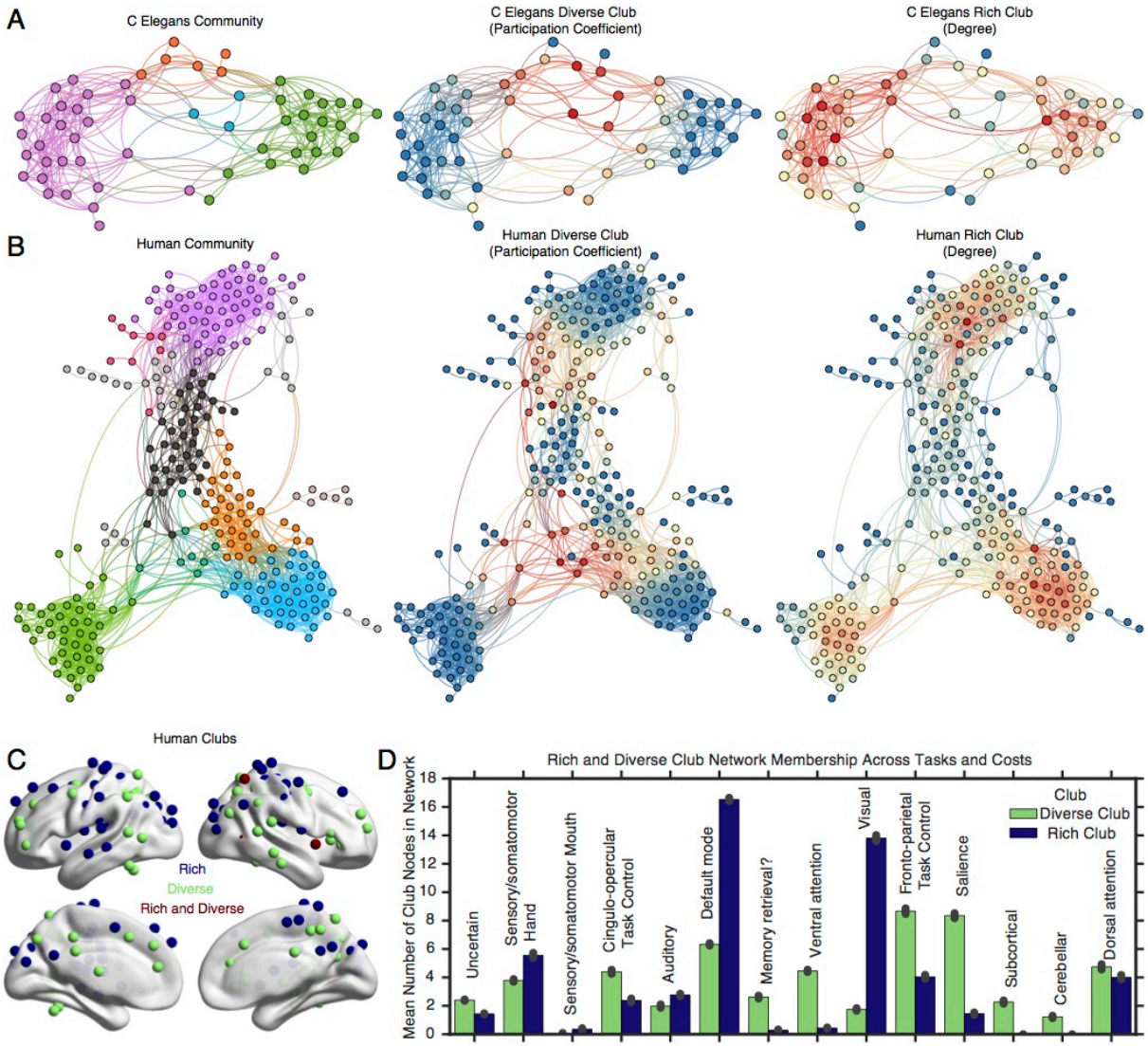


Figure 2 | Topology of the diverse and rich clubs. a, Visualization of a single c. elegans functional network, labeled according to the community affiliation, the diverse club, and the rich club. Nodes in red represent the maximum value for the given metric, yellow is median, and blue is the minimum. Edges are colored by the mix between the two nodes each edge connects. **b**, Visualization of the human resting-state network labeled according to the community affiliation, the diverse club, and the rich club. In both networks, the diverse club clusters in the center of the layout, while the rich club forms clusters on the periphery. **c**, The rich club and the diverse club (human resting-state), along with nodes that are members of both clubs, are shown on the cortical surface of the human brain. **d**, The mean number of club nodes (across costs and tasks) in each functional community identified in the human fMRI data. Community names and division from a previous publication(56).

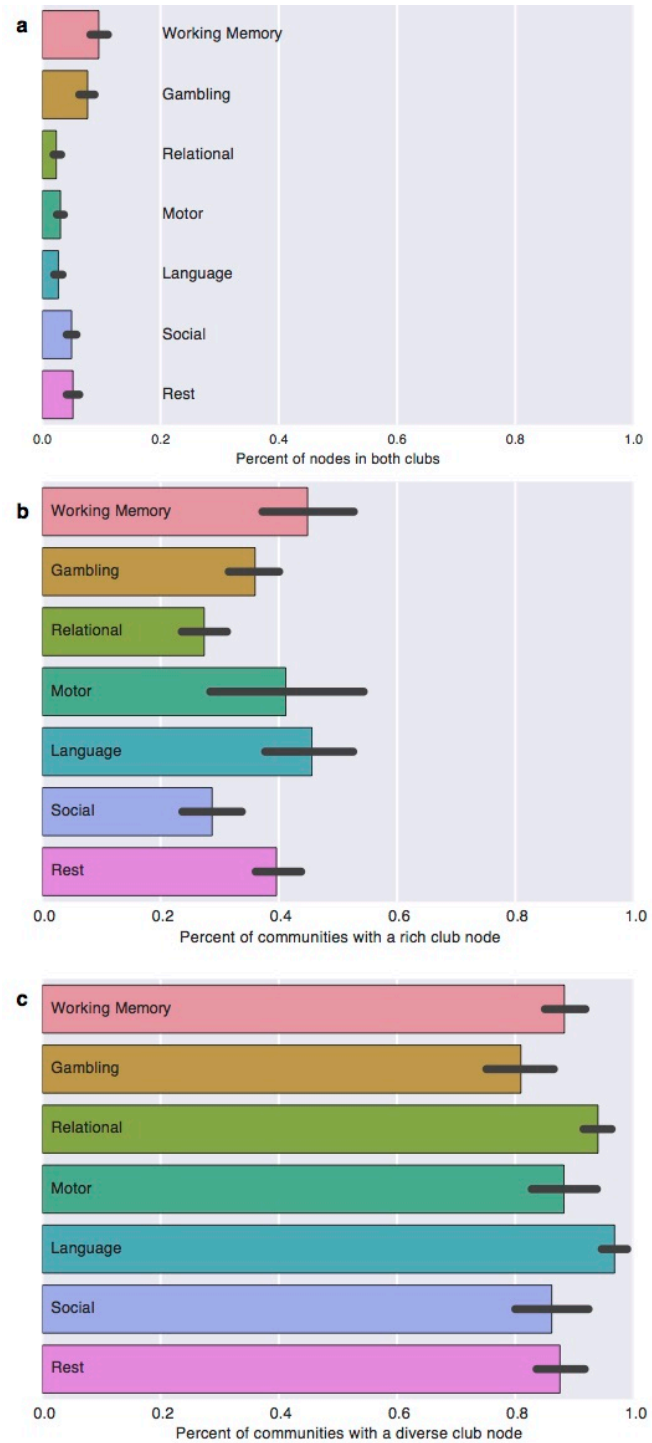


Figure 3 | Distribution of nodes in clubs and communities. **a**, The percentage of nodes in human functional networks that are in both the rich and diverse clubs. Zero percent represents that no nodes were members of both clubs, and 100 percent represents that the clubs are identical. No larger than 11 percent overlap was found across costs (0.05-0.20) and networks. The percentage of communities that contain a node from the rich club (**b**) or the diverse club (**c**). 100 percent represents that every community contains at least one node from that club.

Given that the diverse club appears to be in the topological center of complex networks, and an integrative club of nodes should have members in many different communities, we tested how many communities has a member of each club. Across all networks, a higher percentage of communities contained a node in the diverse club than the rich club (Figure 3b,c, Extended Data Figure 6).

We next tested if the betweenness centrality—the number of shortest paths between pairs of nodes that pass through a node—of the diverse club is higher than that of the rich club. Across all networks we analyzed, the betweenness centrality of the diverse club was either significantly higher than the rich club or there was no significant difference (Extended Data Figure 7). Betweenness centrality, however, does not capture if the network's shortest paths traverse edges between nodes in the rich or diverse club. Thus, we measured the edge betweenness—how many shortest paths between pairs of nodes traverse a particular edge—of the edges between members of the rich club or the diverse club. With this calculation, in almost all networks, the edge betweenness was significantly higher for the diverse club than the rich club. It was not significantly higher in the structural c. elegans and macaque networks (Extended Data Figure 7). Moreover, in the air traffic network, until the clubs reached a size of 356 airports (approximately 10 percent of all airports), the diverse club had more international airports in it than the rich club. Furthermore, flights between airports not in the diverse club are predominately domestic, while international flights were mainly between diverse club airports; this was not the case for the rich club and non-rich club airports (Extended Data Figure 3). These analyses demonstrate that, relative to the rich club, the diverse club is represented in more communities and more shortest paths between nodes pass through the diverse club. These are two properties that are likely critical for global network integration and communication.

Diverse and rich club activity during cognitive tasks

Previously, using the BrainMap database, we demonstrated that the diverse club (nodes with a high participation coefficient) exhibits increased activity in tasks that engaged more cognitive components or communities (see(140, 151) for detailed descriptions). Using the Human Connectome Database resting-state network studied here, we replicated these findings—increased activity of the diverse club was correlated with the number of cognitive components ($r=0.41$, $p=0.0002$) and communities ($r=0.40$, $p=0.0003$) a task engaged. For the rich club, nodes exhibited significantly decreased activity as more cognitive components ($r=-0.53$, $p=1e-7$) or communities were engaged in a task ($r=-0.33$, $p=0.003$). Thus, the diverse club, not the rich club, exhibits increased activity when more communities are engaged in a task, which likely occurs when more integration across and coordination between communities is required.

Targeted attacks of intra-club connections

To further investigate the importance of the diverse and rich clubs for efficient global communication in a network, we simulated lesioning intra-club connections. For each network, in 10,000 iterations, we removed between (randomly) 50 and 90 percent of edges from either club (skipping edges that disconnected the graph into two sub-graphs). We then calculated the increase in the sum of shortest paths, which indicates decreased global efficiency. In every network, removing edges between diverse club nodes increased the sum of shortest paths to a greater extent than removing edges between rich club nodes (Extended Data Figure 8). This demonstrates that the edges in the diverse club, rather than the rich club, are more critical to efficient global communication.

A generative model of the diverse club

Evolutionary pressures have selected networks with rich clubs and diverse clubs. Thus, the final distinction between the diverse club and the rich club we sought to make is if these clubs were potentially naturally selected for different reasons. One of the first observations in neuroscience—Cajal's conservation principle—was that the brain is organized by an economic trade-off between minimizing the number of connections in the network and adaptive topological patterns(154). One topological pattern that might be adaptive is modularity, which is how sparse the connectivity is between communities relative to the connectivity within communities. Another potentially adaptive topological pattern is efficiency, which is the inverse of the sum of shortest paths between all nodes, and thus measures how efficiently signals can be integrated across the network. For example, in brain networks, efficiency is used as a measure of the overall capacity for parallel information transfer and integrated processing (106). Networks that are modular (i.e., exhibit high clustering) and efficient are described as "small world"(58). Thus, we asked the question: do evolutionary pressures that select high modularity and efficiency given a limited number of connections generate a network topology that contains a rich club or a diverse club? In other words, is one of the clubs nature's solution to efficient integrative processing in a modular network?

To answer this question, we developed a generative graph model that maximizes Q and efficiency. The model starts with a graph of 100 nodes that are randomly connected, with 5 percent of all possible edges (272 binary edges). To simulate natural selection of high Q and efficiency, we found edges that, when removed, increase Q and decrease efficiency the least. We remove these edges and then randomly place them back in the network, thus artificially selecting edges in the graph that maximize Q and efficiency. We also ran the same model, except we randomly selected the edges, removed them, and then randomly placed them in the network. This allowed us to decipher if the model selects a network with a diverse club that is more highly interconnected than if random selection had occurred. Our hypothesis was that, if the diverse club is nature's solution to efficient integrative processing in a modular network, a highly interconnected diverse club, but not a rich club, will emerge when networks are selected based on maximizing modularity and efficient integration.

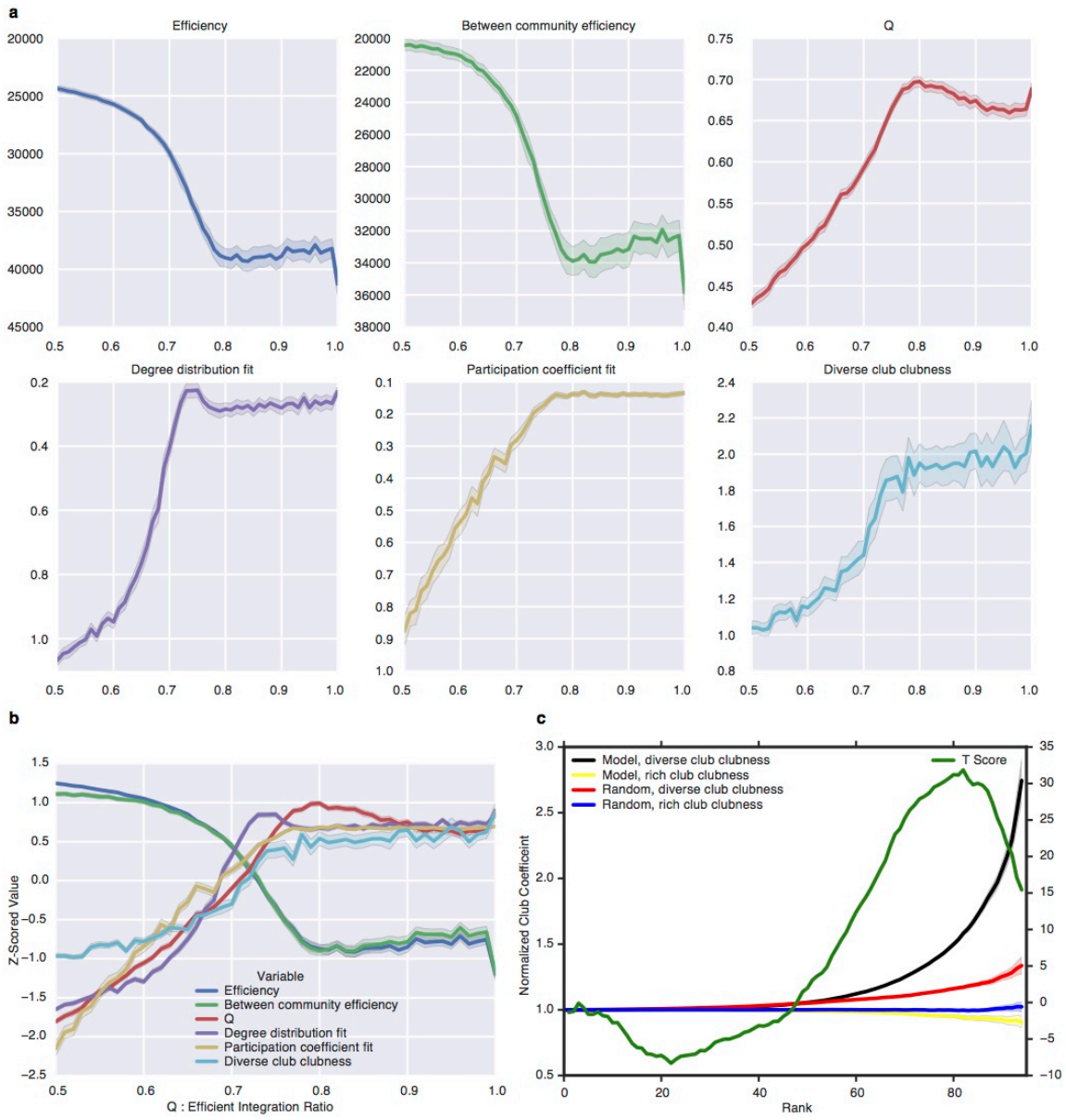


Figure 4 | A generative model of the diverse club. **a**, Six features of the network were analyzed across different ratios of maximizing Q and efficiency (inverse of the sum of shortest paths between all nodes). 100 models were run at each ratio in 0.01 steps. Each value's mean and 95 percent confidence intervals (shaded) are shown. **b**, At ratios of 0.7 to 0.8 between weighting modularity and weighting efficiency, a balance between these six variables was achieved. **c**, The average clubness across 1000 iterations for each rank for the diverse and rich clubs in the generative model at a ratio of 0.75 and the random model, as well as the t-test at each rank between the clubness of the diverse club in the model and the random model (similar results from ratios of 0.70 and 0.80 are shown in Extended Data Figure 9). Only the diverse club in the model has a high normalized club coefficient.

We varied the amount of importance Q or efficiency played in the selection of edges. A ratio of 0.5 equally maximized both Q and efficiency, while 1 maximized only Q. We found that, at a ratio of 0.75 (Q) to 0.25 (efficiency), a balance was achieved with high efficiency, high between community efficiency (the inverse of the sum of shortest paths between nodes in different communities), high modularity, a high correspondence between the degree distribution of the model network and the human brain network (resting-state), a high correspondence between the participation coefficient distribution of the model network and the human brain network (resting-state), and a high clubness of the diverse club. Figure 4 shows these values individually (a) and together (b) across different ratios of maximizing Q versus efficiency.

Using a ratio of 0.75, we ran 1000 iterations of the model and 1000 iterations of the random model. We then calculated the clubness of the diverse club in the model and the random model. We found that, at higher ranks, the clubness of the diverse club in the models that maximize Q and efficiency was higher than the clubness of the diverse club in the random models (Figure 4c; ratios of 0.70 and 0.8 led to similar results (Extended Data Figure 9)). This demonstrates that the diverse club's high clubness is a not a mathematically necessity of defining the club based on high participation coefficients, as randomly selected networks do not exhibit a highly interconnected diverse club. Thus, the diverse club's strong interconnectedness is a non-trivial feature of real networks. Moreover, we did not find high clubness of the rich club in the model. Thus, while the diverse club was captured by the generative model, the rich club was not captured by this model. These results suggest that the diverse club, but not the rich club, might be nature's solution to efficient integrative processing in a modular network.

Discussion

Nodes in a network with many edges (i.e. high degree nodes) or with edges that are diversely distributed across communities (i.e. high participation coefficient nodes) have both been suggested to be integrative hubs(40, 54, 78-82, 127, 142-147). Here, we provided evidence that high participation coefficient nodes, which we refer to as the diverse club, have properties that are more characteristic of integrative hubs, as compared to high degree nodes (i.e., the rich club). The diverse club is more interconnected than the rich club in every network we analyzed—the human brain (in 7 different states), the c. elegans, the macaque brain, the United States power grid, and global air traffic. In the human brain, diverse club nodes are up to four times as interconnected as rich club nodes. Importantly, in all networks examined—especially in the functional brain networks, c elegans structural network, and the man-made networks—very few nodes are members of both clubs.

Having established that the diverse club is relatively distinct from the rich club, we further differentiated the functions of these two clubs in a myriad of analyses. The diverse club spans more communities and has higher edge betweenness and betweenness centrality than the rich club. This pattern of connectivity, which is spread

across the entire network and exhibits the most economical route between nodes, is a critical property of nodes that integrate across network communities. Moreover, in humans, these two clubs exhibit different activity patterns as cognitive tasks become more complex. Unlike rich club nodes, diverse club nodes increase activity in response to more communities being engaged by a task, which likely requires more integration across the network's communities. Finally, across all networks, edges between diverse club nodes are more critical to efficient global communication than the edges between rich club nodes. When diverse club edges were removed, the sum of shortest paths between nodes increased significantly more than when rich club edges were removed.

We also investigated if the two clubs might have distinct evolutionary origins. Many of the brain's network properties that are related to integration are heritable and impact its fitness—how likely that brain network architecture is to be naturally selected. Specifically, the brain network's cost-efficiency ratio (high efficiency given a constrained number of connections—the *wiring cost*) is heritable. Moreover, the diverse club's local efficiency (the sum of each node's shortest paths to all nodes) is heritable(155). Efficiency is also behaviorally relevant, making it likely to factor in natural selection. For example, working memory performance is correlated with network efficiency, and individuals with schizophrenia have lower efficiency and working memory performance(156). Also, higher intelligence quotient scores are associated with higher network efficiency and betweenness centrality of the fronto-parietal network (which we found to have the highest number of diverse club nodes)(157-159). However, brain networks are not purely optimized for efficiency, given that they exhibit high modularity, with segregated communities performing distinct functions, at the cost of lower efficiency(108, 160, 161). Modularity likely increases fitness in information processing systems(26, 162, 163) and confers robustness to network dynamics (i.e., information processing) when the connections between nodes are reconfigured, a process necessary for the evolution of a network(164). Modular networks also outperform and evolve faster than non-modular networks(165) with lower wiring costs than non-modular networks(112). Like efficiency, modularity is also behaviorally relevant, and thus potentially naturally selectable. For example, modularity explains intra-individual variation in working memory capacity(166) and predicts how well an individual will respond to cognitive training(167).

As modularity and efficiency are both heritable and impact the fitness of an organism, we probed the possible evolutionary origins of the two clubs by asking if the rich or diverse club was selected to balance efficient global integration without sacrificing modularity. We found that, if we simulate natural selection for a balance between modularity and efficient integration, a highly interconnected diverse club, but not a rich club, emerges. Thus, the diverse club potentially evolved via selective pressures that favored both modularity and efficient integration. This provides further evidence for dissociable functions of these clubs. Additionally, the evolutionary generative model, compared to the random null model, produced significantly higher clubness in the

diverse club. This demonstrates that the high clubness of the diverse club is only a feature of real world networks with a non-random architecture.

The interpretation of many previous network analyses could be dramatically altered in light of our findings, as our results provide a strong motivation for the consideration of both a rich and diverse club in network function. Contrary to previous proposals, we propose that the true integrative core of networks is the diverse club, not the rich club. Thus, we hypothesize that the rich club likely plays an alternative role in network function. One possibility that has been previously suggested is that the rich club maintains the stability of the dynamics of spontaneous activity. In the macaque structural network, rich club nodes exhibit very high in-degree—many white matter connections terminate on these nodes. Thus, autonomous dynamics of the rich club are largely constrained by the summary of strong rhythmic outputs from the entire network—rich club nodes stay closer to the summated and global network oscillations than other nodes and thus promote stability in the network dynamics at slower time-scales(168). An analogy can be made in social networks where members that exhibit a high in-degree, like politicians, are “slaves to their own power”, as they are only able to act in limited, and often slow, ways that mostly reflect the entire social network(168).

The functional connectivity, anatomical location, and cognitive functions of rich club nodes in humans fit with this proposal. The default-mode network (which we found has the highest concentration of rich club nodes), is equidistant and maximally distant from primary sensory and motor networks based on both functional connectivity and anatomical geodesic distance(169). Moreover, a meta-analysis of human brain imaging data showed that the default mode network is involved in tasks unrelated to immediate stimulus input, such as daydreaming or mind-wandering(169). These empirical findings suggest that the function of the rich club may predominately be to maintain stability in the entire network via slow processing, potentially using its high degree to integrate information at slower time scales, in contrast to the diverse club, which may act on shorter time scales. This potential distinction between the rich and diverse club warrants further investigation.

Methods

c. elegans network data. Four c. elegans worms were imaged while executing behavior with calcium imaging, and each neuron’s extracted time series of activity was made publically available(23). In this analysis, each neuron was treated as a node, and the edge weights between nodes *i* and *j* represented the Pearson r correlation (no Fisher transform was applied, as the original paper analyzed raw r values and r values were not averaged across worms) between the time series of nodes *i* and *j*. The worms had 56, 77, 68, and 57 nodes. Each worms’ graph was thresholded at a particular cost, retaining 5 to 20 percent of possible edges in 1 percent intervals. The maximum spanning tree (the set of edges (i.e., path) that connects all nodes with the maximum sum of edge weights possible) for each graph was calculated, and these edges were

not removed in order to keep the graph connected. Community detection was applied at every cost separately, using the InfoMap algorithm. We also analyzed the structural network of the c. elegans(170), where we constructed a binary and undirected network of all 297 neurons and their 2359 axonal connections (i.e., no thresholding).

Human functional MRI (fMRI) data. Human fMRI data from 471 subjects (S500 release) during rest and the performance of six tasks from the Human Connectome Project(171) were used. For the task fMRI data, Analysis of Functional Neuroimages (AFNI)(172) was used to preprocess the images, matching traditional resting-state functional connectivity studies. The AFNI command *3dTproject* was used, passing the mean signal from the cerebral spinal fluid mask, the mean signal from the white matter mask, the mean whole brain signal, and the motion parameters to the "-ort" options, which remove the signals via linear regression. The options "-automask", which generates the mask automatically was used. The "-passband 0.009 0.08" option, which removes frequencies outside 0.009 and 0.08, was used. Finally, the "-blur 6", was used, which smooths the images inside the mask only) with a filter that has a width (FWHM) of 6mm after the time series filtering. We analyzed the working memory (405 timepoints), relational reasoning (232 timepoints), motor (284), social cognition (274 timepoints), mixed math and language (316), and gambling tasks (253 timepoints). Given the short length (176 timepoints, and thus low degrees of freedom during preprocessing) of the Emotion task, it was not included in our analyses. For the resting state fMRI data (1200 timepoints), we used the images that were previously preprocessed with ICA-FIX. The AFNI command *3dBandpass* was used to further preprocess these images. We used it to remove the mean whole brain signal and frequencies outside 0.009 and 0.08 (explicitly, "-ort whole_brain_signal.1D -band 0.009 0.08 -automask").

For each task (both LR and RL encoding directions were used), for each subject, the mean signal from 264 regions in the Power atlas [3] was computed. The Pearson *r* between all pairs of signals was computed to form a 264 by 264 matrix, which was then Fisher z transformed. All subjects' matrices were then averaged. No negative correlations were included in our analyses. This matrix served as the edge weights for the graph for that particular task. The same thresholding and analyses across costs that was applied to the c. elegans functional networks was executed for human networks.

Macaque structural network. The structural network of the macaque cortex is publically available(173). While the c. elegans is a micro-scale network, with individual neurons represented as nodes, the macaque network is a macro-level network, with 71 brain regions modeled as nodes and 438 white matter tracts modeled as edges. Edges were treated as undirected and binary; thus, no thresholding and analyses across costs is required.

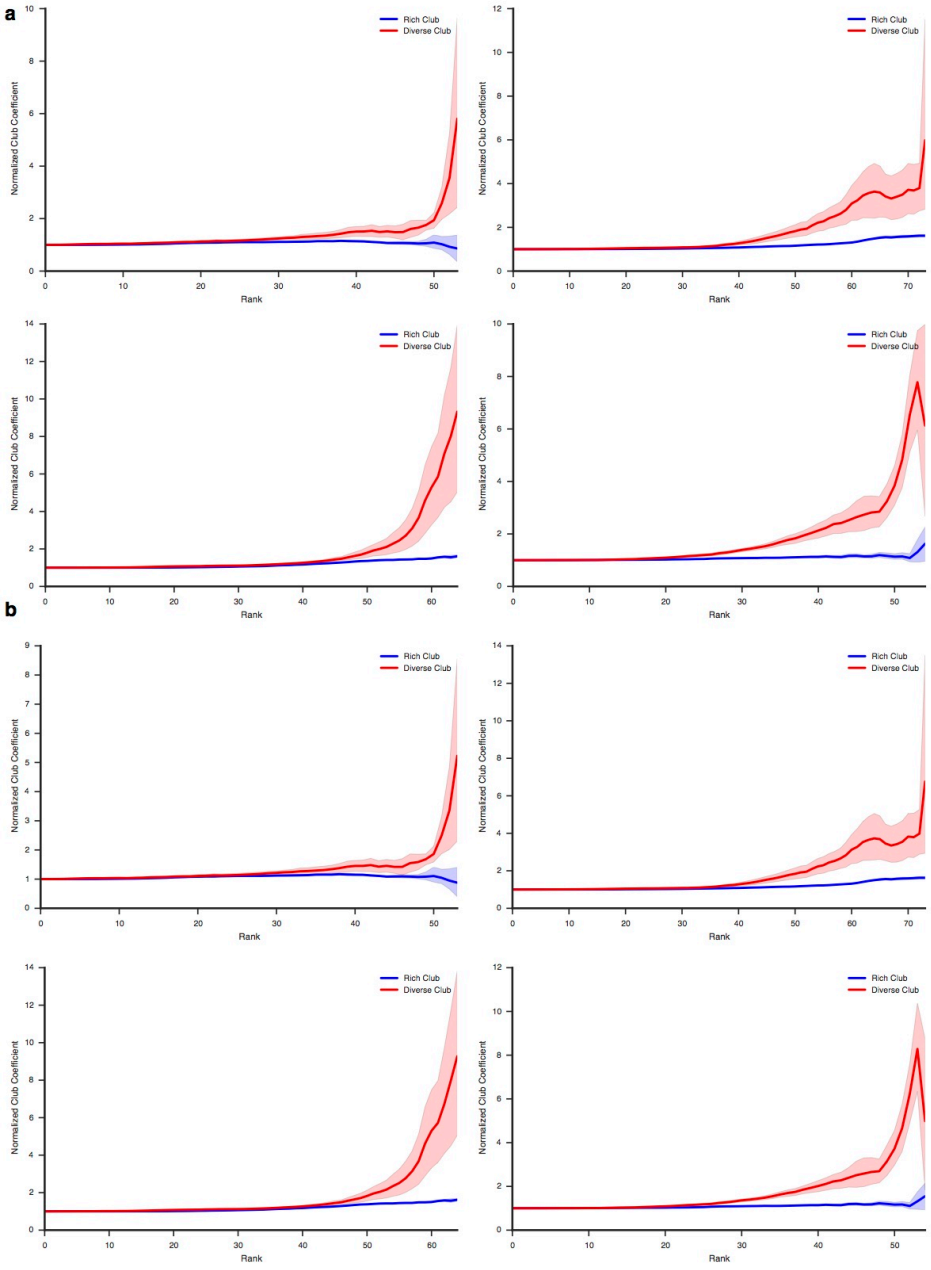
Man-made networks. We analyzed air traffic patterns between 3281 airports and 531 airlines spanning the globe, where a node is an airport, and the edge weight between nodes is the number of airlines flying between them, resulting in 10,924 edges (data

downloaded from OpenFlights.org). We also analyzed the United States power grid, where a node is either a generator, a transformer, or a substation (n=4,941), and an edge represents a power supply line (n=6,594). No thresholding was applied to either network. Data was downloaded from: http://konect.uni-koblenz.de/networks/opsahl-powergrid.

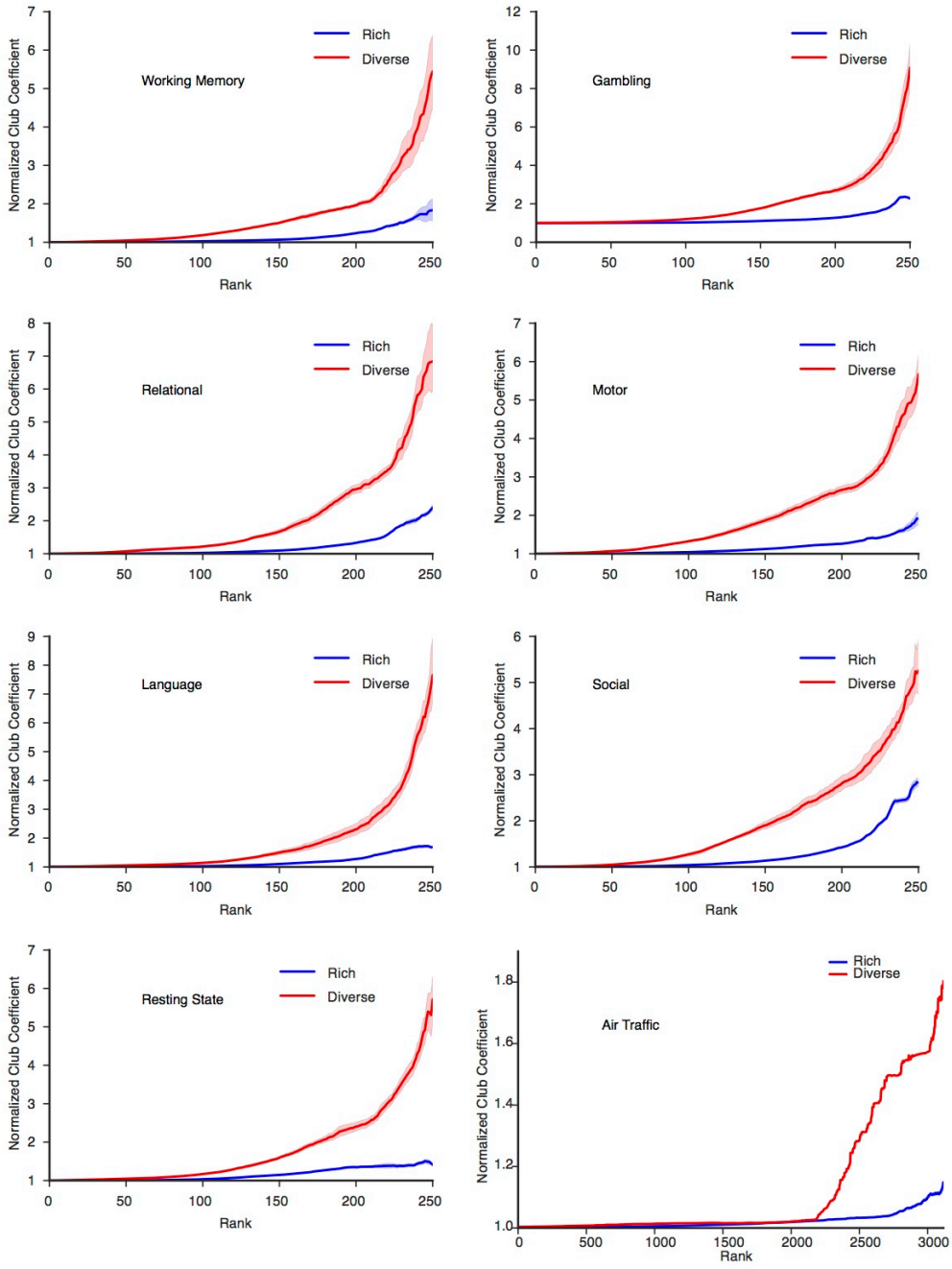
Generative evolutionary model. The model starts with a graph of 100 nodes that are randomly connected, with 5 percent of all possible edges (n=272). Binary edges were used. At each iteration, the change in Q and the change in the sum of shortest paths following the removal of each edge is calculated. Edges are chosen for removal that, when removed, lead to a maximal increase in Q and a minimal increase in the sum of shortest paths. These edges are removed and then randomly placed back in the graph, maintaining a constant density of edges (0.05). At each iteration, 0.25 percent of edges are removed from the graph and then placed randomly back into the network. This process maximizes Q and minimizes the sum of shortest paths. This procedure is repeated for 150 iterations, resulting in 1950 edges being shuffled. At this point, the generative model stops. For a null model, we also ran the generative process, except we randomly selected the edges, removed them, and then randomly placed them in the network.

The degree and participation coefficient fit was measured by the inverse of Kullback-Leibler divergence: sum(pk*log(pk / qk), axis=0), where pk is histrogram of the model network's distribution and qk is the histogram of the human brain network's distribution, both of which have been sorted into 10 bins, where each bin's value is the proportion of nodes in that bin. This was implemented in python as scipy.stats.entropy(model_histogram,human_brain_histogram).

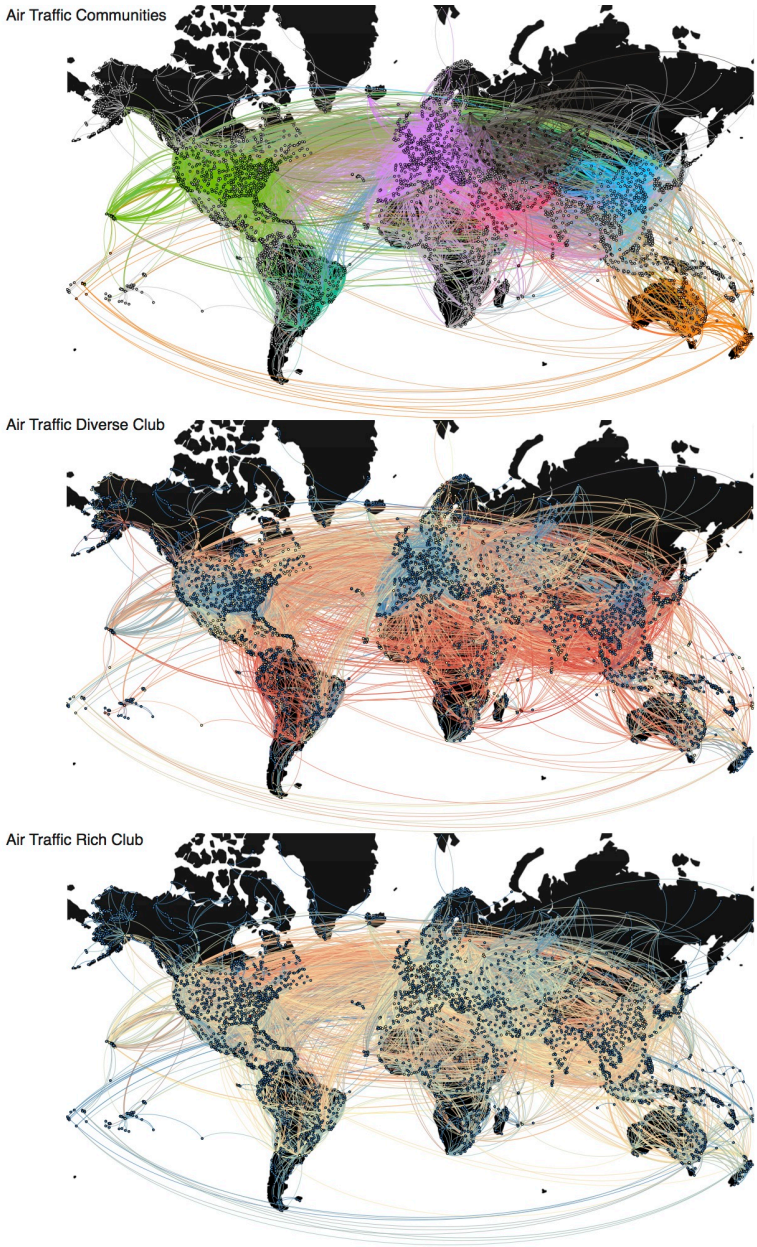
Extended Data Figures



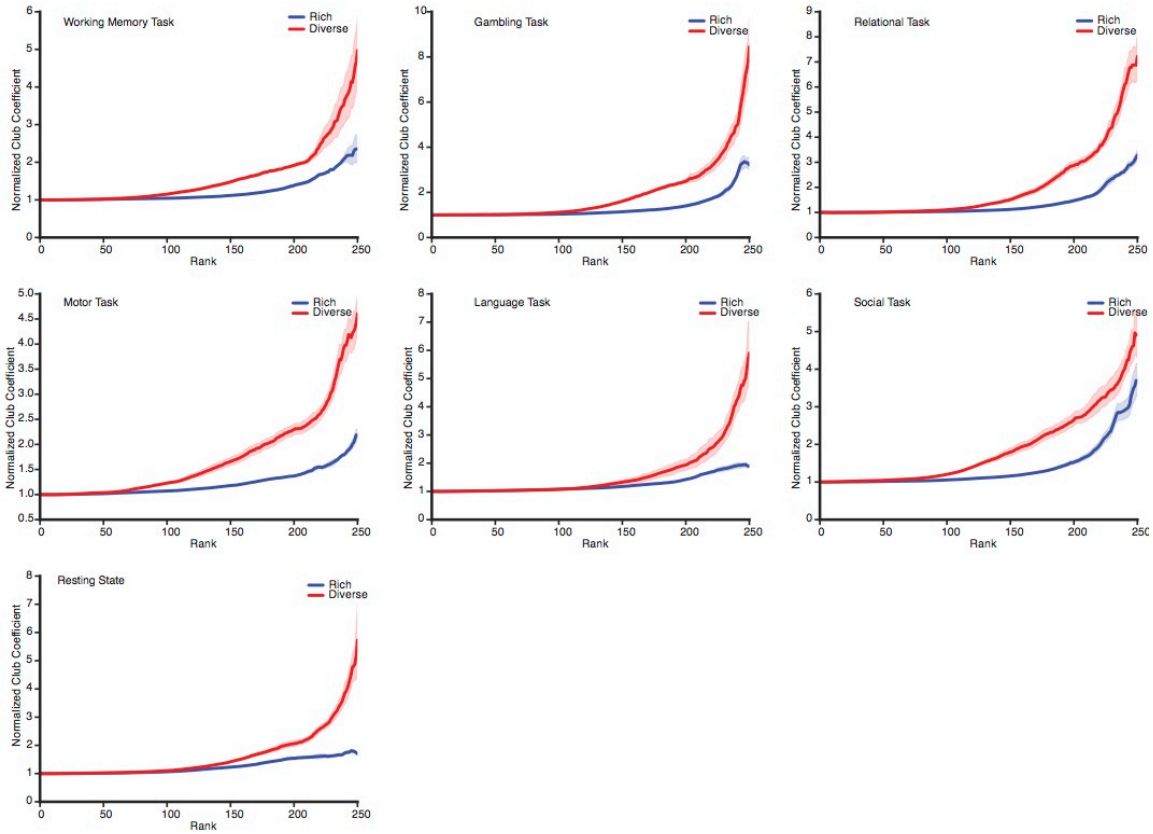
Extended Data Figure 1 | Clubness of the four c elegans functional networks. **a**, Normalized club coefficients are calculated with random graphs that place the edges randomly, but retain each node's degree and sum of weights, which accounts for the contribution of edge placement, but not edge weights, to the normalized club coefficient. **b**, Normalized club coefficients are calculated with random graphs, where all nodes maintain their degree, but the edges are randomly placed and the edge weights are shuffled between nodes with the same degree, which accounts for the contribution of both edge placement and edge weights to the normalized club coefficient. This latter method is the method reported in Figure 1.



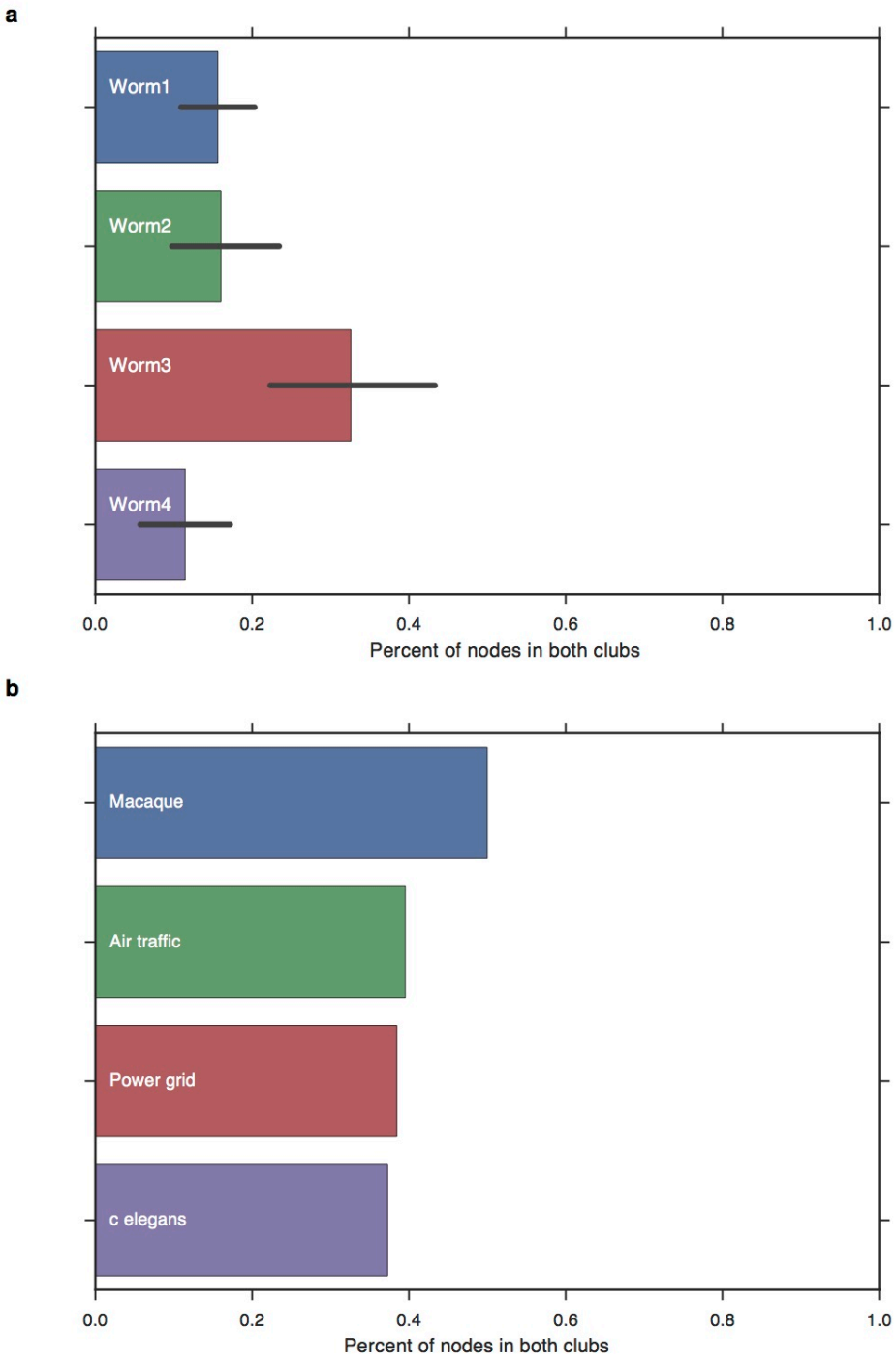
Extended Data Figure 2 | Clubness of the human functional networks and the air traffic network. Normalized club coefficients are calculated with random graphs that place the edges randomly, but retain each node's degree and sum of weights, which accounts for the contribution of edge placement, but not edge weights, to the normalized club coefficient.



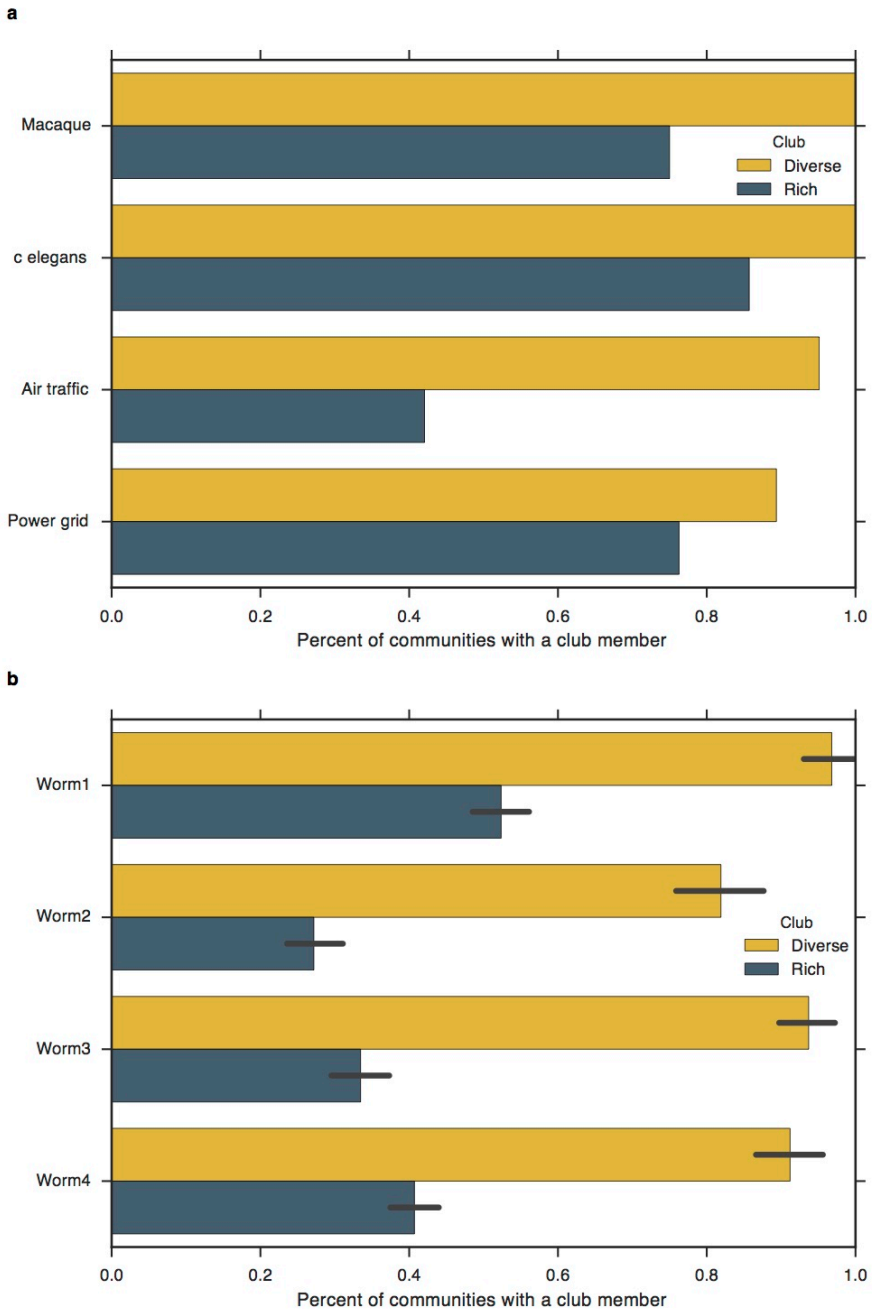
Extended Data Figure 3 l Air traffic communities and clubs. Top, community detection results from the air traffic network. Here, each node is colored according to the community it is in. Middle, diverse club; bottom, rich club. Nodes in red represent the maximum value for the given metric (participation coefficient or degree), yellow is median, and blue is the minimum. Edges are colored by the mix between the two nodes each edge connects. Edges represent a flight route, with red edges being intra club, yellow between a club node and a non-club node, and blue as between two non-club nodes. Note that, only in the diverse club, non-club flights are predominately domestic, with diverse club flights predominately being international. This is in line with our finding that the airports with the highest participation coefficients are all international airports.



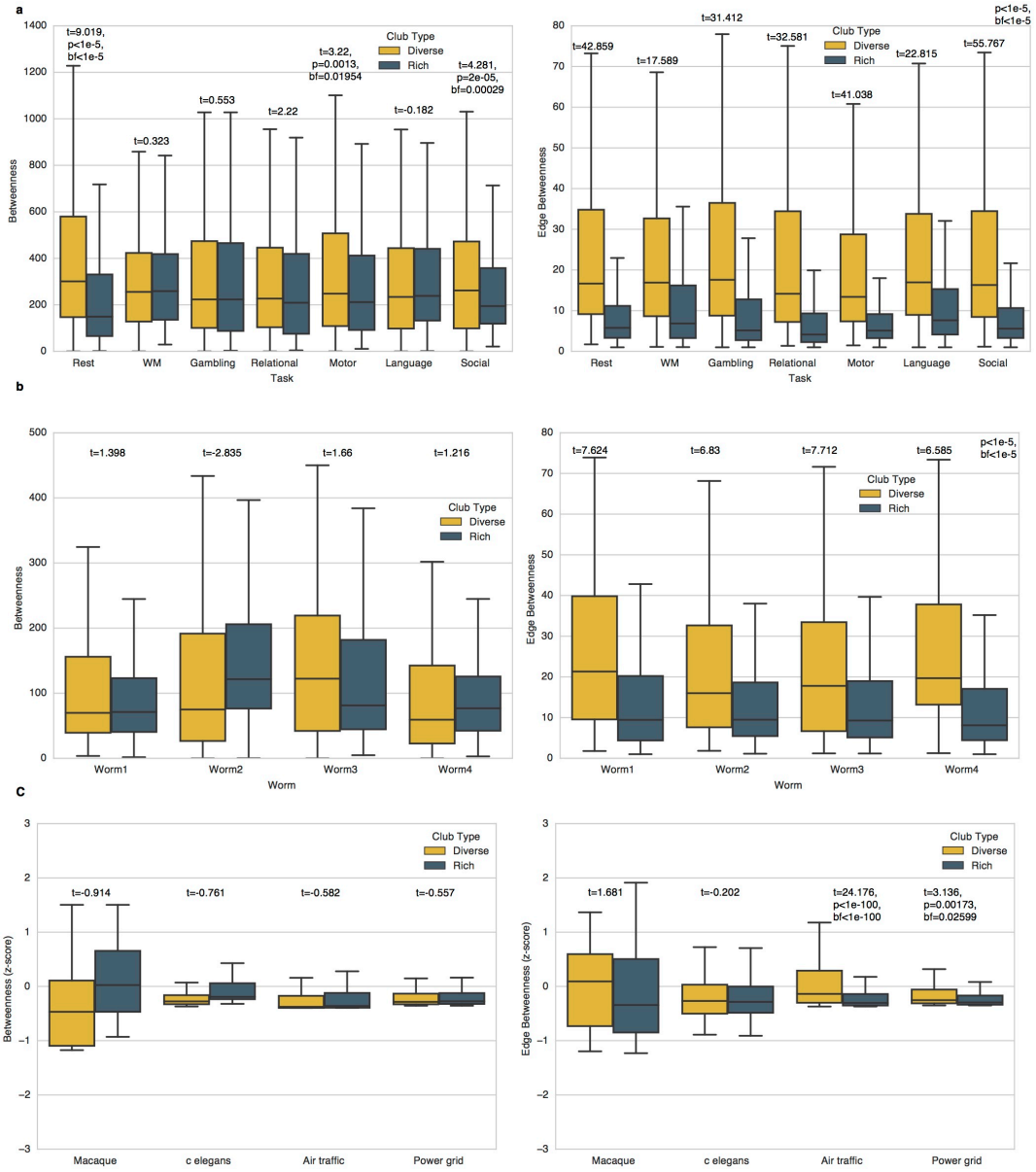
Extended Data Figure 4 | Clubness for all 7 human functional network states. Normalized club coefficients are calculated with random graphs, where all nodes maintain their degree, but the edges are randomly placed and the edge weights are shuffled between nodes with the same degree, which accounts for the contribution of both edge placement and edge weights to the normalized club coefficient. Shaded regions represent 95% confidence intervals across costs.



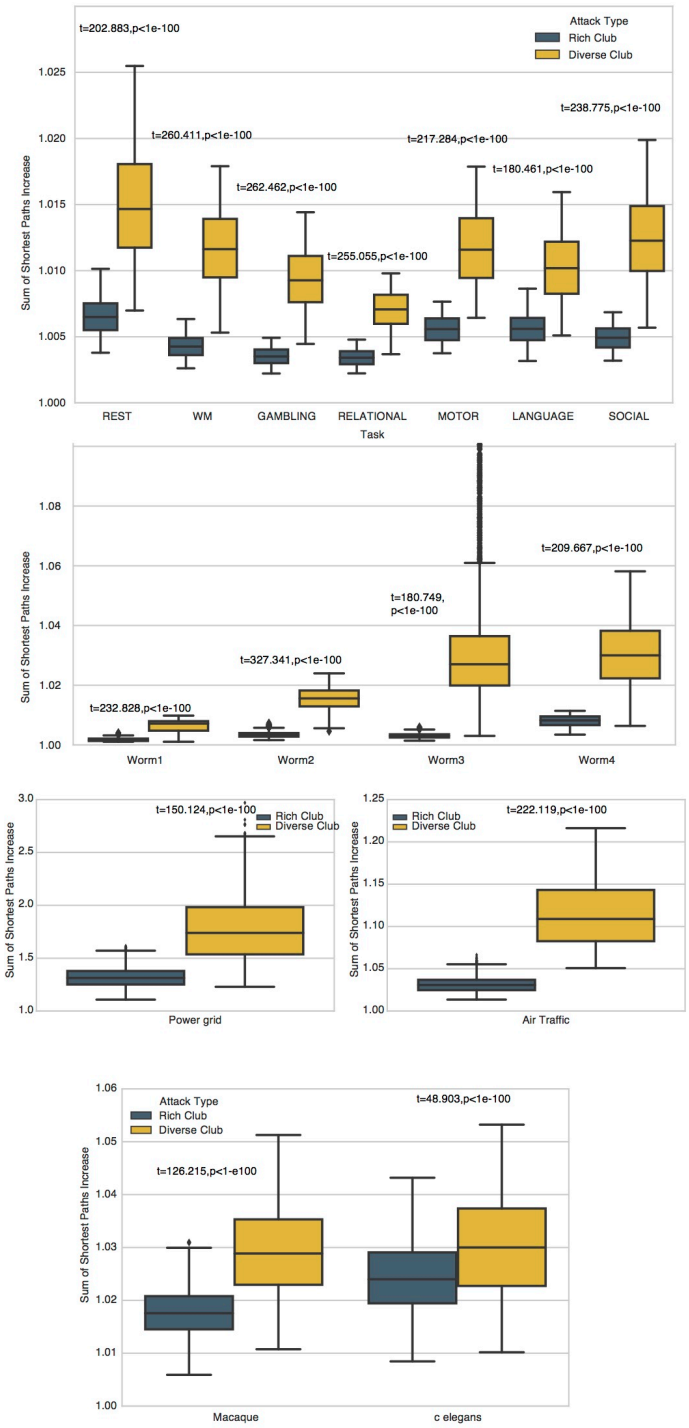
Extended Data Figure 5 | Overlapping members of the clubs. Percentage of nodes that are in both clubs for the four c. elegans functional networks (mean across costs) and structural networks. Zero percent represents that no nodes were members of both clubs, and 100 percent represents that the clubs are identical.



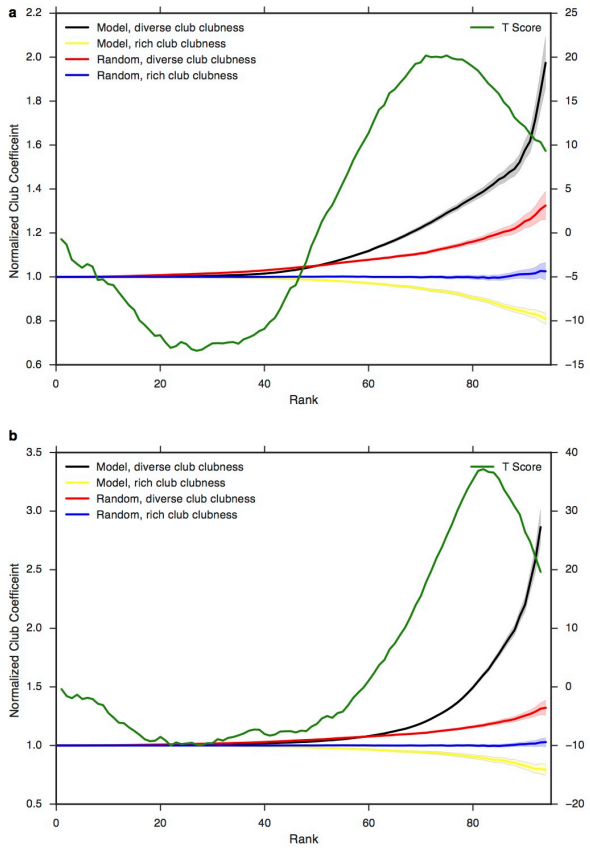
Extended Data Figure 6 | Distribution of clubs' members across communities. For each network, the percentage of communities with a node from the rich club and the diverse club is plotted. 100 percent represents that every community contains at least one node from that club. **a**, structural networks. **b**, functional networks from four c. elegans worms.



Extended Data Figure 7 | Distribution of betweenness centrality and edge centrality scores. Betweenness centrality measures how many shortest paths cross through a particular node, while edge betweenness measures how many shortest paths cross a particular edge. Only edges between club members were included in the calculation for edge betweenness. A significant Bonferroni (*bf*, number of tests=15) corrected p value is shown if the result was significant for that particular network. Pooling betweenness across costs and tasks for the human networks, the diverse club was significantly higher in betweenness than the rich club (t=7.2, Bonferroni, p=2.9e-14). Pooling betweenness across costs and worms for the c. elegans networks, there was no significant difference (t=0.35). Pooling betweenness across structural networks, there was also no significant difference (t=-0.6). In all networks besides the structural macaque and c. elegans the edge betweenness of the diverse club was significantly higher than that of the rich club.



Extended Data Figure 8 | Targeted Attacks. Sum of shortest paths after attacks on the rich club or the diverse club for every network. For each network, over 10,000 iterations, we removed anywhere (randomly) between 50 and 90 percent of edges (skipping edges that disconnected the graph into two sub-graphs) from the rich club or the diverse club. We then calculated the increase in the sum of shortest paths. An increase in the sum of shortest paths indicates decreased global efficiency.



Extended Data Figure 9 | Generative Models. Alternative ratios of Q to efficiency for the generative models (0.70(**a**), and 0.80(**b**))

Acknowledgments. M.A.B and M.D. are supported by NIH Grant NS79698 and the National Science Foundation Graduate Research Fellowship Program under Grant no. DGE1106400. B.T.T.Y. is supported by Singapore MOE Tier 2 (MOE2014-T2-2-016), NUS Strategic Research (DPRT/944/09/14), NUS SOM Aspiration Fund (R185000271720), Singapore NMRC (CBRG14nov007, NMRC/CG/013/2013) and NUS YIA.

Author contributions. M.A.B. devised the concept and study; M.A.B., B.T.T.Y., and M.D. jointly designed the analyses; M.A.B. contributed new reagents/analytic tools; M.A.B. ran the experiments and analyzed the data; M.A.B., B.T.T.Y., and M.D. wrote the paper.

Author Affiliations. M.A.B and M.D., Helen Willis Neuroscience Institute & Department of Psychology, University of California, Berkeley, California, USA. B.T.T.Y. Electrical and Computer Engineering, Clinical Imaging Research Centre, Singapore Institute for Neurotechnology & Memory Networks Programme, National University of Singapore, Singapore.

A Task-General Optimal Network Structure for Cognitive Processing

Maxwell A. Bertolero[1], B.T.T. Yeo[2], Danielle S. Bassett[3,4], Mark D'Esposito[1]

[1]Helen Wills Neuroscience Institute and the Department of Psychology, University of California, Berkeley, CA 94720-3190, USA
[2]Electrical and Computer Engineering, ASTAR-NUS Clinical Imaging Research Centre, Singapore Institute for Neurotechnology & Memory Networks Programme, National University of Singapore, Singapore 119077, Singapore
[3]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA
[4]Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA

Summary

Brain regions are connected in a complex pattern than can be fruitfully studied using the tools of network science. These tools reveal that the brain's network architecture is modular, comprising groups of tightly interconnected nodes referred to as communities. The tools also reveal the presence of local hubs that have many connections within their own communities and connector hubs that have connections diversely distributed across communities. Connector hubs are thought to integrate information and coordinate connectivity across communities, allowing for cohesive processing in an otherwise nearly decomposable system. Yet, direct evidence for the functional role of connector hubs has remained sparse, particularly in the context of individual differences in brain and behavior. Here, we address this gap by determining whether individual differences in a hub's role within the network are predictive of individual differences in brain network modularity and cognitive function. Using Human Connectome Project data, we show that individual differences in the connectivity of connector and local hubs predict individual differences in brain network modularity. Modularity was found to be higher in networks with strong connector and local hubs, and was lower in networks with weak hubs of either type. This modular network structure with strong hubs is associated with higher performance on a battery of five distinct cognitive tasks. Collectively, these results suggest that, when local hubs are strongly connected to their own community and connector hubs are strongly connected to many communities, the brain is able to optimally integrate information while remaining modular, thereby increasing the efficacy of cognitive function.

The human brain is a complex network that can be parsimoniously summarized by a set of nodes representing brain regions and a set of edges representing the connections between brain regions. Edges can represent functional connectivity, defined by statistical similarities in regional time series, or structural connectivity, defined by white matter streamlines linking brain areas. Such network representations can be used to study global and local brain connectivity patterns.

Both structural and functional brain networks display community structure, where groups of nodes are more strongly connected to other members of their group than to members of other groups. These groups are referred to as modules or *communities* (36, 140). Each node within one of these communities can play a distinct role in the overall topology. For example, local hubs display many connections to other nodes within their own communities, as measured by the within-community-degree. Conversely, connector hubs have connections diversely distributed across other communities, as measured by the participation coefficient. Thus connector hubs are ideally wired for integrative processing(54, 78, 80, 81, 105, 140). Local and connector hubs play a role in the modular architecture of networks as diverse as neuronal connections in c elegans(145), inter-areal connections in the mouse brain(82), tract networks in the cat and macaque brain(40), the Internet, air transportation networks, Arabidopsis thaliana, the protein interactome of c elegans, gene co-expression networks(174, 175), and metabolic networks(81).

In addition to their topological roles, local and connector hubs play important functional roles within a network. In human brain networks, connector hubs have a particular cytoarchitecture(83), are implicated in a diverse range of cognitive tasks(20, 176), and are physically located in anatomical areas at the boundaries between many communities(78). Moreover, damage to connector hubs causes widespread cognitive deficits(125) and a decrease in the modular structure of the network(152). Connector hubs also appear to coordinate connectivity changes between other pairs of nodes(67, 71), in that activity in these nodes predicts changes in the connectivity of other nodes, particularly the connectivity between nodes in different communities, during cognitive tasks(150). Finally, connector hubs, but not local hubs, exhibit increased activity if more communities are engaged in a task, which suggests that connector hubs are involved in processes that are more demanding as more communities are engaged(140). A parsimonious explanation of these findings is that connector hubs integrate information and coordinate connectivity between communities. In contrast, local hubs might support narrower processes localized within a single community. Interestingly, because the connections of local hubs are predominantly concentrated within their own community, damage to these hubs tends to cause relatively specific cognitive deficits(125, 152) and does not significantly alter the modular organization of the network(152). Supporting their more targeted role in information processing, it is also important to note that their activity levels do not increase as more communities are involved in a task(140).

We hypothesized that connector hubs integrate information across communities and coordinate connectivity between communities, two functions that serve to maintain the modular structure of the brain's functional network. When connector hubs are connected to many communities (as indicated by high participation coefficients), they can integrate information and coordinate connectivity optimally, allowing other regions to perform more modular local processing. Thus, across subjects, we predicted that increased participation coefficients of connector hubs would be correlated with increases in the modularity of the global network (Figure 1). Moreover, we predicted that increases in modularity would be accompanied by a simultaneous enhancement of segregation between local communities, particularly those that include cortical areas that sub serve primary functions (sensory, motor) more so than those that include cortical areas that sub serve hetero-modal functions (default mode, fronto-parietal). Finally, we hypothesized that the enhanced modular architecture afforded by strong local and connector hubs would be accompanied by enhanced cognitive performance.

To determine if individual differences in hubs' strengths within the network are predictive of individual differences in network modularity and cognitive performance, we leveraged the size and richness of the Human Connectome Project fMRI data (S500 Release). Specifically, we measured regional roles as local hubs using the within-community-degree, and we measured regional roles as connector hubs using the participation coefficient. We examined how well brain network modularity was explained by the participation coefficients of connector hubs or the within-community-degrees of local hubs (Figure 1). We also assessed how well these regional roles explained the modularity of their neighbors' connectivity patterns (Figure 1), as well as individual differences in cognitive performance during five distinct cognitive tasks. Our results offer explanatory mechanisms of individual differences in cognitive function that depend on a network's modular architecture as quantified by the differential roles of local and connector hubs.

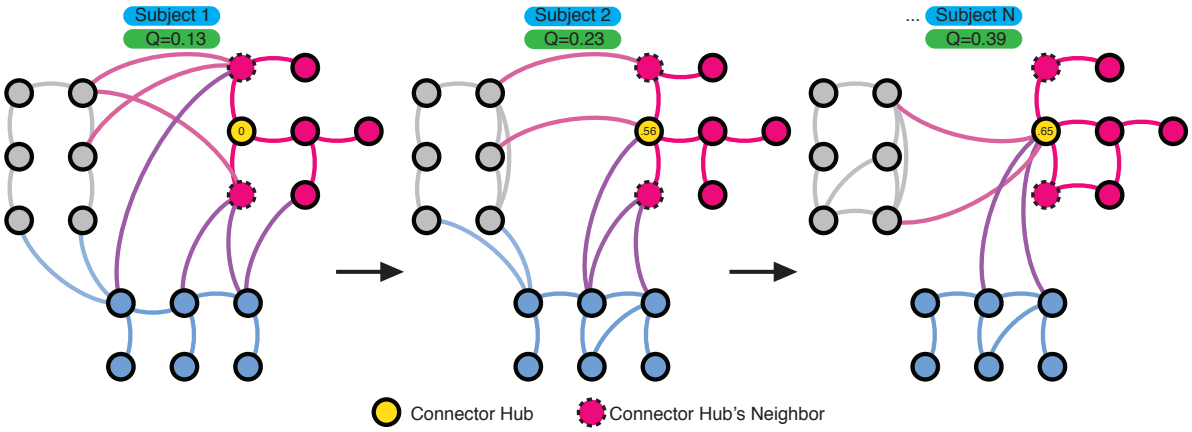


Figure 1. The Foundational Hypothesis and Associated Predictions. We hypothesized that connector hubs integrate information across communities and coordinate connectivity between communities, two functions that serve to maintain the

modular structure of the brain's functional network. When connector hubs are connected to many communities (as indicated by high participation coefficients), they can integrate information and coordinate connectivity optimally, allowing other regions to perform more modular local processing. Thus, across subjects, we predicted that increased participation coefficients of connector hubs would be correlated with increases in the modularity of the global network. This prediction is illustrated above. Each graph represents the network structure of an individual subject. A single connector hub, shown in yellow, is identified based on its consistently high participation coefficient across individuals; subject-level values are given inside the yellow nodes. Local neighborhoods or *communities* are shown in grey, blue, and pink, and the modularity of the network is estimated by the modularity quality function Q; subject-level values of Q are given above each graph. Across subjects, we predict that higher participation coefficients of connector hubs will occur in more modular networks (Figure 2A) in which the neighbors of connector hubs (pink, dashed-outline) display more local connectivity (Figure 2B, 3C). Finally, we predict that connectivity changes related to connector hubs' participation coefficients should be limited to weaker connectivity between communities composed of primary sensory and motor cortex, but not between communities composed of higher-order areas implicated in the brain's default mode and fronto-parietal circuitry (Figure 3A-C).

Results

Diversity and Locality Facilitated Modularity Coefficients

We first sought to understand how individual differences in a hub's role within the network are predictive of individual differences in brain network modularity. Thus, we tested if, across subjects, a node's participation coefficients are positively correlated with the magnitude of the modularity quality index Q. For every node, we calculated the Pearson *r* correlation coefficient between the participation coefficients and Q values across subjects. Intuitively, higher Pearson *r* correlation coefficients indicate that the node's connector hub strength (i.e., the participation coefficient) explains greater variance in the network's modularity. This is an important feature that can be used to distinguish the roles of different brain regions. For ease of presentation, we refer to each node's *r* value as the diversity-facilitated modularity coefficient, as it measures how the diversity of the node's connections facilitates the modularity of the network.

For every node, we also calculated the Pearson *r* correlation coefficient between the within-community-degrees and modularity values across subjects. Intuitively, higher Pearson *r* correlation coefficients indicate that the node's local hub strength (i.e., the within-community-degree) explains greater variance in the network's modularity. We refer to each node's *r* value (between within-community-degrees and Q values across subjects) as the locality-facilitated modularity coefficient, as it measures how the locality of the node's connections facilitates the modularity of the network.

To investigate if the strength of hubs is related to network modularity, we first calculated the Pearson *r* correlation coefficient between each node's mean participation coefficient across subjects (which defines a connector hub) and the node's diversity-facilitated modularity coefficient (Figure 2A). Next, we calculated the Pearson *r* correlation coefficient between each node's mean within-community-degree across subjects (which defines a local hub) and the node's locality-facilitated modularity coefficient (Figure S1). We performed these computations separately for seven distinct cognitive states: the resting state, a working memory task, a gambling task, a relational reasoning task, a motor task, a language and math task, and a social cognition task.

In all cognitive states, we observed a positive Pearson *r* correlation coefficient between a node's mean participation coefficient (which defines a connector hub) across subjects and the node's diversity-facilitated modularity coefficient (Figure 2). Moreover, in all cognitive states, there was a positive Pearson *r* correlation coefficient between a node's mean within-community-degree (which defines a local hub) across subjects and the node's locality-facilitated modularity coefficient (Figure S1). Figure S2 displays brain images, for each state, showing the surface based anatomical locations of the nodes and their corresponding mean participation coefficient (S2A), the diversity-facilitated modularity coefficient (S2B), mean within-community-degree (S2C), and the locality-facilitated modularity coefficient (S2D). These data demonstrate that connector hubs' strong diverse connectivity to many communities and local hubs' strong local connectivity predicts higher brain network modularity, regardless of the state that the subject is in.

To confirm the reliability and reproducibility of our findings, we performed an iterative split-half analysis. Specifically, we estimated the mean within-community-degree or participation coefficient of each node in one half of subjects, and each node's locality and diversity facilitated coefficient in the other half, testing 10,000 random split of subjects. All relationships were reliably observed in every cognitive state (Figure S3). Next, we sought to determine if this relationship was a necessary feature of the underlying mathematics, or whether it was a phenomenon specific to the neurophysiology of brain networks. To address this question, we tested four null model networks and observed that none of them exhibited a significant relationship between mean participation coefficient and diversity-facilitated modularity coefficient (Figure S4). As a third check, we assessed whether the number of communities identified in the network was inadvertently biasing our results. We observed that the number of communities in each network was negatively correlated with the modularity value (Figure S5). After regressing out the number of communities in each network from the modularity value, we observed that our findings remained unchanged (Table S1). Finally, we tested whether the relationships between variables of interest were linear (and therefore appropriate to examine with Pearson *r* correlation coefficients), or nonlinear. To address this question, we analyzed individual 1st, 2nd, and 3rd order curve fits of the relationship between participation coefficients and modularity values. We observed that many relationships were well-captured by a first order fit, with the

connector hub's maximal participation coefficients corresponding to maximal Q indices, with only a few showing a more nonlinear relationship (Figure S6).

Discrete Connectivity Changes Underlying Locality and Diversity Facilitated Modularity Changes

The modularity quality value Q is a single metric summarizing the structure of the entire network. Thus, an interesting question is whether a connector hub's relationship to modularity was driven by a global change in connectivity, a more local change in connectivity, or both. Based on the conceptual model of optimal communication efficiency in modular networks, we predicted that each connector hub directly coordinates the connectivity of its neighbors—other brain regions to which they are strongly connected, either functionally or structurally(64). To investigate this prediction, we wished to determine whether the relationship between modularity and a connector hub's participation coefficient could be explained by increases in the modular connectivity of the hub's neighbors (Figure 1). This relationship should only hold for connector hubs, not local hubs, as connector hubs are the hubs that previous studies suggest coordinate connectivity between communities and maintain a modular structure(20, 67, 71, 140, 150).

To test this prediction, we examined the relationship between a node's participation coefficient and the weight of each edge in the network (see Methods), by computing the Pearson *r* correlation coefficient across subjects. We considered two sets of *i* nodes, based on their positive or negative diversity-facilitated modularity index. Because all nodes with a positive diversity-facilitated modularity index had a high mean participation coefficient, we refer to these as "connector hubs (Q+)", representing the fact that their participation coefficients were positively correlated with modularity values. If connector hubs' (Q+) increased participation coefficients are only modulating their neighbors' connectivity patterns, the metric described above between node's *i* and *j* should be correlated with how strongly functionally connected nodes *i* and *j* are. Thus, for each pair of nodes *i* and *j*, (where *i* is either a connector hub (Q+) or nodes with a negative diversity-facilitated modularity coefficient), we calculated the correlation coefficient between two variables (x) and (y) where (*x*) is the how the strength of hub *i* relates to the modularity of j's connectivity (*y*) is the mean edge weight across subjects between nodes *i* and *j*. In other words, we measured the relationship between how strongly nodes *i* and *j* are connected and the degree to which changes in the participation coefficients of node *i* changes the modular connectivity of node *j*. Across all states, only connector hubs (Q+) exhibited a positive correlation between how strongly nodes *i* and *j* are connected and how much the participation coefficients of node *i* changes the modular connectivity of node *j* (Figure 2B; Figure S7). Nodes with negative diversity facilitated coefficients did not exhibit this relationship (*r* between *-0.32* and *0.15* across cognitive states).

To test if this relationship was unique to connector hubs' (Q+) and participation coefficients, we computed the same metric for within-community-degree. Thus, we divided nodes into two sets based on their positive or negative locality-facilitated modularity index. Because all nodes with a positive locality-facilitated modularity index had a high mean within-community-degree, we refer to these as "local hubs (Q+)", representing the fact that their within-community-degrees were positively correlated with modularity values. The same metrics calculated above was calculated again. However, the metrics now measures how well the within-community-degrees of node *i* correlate with the increased modular connectivity of node *j*. We predicted that, because local hubs' (Q+) increased within-community-degree are not modulating their neighbors' connectivity patterns, this metric should not be correlated with how well functionally connected nodes *i* and *j* are. In line with this prediction, local hubs' (Q+) within-community-degrees did not exhibit this relationship (*r* between *-0.03* and *0.085* across cognitive states), nor did nodes with negative locality facilitated modularity coefficients (*r* between *-0.11* and *-0.07* across cognitive states; Figure S5).

These results demonstrate that if and only if a connector hub (Q+) is strongly functionally connected to a node, that connector hub's increased participation coefficients correlated with the increased modular connectivity of the node's connectivity. This was not true for the participation coefficients of nodes with negative diversity-facilitated modularity coefficients, nor the within-community-degrees of local hubs (Q+) or nodes with negative locality-facilitated modularity coefficients. These results suggest that a connector hub's high diversity-facilitated modularity coefficient does not largely reflect diffuse global connectivity changes. Instead, connector hubs are connected in a way that allows them to directly tune or coordinate the connectivity of their neighbors to be more modular, thereby increasing global network modularity. Supporting this interpretation, we found that, when connector hubs (Q+) have high participation coefficient, local hubs (Q+) have high within-community-degree (Figure S8).

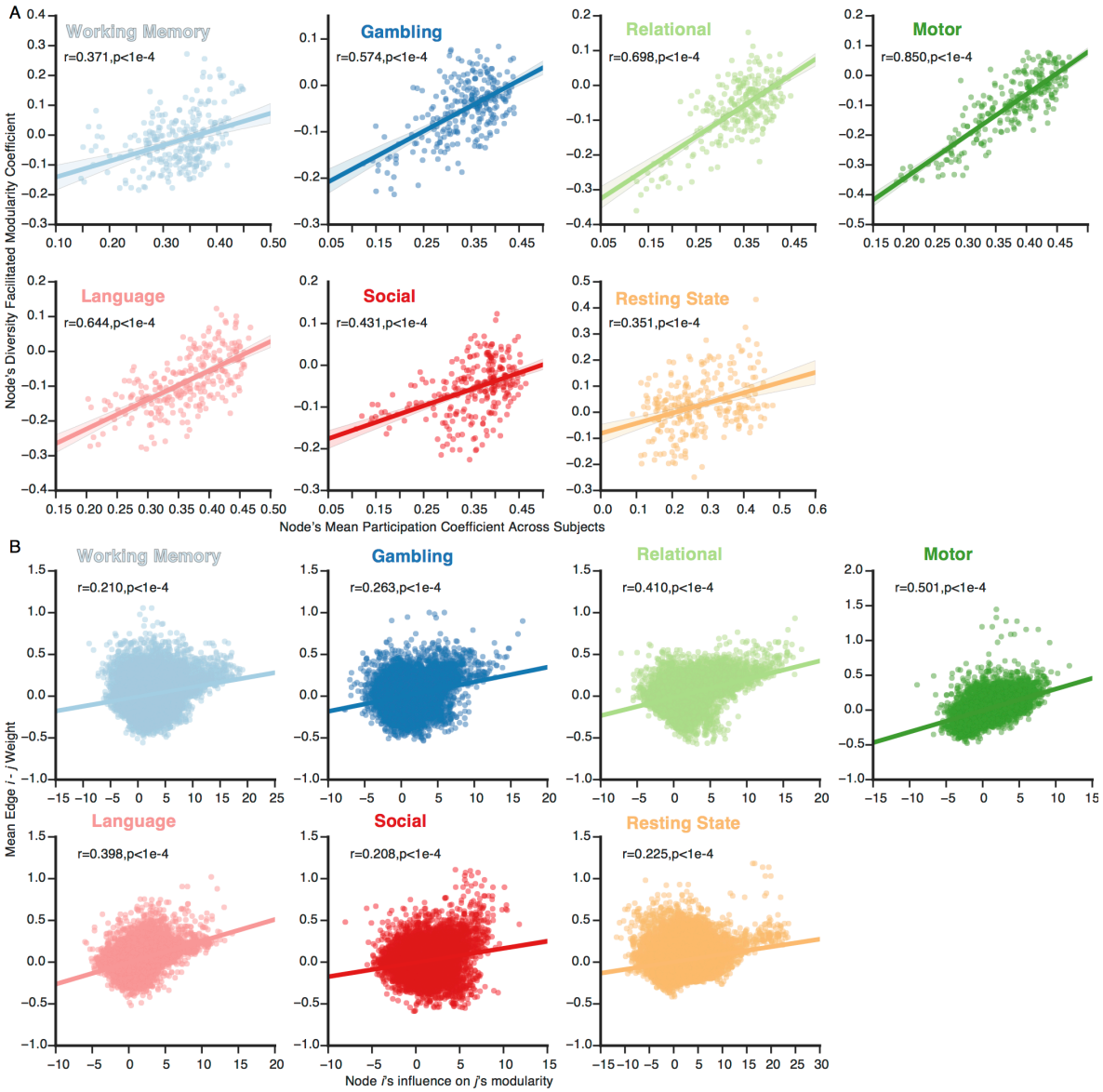


Figure 2. **Connector Hubs' Diversity-Facilitated Modularity Coefficients and their Relationships with Individual Edges' Weights**. (A) Correlations between a node's mean participation coefficient and the node's diversity facilitated modularity coefficient, which we define as the Pearson *r* correlation coefficient between that node's participation coefficients and modularity values (Q) across subjects. Regression plots for data acquired during the performance of each task show the correlation of each node's diversity-facilitated modularity coefficient with that node's mean participation coefficient across subjects. Each dot represents a node in the brain's functional network. Shaded areas represent 95 percent confidence intervals. In all cognitive states, there was a positive correlation between a node's mean participation coefficient and the node's diversity-facilitated modularity coefficient. (B) For each task, for each edge between a connector hub (Q+) *i* and node *j*, we calculated the Pearson correlation coefficient between (i) the mean edge weight between node *i* and node *j* shown along the y-axis, and (ii) how strongly the participation coefficients of node *i* were correlated

with the modular connectivity of node *j*. We observed that if and only if a connector hub (Q+) was strongly functionally connected to a node, that connector hub's increased participation coefficients were correlated with the increased modular nature of the node's connectivity.

Community Level Connectivity Changes of Locality- and Diversity-Facilitated Modularity Changes

We sought to further investigate changes in edge weights in the network—do the edge changes related to increases in modularity (Q) fall within or between particular communities? First, we assessed, for each edge in the network, how its weights related to modularity values (Q) across subjects (Figure 3A). We observed that higher modularity values were associated with decreased connectivity between the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular communities, and increased connectivity within those same communities (Figure 3B). Next, we visualized how well each connector hub's (Q+) participation coefficients correlate with each edge's weights across subjects (Figure 3B, see methods). We observed that increases in connector hubs' (Q+) participation coefficients were also accompanied by decreases in the connectivity between the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular communities. Within each of those communities, some edges were increased, while others were decreased. However, in general, the between community edge weight decreases associated with higher participation coefficients of connector hubs (Q+) were also associated with increased Q. Thus, connector hubs' increased connectivity to many communities correlates with increased segregation between sensory, motor, and attentional communities; the segregation of these communities from each other is explained in part by the variance in modularity across subjects.

We next sought to test if the increased segregation between sensory, motor, and attentional communities mediate the relationship between connector hubs' increased participation coefficients and modularity (Q). Specifically, we performed a mediation analysis for each connector hub (Q+), with an edge weight mediating the relationship between the connector hub's (Q+) participation coefficients and the Q indices of the network across subjects. An edge's mediation value of a connector hub's (Q+) participation coefficients and Q indices is the (regression coefficient of the edge's weights by the hub's participation coefficients across subjects) multiplied by (regression coefficient of Q indices by the edge's weights, controlling for the hub's participation coefficients, across subjects). Each edge's mean mediation value across connector hub's (Q+) is shown in Figure 3C. We observe that edges between the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular mediate the relationship between connector hubs' participation coefficients and Q indices. Thus, connector hubs (Q+) increased connectivity to many communities increases segregation between sensory, motor, and attentional

communities, which increases the global modularity value of the network. This suggests that strong connector hubs lead to very specific connectivity changes, which, in turn, increases the modularity of the network.

Finally, we further tested if the relationship between a connector hub's (Q+) participation coefficients and Q indices is mediated primarily by that hub's neighbors' edge pattern increasing Q (per our prediction and evidenced by the analysis in Figure 2B). For each connector hub (Q+) node *i*, we have an array of mediation values of node *i*'s neighbors' edges and an array of the absolute mediation values of node *i*'s non-neighbors' edges. Mediation values were based on the edge mediating between node *i*'s participation coefficients and Q values. Thus, if a connector hub is primarily modulating Q via the mediation of its neighbors' edges, the mediation values in the neighbors' array should be greater than the mediation values in the non-neighbors array. Neighbors were defined based on edges present between the two nodes in a graph at a density of 0.15 (as it was our densest cost explored); edges of node *i* were ignored, as we were only interested in how the participation coefficients of node *i* modulate Q via the mediation of *j*'s connectivity to the rest of the network, not *j*'s connectivity to node *i*. The distribution of *t*-values is shown in Supplemental Figure 9. The mediation values were consistently and significantly higher for connector hubs' (Q+) neighbors' edges than non-neighbors. This result further confirms that each connector hub's increased participation coefficients increase Q indices via the connector hub's neighbors' edge patterns becoming more modular and mediating an increase in the global Q value.

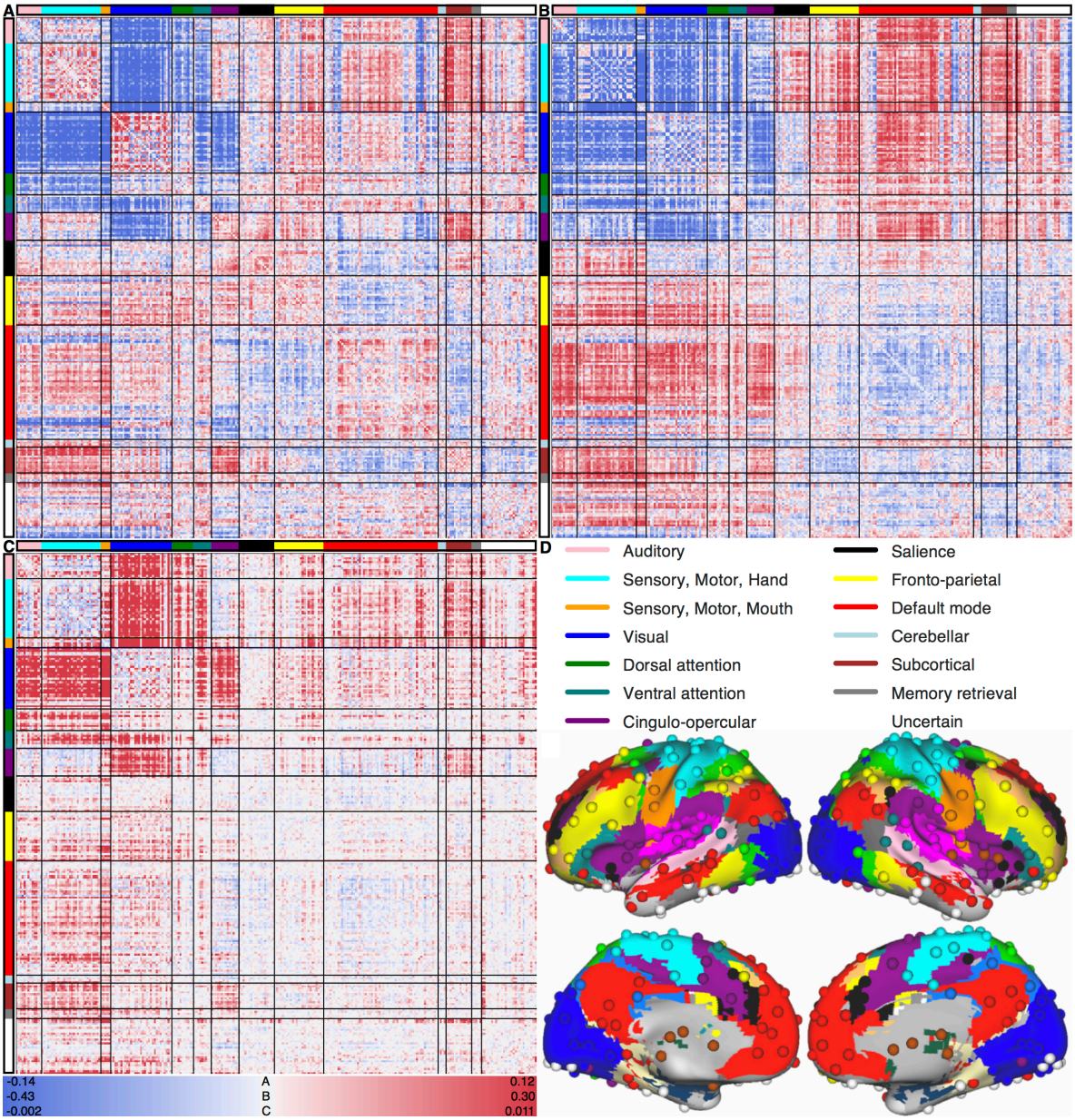


Figure 3. **Participation Coefficients, Edge Weights, and the Modularity Quality Indices**. *(A)* Each entry is the Pearson *r* correlation coefficient, across subjects, between the modularity quality indices Q and that edge's weights. Decreased connectivity between and increased connectivity within the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular communities are correlated with increased Q indices. *(B)* For each connector hub (Q+), we calculated the Pearson correlation coefficient *r* between the hub's participation coefficients and each edge's weights across subjects. This procedure resulted in a 264x264 matrix for each connector hub, to which we applied a Fisher's *r-z* transform. The matrix in *(B)* is the mean of those matrices across connector hubs (Q+); connector hubs' increased participation coefficients correlate with decreased connectivity *between* the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular communities. To investigate the relationship between

connector hubs' participation coefficients, edge weights, and Q, we performed a mediation analysis for each node, with an edge's weights mediating the relationship between the node's participation coefficients and Q indices. Each edge's mean mediation value between connector hub's participation coefficients and Q indices is shown in panel *(C)*. Edges between the visual, sensory/motor hand, sensory/motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular mediate the relationship between connector hubs' participation coefficients and Q indices. (D) The anatomical locations of each node and community on the cortical surface(20, 56).

A Task-General Optimal Network Structure for Cognitive Processing

We next asked whether individual differences in the connectivity of connector and local hubs predict task performance. To address this question, we calculated the correlation between a node's participation coefficients and task performance in the tasks for which performance measures were available: working memory, language/math, relational reasoning, and social cognition (see Methods for details on task performance quantification). We performed this analysis for every node, generating a Pearson *r* correlation coefficient between a node's participation coefficients and task performance values. A positive *r* at a node indicates that a subject with a higher participation coefficient at that node tends to perform better on the task. We call this *r* value the node's diversity-facilitated performance coefficient. The same analysis was executed using each node's within-community-degrees, and we call this *r* value the node's locality-facilitated performance coefficient.

To test whether the participation coefficients of connector hubs was correlated with performance, we calculated the Pearson *r* correlation coefficient between each node's diversity-facilitated performance coefficient and each node's mean participation coefficient. All correlations were positive and significant (Figure 4A). Thus, for connector hubs, a higher participation coefficient is associated with higher task performance. To test if the within-community-degrees of local hubs correlate with performance, we calculated the Pearson correlation coefficient between each node's locality facilitated performance coefficient and each node's mean within-community-degree. All correlations were positive and significant (Figure S10). Thus, for local hubs, a higher within-community-degree is associated with higher task performance.

We next tested whether there was a significant correlation between a node's diversity facilitated *modularity* coefficient and a node's diversity facilitated *performance* coefficient. All correlations were positive and significant (Figure 4B). We next tested whether there was a significant correlation between a node's locality facilitated *modularity* coefficient and a node's locality facilitated *performance* coefficient. All correlations were positive and significant (Figure S10). Thus, across subjects, a node's participation coefficients and within-community-degrees are similarly associated with modularity and task performance.

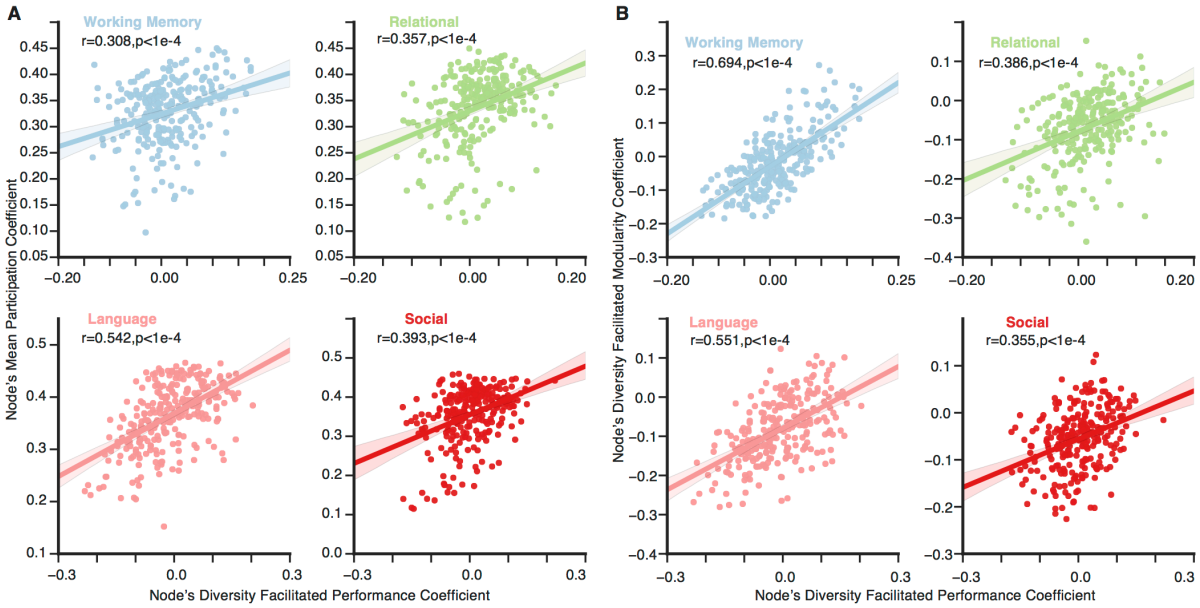


Figure 4. **Participation coefficients, Q indices, and Performance**. (A) The relationship between (x-axis) Pearson *r* value between a node's participation coefficients and task performance—the node's diversity facilitated performance coefficient—with (y-axis) the node's mean participation coefficient. (B) The relationship between (x-axis) the node's diversity facilitated performance coefficient and (y-axis) the node's diversity facilitated modularity coefficient. An identical figure for within-community-degree is shown in Figure S10.

We next built ordinary least squares linear predictive models to predict performance in subjects using a leave one out procedure, fitting each predictive model for all the subjects except the subject for which performance is predicted. To test the prediction accuracy, we calculated the Pearson *r* correlation coefficient between the predicted performance for each subject with the real performance of each subject. Our hypothesis is that networks with strong connector and local hubs are highly modular, and this network state is optimal for task performance. Thus, our predictive model consisted of two features, (1) the mean participation coefficient of connector hubs (Q+) minus the mean participation coefficient of all other nodes, and (2) the mean within-community-degree of local hubs (Q+) minus the mean within-community-degree of all other nodes. Thus, our predictive model measures the strength of connector and local hubs—based on participation coefficients and within-community-degree, respectively—in the network, relative to other nodes in the network. This predictive model significantly ($p<=0.01$) predicted performance in all five task states (Figure 5A).

For comparison, we built another predictive model based on each node's participation coefficient and each node's within-community-degree in each subject. As there are 264 nodes, this predictive model has 528 features. This predictive model was not able to predict performance in any task (Figure 5B). For another comparison, and to test whether the strength of connector hubs (Q+) and local hubs (Q+) in the first model provided information above and beyond that housed in the Q value, we built a predictive

model of performance with the Q value as the only feature (Figure 5C). This predictive model only significantly (p<=0.01) predicted performance in three of the five tasks, and underperformed the first predictive model in all five tasks. While Q alone is correlated with higher performance, strong connector and local hubs drive Q higher and the strength of these hubs most effectively predict task performance. These results suggest that a high Q value that is afforded by connector hubs having strong diverse connectivity across communities and local hubs having strong connectivity within their own communities is what increases performance, consistent with our predictions laid out in Figure 1.

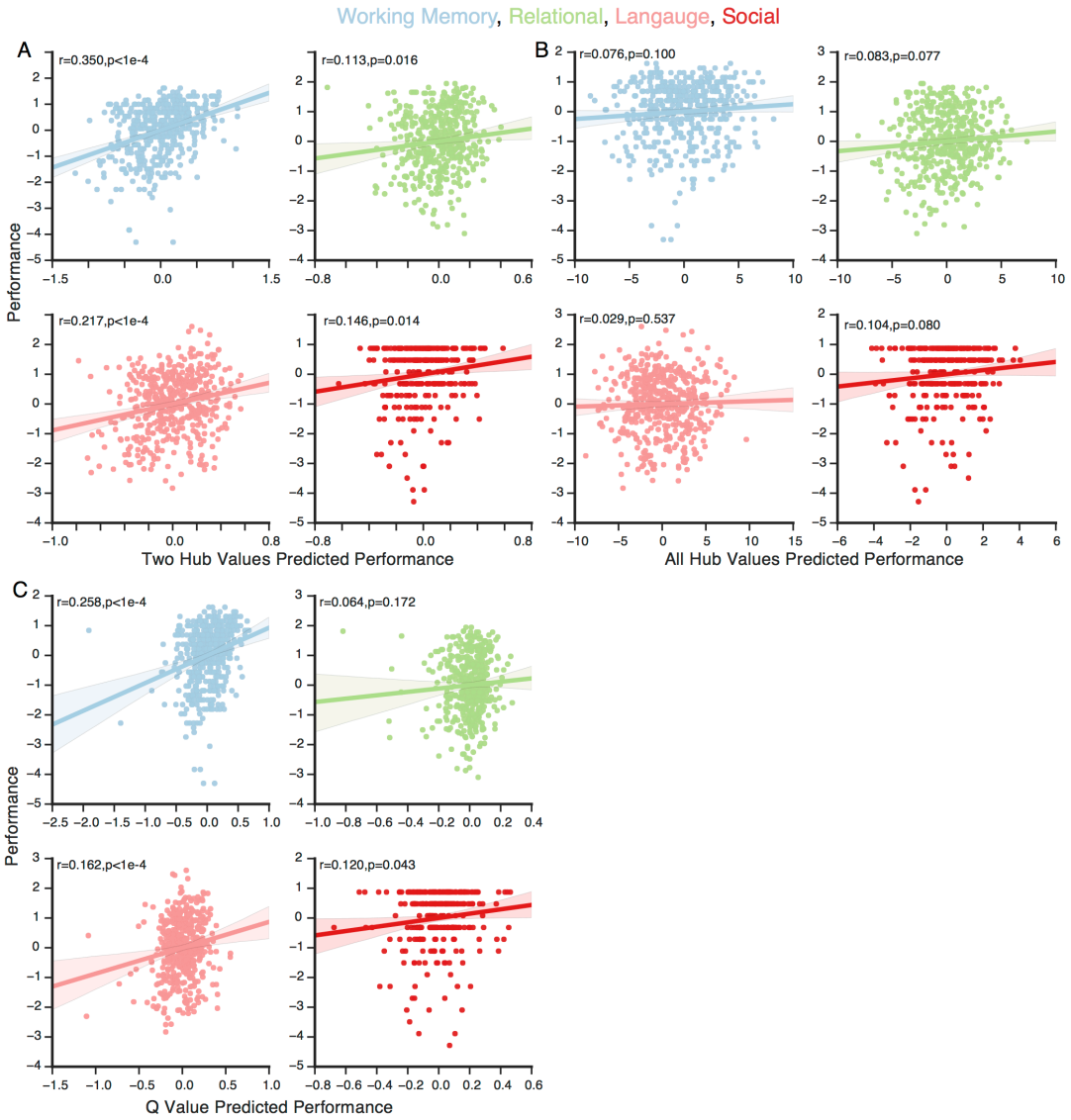


Figure 5. **Performance Models**. Three cross-validated (leave-one-out) predictive models to predict performance in tasks. Each data point represents the (x-axis) predicted performance of the subject and the (y-axis) true performance of the subject. (A) Predictive model with two features: the strength of connector and local hubs in the subject. (B) Predictive model with 528 features—each node's participation coefficient

and each nodes within-community-degree in the subject. (C) Predictive model with the subject's Q value as the only feature. In summary, we observe that the predictive model that uses the strength of connector and local hubs most accurately predicts performance across all five tasks.

Moreover, previous studies have suggested that connector hubs maintain modularity when many communities are engaged in a task(140). Thus, we then tested if the maintenance—the value remaining stable—of a high modularity value from resting-state to a task-state was correlated with the maintenance of high participation coefficients of connector hubs (Q+) from resting-state to task-state. We found that the change in Q was predicted by the changes in connector hubs' (Q+) participation coefficients from resting-state to task-state: if the participation coefficients of connector hubs (Q+; defined based on the task) remained high from rest to task, then the modularity value (Q) also remained high (Figure 6A). The same relationship held for local hubs and within-community-degrees: if the within community degrees of local hubs (Q+; defined based on the task) remained high, the modularity value also remained high Figure S11). These results provide additional evidence that connector hubs effectively serve to maintain modularity in the network: when connector hubs stay optimally connected to integrate and coordinate connectivity across communities (i.e., high participation coefficient), changes in modularity values from rest to task are minimized.

To test if the maintenance of high Q values increases task performance, we correlated task performance with the difference between a subject's resting-state modularity values and task-state modularity values. While the modularity values (Q) decreased in all task-states relative to the resting-state, the task-related Q value minus resting-state Q value (i.e., a smaller decrease in or maintenance of Q) was significantly correlated with an increase in task performance (Figure 6B). Thus, the maintenance of a high Q index from the resting-state to a task-state, which is partly driven by connector hubs (Q+) participation coefficients remaining high (Figure 6A), was associated with higher task performance. This result extends previous work showing that high-performing subjects exhibit smaller changes in connectivity between resting state and task-states(177).

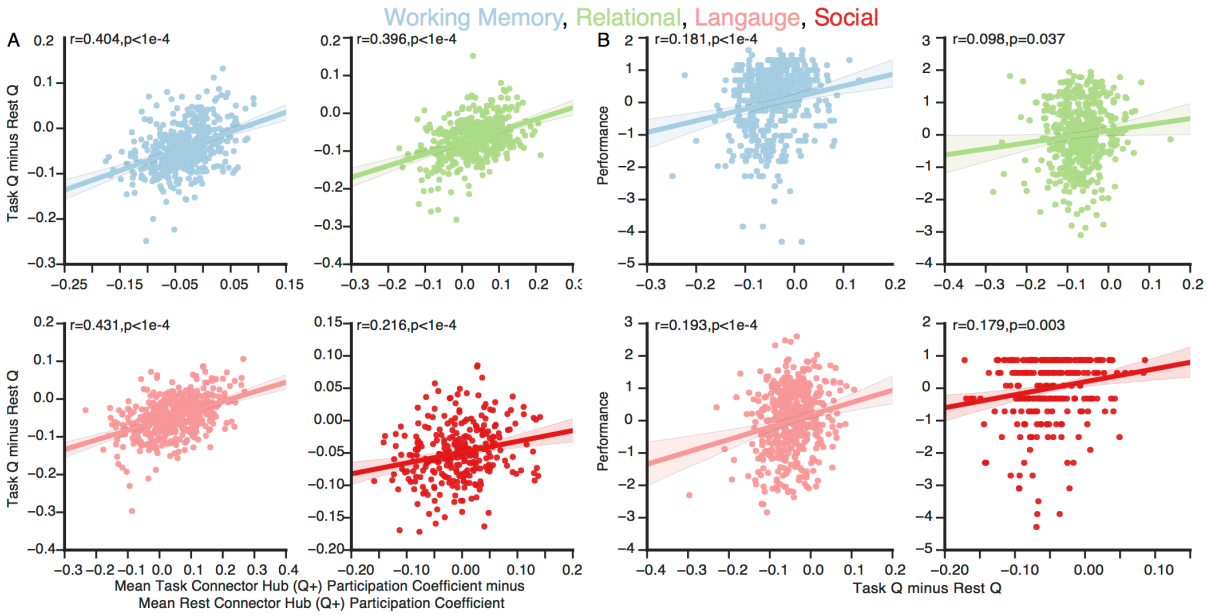


Figure 6. **The Maintenance of Connector Hubs and Modularity**. (A) The difference between task and rest modularity values (defined as task Q minus resting Q) was predicted by the difference in the strength of connector hubs—the mean participation coefficient of connector hubs (Q+) minus the mean participation coefficient of other nodes—between task and rest. If connector hubs remained strong from rest to task, then the modularity values (Q) also remained high from rest to task. Moreover, the difference between task and rest modularity values (defined as task Q minus resting Q) was predicted by the difference in the mean participation coefficient of connector hubs' (Q+) between task and rest (i.e., not relative to the mean participation coefficient of other nodes in the network, as in A; Figure S12). (B) The difference between task and rest modularity values (defined as task Q minus resting-state Q) is significantly correlated with enhanced performance in all tasks.

Discussion

Connector hubs' play a critical role in the overall network topology of the human brain—many connections spread across distant communities. Their expected function is to integrate information across communities and coordinate connectivity between communities. We hypothesized that this function allows for the brain network to be modular. We further hypothesized that a modular brain network with strong connector hubs is a cognitively optimal network state. To test these hypotheses, we leveraged individual differences in functional brain networks across a large cohort of healthy adult humans. Our approach addresses the question: when connector hubs are strong in an individual's network, is the network more modular and is cognitive performance enhanced? Our results demonstrate that strong connector hubs allow for a more modular and optimal network structure. Each connector hub's increased strength of connectivity to many networks altered the connectivity of that hub's neighbors to be more modular. Collectively, networks with strong connector hubs were more modular. Increases in modular connectivity were also accompanied by weaker connectivity

between the visual, sensory and motor hand, sensory and motor mouth, auditory, ventral attention, dorsal attention, and cingulo-opercular communities. Moreover, we uncover evidence that this local segregation and more distributed integration(108) supports optimal network performance—strong hubs and consequent high modularity accompany higher task performance in 5 different tasks.

Modularity is ubiquitously observed in Nature, across both organisms and spatial scales of architecture. For example, the genotype-phenotype map is modular, forming groups of traits that are co-affected by groups of genes(44, 174, 175, 178) and map non-trivially onto the organization of human brains(179, 180). Moreover, a modular structure is observed consistently across the brains of very different species, from *Caenorhabditis elegans* to humans(47). Modularity in these brain networks occurs when groups of nodes are more densely connected to other nodes in their same group than to nodes in other groups. Given its ubiquity, one could argue that Nature has selected for this modular network organization(181), perhaps due in part to the dependence of natural selection on considerations of metabolic cost. The maintenance of functional connectivity(182), as well as growth and utilization of white matter connectivity, is metabolically expensive(181-185). A modular architecture with anatomically localized and functionally specialized communities conserves white matter by reducing the average length and number of axonal projections, a quantity that is often referred to as the network's *wiring cost*. The modular solution to this cost dilemma has the added benefit that it provides the system with the ability to flexibly evolve and quickly adapt to new problems in a new environment(110, 111).

Yet, a complete separation of modules can be detrimental to the system's function, because it destroys the system's ability to transmit information between independent modules, a feature often necessary for an organism's survival. Thus, a natural question to ask is: how are these mostly segregated groups connected? Recent research has suggested that connector hubs integrate information(78, 140), coordinate connectivity between nodes in different communities(20, 67, 71, 140, 150), and maintain modularity(140, 152). However, given the amount of wiring required to link to many different and distant communities, connector hubs' connectivity pattern dramatically increases wiring costs(181). Moreover, the addition of connector hubs to a network with no connections between communities decreases the modular architecture of the network. However, connector hubs provide a *necessary* topological feature—a substrate for information to pass between or be integrated across modular communities. When connector hubs are damaged, the modular architecture of the brain is decreased(152) and cognitive processing suffers widespread consequences(125). Thus, despite their cost, connector hubs might be critically necessary for complex modular neural networks.

In the series of analyses we report here, we explicitly and comprehensively tested this principle in healthy adult individuals. We found that, when connector hubs are strongly and diversely connected to many communities, the modular architecture of the brain is

increased and cognitive performance across a broad range of tasks is enhanced. Thus, connector hubs appear to maintain an optimal modular architecture during integrative cognition—cognition that requires multiple communities(140). The presence of connector hubs might allow networks to remain modular and efficient while becoming increasingly fit for cognition: executing more complex functions, or novel functions, without greatly increasing the wiring cost or decreasing the modularity(106, 186). These data inform several hypotheses that could be addressed in future work. In particular, it is intuitively plausible that fully modular networks were selected early on in evolution and development to solve (relatively) simple and discrete problems, such as tactile sensation(106, 110, 186). In contrast, the requirements to solve new or more complex and general problems, like grasping, language, or social cognition, drive the system towards complex combinatorial processing across modular communities. The human brain appears to maintain modular structure while simultaneously solving the problem of integration with a select number of connector hubs that have their connections diversely spread across network communities. Supporting this reasoning, we found that high modularity alone was weakly predictive of task performance, while high modularity specifically afforded by strong connector and local hubs was the highly predictive of task performance. Thus, it is critical that modularity is maintained during integrative processing via a very specific solution—strong connector hubs.

It is important to note that there are other potential solutions for integration across communities, and thus the brain's solution is not necessary. For example, the addition of many "bridge" edges between communities can provide some degree of integration but can also diminish the network's fitness. Specifically, bridging connections provide a diffuse integration that is not localized to a few connector hub areas, thereby leading to a longer path length (how many edges a signal traverses) between communities. Indeed, to achieve the same level of integration and short path length characteristic of networks with strong connector hubs, a network with bridges would need to add a plethora of edges, which would have the detrimental effect of decreasing modularity and increasing wiring costs. Future empirical and computational work that investigates precisely how and why complex modular networks with connector hubs evolved and were selected will be valuable.

Finally, the interdependence between modular communities and connector hubs is a modern rendition of one of the first observations in neuroscience—Cajal's conservation principle. The brain has been naturally selected and is thus organized by an economic trade-off between minimizing the wiring cost of the network, which leads to modularity, and more costly connectivity patterns that increase fitness, like the functions afforded by connector hubs' diverse connectivity(105, 186). While modularity and integration are intuitively juxtaposed, connector hubs might be Nature's most modular and cheapest solution to integration.

Methodological Considerations

Network analyses require many assumptions and analytical choices. First, a fundamental assumption in our analysis is our selection of nodes. Many different parcellations of the brain into nodes exist. However, the parcellation we used was the only publically available one that fit all of our requirements (see Methods). Second, we chose Pearson *r* values to represent functional connectivity (i.e., edges) between nodes, for its simplicity in interpretation and ubiquity in human network neuroscience(187). However, more complex statistical measures could be employed, including measures that attempt to estimate the direction of each edge. Third, while we analyzed the resting state and 7 different tasks states, there are many other tasks, including more naturalistic tasks(8), that our analyses could be extended to. Fourth, we detected communities in each subject with InfoMap(129). While this method has been shown to be highly accurate, it is still a heuristic, as modularity maximization is NP-hard(188). In fact, the heuristic does not explicitly maximize Q(129). Regardless, we found that it detects a community structure with Q values highly similar to other methods (Supplemental Figure 13). Fifth, we used the participation coefficient to measure the strength of connector hubs. However, the participation coefficient only measures the hubs strength with communities defined at a single hierarchical level, not across hierarchical levels(189). While we integrated participation coefficients across network densities, which vary the granularity of the communities detected (lower density networks tend to have more communites(56), this is not explicitly hierarchical. Moreover, our brain networks only contained 264 nodes, while fMRI can study the human brain with upwards of 30,000 nodes (i.e., each voxel as a node). Thus, it would be interesting to extend this work across hierarchical and spatial scales in the network. Sixth, effects of motion on the connectivity structure of the network were aggressively addressed (Supplemental Tables 2-4). Finally, we did not regress out stimulus or task effects from the time series of each node, because how nodes' low frequency oscillations respond to stimulus or task effects is meaningful; however, the stability of our analyses across different tasks and resting state suggest that our results are not driven by stimulus effects or the specifics of a certain task. Other investigators have noted that task effect regression has minimal effects(21). Regardless, a similar analysis of the "background" functional connectivity not related to the task would be interesting.

Conclusion

In summary, when an individual's network has connector hubs that are diversely connected to many different communities in the brain, they are optimally connected to integrate information and coordinate connectivity across communities, which allows for both integrative and modular processing. We discovered that this optimal wiring increases an individual's global network modularity, the strength of their network's local hubs, and the individual's task performance. These results held across all tasks in the Human Connectome Project dataset, suggesting that there might be a universal optimal network structure for cognitive processing.

Methods

Data and Preprocessing. We used tfMRI and rsfMRI from the Human Connectome Project. For the tfMRI data, we used AFNI to preprocess the images, matching traditional resting-state functional connectivity studies. The AFNI command *3dTproject* was used, passing the mean signal from the cerebral spinal fluid mask, the mean signal from the white matter mask, the mean whole brain signal, and the motion parameters to the "-ort" options, which remove the signals via linear regression. The options "-automask", which generates the mask automatically was used. The "-passband 0.009 0.08" option, which removes frequencies outside 0.009 and 0.08, was used. Finally, the "-blur 6", was used, which smooths the images
(inside the mask only) with a filter that has a width (FWHM) of 6mm after the time series filtering. Because of the short length of the Emotion task, it was not included in our analyses. For the rsfMRI data, we used the images that were previously preprocessed with ICA-FIX. We also used the AFNI command *3dBandpass* to further preprocess these images. We used it to remove the mean whole brain signal and frequencies outside 0.009 and 0.08 (explicitly, "-ort whole_brain_signal.1D -band 0.009 0.08 -automask"). As subject motion during fMRI can impact functional connectivity estimates and has been shown to bias brain-task performance relationships(190), all analyses were executed with scrubbing executed on frames with frame-wise displacement greater than 0.2 millimeters, including the frame before and after the movement. Frame-wise displacement measures movement of the head from one volume to the next, and was computed as the sum of the absolute values of the differentiated rigid body realignment estimates (translation and rotation in x, y, and z directions) at every time point with rotation values evaluated with a radius of 50 mm(190). Frames were removed after all preprocessing was executed. Subjects with more than 75 percent of frames removed were not analyzed. Moreover, we executed all analyses after regressing out mean frame-wise displacement from the Q values and from task performance. None of our results were drastically altered following this regression, and all analyses remained statistically significant (at $p<0.05$), except for the prediction of performance on the Relational Reasoning Task. Supplemental Tables 2-4 contain these results.

Graph Theory Analyses. We used the Power atlas(56) to define the nodes in our graph because it was the only atlas that met all our requirements. First, because the homogeneity of nodes is high (all are 4 mm spheres) and the nodes do not share physical boundaries, the atlas will not overestimate the local connectivity of regions, and will capture long range connectivity and communities that are distributed across the brain. Second, the atlas is the only popular atlas that is defined based both on functional connectivity and task activation foci, making it highly principled for use in our case—where we aim to examine functional connectivity during both the resting-state and during task performance. Third, the atlas has been shown to accurately capture functional network organization (i.e., communities). This division of nodes into communities has been found with other approaches (e.g., at the voxel level), and this

division has been used in many studies(20, 56, 78, 191). Moreover, we make use of this division to calculate within and between community edge weight changes across subjects. Finally, this atlas has coverage of cortical, sub-cortical, and cerebellum areas.

All graph theory analyses were executed with our own custom python code (participation coefficient.github.com/mb3152/brain_graphs) that uses the iGraph library. All analysis code is also publically available (github.com/mb3152/dynamic_mod/dynamic_mod.py). For each task (both LR and RL encoding directions were used) and for each subject, the mean signal from 264 regions in the Power atlas was computed. The *r* between all pairs of signals was computed to form a 264 by 264 matrix, which was then Fisher transformed. The LR and RL matrices were then averaged. The mean matrix was then thresholded, retaining edge weights, at a range of cost (0.05 to 0.15 at 0.01 intervals), a common range and interval in graph theory analyses(56, 78, 140, 152). No negative correlations were included in our analyses. The matrix was then normalized to sum to a common value across subjects, and was used to represent the edges in the graph. Thus, all graphs had the same number of edges and sum of weights. For each cost, the InfoMap algorithm was run, and participation coefficients, within-community-degrees, and Q were calculated at each cost. Each subject's participation coefficients, within-community-degrees, and Q were the average of those values across the range of costs. We used the within-community-degree that z-scores the values based on the within-community-degrees of the community the node is in. All analyses were executed and all prediction models were fit separately for each task.

While InfoMap(129) does not explicitly maximize Q, it has been shown to estimate community structure accurately in several test cases(130), rendering the Q value, the participation coefficients, and within-module-degrees computed based on the community structure accurate and valid. Moreover, in biological networks, InfoMap achieves Q values that are similar to algorithms that maximize Q(192); in the current resting-state data, InfoMap Q values and Fast-Greedy Q(193) values were correlated at r=0.87 (p<1e-10); InfoMap Q values were higher than Fast-Greedy Q values (t=16, p<1e-10). InfoMap Q values and Louvain Q(38) values were correlated at r=0.98 (p<1e-10); Louvain Q values were higher than InfoMap Q values. When comparing InfoMap Q values to the distribution that includes both Louvain and Fast-Greedy, two algorithms that explicitly maximize Q, InfoMap Q values are higher (t=5,p<1e-5). These analyses are shown in Figure S13.

Edgewise effects of hub strength. For each node *i* we have a 264 x 264 matrix, where the *j-k* th entry is a Fisher z-transformed Pearson *r* correlation coefficient that captures how well the participation coefficients of node *i* correlate with the edge weights between nodes *j* and *k* in the network across subjects. These Pearson *r* values allowed us to test whether a node's participation coefficients correlate positively with its neighbors increased connectivity to its own community and decreased connectivity to other communities. For each pair of nodes *i* and *j*, we use the *r* values between node *i*'s

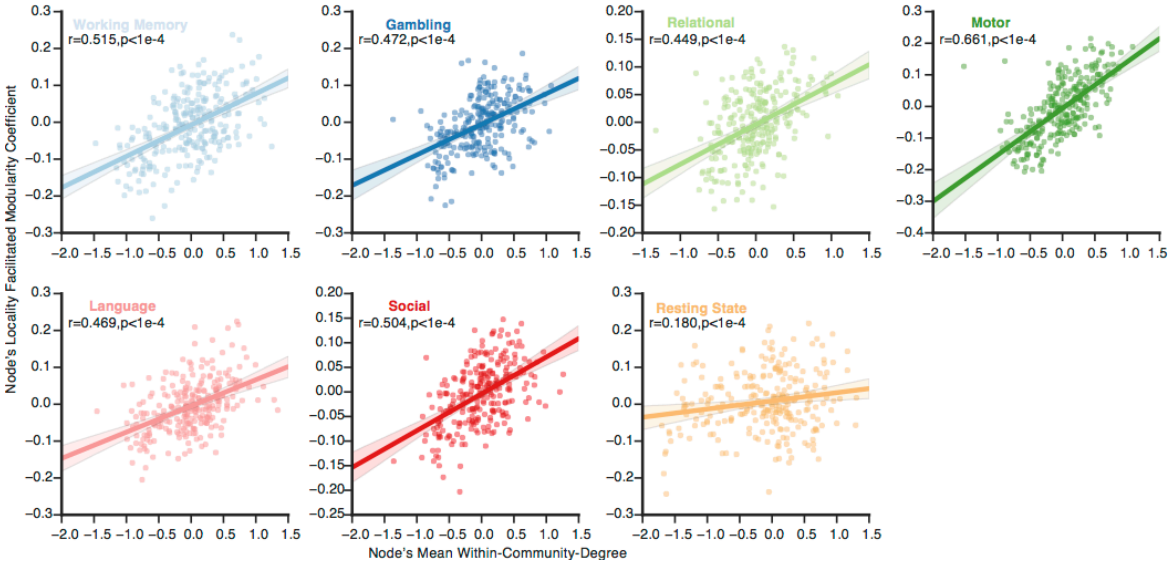
participation coefficients and all of node *j*'s edge weights across subjects. We subtract the sum of *r* values corresponding to node *i*'s participation coefficients and node *j*'s between-community edge weights from the sum of *r* values corresponding to node *i*'s participation coefficients and node *j*'s within-community edge weights. This metric measures how well the participation coefficients of node *i* are correlated with the increased modular connectivity of node *j*. We used the partition of nodes into communities that was created along with the nodes themselves (Figure 3D)(56). Edges between node *i* and node *j* were ignored in this calculation, as the participation coefficients of node *i* is likely highly correlated with the edge weights between node *i* and node *j*, and we were only interested in how the participation coefficient of node *i* modulates node *j*'s connectivity to the rest of the network, not node *j*'s connectivity to node *i*. Edges that were not positive on average across subjects were not included in this analysis, as the interpretation of negative edges in fMRI-based networks is not obvious (results were similar with a more stringent threshold of the top 25% strongest connections; Figure S7). The mean of this matrix for connector hubs(Q+) is shown in Figure 3B. Another 264 x 264 matrix for each node *i*, where the *j-k* th entry is a Fisher z-transformed Pearson *r* value that captures how well the within-community degrees of node *i* correlate with the edge weight between nodes *j* and *k* in the network across subjects.

Performance measures. All performance measures were chosen *a priori*. All performance measures were z-scored within each task. In the working memory tasks, we used the mean accuracy across all 2-back conditions (face, body, place, tool). In the relational task, we used mean accuracy across both the matching and the relational conditions. For the language task, we took the mean average difficulty level that the subject achieved in both the math and language conditions. We did not use accuracy, because the task varies in difficulty based on how well the subject is doing, making accuracy an inaccurate measure of performance for these tasks. Because level achieved might not be a precise accuracy or performance measure for this task, we included age adjusted accuracy measures from the behavioral test for fluency ('ReadEng_AgeAdj') and vocabulary ('PicVocab_AgeAdj'). These two measures were chosen simply because they are the only two behavioral measure of language performance that were collected, and they were age adjusted, which gave us a very precise measure of performance. No analogous mathematics measures were collected. The use of these two measures slightly increased prediction accuracy (*r* of 0.175 to 0.217), but the difference was not significant (z=0.64, p=0.5222). For the social cognition task, we used the mean percent of accurate categorizations of mental interaction or random interactions. Because there was a performance ceiling of 100 percent accuracy, we did not include subjects at ceiling. This slightly increased prediction accuracy (*r* of 0.101 to 0.145), but the increase was not significant (z=0.69, p=0.4902).

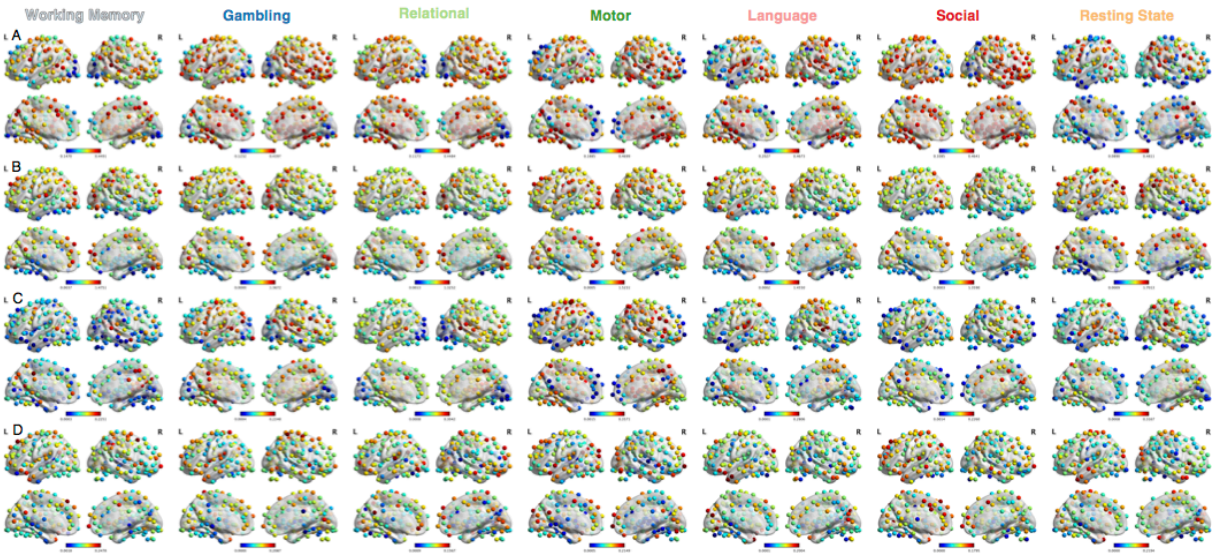
Acknowledgements. This work was supported by NIH Grant NS79698 and the National Science Foundation Graduate Research Fellowship Program under Grant no.

DGE 1106400 to MAB and MD. DSB would also like to acknowledge support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan 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 Health (2-R01-DC-009209-11,1R01HD086888-01, R01-MH107235, R01-MH107703, and R21-M MH-106799), the Office of Naval Research, and the National Science Foundation (BCS-1441502, PHY-1554488, and BCS-1631550). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

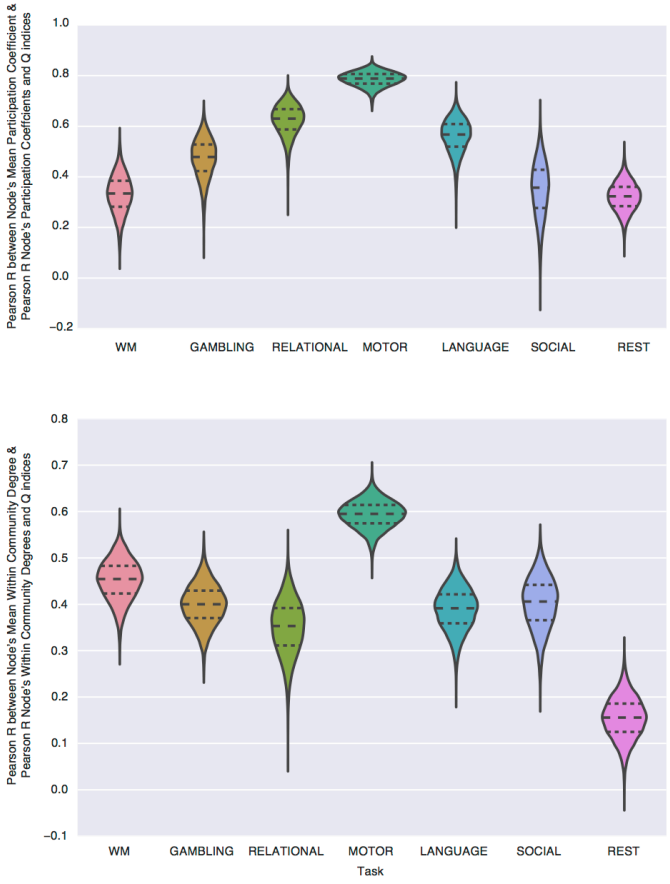
Supplemental Figure Legends



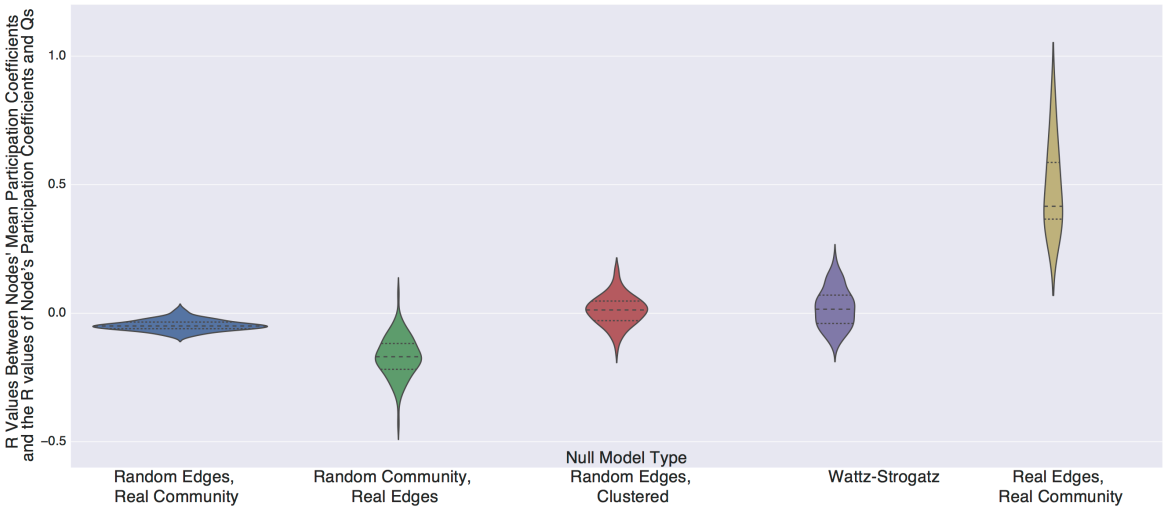
Supplemental Figure 1. **Locality Facilitated Modularity Coefficients**. Correlations between a node's mean within-community-degree and the Pearson *r* between that node's within-community-degrees and Q indices—the node's locality facilitated modularity coefficient. Regression plots for each task show the correlation of each node's locality facilitated modularity coefficient with that node's mean within-community-degree across subjects. Each dot represents a node. Shaded areas represent a 95 percent confidence interval. In all states, there was a positive correlation between a node's mean within-community-degree the node's locality facilitated modularity coefficient.



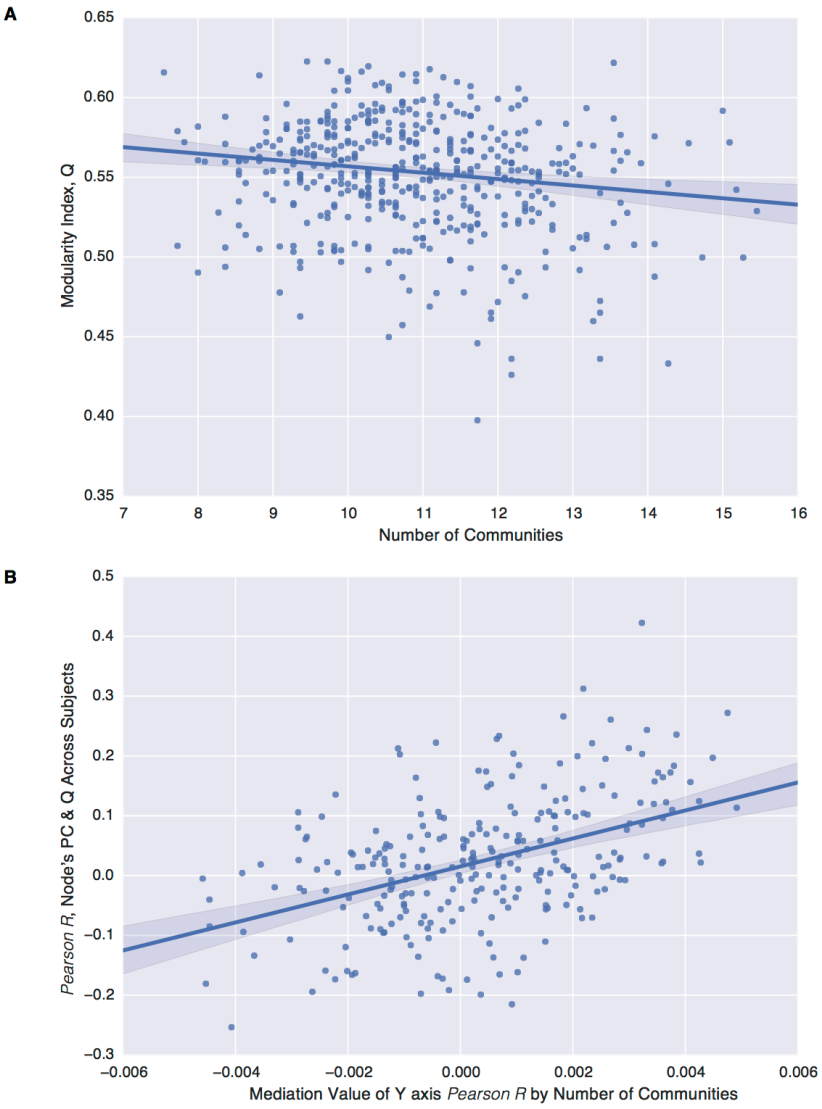
Supplemental Figure 2. **Anatomical Locations of Nodes and Node's Graph Theory Metrics**. Brain images showing the surface based anatomical locations of the nodes and their corresponding mean participation coefficient (A), diversity facilitated modularity coefficient (B), mean within-community-degree (C), and locality facilitated modularity coefficient (D).



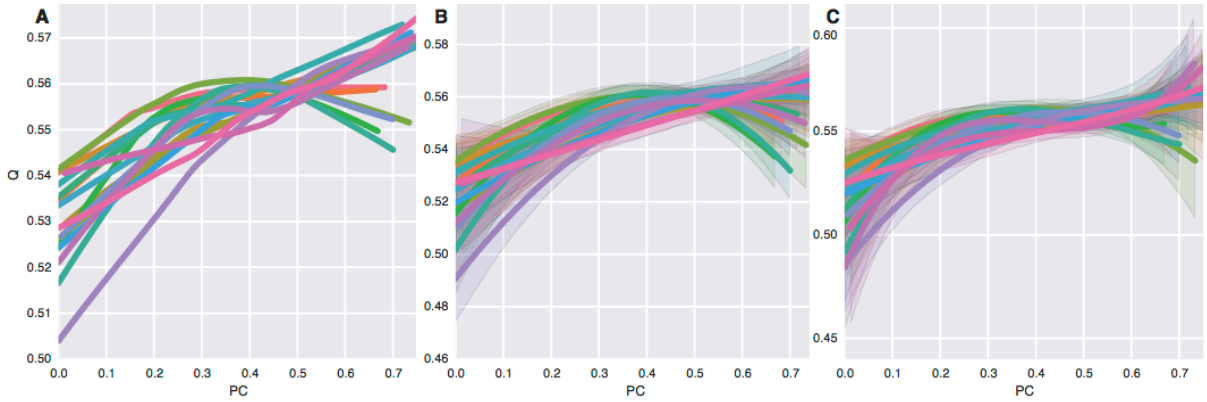
Supplemental Figure 3. **Cross Validation of Diversity and Locality Facilitated Modularity Coefficients**. Connector Hubs and Local Hubs are defined by a higher participation coefficient and within-community-degree, respectively, across subjects on average. To avoid potential dependencies, we estimated the mean within-community-degree or participation coefficient of each node in one half of subjects, and the correlation between a nodes' within-community-degrees or participation coefficients and Q indices—the node's locality and diversity facilitated modularity coefficients—in the other half, testing every possible split of subjects. On average, correlations were still significant in every state. A, the distribution of Pearson *r* values between a node's mean participation coefficient (defined in half the subjects) and the diversity facilitated modularity coefficient (defined in the other half of the subjects). B, the distribution of Pearson *r* values between a node's mean within-community-degree (defined in half the subjects) and the locality facilitated modularity coefficient (defined in the other half of the subjects). Resting-State data was used for this analysis.



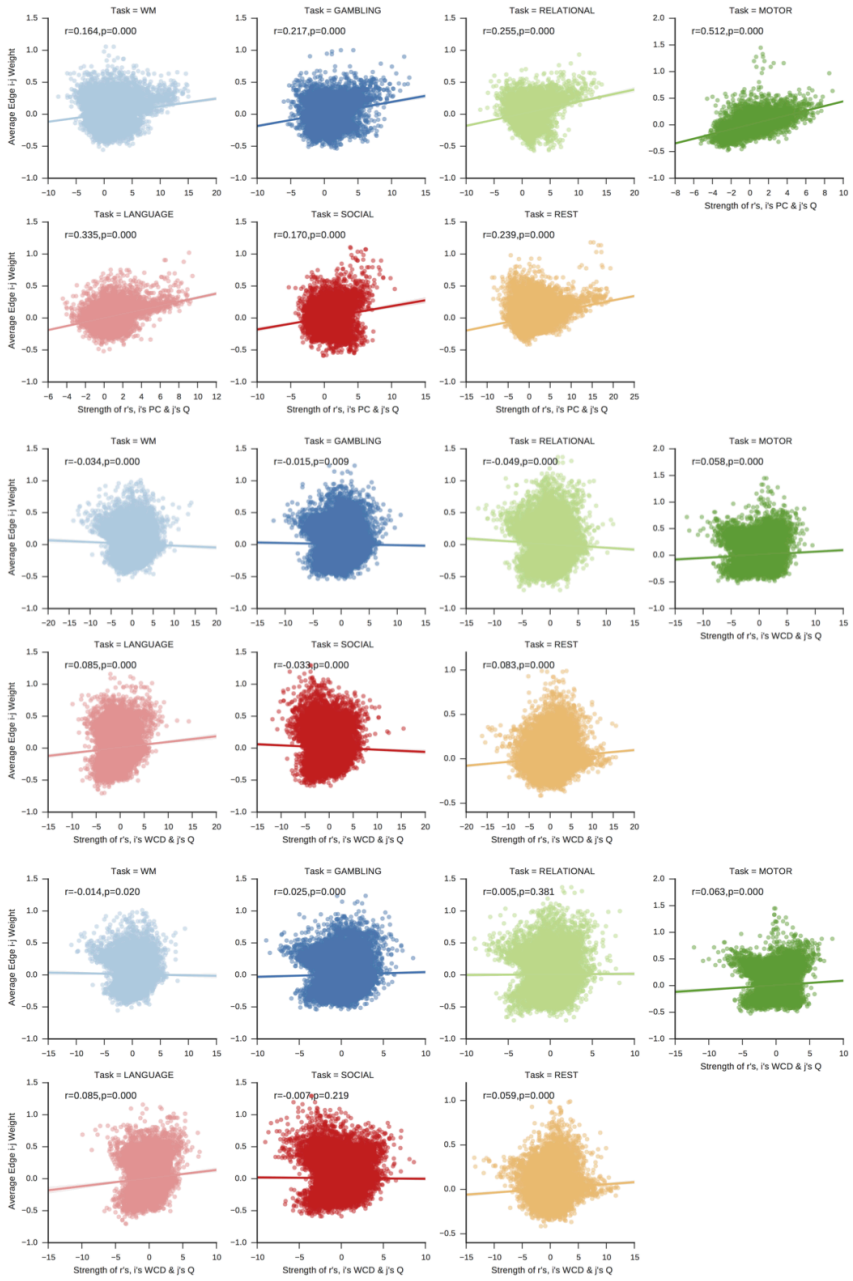
Supplemental Figure 4. **Null Models**. We tested four null models to evaluate the relationship between (x) a node's average participation coefficient across graphs and (y) the Pearson correlation coefficient *r* between that node's participation coefficients and the graphs' Q values—the node's diversity facilitated modularity coefficient—across the ensemble of individual subjects' graphs studied. (1) *Random Edges, Real Community* utilizes the true partition of nodes into communities that was uncovered by the application of community detection to each subject's intact resting-state graph (thresholded at a weighted graph density of 0.10, chosen based on this cost being the median cost from our original analyses), but we randomly permuted the edges uniformly after the partition was identified. (2) *Random Community, Real Edges* utilizes the partition of nodes into communities that was uncovered by the application of community detection to each subject's intact resting-state graph (thresholded at a weighted graph density of 0.10), but we then randomly permuted the assignment of nodes to communities (this retains the same number of communities and sizes of communities as the true partition). The edges remain in their true locations. (3) *Random Edges, Clustered* permutes the edges in each subject's resting-state graph uniformly at random. We then applied community detection to this permuted graph to identify a partition of nodes into communities, and then calculated the participation coefficient of each node based on that partition. We used a cost of 0.05, as denser random graphs result in just one community, which results in trivial participation coefficients of 0. Finally, (4) we generated *Wattz-Strogatz* small world graphs. Each graph had 264 nodes with each node initially connected to its 7 neighbors in the lattice. We set the rewiring probability to 0.25. This results in Q values of roughly 0.40 and a binary density of roughly 0.05. We ran 100 instantiations of each model. For each instantiation of each of the four models, we generated a graph for each subject using the subject's original graph and that particular null model. We then calculated the Pearson correlation coefficient, *r*, between (x) and (y) across the subjects' graphs (i.e., the same analysis as Figure 2a). Violin plots are shown for the distribution of Pearson correlation coefficient *r* values across the 100 instantiations of each model. For comparison, the distribution from the original analysis across tasks (*Real Edges, Real Community*) is also shown.



Supplemental Figure 5. **Participation Coefficients, the Number of Communities in the Network, and the Modularity Quality Index (Q)**. (A), negative correlation between the number of communities in subjects' graphs (averaged across graph densities in the range 0.05-0.15) and the modularity index Q of the graphs (also averaged across graph densities in the range 0.05-0.15). (B), we performed a mediation analysis between each node's participation coefficients, the number of communities in the graphs, and the Q values of the graphs, with the number of communities being the mediator. Mediation values are plotted for each node on the x-axis. The y-axis is the Pearson correlation coefficient *r* between each node's participation coefficients and Q indices across subjects' graphs—the diversity facilitated modularity coefficient.

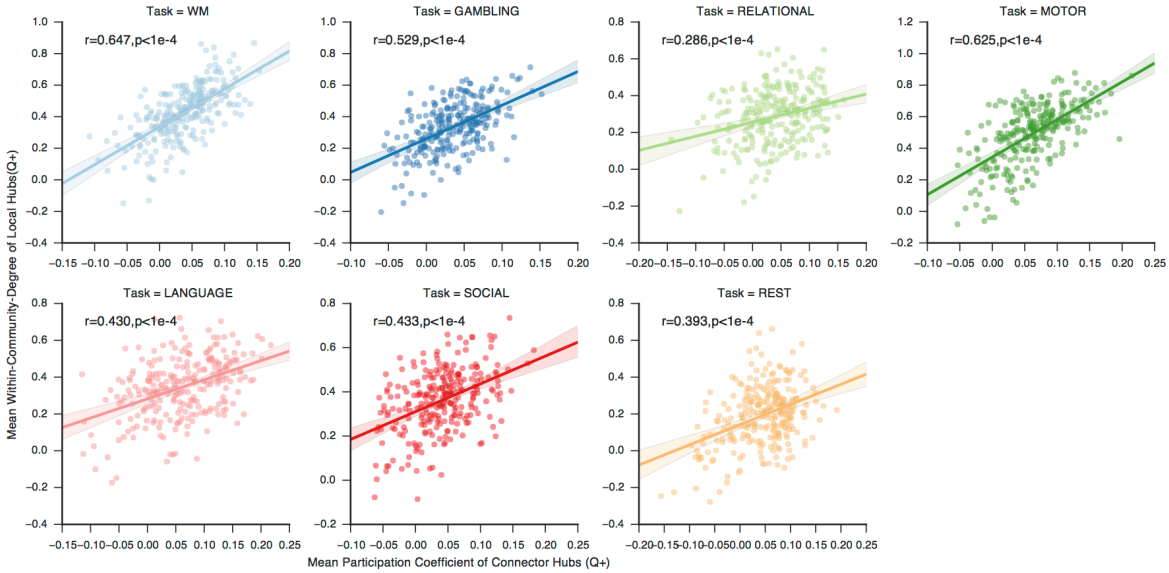


Supplemental Figure 6. **Relationships Between Individual Connector Hubs' Participation Coefficients and the Modularity Quality Index Q**. To test if the relationship between a connector hub's participation coefficients and modularity indices (Q)—the connector hubs' diversity facilitated modularity index—was linear, we fit three regression models to individual connector hubs' participation coefficients and Q indices across subjects. In each plot, each line is the relationship between a single connector hub's participation coefficients and Q indices across subjects. Only nodes with positive Pearson *r* values are shown. A, locally weighted scatter-plot smoother fit. B, 2nd order fit. C, 3rd order fit. The relation between many connector hubs' participation coefficients and Q indices across subjects is well captured by a first order fit, with the hub's maximal participation coefficients (0.7; the mathematically upper limit is $0.9\overline{9}$) corresponding to maximal Q. However, for example, some connector hubs' participation coefficients correspond to the maximal Q at a participation coefficient of 0.4-0.5, and then Q decreases at higher participation coefficients of that connector hub.

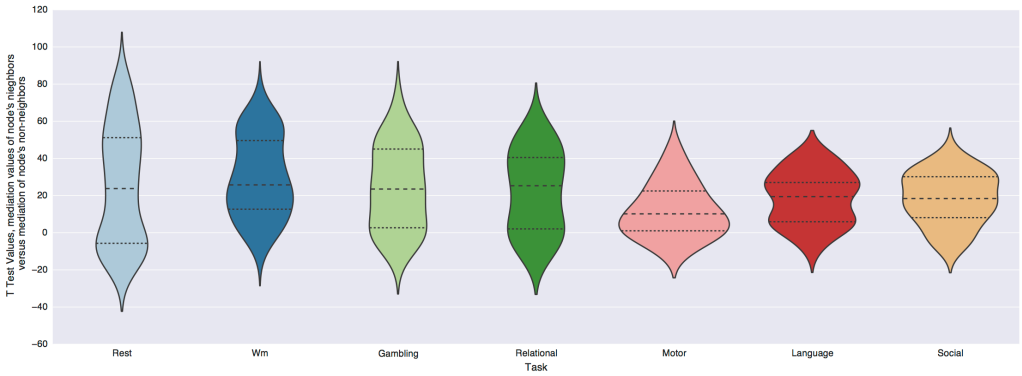


Supplemental Figure 7. **Nodes' Relationships with Individual Edges' Weights**. A, for each task, for each pair of nodes *i* and *j*, we correlated (x) the sum of Pearson r values between the participation coefficients of *i* and the within community edges of *j* minus the sum of Pearson r values between the participation coefficients of *i* and the between community edges of *j* with (y) the average connectivity weight between *i* and *j*. A, in contrast to our original analysis (Figure 2B), this analysis only includes the strongest 25 percent of edges. B, C, for each task, for each pair of nodes *i* and *j*, we correlated (x) the sum of Pearson *r* values between the within-community-degrees of *i* and the within community edges of *j* minus the sum of Pearson *r* values between the within-community-degrees of *i* and the between community edges of *j* with (y) the average

connectivity edge weight between *i* and *j*. B includes all positive edges (as in the main analysis for connector hubs' (Q+) participation coefficients), while C only includes the strongest 25 percent of edges.

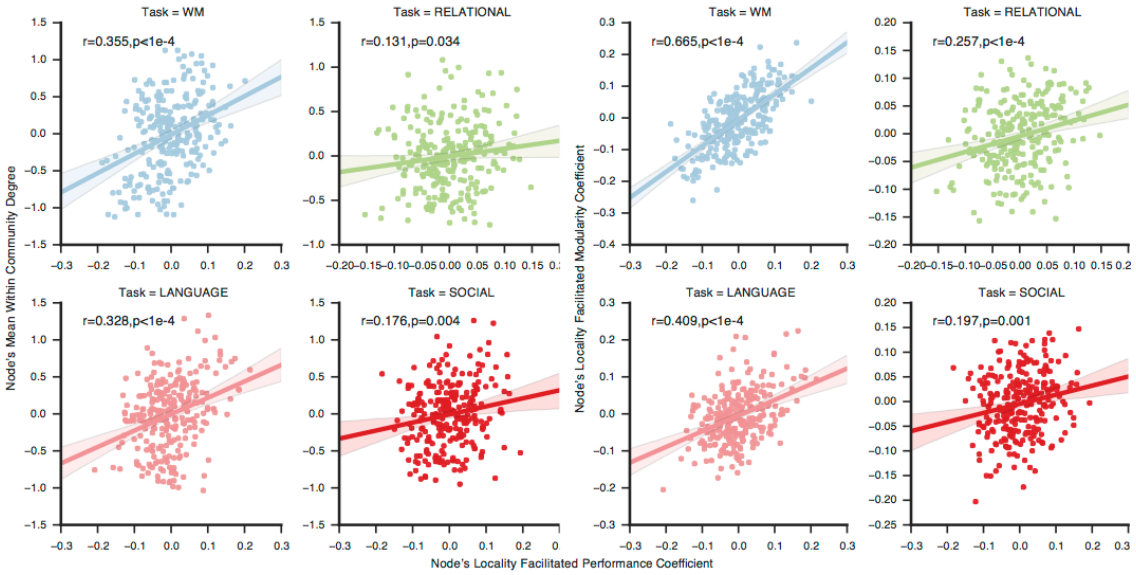


Supplemental Figure 8. **Connector Hubs (Q+) and Local Hubs (Q+) are Interrelated**. For each state, the correlation between connector hubs' (Q+) mean participation coefficient and local hubs' (Q+) mean within-community-degree.

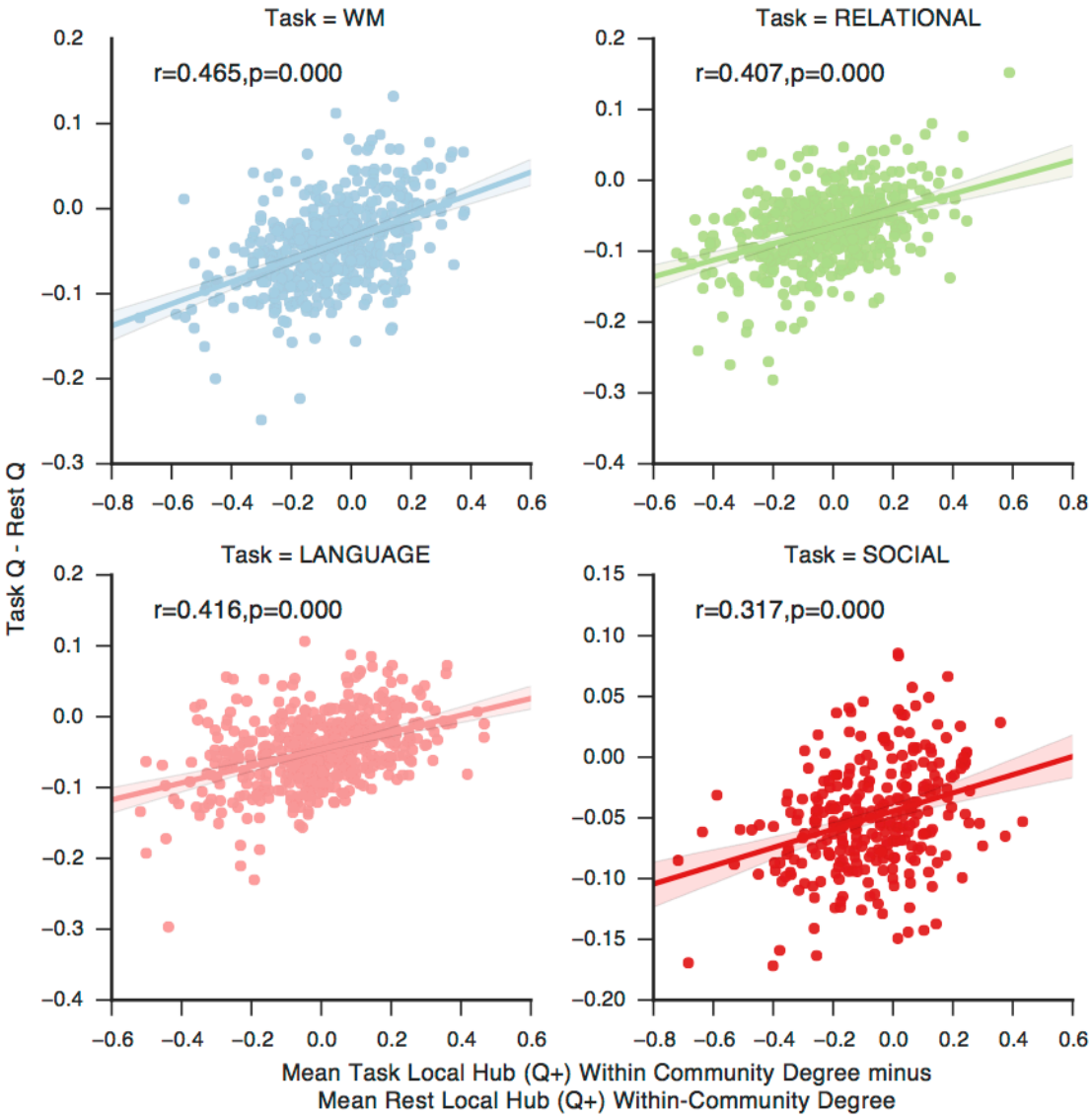


Supplemental Figure 9. **Connector Hubs' Mediation of Neighboring Nodes' Edges**. To investigate if the relationship between a connector hub's participation coefficient and Q is mediated primarily by that connector hub's neighbors' edge pattern increasing Q (per our prediction), we executed, for each connector hub node (i), a t-test between the absolute mediation values of node (i)'s neighbors' edges versus the absolute mediation values of node(i)'s non-neighbors' edges (neighbors were defined based on edges present between the two nodes in a graph at a density of 0.15, which was used because it is the densest density we utilized in our analyses; neighbors' and non-neighbors' edges connecting to the node (i) were ignored). Mediation values were based on the edge mediating between node(i)'s participation coefficients and modularity quality indices (Q). The distribution of t-values for connector hubs (Q+) is shown for each task. Connector hubs (Q+) showed significantly higher mediation values with

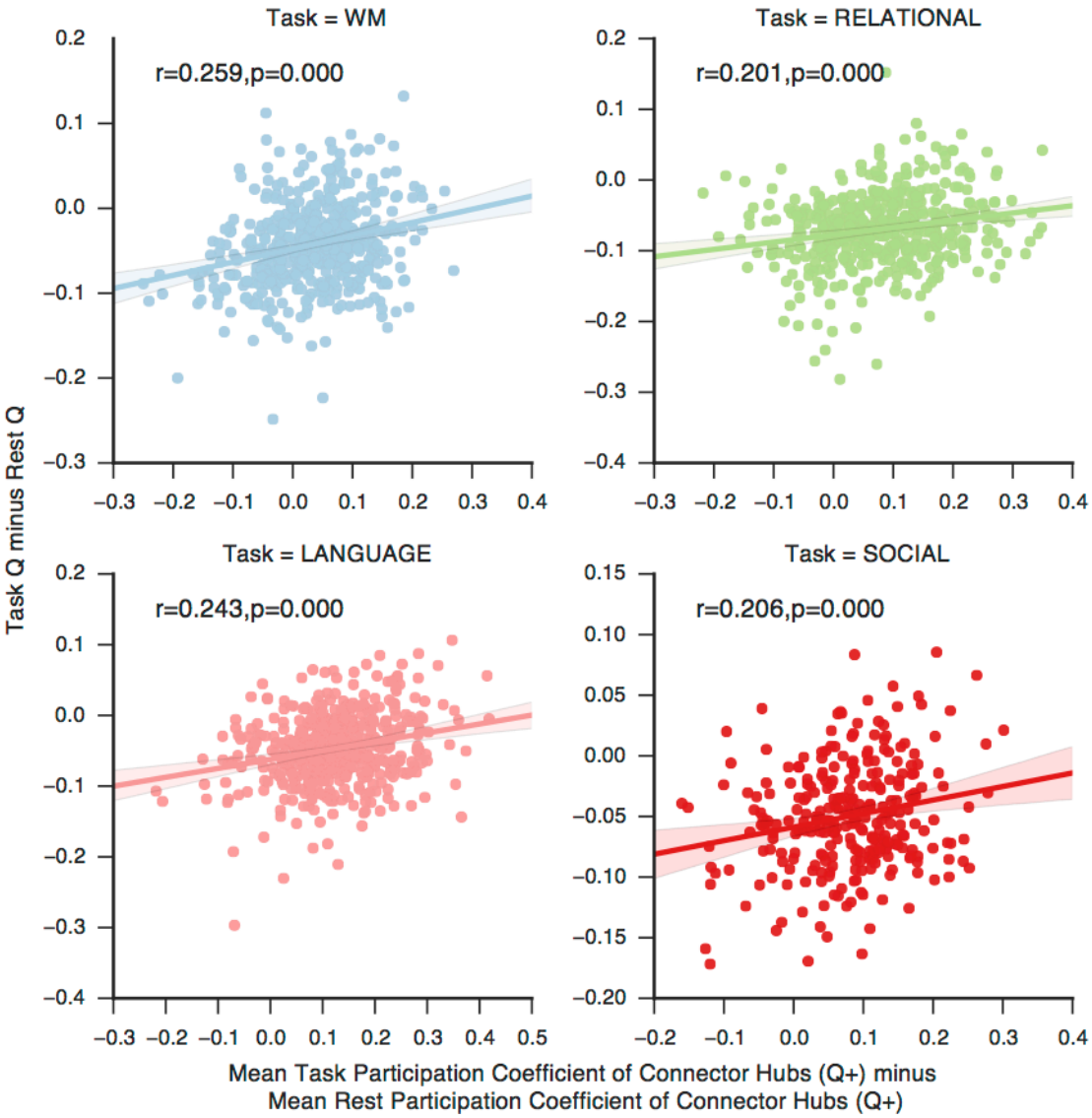
edges of its neighbors compared to the edges of its non-neighbors. An alternative analysis (Figure 2B) corroborates this analysis.



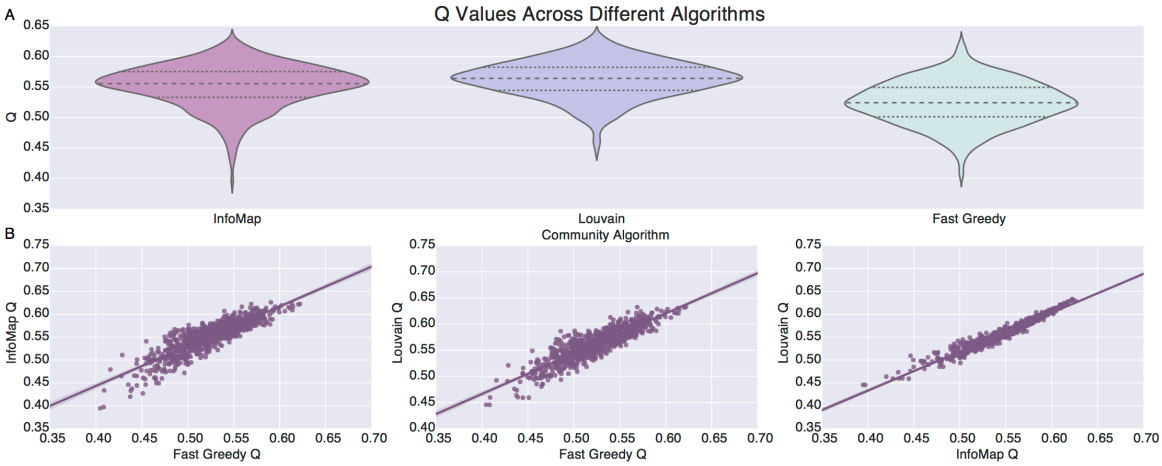
Supplemental Figure 10. **Within-community-degree, Q, and Performance**. A, the correlation between (x-axis) Pearson *r* value between a node's within-community-degree and task performance—the node's locality facilitated performance coefficient—with (y-axis) the node's average within-community-degree. B, the correlation between (x-axis) Pearson *r* value between a node's within-community-degree and task performance—the node's locality facilitated performance coefficient—with (y-axis) the node's locality facilitated modularity coefficient. An identical figure for participation coefficients is shown in Figure 4.



Supplemental Figure 11. **Maintenance of Q and Local Hubs' Within-Community-Degrees from Rest to Task**. Task Q – Resting-State Q was predicted by the change in the mean within-community-degree (Q+; relative to nodes for which their within-community was negatively correlated with Q) from rest to task. If the within-community-degrees of local hubs (Q+) remained high, Q remained high.



Supplemental Figure 12. **Maintenance of Q and Connector Hubs' Participation Coefficients from Rest to Task**. A, Task Q – Resting-State Q was predicted by the change in the mean participation coefficient of connector hubs (Q+; in contrast to Figure 6, this is not relative to nodes for which their participation coefficient was negatively correlated with Q) from rest to task. If the participation coefficients of connector hubs (Q+) remained high, Q remained high.



Supplemental Figure 13. **Modularity Quality Indices from Three Community Detection Algorithms**. We used Q indices repeatedly in our analyses. However, the community detection algorithm that we utilized, InfoMap, does not explicitly maximize Q. To see if this could potentially impact our analyses, we compared Q values from InfoMap to two popular alogrithms, Fast-Greedy and Louvain, that explicitly maximize Q. The mean Q value, as in our analyses, was taken across costs of 0.05 to 0.15. A, distribution, across subjects, of Q values of each algorithm. B, correlation between Q values across algorithms.

Original Performance Model

Task	Analysis	Statistic	Result
WM	Mean Nodal Prediction of Performance, LOO,	R, P	0.35, 0.0
RELATIONAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.113, 0.0163286961
LANGUAGE	Mean Nodal Prediction of Performance, LOO,	R, P	0.217, 2.3398e-06
SOCIAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.146, 0.0140094413

Performance Model, Scrubbed Frames > 0.2 FD

Task	Analysis	Statistic	Result
WM	Mean Nodal Prediction of Performance, LOO,	R, P	0.278, 7.34e-08
RELATIONAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.025, 0.6404642906
LANGUAGE	Mean Nodal Prediction of Performance, LOO,	R, P	0.222, 2.29587e-05
SOCIAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.131, 0.0499293239

Performance Model, Mean FD Regressed out from Performance

Task	Analysis	Statistic	Result
WM	Mean Nodal Prediction of Performance, LOO,	R, P	0.275, 1.2e-09
RELATIONAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.053, 0.2624622098
LANGUAGE	Mean Nodal Prediction of Performance, LOO,	R, P	0.186, 5.24167e-05
SOCIAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.122, 0.0396666447

Performance Model, Number of Communities Regressed out from Performance

Task	Analysis	Statistic	Result
WM	Mean Nodal Prediction of Performance, LOO,	R, P	0.35, 0.0
RELATIONAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.128, 0.0064493856
LANGUAGE	Mean Nodal Prediction of Performance, LOO,	R, P	0.152, 0.0010245281
SOCIAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.115, 0.0530236633

Performance Model, Mean FD Regressed out from Performance

Task	Analysis	Statistic	Result
WM	Mean Nodal Prediction of Performance, LOO,	R, P	0.275, 1.2e-09
RELATIONAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.053, 0.2624622098
LANGUAGE	Mean Nodal Prediction of Performance, LOO,	R, P	0.186, 5.24167e-05
SOCIAL	Mean Nodal Prediction of Performance, LOO,	R, P	0.122, 0.0396666447

Analyses Regressing out Number of Communities from Performance / Q

Task	Analysis	Statistic	Result
WM	r, Mean PC, r(PC,Q)	R, P	0.371, 5e-10
WM	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.355, 2.8e-09
WM	r, Mean PC, r(PC,Performance):	R, P	0.308, 3.389e-07
WM	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.305, 4.335e-07
GAMBLING	r, Mean PC, r(PC,Q)	R, P	0.574, 0.0
GAMBLING	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.404, 0.0
RELATIONAL	r, Mean PC, r(PC,Q)	R, P	0.698, 0.0
RELATIONAL	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.578, 0.0
RELATIONAL	r, Mean PC, r(PC,Performance):	R, P	0.357, 2.4e-09
RELATIONAL	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.387, 1e-10
MOTOR	r, Mean PC, r(PC,Q)	R, P	0.85, 0.0
MOTOR	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.768, 0.0
LANGUAGE	r, Mean PC, r(PC,Q)	R, P	0.644, 0.0
LANGUAGE	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.482, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance):	R, P	0.542, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.49, 0.0
SOCIAL	r, Mean PC, r(PC,Q)	R, P	0.431, 0.0
SOCIAL	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.332, 3.34e-08
SOCIAL	r, Mean PC, r(PC,Performance):	R, P	0.4, 0.0
SOCIAL	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.336, 2.23e-08
REST	r, Mean PC, r(PC,Q)	R, P	0.351, 4.5e-09
REST	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.341, 1.26e-08

Analyses after Scrubbing frames > .2 FD

Task	Analysis	Statistic	Result
WM	r, Mean PC, r(PC,Q)	R, P	0.512, 0.0
WM	r, Mean PC, r(PC,Performance):	R, P	0.347, 7e-09
GAMBLING	r, Mean PC, r(PC,Q)	R, P	0.291, 1.5162e-06
RELATIONAL	r, Mean PC, r(PC,Q)	R, P	0.5, 0.0
RELATIONAL	r, Mean PC, r(PC,Performance):	R, P	0.246, 5.4445e-05
MOTOR	r, Mean PC, r(PC,Q)	R, P	0.805, 0.0
LANGUAGE	r, Mean PC, r(PC,Q)	R, P	0.605, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance):	R, P	0.52, 0.0
SOCIAL	r, Mean PC, r(PC,Q)	R, P	0.279, 4.2262e-06
SOCIAL	r, Mean PC, r(PC,Performance):	R, P	0.321, 9.31e-08
REST	r, Mean PC, r(PC,Q)	R, P	0.483, 0.0

Correlations, Regressing out Mean FD from Q and Performance

Task	Analysis	Statistic	Result
WM	r, Mean PC, r(PC,Q)	R, P	0.371, 5e-10
WM	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.276, 5.4592e-06
WM	r, Mean PC, r(PC,Performance):	R, P	0.308, 3.389e-07
WM	r, Mean PC, r(PC,Performance), Motion Controlled r:	R, P	0.191, 0.0018233537
GAMBLING	r, Mean PC, r(PC,Q)	R, P	0.574, 0.0
GAMBLING	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.547, 0.0
RELATIONAL	r, Mean PC, r(PC,Q)	R, P	0.698, 0.0
RELATIONAL	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.655, 0.0
RELATIONAL	r, Mean PC, r(PC,Performance):	R, P	0.357, 2.4e-09
RELATIONAL	r, Mean PC, r(PC,Performance), Motion Controlled r:	R, P	0.256, 2.46314e-05
MOTOR	r, Mean PC, r(PC,Q)	R, P	0.85, 0.0
MOTOR	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.833, 0.0
LANGUAGE	r, Mean PC, r(PC,Q)	R, P	0.644, 0.0
LANGUAGE	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.622, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance):	R, P	0.542, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance), Motion Controlled r:	R, P	0.543, 0.0
SOCIAL	r, Mean PC, r(PC,Q)	R, P	0.431, 0.0
SOCIAL	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.413, 0.0
SOCIAL	r, Mean PC, r(PC,Performance):	R, P	0.4, 0.0
SOCIAL	r, Mean PC, r(PC,Performance), Motion Controlled r:	R, P	0.386, 1e-10
REST	r, Mean PC, r(PC,Q)	R, P	0.351, 4.5e-09
REST	r, Mean PC, r(PC,Q), Motion Controlled r:	R, P	0.356, 2.5e-09

Analyses Regressing out Number of Communities from Performance / Q

Task	Analysis	Statistic	Result
WM	r, Mean PC, r(PC,Q)	R, P	0.371, 5e-10
WM	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.355, 2.8e-09
WM	r, Mean PC, r(PC,Performance):	R, P	0.308, 3.389e-07
WM	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.305, 4.335e-07
GAMBLING	r, Mean PC, r(PC,Q)	R, P	0.574, 0.0
GAMBLING	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.404, 0.0
RELATIONAL	r, Mean PC, r(PC,Q)	R, P	0.698, 0.0
RELATIONAL	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.578, 0.0
RELATIONAL	r, Mean PC, r(PC,Performance):	R, P	0.357, 2.4e-09
RELATIONAL	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.387, 1e-10
MOTOR	r, Mean PC, r(PC,Q)	R, P	0.85, 0.0
MOTOR	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.768, 0.0
LANGUAGE	r, Mean PC, r(PC,Q)	R, P	0.644, 0.0
LANGUAGE	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.482, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance):	R, P	0.542, 0.0
LANGUAGE	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.49, 0.0
SOCIAL	r, Mean PC, r(PC,Q)	R, P	0.431, 0.0
SOCIAL	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.332, 3.34e-08
SOCIAL	r, Mean PC, r(PC,Performance):	R, P	0.4, 0.0
SOCIAL	r, Mean PC, r(PC,Performance), Number of Communities Controlled r:	R, P	0.336, 2.23e-08
REST	r, Mean PC, r(PC,Q)	R, P	0.351, 4.5e-09
REST	r, Mean PC, r(PC,Q), Number of Communities Controlled r:	R, P	0.341, 1.26e-08

The Fronto-Parietal-Thalamic Network

Max Bertolero, Kai Hwang, Mark D'Esposito

Abstract

The fronto-parietal network and the thalamus have both been identified as integrative or coordinative brain structures. However, the connectivity properties that allow for their function, and if these two structures interact, remains underspecified. Here, we find that every cortical network (n=15) has a distinct sub-region in the fronto-parietal network for which it is strongly functionally connected. These distinct sub-regions become interdigitated in the fronto-parietal network, forming a gradient towards convergent sub-regions where the fronto-parietal network is functional connected to many different cortical networks. However, we also found areas of the fronto-parietal network with strong functional connectivity to a single network. We also found that functional connectivity from the 15 cortical networks to the fronto-parietal network can be accurately reduced to three sets of cortical networks, with the cortical networks in each set exhibiting very similar functional connectivity to the fronto-parietal network. Moreover, for each cortical network's sub-region in the fronto-parietal network, there is a corresponding sub-region in the thalamus that that network is strongly functionally connected to. For each network, these two sub-regions (fronto-parietal and thalamic) are functional connected. Finally, functional connectivity between the fronto-parietal, the thalamus, and cortical networks predicts functional connectivity between cortical networks across individuals. These results reveal the underlying distinct and convergent functional connectivity of the fronto-parietal network, as well as its functional connectivity with the thalamus, that potentially allows for these brain structures to be integrative or coordinative.

Introduction

The human cortex can be divided into distinct networks, typically 7 to 17 (141) networks, by defining networks based on maximizing the similarity of connectivity among brain regions in the same network. Connectivity between two brain regions is typically defined based on the similarity of two regions' oscillatory patterns, measured with the Pearson correlation between the two regions' time-series of fMRI activity while the subject is not engaged in any cognitive task. This type of connectivity is termed "functional". Each network has been proposed to have a specific function (140). The fronto-parietal network is consistently defined based on functional connectivity and shows diverse connectivity to other networks(91, 140, 141, 194). It also exhibits the strong interconnectivity—it is a prominent part of the diverse club, a set of highly interconnected regions with diverse connectivity across the brain. The fronto-parietal network (Fig. 1a) has thus been suggested to integrate information across the other networks, as well as coordinate connectivity between other networks (140). The fronto-parietal network has also been suggested to be a "common end point" of diverse

information integration (78, 195) . Finally, the fronto-parietal network exhibits high fractionation; it can be subdivided into many distinct sub-networks of itself (196).

While sub-cortical structures have mostly been ignored in functional connectivity and network studies, recent studies suggest that thalamus is also an integrative structure(197). The thalamus also shows strong and diverse functional connectivity to all cortical networks, including the fronto-parietal network(197). Moreover, without appropriate thalamocortical inputs, the development of distinct cortical networks does not occur(198). Finally, damage to the regions of the thalamus or the fronto-parietal network that are functionally connected to many networks decreases the modularity of the entire network (152, 197).The connectivity properties that allow for the integrative and coordinative function of the fronto-parietal networks and the thalamus, and if these two structures interact, remains underspecified.

The fronto-parietal network appears to have distinct sub-regions that connect to other cortical networks(195, 199-202). However, other sub-regions in the fronto-parietal network are connected to many different modalities, including parts of posterior parietal cortex (98), superior temporal polysensory area (203), and prefrontal cortex(204-207). The fronto-parietal network participates in canonical and non-canonical processing. Canonical hierarchical information convergence in the fronto-parietal network involves information flow that is mostly linear and unidirectional, becoming increasingly integrated with other information(205). For example, information from early retinotopic visual areas, which are hierarchically organized, projects to specialized extrastriate fields (MT) that represent direction and speed, and this information is then passed from the lateral intraparietal sulcus to the frontal eye field (two regions in the fronto-parietal network), which guides eye movements(208, 209). However, the fronto-parietal network is mostly defined by non-canonical reciprocal connectivity and processing. There is strong interconnectivity between fronto-parietal regions, as well as with the thalamus; the posterior parietal cortex and frontal cortex are reciprocally connected to each other and also connected to at least 15 other cortical areas, as well as the medial pulvinar of the thalamus(195, 210-213).

A complete model of the fronto-parietal network's and the thalamus's function and connectivity explains and fits with all of these findings. One potential connectivity structure of the fronto-parietal network, the thalamus, and cortical networks is that both the fronto-parietal network and the thalamus have sub-regions that are strongly connected to a single network, and these two sub-regions (e.g., the two sub-regions selective for the motor network in the thalamus and in the fronto-parietal network) are also highly connected. Here, we test this model by parcellating the fronto-parietal network and the thalamus based on their functional connectivity to cortical networks, and examining the connectivity structure between the fronto-parietal network, the thalamus, and cortex.

Methods

Data

We utilized the resting-state fMRI data in the The Enhanced Nathan Kline Institute Rockland Sample (NKI) and the Human Connectome Project (HCP; S500 release (171)). The frontal-parietal network and the thalamus were parcellated using the NKI data. All futher analyses of functional connectivity using these parcels utilized the HCP Database. This procedure ensures the reproducibility of the results and ensures that the parcellation is not overfitting the data. In other words, that we are not simply fitting, and then analyzing, noise. The parcellations were fit using the NKI data and further analyses of the functional connectivity were executed utilizing the HCP data, as the NKI dataset does not have enough subjects to fit the predictive models that were executed. Moreover, the temporal signal-to-noise of the NKI data was higher than the HCP data, which is especially important for analyses of sub-cortical structures, where temporal signal-to-noise is typically lower than in cortex.

Preprocessing

NKI Dataset. We utilized the 645 millisecond TR resting-state scan (900 timepoints). All subjects aged 18-35 were analyzed (number of subjects=64). Image preprocessing was performed with the software Configurable Pipeline for the Analysis of Connectomics(214). First, brain images were segmented into white matter (WM), gray matter, and cerebral spinal fluid (CSF). Rigid body motion correction was then performed to align each volume to a temporally averaged volume, and a boundary-based registration algorithm was used to register the EPI volumes to the anatomical image. Advanced Normalization Tools (ANTS) was used to register the images to MNI152 template using a nonlinear normalization procedure(215). All images were spatially resampled to 2mm voxel resolution. We then performed nuisance regression to further reduce non-neural noise and artifacts. To reduce motion-related artifacts, we used the Friston-24 regressors model during nuisance regression. WM and CSF signals were regressed using the anatomical CompCor approach with five components for each tissue class(216). Linear and quadratic drifts were also removed. After regression, data were bandpass filtered from 0.009–0.08 Hz, and scaled to a whole-brain mean value of 10000. Given that the thalamus is a relatively small structure, to avoid signal blurring we did not perform any spatial smoothing.

HCP Dataset. Human fMRI data from the 150 subjects with the smallest mean framewise displacement across all four resting-state scans during rest were analyze. We used the images that were previously preprocessed with ICA-FIX. Analysis of Functional Neuroimages (29) (AFNI) was used to further preprocess the images. The AFNI command 3dBandpass was used to remove the mean whole brain signal and frequencies outside 0.009 and 0.08 (explicitly, “-ort whole_brain_signal.1D -band 0.009 0.08 -automask”). Both LR and RL encoding directions and both sessions (REST1 and REST2) were used, resulting in 4800 timepoints.

Fronto-parietal parcellation

A parcellation method was utilized that is identical to a previous parcellation of the thalamus, which assigns each voxel to the cortical network is has the strongest functional connectivity with (Hwang et al., 2016). Here, for each voxel in the fronto-parietal network, its functional connectivity to each cortical network is calculated. We used the cortical networks defined by a previous publication (141). The mean signal from each cortical network's voxels is utilized. As two fronto-parietal networks were found in that study, we merged them (Figure 1a). For each correlation between the fronto-parietal voxel and a given cortical network, signals from the other cortical networks are regressed out from the cortical network's signal (i.e., partial correlations were computed). Partial correlation were used for two reasons. First, we also wanted to explore thalamic connectivity, and the thalamic parcellation utilized partial correlations (197). Second, for each voxel in the fronto-parietal network, the partial correlation measures the amount of variance that voxel's activity explains in a given cortical network's activity that cannot be explained by other cortical networks' activity. This measure thus captures the unique connectivity between the frontal-parietal network and each cortical network. This analysis was executed for each NKI subject, and the mean functional connectivity value, across subjects, was computed. Next, each voxel is assigned to the cortical network it has the strongest functional connectivity to, creating a parcel for each cortical network in the fronto-parietal network. The same analysis was executed for voxels in the thalamus. This thalamus parcellation method is thus identical to previous work(197).

The same correlations between fronto-parietal and cortical networks and between thalamic voxels and cortical networks were computed using the Human Connectome Project dataset. To analyze the accuracy of the fronto-parietal parcels defined in the NKI data, we visualized the mean functional connectivity, calculated using the HCP data, of each cortical network to each voxel in the fronto-parietal network, with the parcel for that network shown (Figure 1b). Additionally, we measured the mean functional connectivity of each parcel to the parcel's cortical network. We also measured the homogeneity of each parcel—the mean of all Pearson correlations between the time-series of voxels in each parcel. Additionally, we compared the distribution of those Pearson correlations between the time-series of voxels in a parcel to the distribution of Pearson correlations between the time-series of voxels in that parcel and voxels in other parcels in the fronto-parietal network.

Finally, we measured how strongly connected each parcel in the fronto-parietal network is to each parcel in the thalamus. Here, the correlation between each parcel in the thalamus and each parcel in the fronto-parietal network was calculated; signals from the parcels in the fronto-parietal network are regressed out from the fronto-parietal network's parcel's signal (i.e., partial correlations were computed).

Clustering of Functional Connectivity to the Fronto-Parietal Network

Utilizing the NKI Data, we calculated the similarity of cortical networks' connectivity to the fronto-parietal network. Each cortical network has a functional connectivity value to each voxel in the fronto-parietal network. Thus, the similarity of two cortical networks' connectivity to the fronto-parietal network was measured by the Pearson correlation of the cortical networks' functional connectivity to the voxels in fronto-parietal network. This measure is computed for each pair of cortical networks. Given this matrix, a graph of nodes and edges can be formed, where each node is a cortical network, and the edge is the similarity metric. Negative correlations were ignored, resulting in a graph with a density of 0.419. Community detection can then be applied. We used InfoMap to identify communities. The mean connectivity of each community was then calculated using the HCP data.

Measures of Discrete and Convergent Connectivity in the Fronto-Parietal Network

We first sought to measure the discreteness of each fronto-parietal voxel's connectivity to cortical networks. For each voxel in the fronto-parietal network, we had previously calculated its mean connectivity (across subjects) to each cortical network. Using these 15 values (1 for each cortical network, NKI dataset), we measured the discreteness of that voxel's connectivity by computing the difference in the strongest value and the second strongest value. Thus, if a voxel has strong connectivity to a single network, and weak connectivity to other networks, it has a high discreteness score. If a voxel is equally connected to two cortical networks, it will have a very low discreteness score.

We then sought to measure the convergence of cortical network signals at each voxel in the fronto-parietal network. Using the same 15 values above, we calculated the standard deviation of these values. A small standard deviation demonstrates that that voxel is equally connected to all cortical networks. Thus, our convergence measure is simply 1 minus the standard deviation of the functional connectivity scores of that voxel to all cortical networks.

Predictive Models of Cortical to Cortical Functional Connectivity

We first sought to predict the functional connectivity between cortical networks using the functional connectivity between parcels in the fronto-parietal network and cortical networks. For each parcel in the fronto-parietal network, functional connectivity between the mean signal of voxels in that parcel and each cortical network (excluding the fronto-parietal network) was calculated (the signals from other cortical networks were regressed out from the cortical network's signal), resulting in a 15x15 matrix with 105 unique values for each subject. We also calculated the functional connectivity between all cortical networks, resulting in a 16x16 matrix with 120 unique values for each subject. Full correlations, not partial, were used for this matrix. We used the fronto-parietal parcel to cortical matrix functional connectivity values to predict each cortical to cortical matrix functional connectivity value.

We then sought to predict the connectivity between cortical networks using the functional connectivity between parcels in the thalamus and cortical networks. For each parcel in the thalamus, functional connectivity between the mean signal of voxels in that parcel and each cortical network (excluding the fronto-parietal network) was calculated (the signals from other cortical networks were regressed out from the cortical network's signal), resulting in a 15x15 matrix with 105 unique values for each subject. We then used this matrix to predict each cortical to cortical matrix functional connectivity value.

Finally, we sought to predict the connectivity between cortical networks using the functional connectivity between parcels in the fronto-parietal network and parcels in the thalamus. For each parcel in the thalamus, functional connectivity between the mean signal of voxels in that parcel and the mean signal of each parcel in the fronto-parietal network was calculated (the signals from the other fronto-parietal parcels were regressed out from the fronto-parietal parcel's signal), resulting in a 15x15 matrix with 105 unique values for each subject. We then used this matrix to predict each cortical to cortical matrix functional connectivity value.

For all three prediction models, we used a leave-one-out cross validation technique. For each subject in the HCP data, we fit the model on all subjects in the HCP data except for that subject, and used the model to predict each cortical to cortical functional connectivity of that subject. To measure the performance of the model, we then correlated the real cortical to cortical functional connectivity value with the predicted cortical functional connectivity value. We then calculated how similar the models' performance were; in other words, we measured if the models were similar in which cortical to cortical functional connectivity values they predicted well. This was done by calculating the Pearson correlation of the prediction values across all cortical to cortical networks (e.g., the Pearson correlation between the two matrices, a and b, in Figure 5).

Results

Fronto-parietal parcellation

Every cortical network had voxels in the fronto-parietal network for which that cortical network exhibited the strongest functionally of all cortical networks (Figure 1b). Moreover, while the parcels were defined based on the NKI data, Figure 1b shows the parcels overlaid on top of functional connectivity, calculated in the HPC data, between each network and the fronto-parietal voxels. For all results, the parcels were calculated using the NKI data, and the analyses were executed using the HCP data.

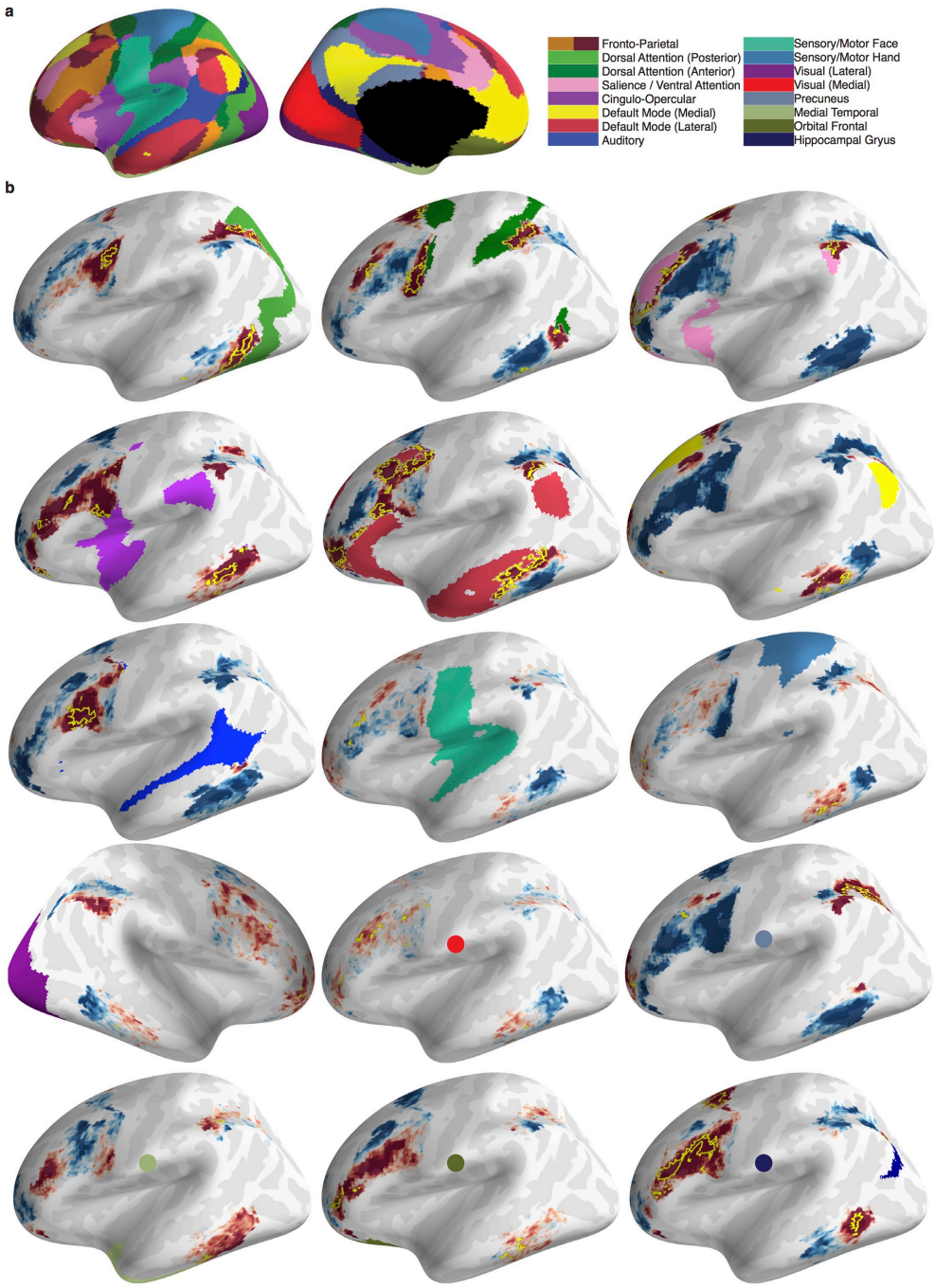
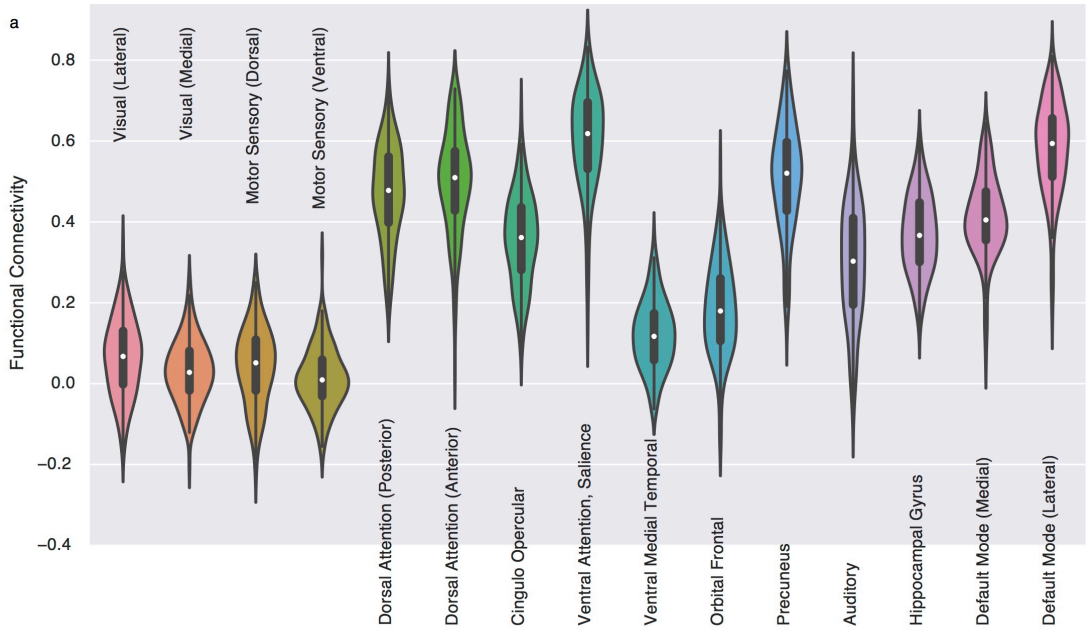


Figure 1. a, Cortical networks. b, Using the NKI Data, for each cortical network, the voxels in the fronto-parietal network for which that network was most strongly connected to (i.e., that network's parcels) are outlined in yellow. The mean, across HCP subjects, functional connectivity between that cortical network and every voxel in the fronto-parietal network is shown from blue to red.

The voxels in each parcel, on average, all showed positive functional connectivity to the cortical network it was assigned to (Figure 2a). Moreover, for all 15 cortical networks,

the voxels in the cortical network's parcel in the fronto-parietal network were significantly more strongly connected to other voxels in that parcel than to voxels in other cortical networks' parcels in the fronto-parietal network (Figure 2b). The homogeneity of all parcels was positive (Figure 2c). We also found that, across individual HCP subjects, the functional connectivity between thalamic parcels and fronto-parietal parcels of the same cortical network was significantly stronger than the functional connectivity between thalamic parcels and fronto-parietal parcels of the different cortical networks (t=16.89,pvalue=9.6e-64, df=2100). We also analyzed individual cortical networks' parcels in the thalamus and fronto-parietal network. For all but four cortical networks, across individual HCP subjects, the cortical network's parcels, one in the fronto-parietal network and one in the thalamus, were more strongly functionally connected to each other than to other parcels in the thalamus and the fronto-parietal network (Figure 2d). These analyses demonstrate that all cortical networks have sub-regions of strong functional connectivity with the fronto-parietal network, as well as with the thalamus, and those two parcels are usually strongly functionally connected.

a



b

Network	Mean T	Mean P
Visual (Lateral)	54.3240920342	1.58426085456e-15
Visual (Medial)	31.4319580593	6.74159901377e-19
Motor Sensory (Dorsal)	70.0020341815	2.15368523371e-61
Motor Sensory (Ventral)	26.4709740267	5.58243642136e-15
Dorsal Attention (Posterior)	377.009989465	0.0
Dorsal Attention (Anterior)	500.437476841	0.0
Cingulo Opercular	114.755910292	5.5875208967e-218
Ventral Attention, Salience	323.54361653	1.93590409987e-36
Ventral Medial Temporal	59.4085012251	3.5358665227e-72
Orbital Frontal	35.2659444176	0.0035684107544
Precuneus	451.785261581	0.0
Auditory	223.185557415	0.0
Hippocampal Gyrus	212.867762247	0.0
Default Mode (Medial)	550.360222411	0.0
Default Mode (Lateral)	414.158124482	0.0

c

Network	Mean
Visual (Lateral)	0.0607831082226
Visual (Medial)	0.0468740717131
Motor Sensory (Dorsal)	0.0597160107384
Motor Sensory (Ventral)	0.0601897890917
Dorsal Attention (Posterior)	0.0467075372158
Dorsal Attention (Anterior)	0.0565952109247
Cingulo Opercular	0.037034110143
Ventral Attention, Salience	0.0313503980913
Ventral Medial Temporal	0.0419389369115
Orbital Frontal	0.0273084966189
Precuneus	0.0729903952736
Auditory	0.0313442221851
Hippocampal Gyrus	0.0431573986476
Default Mode (Medial)	0.0474325029652
Default Mode (Lateral)	0.041124040923

d

Network	T	p value
Visual (Lateral)	-1.84441484574	0.0651931169477
Visual (Medial)	2.34724218373	0.0189591668952
Motor Sensory (Dorsal)	5.93657486577	3.14549608161e-09
Motor Sensory (Ventral)	0.85772078621	0.391095623666
Dorsal Attention (Posterior)	5.63268303831	1.89033453764e-08
Dorsal Attention (Anterior)	7.95097811893	2.35958620135e-15
Cingulo Opercular	8.63851090927	7.97333829318e-18
Ventral Attention, Salience	6.30040117751	3.27412330097e-10
Ventral Medial Temporal	3.95062270241	7.92280693306e-05
Precuneus	8.63006991923	8.57327201988e-18
Auditory	1.42504872833	0.154217455612
Hippocampal Gyrus	1.74200797591	0.0815802864508
Default Mode (Medial)	4.07470394481	4.69241582351e-05
Default Mode (Lateral)	6.83201174855	9.5764362004e-12

Figure 2. a, Violin plots of the functional connectivity values, across HCP subjects, from each cortical network's parcel in the fronts-parietal network to that cortical network. b, For each network's parcel's voxels, a t-test was calculated between the pairwise functional connectivity values of all voxels in that cortical network's parcel versus the functional connectivity values between voxels in that parcel and voxels other cortical networks' parcels. The mean T and P, across HCP subjects, are shown. c, The mean pairwise functional connectivity values, across HCP subjects, of all voxels in each cortical network's parcel. d, For each network's fronto-parietal parcel and thalamus parcel, a t-test was calculated between the functional connectivity of those two parcels across HCP subjects, and the functional connectivity, across subjects, of those two parcels to parcels belonging to other networks.

Clustering of Functional Connectivity to the Fronto-Parietal Network

Utilizing the NKI Data, we detected “communities” of cortical connectivity to the fronto-parietal network, where all cortical networks in each community have similar connectivity to the fronto-parietal network. This analysis revealed three distinct communities and a graph with a Q value of a 0.398. Figure 3 displays these three communities, as well as the pair-wise similarity of all cortical networks’ connectivity to the fronto-parietal network (NKI data). The Q value of the graph built using the HCP data, but the community detection from the NKI data, was 0.221. We also visualized the community detection from the NKI data (i.e., which cortical networks are in the same community) with the mean functional connectivity, calculated with the HCP data, of the cortical networks in that community (Figure 3).

The first community is positively correlated with the posterior regions of the frontal and temporal aspects of the fronto-parietal network. The second community is negatively functionally connected with the entire fronto-parietal network. This community of cortical networks exhibited the highest levels of similarity in their functionally connected to the fronto-parietal network. Similar to the first community, the third community is positively functionally connected with the posterior regions of the fronto and temporal aspect of the fronto-parietal network. However, it is also positively functionally connected with the dorsal section of the parietal aspect of the fronto-parietal network, as well as the more anterior section of the frontal aspect of the fronto-parietal network. This analysis demonstrates that there are three distinct sets of cortical functional connectivity to the fronto-parietal cortex.

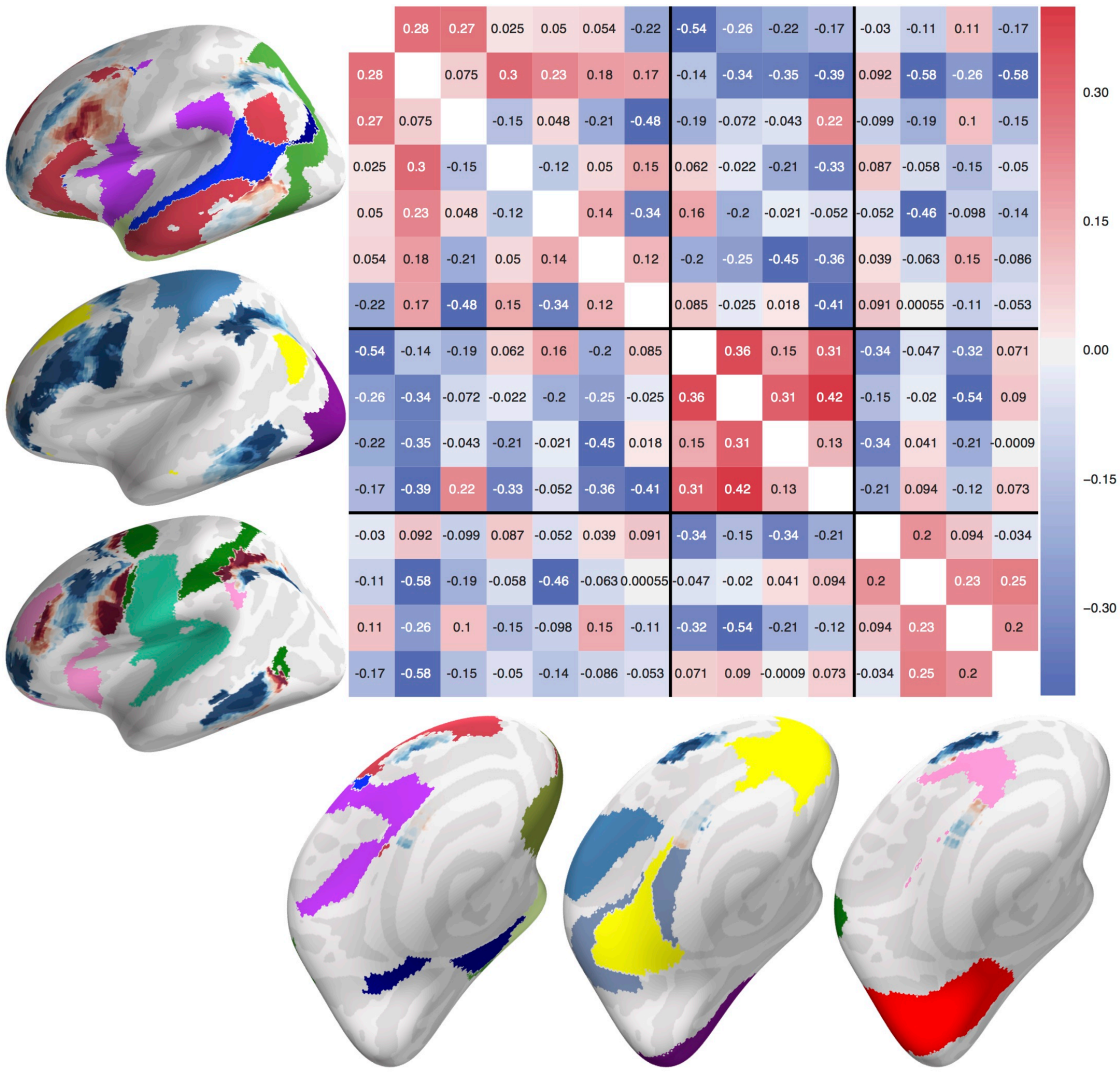


Figure 3. Community detection of cortical network connectivity to the fronts-parietal network. Each value in the matrix is the similar of the two cortical networks' functional connectivity to the fronts-parietal network. Three distinct communities, each on the cortical surfaces of each brain, were found. The mean, across HCP subjects, functional connectivity of each community is shown on the cortical surface.

Measures of Discrete and Convergent Connectivity in the Fronto-Parietal Network

We measured the discreteness of each fronto-parietal voxel's connectivity to cortical networks. We measured the discreteness of that voxel's connectivity by computing the difference in the strongest value and the second strongest value. We also measure the convergence of cortical network signals at each voxel in the fronto-parietal network by calculating 1 minus the standard deviation of the functional connectivity scores of that voxel to all cortical networks. The fronto-parietal shows apexes of high flexibility and discreteness in all 4 sections—the frontal, parietal, temporal, and medial (Figure 4). However, these two measures were negative correlated (r=-0.271, p=2e-315). These analyses demonstrate that different sub-regions of the fronto-parietal network exhibit

functional connectivity that can support either discrete or convergent information processing.

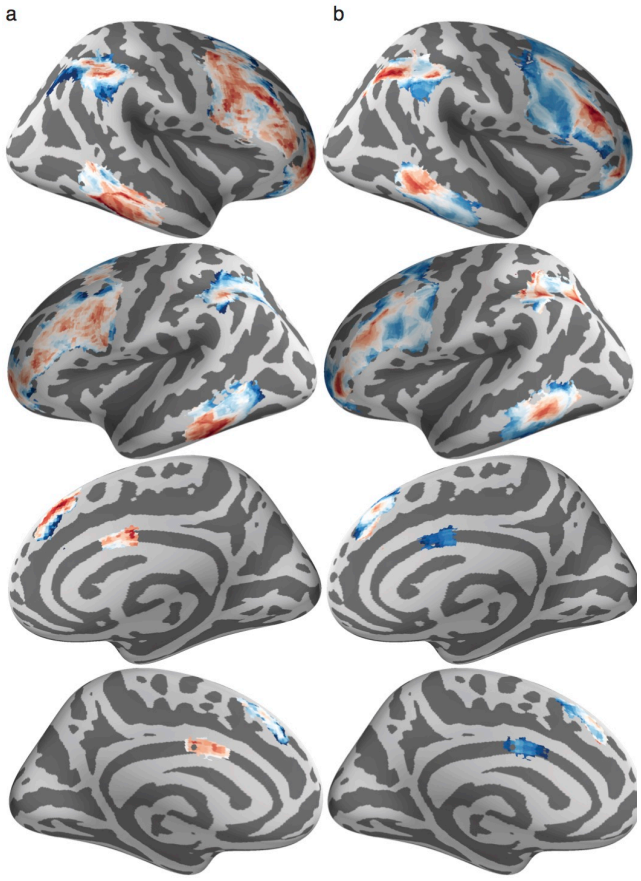


Figure 4. a, Convergent connectivity in the fronto-parietal network. b, Discrete connectivity in the fronto-parietal network.

Predictive Models of Cortical to Cortical Functional Connectivity

We predicted the connectivity between cortical networks using the functional connectivity between parcels in the fronto-parietal network and cortical networks (Figure 5a), the functional connectivity between parcels in the thalamus and cortical networks (Figure 5b), and the functional connectivity between parcels in the fronto-parietal network and parcels in the thalamus (Figure 5c).

Figure 5. a,b,c, Prediction accuracy of three models. d, Similarity of prediction accuracy across the three models.

Prediction accuracy was measured by the correlation between the predicted cortical to cortical functional connectivity value and the real cortical to cortical functional connectivity value. Values higher than r=0.282 are significant at p<0.05 after Bonferroni

correction for multiple comparisons (df=148, multiple comparisons=120). Significant correlations are shown from pink to purple with figure, with any correlation for which the p value did not pass multiple comparisons correction in beige. We also measured the similarity of the accuracies of the predictions of each model. All of the models were similar in which cortical network to cortical network functional connectivity they were able to accurately predict and which cortical network to cortical network functional connectivity they were not able to accurately predict (Figure 5d).

In every predictive model, the cortical networks for which prediction accuracy was highest were the attention networks (dorsal attention networks (anterior and posterior) and the ventral attention / salience network). Auditory, visual, and motor (ventral) network connectivity was also well predicted. The orbital frontal and ventral medial temporal networks' connectivity was poorly predicted, likely do to the low signal-to-noise and dropout in these regions. However, the Precuneus's connectivity was also not well predicted, despite not suffering from those issues. Ranking order for each model, including the mean across models, is shown in Figure 6.

	FP to Cortex	Thalamus to Cortex	Thalamus to FP	Mean
0	Dorsal Attention (Posterior)	Dorsal Attention (Posterior)	Dorsal Attention (Posterior)	Dorsal Attention (Posterior)
1	Ventral Attention, Salience	Dorsal Attention (Anterior)	Ventral Attention, Salience	Dorsal Attention (Anterior)
2	Cingulo Opercular	Auditory	Motor Sensory (Ventral)	Ventral Attention, Salience
3	Dorsal Attention (Anterior)	Ventral Attention, Salience	Dorsal Attention (Anterior)	Auditory
4	Auditory	Motor Sensory (Ventral)	Auditory	Motor Sensory (Ventral)
5	Visual (Medial)	Visual (Lateral)	Visual (Medial)	Visual (Medial)
6	Default Mode (Lateral)	Visual (Medial)	Visual (Lateral)	Visual (Lateral)
7	Visual (Lateral)	Hippocampal Gyrus	Fronto Parietal	Cingulo Opercular
8	Motor Sensory (Ventral)	Motor Sensory (Dorsal)	Cingulo Opercular	Default Mode (Lateral)
9	Default Mode (Medial)	Cingulo Opercular	Motor Sensory (Dorsal)	Hippocampal Gyrus
10	Orbital Frontal	Default Mode (Medial)	Default Mode (Lateral)	Default Mode (Medial)
11	Precuneus	Fronto Parietal	Ventral Medial Temporal	Fronto Parietal
12	Hippocampal Gyrus	Default Mode (Lateral)	Hippocampal Gyrus	Motor Sensory (Dorsal)
13	Fronto Parietal	Ventral Medial Temporal	Default Mode (Medial)	Ventral Medial Temporal
14	Ventral Medial Temporal	Orbital Frontal	Orbital Frontal	Orbital Frontal
15	Motor Sensory (Dorsal)	Precuneus	Precuneus	Precuneus

Figure 6. Rank order of the ability of the models to predict a network's connectivity to other cortical networks

Additionally, the fronto-parietal network's connectivity was not well predicted, despite one of the models actually being the fronto-parietal network's connectivity to other cortical networks. This is an example of how shared signals in the model will not necessarily increase prediction accuracy. The current predictive models depend on variance across subjects, not within subjects. For example, we calculated the matrix of the functional connectivity between all parcels in the fronto-parietal network. Within subjects, this matrix is highly correlated with the cortical to cortical functional connectivity matrix (r=.603, p=0.0). For example, in any single subject, connectivity between the auditory parcel and medial visual parcel and connectivity between the auditory network and medial visual network will be very similar. However, these values are not able to significantly predict any cortical to cortical functional connectivity, as variance across subjects in the parcel to parcel matrix is not predictive of variance in the

cortical to cortical matrix. In fact, the predictions of this model were negatively correlated with the real values (mean r=-0.45).

These analyses suggest that variance in the connectivity of the fronto-parietal network to other cortical networks, the variance in the connectivity of the thalamus to the fronto-parietal network, and the variance in the connectivity of the thalamus to cortical networks are all predictive of variance in cortical to cortical network functional connectivity. Moreover, given the similarity of their prediction accuracies, along with their functional connectivity to each other, these structures potentially interact to control cortical to cortical functional connectivity, or at least do so via a similar mechanism.

Conclusion

Our results suggest a very specific model of fronto-parietal-thalamic connectivity. The fronto-parietal network is typically characterized as an integrative network with diverse connectivity to the rest of cortex. However, similar to the thalamus(197), distinct information from any given cortical network projects to a distinct sub-region or set of sub-regions in the fronto-parietal network. These sub-regions are homogeneous—voxels in each parcel are, on average, positively functional connected, and are significantly more strongly functionally connected to each other than to voxels in other parcels, even though all voxels under consideration are in the same cortical network, an already homogenous set of voxels. Moreover, functional connectivity between cortex and the fronto-parietal network can be reduced to three distinct patterns of functional connectivity. Finally, the fronto-parietal network contains sub-regions that are only strongly functionally connected to a single network.

This aspect of the model is supported by anatomical connectivity. In the macaque, somatic, visual, and auditory regions project to different parts of parietal-temporal and frontal aspects of the fronto-parietal network (199). Even within a single spatially contiguous area in the fronto-parietal network, distinct connectivity patterns are found. For example, the most caudal region of the posterior parietal area (7a) is connected to anterior and posterior cingulate, the dorsal parahippocampus, V3-V5, and the superior temporal sulcus(195, 200-202). While the intraparietal sulcus (7ip) is also connected to those regions, area 7ip is connected to the more posterior regions of V3-V5 and the more ventral area (MT) of the superior temporal sulcus. Moreover, 7a and 7ip, but not the more rostral and lateral area (7b) and a medial region (7m), have strong limbic connections(195). Finally, 7b and 7m are biased toward inputs from somatosensory areas (S1 & SII) or visual areas (medial area 19), respectively. A distinction in posterior parietal cortex is also supported by electrophysiological data. Area 7b is related to more somatic stimuli, while 7a is more related to visual and oculomotor mechanisms(217-219).

However, we also found that the fronto-parietal network contains sub-regions that are diversely connected to many different cortical networks. Integrated information is likely processed in these convergent sub-regions of the fronto-parietal network and via non-

canonical reciprocal wiring to integrate across sub-regions with converging projections from many cortical networks and sub-regions that each have a distinct projection from a single cortical network. This aspect of the model is also supported by anatomical connectivity. For example, while somatic, visual, and auditory regions project to different sub-regions in the fronto-parietal network, these modalities also have converging projections in the superior temporal sulcus (which is part of the fronto-parietal network; Fig 1a), and sub-regions of convergence in the superior temporal sulcus are interconnected with each other(199).

The combination of discrete and convergent connectivity in the fronto-parietal network allows information flow to be bi-directional. Information from all cortical networks can be transferred to the fronto-parietal network via distinct pathways that exists between cortical networks and the fronto-parietal network. Information is then likely integrated in the convergent regions of the fronto-parietal network, and can then be sent back to specific cortical networks via the distinct pathways that exist between the fronto-parietal network and specific cortical networks.

The thalamus is likely a critical aspect of this information flow, potentially mediating connectivity between the fronto-parietal network and other networks, aiding in information integration and signal propagation across cortex. The fronto-parietal network is distinctively connected to the thalamus. For all but four cortical networks, the cortical network's parcels, one in the fronto-parietal network and one in the thalamus, were more strongly functionally connected to each other than to other parcels in the thalamus and the fronto-parietal network. Moroever, functional connectivity between the fronto-parietal network to cortex, the thalamus to cortex, and the fronto-parietal network to the thalamus were all significantly and similarly predictive of cortical to cortical functional connectivity.

More speculatively, it has been suggested that thalamocortical projections provide synchronized signals that coordinate the development of cortical networks, and these signals aid in organizing connectivity across the massively expanded human cortex(220). The fronto-parietal-thalamic connectivity found here would support this idea—the thalamus is connected to the fronto-parietal network such that each sub-region in the fronto-parietal network that is connected to a specific cortical network is also connected to the area of thalamus that is dedicated to that network. This positions the thalamus as being connected both to where a given cortical network is represented in the fronto-parietal network and to the network itself, allowing it to control how each pair of networks (fronto-parietal and other networks) become connected.

This connectivity pattern also generates the idea that an evolutionary older integrative structure—the thalamus—might have been "copied" (i.e., it performs similar functions and has similar connectivity patterns) and expanded in the cortex as the fronto-parietal network, and these two integrative structures now function together. Thus, the fronto-parietal network might have been naturally selected because it allows for greater control over the connectivity between cortical networks both during development and ongoing cognition, allowing for more complex computations to occur, utilizing both canonical and

non-canonical wiring and strong connectivity to the thalamus. It is then no surprise that this network is often referred to as a “control” network and its activity is related to “top down” regulation of other networks. It is also not surprising that similar global network changes (decreased modularity) occur when regions of the thalamus and the fronto-parietal network that are functionally connected to many networks are damaged (152, 197).

The human thalamus is an integrative hub for functional brain networks

Kai Hwang, Maxwell Bertolero, William Liu, Mark D'Esposito

Helen Wills Neuroscience Institute and Department of Psychology, University of California Berkeley, CA, USA

Address correspondence to:
Kai Hwang Ph.D.
132 Barker Hall MC 3190
University of California Berkeley
California, CA 94720
USA
kai.hwang@berkeley.edu

Abstract

The thalamus is globally connected with distributed cortical regions, yet the functional significance of this extensive thalamocortical connectivity remains largely unknown. By performing graph-theoretic analyses on thalamocortical functional connectivity data collected from human participants, we found that most thalamic subdivisions display network properties capable of integrating multimodal information across diverse cortical functional networks. From a meta-analysis of a large dataset of functional brain imaging experiments, we further found that the thalamus is involved in multiple cognitive functions. Finally, we found that focal thalamic lesions in humans have widespread distal effects, disrupting the modular organization of cortical functional networks. This converging evidence suggests that the human thalamus is a critical hub region that could integrate diverse information being processed throughout the cerebral cortex, as well as maintain the modular structure of cortical functional networks.

Significance

The thalamus is traditionally viewed as a passive relay station of information from sensory organs or subcortical structures to the cortex. However, the thalamus has extensive connections with the entire cerebral cortex, which can also serve to integrate information processing between cortical regions. In this study, we demonstrate that multiple thalamic subdivisions displays network properties capable of integrating information across multiple functional brain networks. Moreover, the thalamus is engaged by tasks requiring multiple cognitive functions. These findings support the idea that the thalamus is involved in integrating information across cortical networks.

Introduction

The mammalian brain can be conceptualized as a thalamocortical system. Every cortical region receives projections from the thalamus, and in turn sends outputs to one or multiple thalamic nuclei(221). Thalamocortical projections relay nearly all incoming information to the cortex, as well as mediate cortico-cortical communication(222). Thus, full insight into brain function requires knowledge of the organization and properties of thalamocortical interactions.

The thalamus can be divided into two types of nuclei: first order and higher order thalamic nuclei(223). First order thalamic nuclei, such as the lateral geniculate nucleus (LGN) and the ventral posterior (VP) nuclei, receive inputs from ascending sensory pathways or other subcortical brain regions. In contrast, higher-order thalamic nuclei, such as the mediodorsal (MD) and the pulvinar nuclei, receive inputs predominately from the cortex. Higher order thalamic nuclei have both reciprocal and non-reciprocal connections with multiple cortical regions(211, 224, 225), suggesting that in addition to relaying external sensory information *to* the cortex though first order nuclei, another principle function of the thalamocortical system is to mediate cortico-cortical information transfer *within* the cortex(226).

Graph-theoretic network analysis of resting-state functional MRI (rs-fMRI) data is well suited for exploring the organization and functional properties of the

thalamocortical system. Previous functional connectivity analyses of rs-MRI data have demonstrated that the human cerebral cortex can be decomposed into several modular functional networks(56, 141), each of which is potentially involved in executing a discrete set of cognitive functions(140, 227). Graph-theoretic measures can be further used to characterize each region's topological properties(40). For instance, a brain region with many within-network connections has a strong "provincial hub" property(79), presumably to promote within-network interactions for executing specialized functions of the network. In contrast, a brain region with many between-network connections has a strong "connector hub" property, presumably to mediate interactions between functional networks. Cortical connector and provincial hubs have distinct contributions to cortical networks' modular organization(152), and cortical connector hubs have been found to be involved in multiple cognitive tasks(20, 125, 140, 176).

The topographic properties of the thalamocortical system are largely unknown. The thalamus and its constituent nuclei have global and extensive connectivity with multiple brain structures, and are likely major hub regions for functional brain networks(74, 228). Thalamic nuclei are traditionally hypothesized to function as modality-specific relays. This view of thalamic function would predict that thalamic nuclei should exhibit strong provincial hub properties to support information communication within relatively encapsulated processing channels(229). However, it is not known if all thalamic nuclei serve as provincial hubs, or if some nuclei, such as the higher order nuclei, further support between-network interactions, serving as connector hubs. It has been shown that a single thalamic nucleus could have connections with multiple anatomically segregated cortical regions(211, 224). If these cortical regions innervated by a single thalamic nucleus belong to the same functional network, then this thalamic nucleus should exhibit strong provincial hub properties for supporting modality-selective processes (Figure 1A). Alternatively, if these cortical regions belong to different functional networks, then this thalamic nucleus is capable of integrative processes mediating interactions across multiple cortical networks, and should exhibit strong connector hub properties (Figure 1B). These hypotheses are not mutually exclusive—the thalamus could contain nuclei that act as "global kinless" hubs(79, 81), not only strongly associated with one but multiple functional networks, involved in both modality-selective and multimodal integrative processes.

The goal of this study was to elucidate the network topological roles of thalamic nuclei in functional brain networks. To measure network properties of thalamocortical functional connectivity, we performed graph theoretic network analyses on rs-fMRI data collected from healthy human participants. To relate network topology to cognitive functions, we analyzed task-related activity using a meta-analysis of 10,449 functional neuroimaging experiments from the BrainMap database(176, 230). Finally, we examined the contribution of thalamic nuclei to cortical network organization by analyzing rs-fMRI data from human patients with focal thalamic lesions.

Methods

Datasets

For the main analyses, we analyzed publically available resting-state fMRI (rs-fMRI) data from 303 subjects (mean age = 21.28, SD = 2.64, age range =19-27, 128 males) that were acquired as part of the Brain Genomics Superstruct dataset(231). For each subject, two runs (6.2 minutes each) of rs-fMRI data were collected using a gradient-echo echo-planar imaging sequence with the following parameters: relaxation time (TR) = 3000 ms, echo time (TE) = 30 ms, flip angle = 85 degrees, 3 mm^3 isotropic voxels with 47 axial slices. Structural data were acquired using a multi-echo T1-weighted magnetization-prepared gradient-echo (MPRAGE) sequence (TR = 2,200 ms, TE= 1.54 ms for image 1 to 7.01 ms for image 4, flip angle = 7 degree, 1.2 mm^3 isotropic voxel). We replicated our main analyses with another publically available rs-fMRI dataset from 62 healthy adults (mean age = 23.96, SD = 5.24, age range =18-37, 27 males) that were acquired as part of the NKI-Rockland sample(232). For each subject, 9 minutes and 35 seconds of rs-fMRI data were acquired using a multiband gradient-echo echo-planar imaging sequence with the following parameters: TR = 1400 ms, echo time = 30 ms, multiband factor = 4, flip angle = 65 degrees, 2 mm^3 isotropic voxels with 64 axial slices. Structural data were acquired using a MPRAGE (TR = 1900 ms, TE= 2.51 ms, flip angle = 9 degree, 1 mm^3 isotropic voxel). For both datasets, subjects were instructed to stay awake and keep their eyes open.

For the lesion analyses, we analyzed rs-fMRI data from four patients with focal thalamic lesions (ages of the four patients: S1 = 81 years, S2 = 50 years, S3 = 55 years, S4 = 84 years, all males, all were scanned at least 6 months after their stroke). Patient S1 has bilateral lesions, and all other patients have unilateral lesions. Two runs of rs-fMRI data were collected (10 minutes each; TR = 2000 ms, echo time = 30 ms, flip angle = 72 degrees, 3.5 mm^2 in plane resolution with 34 axial 4.2 mm slices). Structural images were acquired using a MPRAGE sequence (TR = 2,300 msec, TE = 2.98 msec, flip angle = 9°, 1 mm^3 voxels). Patients were instructed to stay awake and keep their eyes open. Informed consent was obtained from all patients in accordance with procedures approved by the Committees for Protection of Human Subjects at the University of California, Berkeley.

Functional MRI Data Preprocessing

Image preprocessing was performed with the Configurable Pipeline for the Analysis of Connectomics software(214). First, brain images were segmented into white matter (WM), gray matter, and cerebral spinal fluid (CSF). Rigid body motion correction was then performed to align each volume to a temporally averaged volume, and a boundary-based registration algorithm was used to register the EPI volumes to the anatomical image. Advanced Normalization Tools (ANTS) was used to register the images to MNI152 template using a nonlinear normalization procedure(233). All images were spatially resampled to a 2 mm voxel resolution. We then performed nuisance regression to further reduce non-neural noise and artifacts. To reduce motion-related artifacts, we used the Friston-24 regressors model during nuisance regression(234). WM and CSF signals were regressed using the anatomical CompCor approach with five components for each tissue class(235). Linear and quadratic drifts were also removed. Because the physical proximity between the thalamus and the ventricles could result in

blurring of fMRI signal, we further regressed out the mean signal from CSF, WM, and gray matter that were within 5 voxels (10 mm) from the thalamus. After regression, data were bandpass filtered from 0.009–0.08 Hz, and scaled to a whole-brain mean value of 10000. Given that the thalamus is a relatively small structure, to avoid signal blurring we did not perform any spatial smoothing.

Identifying Cortical Functional Networks

Following preprocessing, mean rs-fMRI time-series were extracted from 333 cortical regions of interests(91), and concatenated across runs for subjects with multiple rs-fMRI scans. Cortico-cortical functional connectivity was assessed in each subject by computing Pearson correlations between all pairs of cortical ROIs, resulting in a 333 x 333 correlation matrix. This correlation matrix was then thresholded to retain the strongest functional connections. To identify the modular structure of cortico-cortical functional connectivity, we performed an InfoMap algorithm(129) to partition the correlation matrix into putative modular functional networks. InfoMap is one of the best performing sub-network detection algorithms available, and has been successfully used to identify cortical network organization(56, 140). Specifically, for each subject, we first identified functional networks using the InfoMap algorithm at a threshold of density = 0.15 (keeping 15 percent of connections out of all possible cortico-cortical connections). Based on this partition result, we then constructed a consensus matrix that consisted of values of 1 where the two ROIs are in the same module, and values of 0 elsewhere(236). We then re-thresholded the correlation matrix by decreasing the density in steps of 0.001, and ran the InfoMap algorithm again to obtain a new partition. The consensus matrix was then updated for the new partition, except for rows and columns for which the ROI had no connections that survived the new threshold, or the ROI was in a fragmented module of less than 5 ROIs. The consensus matrix was continued to be updated to the minimal threshold of density = 0.01. Thus, this subject specific consensus matrix represents the modular network assignment for each pair of nodes at their sparsest level possible (i.e., before they become disconnected from the graph). We then aggregated individual subjects' module organization by averaging consensus matrices across subjects. This averaged consensus matrix was then submitted to the same recursive method described above to identify group-level cortical functional networks. Following methods reported in previously published studies(56, 91), networks with 5 or fewer ROIs were eliminated from further analyses, and ROIs that were not clustered into networks were excluded from further analyses.

Thalamus Parcellation

Previous studies of functional brain networks using graph theory often excluded subcortical structures, or examined the thalamus with gross or no subdivisions. Given the thalamus's complex structure with multiple distinct nuclei, the thalamus is likely not uniformly interacting with the cortex. Instead, different thalamic subdivisions have distinct structural connectivity with the cortex, and thus functional connectivity with the cortex(237-240). We therefore examined thalamocortical network

properties of using three different thalamic parcellations. To localize the thalamus, the Morel Atlas(241) was used to define its spatial extent (2227 2 mm^3 voxels included in the atlas, registered to the MNI152 template). To identify thalamic subdivisions, three different thalamic atlases were utilized. We first performed a custom winner take all functional parcellation using rs-fMRI data. We calculated partial correlations between the mean BOLD signal of each cortical functional network and the signal in each thalamic voxel, while removing signal variance from other functional networks. Partial correlations were then averaged across subjects, and each thalamic voxel was labeled according to the cortical network with the highest partial correlation. The Morel atlas identified thalamic nuclei based on cyto- and myelo-architecture in stained slices of post-mortem tissue collected from five postmortem brains(242), and subsequently transformed to MNI space(241). The Oxford-FSL thalamic structural connectivity atlas defined thalamic subdivisions based on its structural connectivity with different cortical regions estimated from diffusion imaging data(239). Each of these atlases is sensitive to a specific type of anatomical or functional information, therefore using all three atlases allows us to derive a more complete characterization of thalamocortical network properties.

Thalamic and Cortical Nodal Properties

To formally quantify the network properties of thalamocortical functional connectivity, for each subject, we first extracted signal from the thalamus by either using each voxel's preprocessed signal or averaged voxel-wise BOLD signal within each thalamic subdivision. We then calculated the partial correlation between the mean BOLD signal of each cortical region of interest (ROI) and thalamic voxel or subdivision. Partial correlation was calculated by removing signal variance from all other cortical ROIs. Given the large number of cortical ROIs, a dimension reduction procedure using principal component analysis was performed on signals from cortical ROIs not included in the partial correlation calculation, and eigenvectors that explained 95% of variance were entered as additional nuisance regressors in the model. Note that no correlations were calculated between thalamic voxels. This resulted in two thalamocortical connectivity matrices: voxel-wise and subdivision-wise. For voxel-wise matrices, 2227 thalamic voxels were included. For the functional parcellation, Oxford-FSL, and Morel atlases, 9, 7, and 15 thalamic subdivisions were included, respectively. Note that each thalamic voxel and thalamic subdivision all had the same number of functional connections with cortical ROIs (total 333). Thalamocortical connectivity matrices were then averaged across subjects. For cortical ROIs, we calculated Pearson correlations between all pair-wise cortical ROIs to obtain cortico-cortical connectivity matrices (matrix size 333 x 333). Matrices were then averaged across subjects. Following recommendations from a previous study(56), we explored network properties across a range of network density thresholds (density = 0.01 to 0.15), and submitted these to graph analyses. Note that, because we did not perform global signal regression, no negative correlations entered our graph analyses.

For estimating thalamocortical functional connectivity, we chose partial correlations over full correlations because past studies have shown that detailed

thalamocortical connectivity patterns could be obscured without accounting for shared variance between cortical regions(243). We compared graph matrices of thalamic voxels calculated using partial or full correlations, and found that results were independent of whether full or partial correlations were used. We also calculated graph metrics at the level of each thalamic subdivision by averaging voxel-wise signals within each thalamic subdivision before graph analyses, and found consistent results when compared to voxel-wise metrics.

For each thalamic voxel, subdivision, and cortical ROI, we calculated the participation coefficient (PC), a measure of the connector hub property, and the within module degree z-score (WMD), a measure of the provincial hub property(79). To ensure that our results were not biased by a single specific threshold, all graph metrics were calculated across a range of thresholds (density = 0.1 to 0.15). We report results that were averaged across thresholds.

To calculate WMD, correlation matrices were first binarized by setting weights above the density threshold to 1. WMD values were calculated across a range of density thresholds (edge density of 0.1 to 0.15) and averaged across thresholds. Weights were binarized to equate the connectivity weights between thalamocortical and cortico-cortical networks. WMD is calculated as $WMD = \frac{K_i - \overline{CW_s}}{\sigma CW_s}$, where $\overline{CW_s}$ is the average number of connections between all cortical ROIs within cortical network *s,* and σCW_s is the standard deviation of the number of connections of all ROIs in network *s.* K_i is the number of connections between *i* and all cortical ROIs in network *s*. Because our goal was to understand the thalamus's contribution to cortical network organization, WMD scores of thalamic subdivisions were calculated using the mean and standard deviation of within-network degree (number of intra-network connections) calculated from each cortical functional network. For comparison, we also calculated WMD using fully weighted matrices without thresholding, and found similar results. Note that because WMD is normalized by the number of regions within the associated functional network, unlike degree-based centrality measures, it is not influenced by the size of functional networks. Given that we assigned each thalamic voxel to a cortical functional network in the functional parcellation atlas, higher WMD values reflect more within-network connections of the thalamic voxel with the network it was assigned to.

The PC value of region *i* is defined as: $PC = 1 - \sum_{s=1}^{N_M} \left(\frac{K_{is}}{K_i}\right)^2$, where $K_i$is the sum of the connectivity weight of *i* , $K_{is}$is the sum of the connectivity weight between *i* and the cortical network *s,* and N_M is the total number of networks. If a region has connections uniformly distributed to all cortical networks, then its PC value will be close to 1; otherwise, if its connectivity is concentrated within a specific cortical network, its PC value will be close to 0. Given that we identified 9 cortical functional networks, the maximum possible PC value would be 0.89. We therefore further divided PC values by this theoretical upper limit, so that the highest possible PC value given the network architecture would be 1. For comparison, we also calculated PC using binary networks by setting weights above the threshold to 1 and found similar results. Because PC is independent of the node's degree, it is less biased by the number of ROIs within each functional network, resulting in a superior measure of hub properties(78). It is also

important to note that a PC value is independent of the functional network a ROI belongs to; therefore our thalamic PC calculations are not influenced by which cortical network thalamic voxels or subdivisions are assigned to.

Patients with focal thalamic lesions

Lesion masks were manually traced in the native space according to visible damage on a T1-weighted anatomical scan, and further guided by hyperintensities on a T2-weighted FLAIR image. Lesion masks were then warped into the MNI space using the same non-linear registration parameters calculated during preprocessing. For comparing the effects of local thalamic lesions, rs-fMRI volumes with framewise displacement (FD) that exceeded 0.5 mm were removed from further analysis (scrubbed) before band-pass filtering. More volumes were scrubbed for patients when compared to healthy controls (mean percentage of frames scrubbed for patients = 11.04%, SD = 10.63%; mean percentage of frames scrubbed for healthy controls = 1.02%, SD =1.62%; no subjects excluded).

Modularity

Modularity can be measured by Newman's modularity Q (244), defined as: $Q = \sum_{i=1}^{m}(e_{ii} - a_i^2)$, where e_{ii} is the fraction of the connectivity weight connecting ROIs within a cortical functional network *i*, a_i is the fraction of the connectivity weight connecting ROIs in cortical functional network *i* to other cortical networks, and *m* is the total number of cortical functional networks. Network partition from Figure 2A was used. Modularity was calculated for the whole-brain (including all cortical ROIs). Note that no thalamic subdivision was included into this analysis. For lesions analyses, the mean and standard deviation of Q calculated from all healthy controls were used to normalize Q values from each patient(152): $Z = \frac{Q-\bar{Q}}{\delta}$, where Q is the patient's modularity score, $\bar{Q}$ is the average modularity score of healthy controls, and δ is the standard deviation in modularity of healthy controls. Modularity calculations were performed separately for each density threshold (density = 0.01 to 0.15) and results are presented across thresholds.

Within-Network and Between-Network Connectivity Strength

We further analyzed how focal thalamic lesions could affect connectivity between and within cortical functional networks. For each patient, we first mapped the cortical functional networks their thalamic lesions were connected to according to the functional parcellation atlas from healthy subjects (see Thalamic Parcellation section above). For each patient and healthy control subject, we further normalized each functional connection by subtracting the mean and divided by the standard deviation of all connections within each subject. For the affected cortical networks, we then summed its total connectivity strength with other cortical networks (between-network) and with its own respective networks (within-network). The mean and standard deviation of connectivity strength for healthy controls were then used to normalize the patient's values to z-scores.

Meta-Analysis of Functional Neuroimaging Experiments in the BrainMap Database

We reanalyzed data presented in a previously published meta-analysis of the BrainMap database(176). In the meta-analysis, a hierarchical Bayesian model was used to derive a set of 12 cognitive components that best describe the relationship between behavioral tasks and patterns of brain activity in the BrainMap database(230). Specifically, each behavioral task (e.g., Stroop, stop-signal task, finger tapping) engages multiple cognitive components, and in turn each cognitive component is supported by a distributed set of brain regions. To determine whether or not a thalamic voxel is recruited by a cognitive component, a threshold of $p = 10^{-5}$ was used. This is an arbitrary yet stringent threshold that was used in two prior studies(140, 176). Critically, there is potential spatial overlap between components. Therefore, brain regions that can flexibly participate in multiple cognitive components could be identified by calculating the number of cognitive components each brain region engages. The number of cognitive components was summed for each voxel and cortical ROIs, and defined as a "cognitive flexibility" score(176).

Results

Identification of Cortical Networks

To identify cortical functional networks, we first measured functional connectivity matrices between 333 cortical ROIs(91), then performed a network partition analysis to estimate cortical network organization (see Methods). We found that the cerebral cortex can be decomposed into 9 functional networks (Figure 2A), an organization scheme largely similar to previous studies(91, 140, 141, 194).

Parcellation of the Thalamus

Given that the thalamus can be subdivided using different approaches, we performed our analyses using three different atlases based on data from rs-fMRI, diffusion tensor imaging (DTI), and postmortem histology (Figure 2 A-C; see Methods for details). Using RS-fcMRI data, we identified thalamic subdivisions that demonstrated the strongest functional connectivity with the different cortical functional networks reported above (Figure 2A; henceforth referred to as the functional parcellation atlas). We further replicated these results with an independent dataset, and found high correspondence between datasets (Figure 2A; normalized mutual information = 0.69, z-scored Rand coefficient =144.13, $p < 10^{-5}$). The Oxford-FSL thalamocortical structural connectivity atlas (Figure 2B) subdivides the thalamus based on structural connectivity (estimated using probabilistic diffusion tractography on DTI data) to seven large cortical areas: primary motor, primary somatosensory, occipital, premotor, prefrontal (including medial and orbitofrontal cortices), parietal, and temporal cortices(239). The Morel atlas (Figure 2C) subdivides the thalamus into smaller nuclei based on cyto- and myelo-architecture information from five postmortem brains(241, 242). We further classified each thalamic nucleus from the Morel atlas into first order or higher order thalamic

nuclei(222, 226). Our parcellation results and numerous replicable human MRI-based parcellations(237-240, 243, 245-248) suggest that distinct signals from different thalamic subdivisions can be reliably detected with a conventional 2 to 4 mm voxel resolution.

Network Properties of Thalamocortical Functional Connectivity

To determine each thalamic subdivision's network property, we estimated functional connectivity between each thalamic voxel and every cortical ROI (see Methods) to generate a thalamic voxel to cortical ROI thalamocortical graph. Graph metrics were calculated for every thalamic voxel, and pooled across voxels for different categories of thalamic subdivisions (i.e., first order and higher order nuclei, or subdivisions within the three different thalamic atlases). Our goal was to examine the thalamus's network topological properties in the context of functional brain networks. For example, if a thalamic subdivision is found to have strong connector hub properties, how does it compare to cortical ROIs that are connector hubs for cortical functional network? Therefore, we further calculated graph metrics for each cortical ROI by analyzing patterns of cortico-cortical functional connectivity.

Within-Module-Degree Analyses. Provincial hub properties can be measured by the graph theory metric called within module degree (WMD), which measures the number of within-network connections of a region, z-scored by the mean and standard deviation of within-network connections of all regions in that network(79). Higher values reflect more within-network connections. We found that thalamic voxels in both first order and higher order thalamic nuclei exhibited higher WMD values than most cortical ROIs (Figure 3A). A two-sample Kolmogorov-Smirnov test showed that on average thalamic voxels had significantly higher WMD values when compared to cortical ROIs (mean WMD for thalamic voxels = 1.62, SD = 1.9; mean WMD for cortical ROIs = 0, SD = -.89; $D_{333,2227} = 0.45$, $p < 10^{-10}$). We further compared thalamic voxels' WMD values to cortical provincial hubs. Because both cortical and thalamic WMD values exhibited a unimodal distribution, we arbitrary defined cortical provincial hubs as cortical ROIs with WMD values greater than 90% of all cortical ROIs (threshold: WMD = 1.04). We found that on average thalamic voxels exhibited WMD values that were comparable to cortical provincial hubs (mean WMD for cortical provincial hubs = 1.35, SD = 0.26; mean WMD for voxels within first order thalamic nuclei = 1.37, SD = 1.79, mean WMD for voxels within higher order thalamic nuclei mean WMD = 1.75, SD = 1.94).

Participation Coefficient Analyses. Connector hub properties can be measured by the graph theory metric called participation coefficient (PC), which is a measure of the strength of inter-network connectivity for each region normalized by their expected value(79). Higher values reflect more inter-network connections. For each thalamic voxel, we calculated its PC value based on its inter-network connectivity pattern to all cortical ROIs. We found that thalamic voxels in both first order and higher order thalamic nuclei exhibited higher PC values than most cortical ROIs (Figure 3C). A two-sample Kolmogorov-Smirnov test showed that thalamic voxels had significantly higher PC values compared to cortical ROIs (mean PC for thalamic voxels = 0.76, SD = 0.12; mean PC for cortical ROIs = 0.36, SD = 0.22; $D_{333,2227} = 0.65$, $p < 10^{-10}$). We further

compared thalamic voxels' PC values to cortical connector hubs. Because PC values also exhibited a unimodal distribution, we arbitrary defined cortical connector hubs as cortical ROIs with PC values greater than 90% of all cortical ROIs (threshold: PC = 0.64). We found that on average voxels in both first order and higher order thalamic nuclei exhibited high PC values that were comparable to cortical connector hubs (mean PC for cortical connector hubs = 0.72, SD = 0.05; mean PC for voxels in first order thalamic nuclei = 0.74, SD = 0.13, mean PC for voxels in higher order thalamic nuclei = 0.77, SD = 0.11).

Spatial Distribution of PC and WMD Values. We found that cortical ROIs in the precuneus, medial frontal, inferior parietal, insular and middle frontal cortices exhibited high WMD values, whereas ROIs in anterior and inferior frontal, superior precentral sulcus, intraparietal sulcus, and lateral occipital cortices exhibited high PC values (Figure 4A). High PC and WMD values were found throughout the thalamus (Figure 4B-C). To determine differences in the spatial distribution of connector and provincial hub properties in the thalamus, we identified thalamic voxels that exhibited WMD or PC values greater than cortical connector and provincial hubs (cortical provincial hub threshold WMD = 1.04; cortical connector hub threshold: PC = 0.64; hubs arbitrary defined as cortical ROIs with PC or WMD values greater than 90% of all cortical ROIs). We found that anterior, medial, posterior, and dorsal parts of the thalamus exhibited both high PC and WMD values. In addition, portions of the lateral thalamus also exhibited only high PC values, and a small portion of the ventral thalamus only exhibited strong high WMD values (Figure 4D).

PC and WMD Values of Each Thalamic Subdivision. We calculated the median WMD and PC values across voxels for each thalamic subdivision, and compared those values to cortical connector and provincial hubs (see definition above). Based on the functional parcellation atlas, thalamic subdivisions that showed dominant functional coupling with CO, DM, FP, mT, and sFP networks exhibited high WMD values numerically comparable to cortical provincial hubs (Fig 5A). Based on the Oxford-FSL thalamocortical structural connectivity atlas, thalamic subdivisions with dominant structural connectivity with the prefrontal cortex and temporal cortices showed high WMD values comparable to cortical provincial hubs (Figure 5B). Based on the Morel histology atlas, thalamic subdivisions with high WMD values comparable to cortical provincial hubs included the anterior nucleus (AN), LGN, the ventral lateral nucleus (VL), intralaminar nuclei (IL), lateral posterior nucleus (LP), MD, medial pulvinar (PuM), and ventral anterior nucleus (VA) (Figure 5C). We found that all thalamic subdivisions exhibited high PC values comparable or higher than cortical connector hubs (Figure 5D-F).

Connectivity Patterns of Specific Thalamic Nuclei. Based on the Morel histology atlas, we found that AN, LGN, VL, IL, LP, MD, PuM, and VA exhibited both high PC and WMD values comparable to cortical connector and provincial hubs. To further probe their connectivity patterns, for each nucleus we calculated its mean functional connectivity strength with each of the 9 cortical functional networks (using partial correlations, see Methods), and divided by the nucleus's summated total connectivity strength with all networks. If a nucleus is diffusely interacting with all functional

networks, then it should devote ~11% (1/9 = 0.11) of its total connectivity for each network. In contrast, if a nucleus only interacts with a selective network, the majority of its connectivity strength should be devoted to that network, while connectivity with other networks should be considerably lower. Consistent with high PC values we observed in these nuclei, we found that each of these thalamic nuclei exhibited a diffuse functional connectivity pattern, with strong connectivity (>11% of its total connectivity strength) with multiple cortical functional networks (Figure 6). Specifically, we found that multiple nuclei showed strong connectivity with both the CO and DM networks (AN, MD, VA, IL, LP, and VL), and most of these nuclei had strong connectivity with at least one other cortical functional network (e.g., AN, VA, LP, VL). Altogether, every nucleus had strong functional connectivity with at least three or more cortical functional networks, and many cortical functional networks had strong functional connectivity with multiple thalamic nuclei.

Meta-Analysis of the BrainMap Database

We found that multiple thalamic subdivisions exhibited both strong WMD and PC values, suggesting that the thalamus is capable of mediating information communication both within and between multiple functional brain networks. Given that individual cortical functional networks are putatively associated with distinct cognitive functions(140, 227), it is likely that any individual thalamic nucleus with distributive connectivity with multiple cortical functional networks is involved in multiple cognitive functions. We tested this hypothesis by analyzing results from a published meta-analysis of 10,449 functional neuroimaging studies(176). This published meta-analysis derived latent variables—an ontology of cognitive functions or "cognitive components"—that best described the relationship between 83 behavioral tasks and corresponding brain activity. From this data, a "cognitive flexibility score" can be estimated by summing the number of cognitive components that are engaged by every brain region(176). Here, a brain region with high cognitive flexibility score is assumed to be involved in more cognitive functions. We have previously demonstrated that cortical connector hubs exhibit high cognitive flexibility scores(140). We hypothesize that thalamic subdivisions that are recruited by multiple cognitive components likely serve as integrative connector hubs, whereas thalamic subdivisions that are recruited by a limited number of specific functions are likely involved in domain general processes.

Consistent with our hypothesis, both first order and higher order thalamic nuclei were found to be involved in multiple cognitive components (Figure 7A). A two-sample Kolmogorov-Smirnov test showed that thalamic voxels have higher cognitive flexibility scores when compared to cortical ROIs (mean for thalamic voxels = 3.67, SD = 1.71; mean for cortical ROIs = 0.36, SD = 0.22; $D_{333,2227} = 0.54$, $p < 10^{-10}$). We further examined the specific cognitive components (C1-C12, see Yeo et al., 2015 for details) that recruited each of the thalamic subdivisions. As an example, VL, with projections to motor and premotor cortices(229), is recruited by components C1 and C2 that predominately recruit motor cortices (Figure 7B). However, VL also participates in other cognitive components that recruit lateral prefrontal, medial prefrontal, and parietal cortices (C8, C9, C12, Figure 7B). Note that most of these cognitive components not

only recruit VL, but also engage other parts of the thalamus (Figure 7B). This observation is consistent with results we presented in Figure 6, showing that multiple thalamic subdivisions could be associated with the same cognitive component.

Thalamic lesions have Global and Distal Effects on Cortical Network Organization

Modularity is a metric that quantifies the extent to which the brain is differentiated into separable sub-networks, and is an essential property found in many complex systems(249). Previous studies suggest that connector and provincial hubs have distinct contributions to modular organization. For example, a lesion study showed that damage to connector hubs, but not provincial hubs, causes more severe disruption of the network's modular organization(152), suggesting that focal lesions to connector hubs can have a widespread impact on network organization when between-network connections are disrupted. Based on these findings, we predict that if thalamic subdivisions serve as connector hubs for functional brain networks, lesions to those subdivisions should reduce modular organization of these networks. Modularity can be measured by Newman's modularity Q(244), a comparison between the number of connections within a module to the number of connections between modules. In four patients with focal thalamic lesions (Figure 8A; one patient has bilateral lesions (S1), three have unilateral lesions (S2-S4)), we examined the effect of a thalamic lesion on cortical modularity across the whole cerebral cortex (Figure 8B). Each patient's Q score was converted to a z-score using the mean and standard deviation of healthy controls (see Methods). In all four patients, whole-brain modularity was lower (as indicated by negative z-scores).

Reduction in modularity could be caused by increased between-network connectivity or decreased within-network connectivity. Thus, we further quantified how focal thalamic lesions affect between- and within-network connectivity strength. For each patient, we first mapped the lesioned thalamic voxels, and then we identified cortical networks that are connected to these lesioned voxels using data from healthy control subjects (Figure 2A; Figure 8A right panel). For each lesion's connected networks, we calculated the change of between- and within-connectivity strength relative to healthy controls (Figure 8C; see Methods for the z-scoring procedure). In all four patients with thalamic lesions, between-network connectivity strength was higher than healthy controls (as indicated by positive z-scores); in three patients, within-network connectivity strength was lower than healthy controls (as indicated by negative z-scores). Overall, increased between-network connectivity strength (mean z-score = 2.79, SD = 1.73) was more prominent than decreased within-network connectivity (mean z-score = -0.71, SD = 1.10), suggesting that focal thalamic lesions increase between-network connectivity between cortical regions.

Secondary Analyses

Replication. We replicated the WMD and PC analyses using an independent rs-fMRI dataset with higher native voxel resolution (2 mm^3); the spatial correlation values across both cortical ROIs (333 ROIs) and thalamic voxels (2227 voxels) between the test and replication datasets for PC and WMD scores were 0.74 and 0.78, respectively. We also replicated our results using a different cortical ROI definition template that consists of 320 cortical ROIs(92), and the thalamic voxel-wise spatial correlation values for PC and WMD scores were 0.63 and 0.78, respectively.

InfoMap versus Spectral Modularity Detection. To ensure that our lesion analysis was not specific to calculating Q with modular partition derived using the InfoMap algorithm, we recalculated Q values for each healthy control subject using the spectral modularity detection algorithm which maximizes Q to derive modular partition [40], and found it to be highly correlated with and slightly lower than Q values calculated using the InfoMap algorithm (at density = 0.05, r = 0.86; mean Q for InfoMap algorithm = 0.57, SD = 0.06; mean Q for spectral algorithm = 0.53, SD = 0.06).

Discussion

In this study, we provide evidence suggesting that the human thalamus is a critical integrative hub for functional brain networks. First, we found that all thalamic subdivisions exhibited strong between-network connectivity, indicating that a single thalamic subdivision not only connects with multiple cortical regions, but also with multiple cortical functional networks. From a meta-analysis of 10,449 neuroimaging experiments, we further found that the thalamus is engaged by multiple cognitive functions, supporting a role in multimodal information processing. Finally, we found that focal thalamic lesions cause a disruption of the modular structure of cortical functional networks, further underscoring the critical contribution of thalamic function to brain network organization.

The human brain is composed of modular functional networks(74, 140) which comprise provincial hubs — brain regions important for within network communication — and connector hubs — brain regions important for communication between networks. Here, we used graph-theoretic measures to estimate provincial and connector hub properties of the thalamus. Consistent with traditional interpretations of the thalamus relays information to cortex, we found that multiple thalamic subdivisions exhibited strong provincial hub properties (high WMD values) supporting modality-selective processes. However, both first order and higher order thalamic nuclei also exhibited prominent connector hub properties (high PC values), an indication of their involvement in integrative, multimodal processes (Figures 3-7). For example, we found that both first order (e.g., LGN in Figure 6) and higher order (e.g., PuM in Figure 6) nuclei showed strong functional connectivity with more than three cortical functional networks. We also found many thalamic nuclei showed strong connectivity with the CO and DM networks, a finding that is consistent with previous studies that have classified the thalamus as part of the CO, saliency, or DM networks (56, 250, 251). However, these nuclei also exhibited strong functional connectivity with at least one additional cortical network. Based on these results, both first order and higher order thalamic nuclei should be more

accurately described as “global kinless” hubs(79, 81), with strong connectivity with not only one functional network but homogenously interacting with multiple networks. Thalamic nuclei’s global hub property could potentially allow the thalamus to send and access information across diverse cortical functional networks. Via convergence of information, the thalamus may serve as a global hub that subserves multiple cognitive functions. Although it has previously been proposed that the thalamus is not simply a relay station, and that higher order thalamic nuclei could further serve to mediate cortical to cortical communication within its anatomical projection zone(226), the notion that the both first order and higher order nuclei also play an integrative role interacting with multiple cortical functional networks has not been demonstrated in humans. In a rodent fMRI study, using a measure that was analogous to PC, it was demonstrated that the rodent thalamus has the most diverse connectivity pattern with multiple functional sub-networks, more diverse than the hippocampus, basal forebrain, DM and lateral cortical networks(82).

Higher order thalamic nuclei, which receive inputs predominately from the cortex, are hypothesized to provide trans-thalamic routes to support cortico-cortical interactions within a functional network that receive its projections(226, 252). For example, the posterior nucleus transfers information from primary to secondary somatosensory areas(253). Likewise, the pulvinar has extensive reciprocal connections with striate and extrastriate visual cortices(254), and is thought to modulate information communication between visual areas(252, 255). In contrast, first order thalamic nuclei, which receive projections from peripheral sensory organs or other subcortical structures, have projections to primary cortices, and are thought to act as modality-selective relays to relay a limited type of afferent signal to the cortex. Our graph-theoretic analyses of thalamocortical functional connectivity provide evidence suggesting that both first order and higher order thalamic nuclei participate in information exchange with multiple cortical functional networks.

How can an individual thalamic nucleus interact with multiple functional cortical networks? One possibility is that thalamic nuclei could have dense reciprocal connections with cortical connector hubs that in turn are connected with multiple cortical functional networks. For example, from our analyses of the Oxford-FSL and functional atlas, we found that thalamic subdivisions that were most strongly connected with the frontal cortex, the temporal cortex and fronto-parietal functional networks (e.g., FP, sFP, CO, DF), showed kinless global hub properties (prominent connector plus provincial hub properties). Previous studies have found that cortical connectors are primarily located in fronto, parietal, and temporal cortices(78, 140). Our findings suggest that, in the context of functional brain networks, many thalamic subdivisions could share similar information processing roles with association heteromodal regions. In addition, both first order and higher order nuclei are known receive non-reciprocal cortico-thalamic innervations from multiple cortical regions(256). Higher order thalamic nuclei are also known to project to more than one area of the cerebral cortex(221). For example, MD projects to the lateral frontal cortex, the insula, the anterior cingulate cortex, temporal cortex, the supplementary motor area(224, 257), and these regions likely span multiple different cortical networks (e.g. FP, DF, CO, and SM networks). Also, thalamic nuclei not only

comprise “core” thalamocortical cells that project to middle layers of the cerebral cortex that are constrained by the borders of a functional area, but also comprise “matrix” thalamocortical cells that diffusely project to superficial layers of the cerebral cortex and are unconstrained by the boundaries of cortical topographic representations(221). The inhibitory thalamic reticular nucleus also receives projections from multiple cortical and subcortical regions, and further projects to multiple thalamic nuclei(222). Thus, the reticular nucleus could receive information from different cortical networks to further modulate activity in both first order and higher order thalamic nuclei. To summarize, individual thalamic nucleus receive projections and project to multiple cortical regions that belong to different cortical functional networks. Thus, the anatomical and functional architecture of the thalamus is capable of simultaneously receiving and transmitting signals between multiple brain regions that belong to different cortical functional networks. Our results further suggest a man-to-many relationship of the thalamocortical system. Specifically, a single thalamic subdivision could be functionally connected with multiple cortical functional networks, and a single cortical functional network could in turn be functionally connected with multiple thalamic subdivisions.

Consistent with its extensive connectivity with multiple cortical functional networks, we found that the thalamus is one of the most “cognitively flexible” brain regions, indicating that the thalamus is involved in a diverse range of behavioral tasks. This observation derived from our meta-analysis of the BrainMap database is further supported by several representative empirical studies demonstrating that the thalamus mediates interactions between high-level cognitive processes (e.g., attention and working memory) and more elementary sensorimotor functions(252, 258, 259). For example, a monkey electrophysiology study found that deactivating the pulvinar reduced the attentional effects on sensory-driven evoked responses recorded in V4(258). Also, optogenetically perturbing thalamic activity in rodents impaired animals’ ability to select between conflicting visual and auditory stimuli(260). Finally, VL lesions in humans impair their ability to utilize a memorized cue during working memory necessary for guiding a visual search of multiple visual stimuli(259). Together, results from our graph analyses of thalamocortical functional connectivity and meta-analysis of task-related thalamic activity patterns suggest that the thalamus participates in interactions between multiple functional cortical networks, networks that are putatively involved in distinct cognitive functions. Based on our empirical results, future task-based studies can test this hypothesis regarding the role of the thalamus in integrative functions.

Previous studies suggest that connector hubs are critical for maintaining the modular architecture of functional brain networks(152). Mathematically, whole brain modularity is inversely related to between-network connectivity, therefore a loss of connector hub should increase modularity. However, our previous work with patients with focal cortical lesions found that damages to cortical connector hubs decreased modularity(152). Furthermore, a TMS study further found that disruption of connector hub function increased between-network connectivity(261). This suggests that, in addition to mediating inter-network interaction, connector hubs could be further involved in maintaining the modular structure of functional networks through decreasing

between-network connectivity. Consistent with these prior studies, we found that thalamic lesions reduce cortical functional networks' modularity and increase between-network connectivity strength. This suggests that thalamic connector hub damage weakens the modular organization of cortical networks by increasing the strength of between-network connectivity. Moreover, since the effect of a thalamic lesion is not constrained only to cortical regions it directly projects to, it can be considered an example of "connectomal diaschisis"(262).

One possible interpretation of our findings from patients with thalamic lesions is that through its extensive between-network connectivity with multiple cortical regions, the thalamus acts as a "gate" to inhibit cortico-coritcal connectivity while maintaining the modular structure of cortical networks. A weakening of "gating" function following a thalamic lesion would disinhibit cortico-cortical connectivity, causing an increase in between-network connectivity and a decrease in modularity. Alternatively, if functional brain networks must maintain a stable between-network connectivity strength for optimal between-network interactions, a compensatory mechanism may lead to increased direct cortico-cortical functional connectivity strength in response to decreased cortico-thalamo-cortical connectivity. To pinpoint the underlying mechanism would require techniques that can simultaneous manipulate and record activities of the thalamocortical system.

Acknowledgements

This work was supported by NIH Grants RO1 NS79698 and F32 NS090757. The Morel atlas was obtained by a written consent with Professor G. Szekely from the Computer Vision Laboratory of ETH Zurich, Switzerland. This project was made possible by a collaborative agreement allowing comprehensive access to the BrainMap database, a copyrighted electronic compilation owned by the University of Texas Board of Regents. BrainMap is supported by National Institutes of Health, National Institute of Mental Health Award R01 MH074457.

Figures

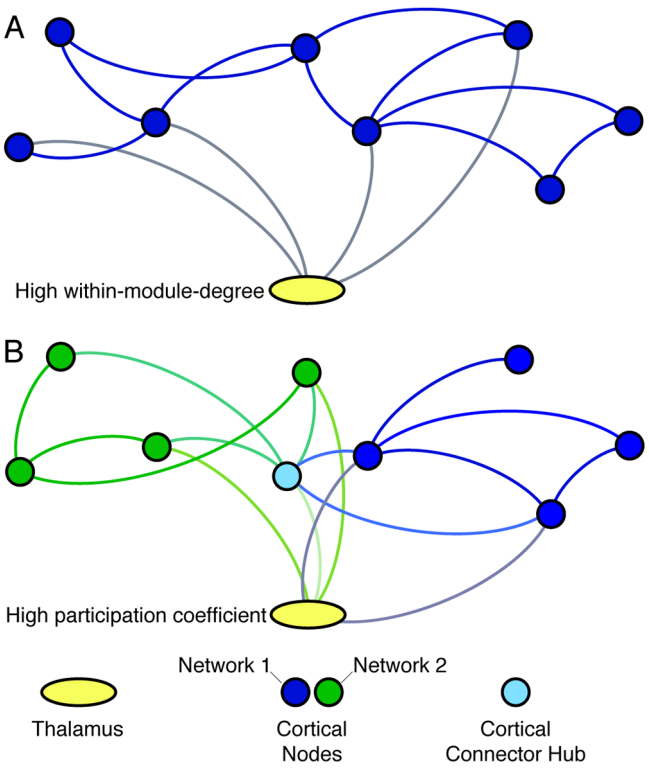


Figure 1. Thalamus mediating cortico-cortical communication for functional brain networks. (A) As a provincial hub, the thalamus is connected with cortical regions that belong to the same cortical functional network (represented in solid blue circles). A provincial hub will have high within-module-degree z-score (WMD). (B) As a connector hub, the thalamus is connected with cortical regions in multiple cortical functional networks (one network colored in blue and the other in green). A connector hub will exhibit high participation coefficient (PC).

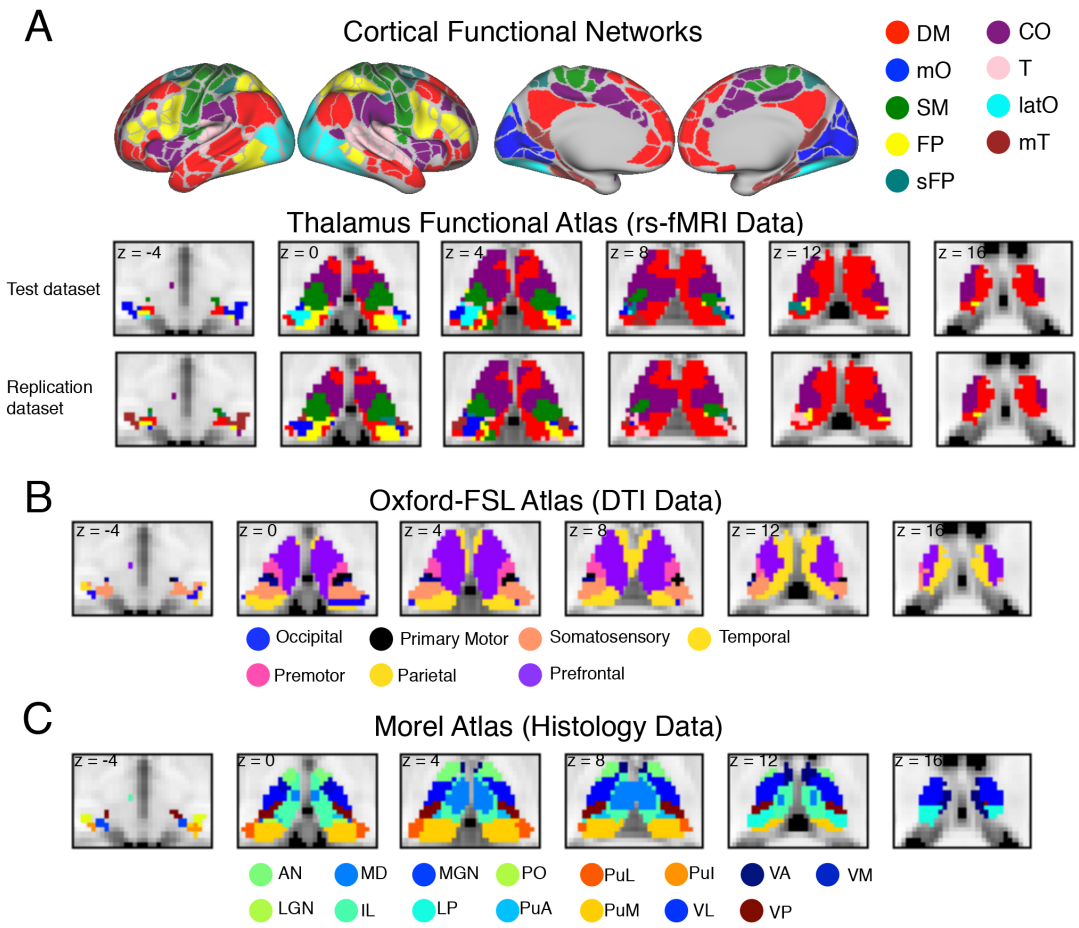


Figure 2. Cortical functional networks and thalamic atlases. (A) Cortical functional networks and thalamic parcellation derived from functional connectivity analyses between the thalamus and each cortical network using rs-fMRI data. Network abbreviations (based on its most predominant anatomical location): default mode (DM), medial occipital (mO), somato-motor (SM), fronto-parietal (FP), superior fronto-parietal (sFP), cingulo-opercular (CO), temporal (T), lateral occipital (latO), medial temporal (mT). (B) Structural connectivity based segmentation of the thalamus using the Oxford-FSL atlas. Each thalamic subdivision was labeled based on the cortical region it is most structurally connected with. (C) Histology based thalamic parcellation using the Morel atlas. Abbreviations for thalamic nuclei: anterior nucleus (AN), intralaminar (IL), lateral posterior (LP), lateral geniculate nucleus (LGN), medial geniculate nucleus (MGN), medial dorsal (MD), medial pulvinar (PuM), inferior pulvinar (PuI), lateral pulvinar (PuL), anterior pulvinar (PuA), posterior (Po) nuclei, ventral posterior (VP), ventral anterior (VA), ventral medial (VM), Ventral lateral (VL).

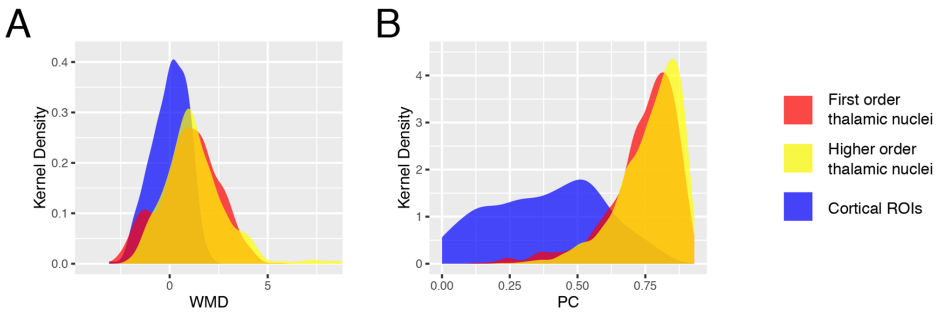


Figure 3. Nodal properties of the thalamus and cortical ROIs. (A) Kernel density plot of WMD values for thalamic voxels and cortical ROIs. Thalamic voxels were categorized into two categories of thalamic nuclei [2,7]. (B) Kernel density plot of cortical WMD values summarized by cortical functional networks. WMD values for each thalamic functional subdivision is represented by the black vertical bar. (C) Kernel density plot of PC values for thalamic voxels and cortical ROIs. Thalamic voxels were categorized into two categories of thalamic nuclei. (D) Kernel density plot of PC values summarized by cortical functional networks. PC values for each thalamic functional subdivision is represented by the black vertical bar. First order thalamic nuclei included AN, LGN, MGN, VL, and VP. Higher order thalamic nuclei included IL, MD, LP, Po, Pulvinar, VA, and VM. Box plot percentiles. All graph metrics averaged across network densities from 0.01 to 0.15.

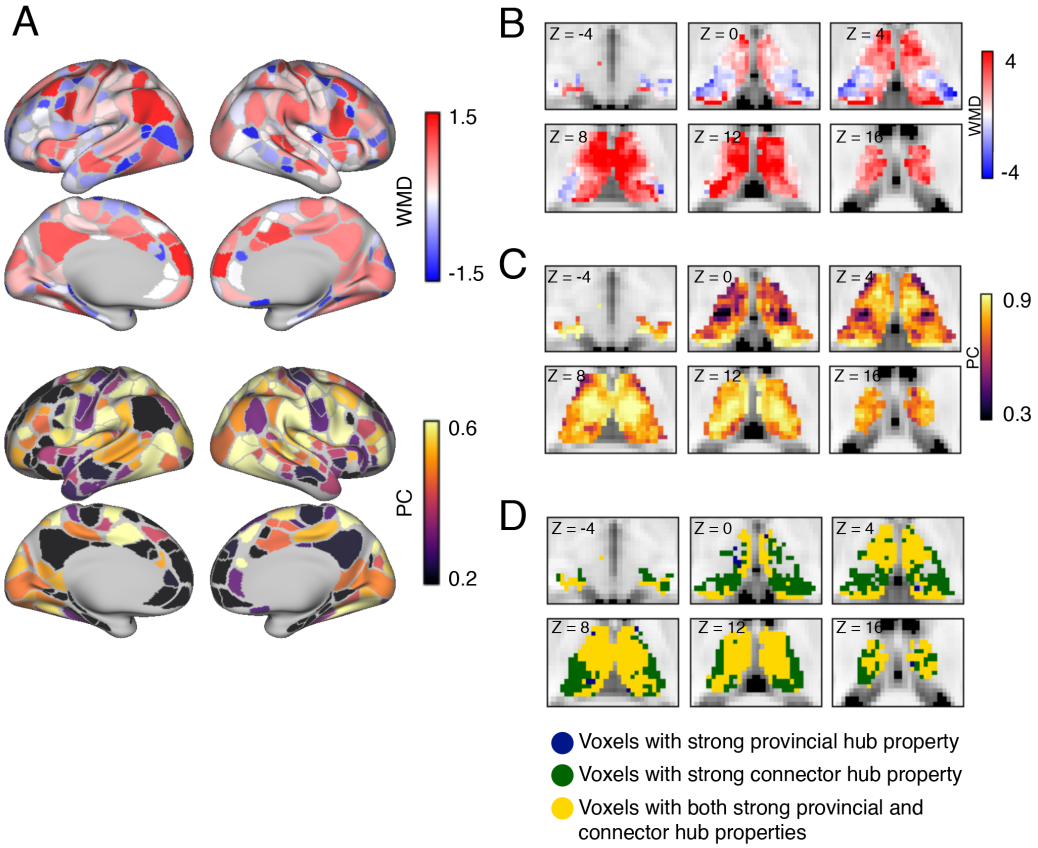


Figure 4. Spatial distribution of network metrics. (A) WMD and PC values of cortical ROIs. (B) WMD values of thalamic voxels. (C) PC values of thalamic voxels. (D) Location of voxels with strong connector (colored in dark navy blue), provincial (colored in dark green), or connector plus provincial hub properties (colored in gold) in the thalamus. In D, Only thalamic voxels that exhibited PC and/or WMD values greater than 90% of all cortical ROIs are displayed. All graph metrics averaged across network densities from 0.01 to 0.15.

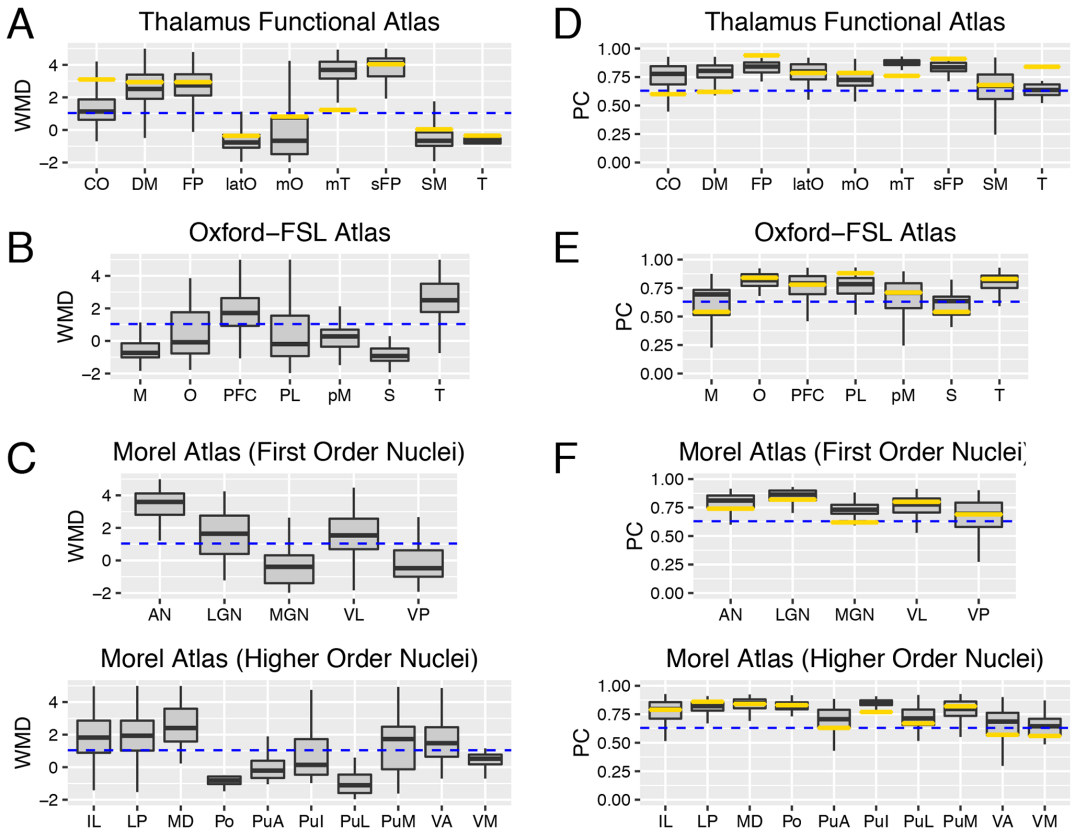


Figure 5. Graph metrics of each thalamic subdivision. (A-C) Box plots summarizing WMD values (A-C) and PC values (D-F) for all thalamic voxels for each thalamic subdivision in each thalamic atlas. The horizontal blue dashed line represents WMD or PC values of cortical provincial or connector hubs (arbitrary defined as greater than 90% of all cortical ROIs). The horizontal gold bar in each individual box plot represents the graph metrics calculated on the level of thalamic subdivision. Abbreviations for the Oxford-FSL atlas: motor (M), occipital (O), prefrontal (PFC), parietal (PL), premotor (pM), somatosensory (S), temporal (T). Box plot percentiles (5^{th} and 95^{th} for outer whiskers, 25^{th} and 75^{th} for box edges) calculated across voxels for each thalamic subdivision. All graph metrics averaged across network densities (see methods).

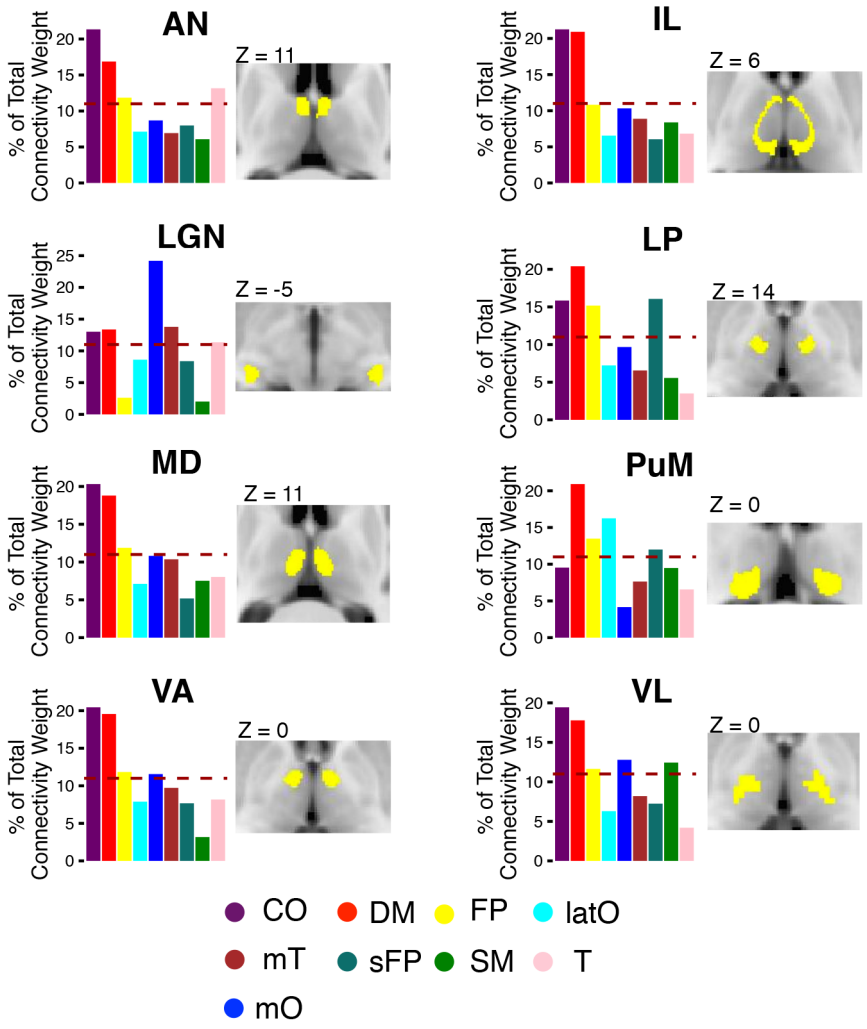


Figure 6. Distributive pattern of thalamocortical connectivity for thalamic nuclei. Cortical functional networks most strongly connected with the following thalamic nuclei: AN, LGN, VL, VA, VM, IL, LP, MD, PuM. Thalamic nuclei (labeled in yellow) are displayed on axial MRI images. The bar graphs represent the distribution of connectivity strength between thalamic nuclei and each of the 9 cortical functional networks. The dashed line represents the expected proportion of total connectivity if connections were equally distributed across networks.

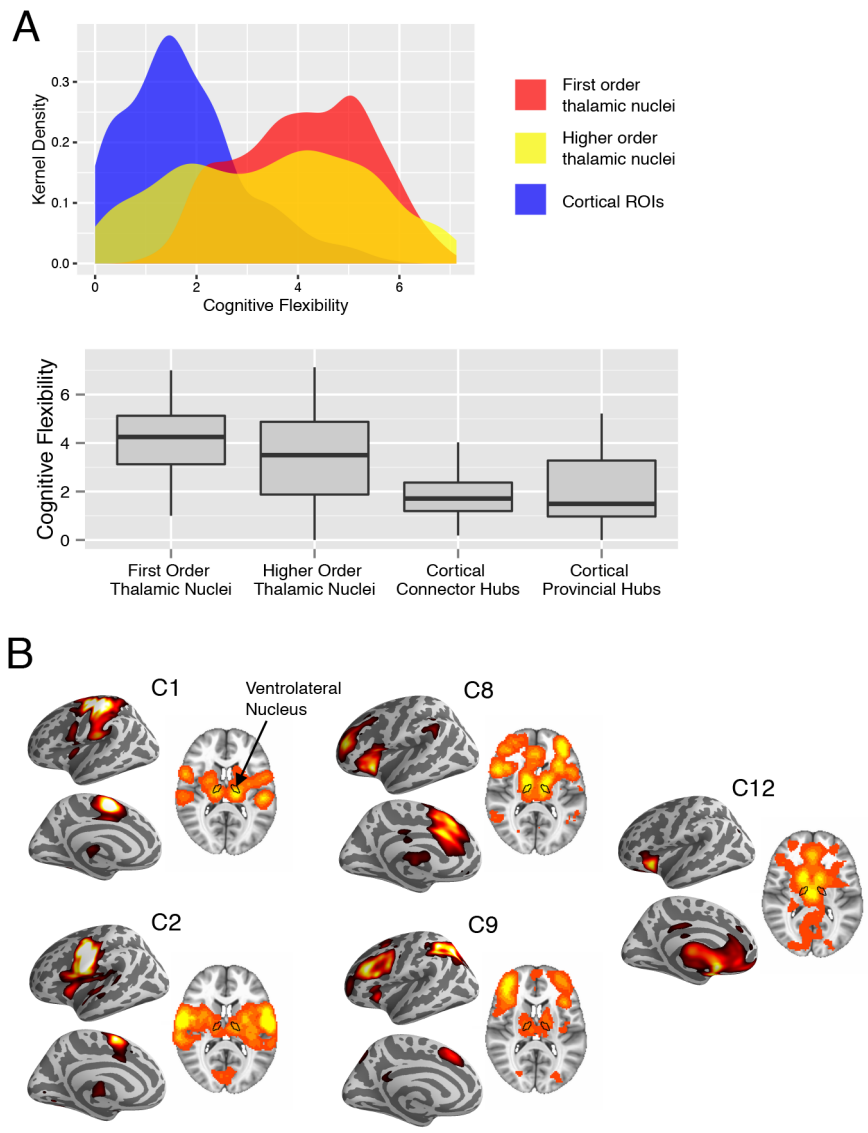


Figure 7. Cognitive flexibility score and cognitive components. (A) kernel density plot and box-plots of cognitive flexibility scores for thalamic voxels and cortical ROIs. Thalamic voxels were categorized into two categories of thalamic nuclei. Box plot percentiles (5^{th} and 95^{th} for outer whiskers, 25^{th} and 75^{th} for box edges) calculated across voxels for each thalamic nuclei type or across cortical ROIs for each cortical ROI type. (B) Spatial distribution of brain activity engaged by each cognitive components recruited by the thalamic nucleus VL.

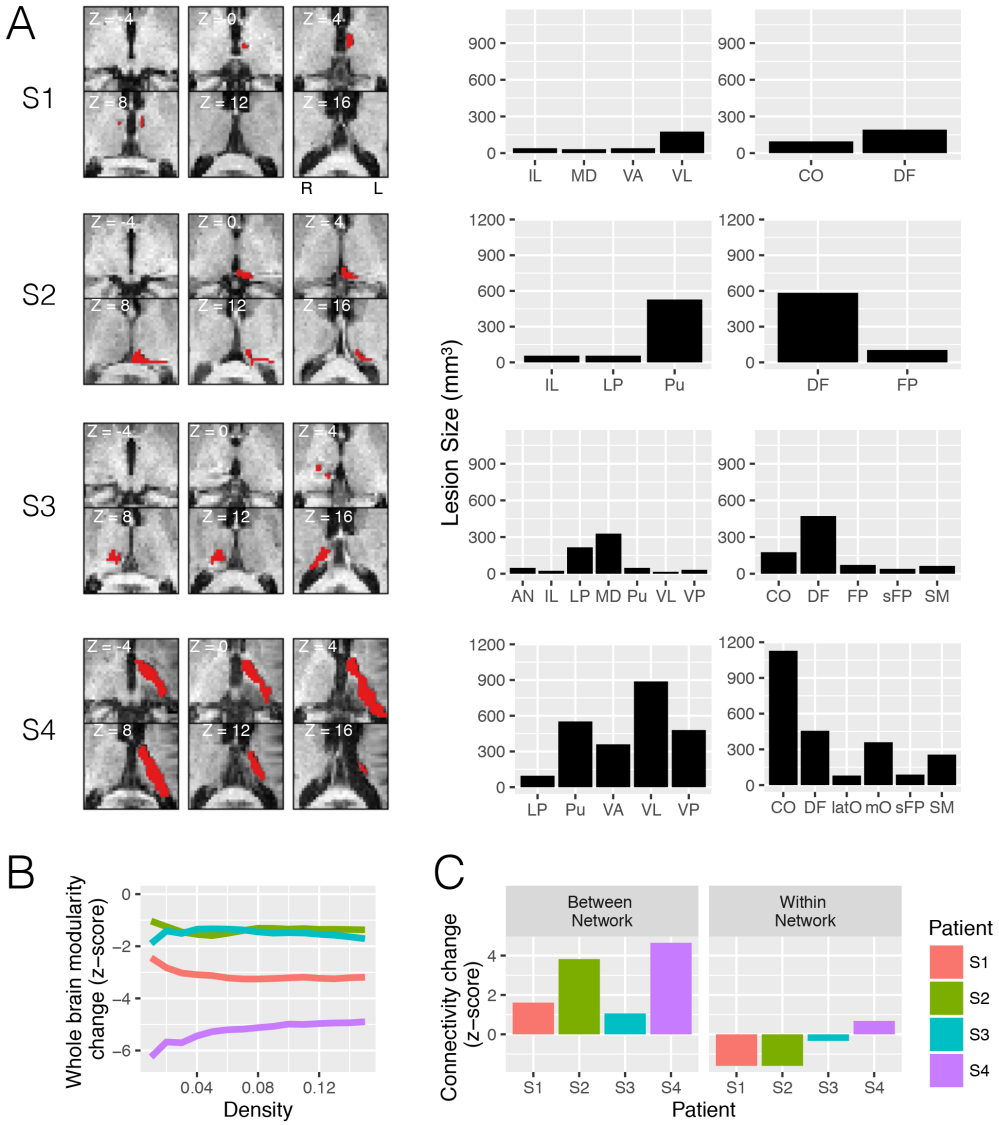


Figure 8. Focal thalamic lesion's effect on cortical network organization. (A). MRI scans of thalamic lesions (marked in red) in four patients (S1-S4). Lesion size for each patient in each thalamic subdivision (left panel based on Morel atlas, right panel based on the functional parcellation atlas) is summarized in bar graphs. (B) Individual patient's normalized modularity score across network densities (x-axis) for the whole cerebral cortex. (C) Individual patient's normalized connectivity strength for between- and within-network connectivity.

Nodeless Networks
M.A. Bertolero & Mark D'Esposito

Abstract

The human brain can be fruitfully investigated by modeling it as a complex network. To build a complex network model of the brain, brain regions are treated as nodes in the network, and connections between brain regions are treated as edges (typically, the correlation of the time series of fMRI activity of the two brain regions, or "functional connectivity"). Thus, the first step in building a complex network model of the brain involves choosing how to parcellate the brain into regions that are treated as the network's nodes. While many methods have been developed to parcellate the brain into nodes, these methods typically attempt to find a single solution that performs well for a large group of subjects. However, boundaries and connectivity patterns vary across subjects. Recently, parcellation methods have been developed for individual subjects. Here, we develop an algorithm that allows for unique parcellations to be built for each subject as well as a high resolution group average network model that utilizes those unique parcellations. We also employ a new community detection method that allows for a precise choice of the number of communities without having to threshold the graph to low densities or have a global resolution parameter. Using these methods, we analyze the communities and integrative brain regions in complex networks of different sizes (100, 200, 300, 600, 900, and 1500 nodes). Compared to more traditional methods, this method captures communities and integrative brain regions that are more similar to a large meta-analysis of task-evoked activity. These methods are well suited for group average complex network analyses, particularly the comparison of two groups (e.g., clinical populations versus healthy controls).

Introduction

One of the most powerful mathematical tools for investigating the functional architecture of the human brain is the network science approach. Here, the brain is described as a graph, which is a set of nodes (i.e., brain regions) and edges (i.e., functional or structural connections) between them. A fundamental issue with this approach is defining how brain regions are represented as nodes. In other words, how do we parcellate the brain into regions that we treat as nodes in the graph? An accurate parcellation is a necessary condition for building an accurate graph theory model of the brain.

Once the brain is parcellated into regions, the time series of fMRI activity from all regions are correlated with each other, forming a matrix that serves as the basis of the graph. The time series of activity is typical measured during the resting-state (i.e., the subject is not doing anything in particular), but has recently been measured during task performance (20, 21, 67, 75, 140, 141, 150, 263-265). The strength of correlation between each pair of regions determines the strength of the edge between the two

nodes that represent those regions. Thus, parcelleting the brain is the first step in converting fMRI data into a network model; any further analyses of the graph will be dependent on the parcellation. However, studies have found consistencies in graph metrics across parcellations(140); regardless, care must be taken to ensure that a particular result is not only present with a particular parcellation.

A wide variety of methods can be utilized to parcellate the brain into nodes. Many graph theory analyses utilize anatomy; for example, the Automated Anatomical Labeling parcellation(266) and the Harvard Oxford parcellation(267) both have nodes defined based on anatomy. Theses parcellations have 90 and 96 of nodes, respectively. Other analyses treat each voxel in the fMRI scan (volumetric-pixels) as a node (56, 268), or group together a small number of voxels arbitrarily(269). These “parcellations” can have upwards of 20,000 nodes. More recent analyses make use of “functional” parcellations(56, 90-92, 270), which, while they vary significantly in implementation, can all be judged on one metric—how similar are the voxels’ time courses of activity within each node? A node should be a homogenous unit of activity. We refer to this as how well a parcellation fits the data. Ideally, all voxels in the node have the same time course of activity. Given this metric, it also becomes clear that not only how one parcellates the brain matters, but also how many nodes are in the brain. A random algorithm for grouping voxels into homogenous units can perform better than a principled algorithm, so long as the random algorithm generates a larger number of nodes(92). At the extreme, a parcellation that defines each voxel as a node fits the data the best, as no reduction in the data has occurred. Moreover, voxels provide the finest-grained resolution possible.

However, voxels are arbitrary units of the brain and nearby voxels share nonbiological signal (causing increased correlation values) due to unavoidable steps in data processing (e.g., reslicing, blurring). Short-distance relationships are especially susceptible to spurious augmentation by subject motion(56, 96, 271, 272). The signal-to-noise of individual voxels is lower than the mean signal of multiple voxels, which is why spatial smoothing is traditionally applied to fMRI images. Moreover, as a practical issue, many graph theory analyses become computationally intractable at the voxel level. For example, community detection via modularity maximization is NP-hard (188). Moreover, it has been shown that group-averaging does not always accurately capture the graph topology of individual subjects(95, 273). Thus, while it is feasible to run community detection on a graph with 20,000 nodes one time, ideally this is done on each subject and the results are averaged across subjects with consensus clustering. Moreover, most community detection algorithms are non-deterministic, so they should be run on each subject many times.

Thus, the use of a parcellation is a practical necessity. While most parcellations are developed to fit a group of subjects, recent analyses suggest that the use of parcels which are standardized across individuals leads to inaccurate calculations of functional connectivity, especially when comparing two groups with heterogeneous

demographics(273). This is not surprising, given that functional boundaries and connectivity patterns vary across subjects(274-277). Thus, there have been recent approaches developed that obtain parcellations that are optimized for individual subjects(273, 278). While these methods allow for thorough analyses of individuals, no method has been proposed for how to obtain a group-average measures of connectivity across individual parcellations. Here, we explore community detection and the participation coefficient, a measure that captures the integrative nature of a node(78, 140). In order to fully utilize individual subject parcellations, we develop a method for averaging across parcellations, resulting in a "nodeless" network in which each voxel belongs to a community and has a participation coefficient. Thus, the graphs are built based on individual subject parcellations with a limited number of nodes. However, the results are obtained at the highest spatial resolution possible—the voxel.

Another fundamental issue is defining the connections. Most studies use diffusion imaging to define structural connectivity or resting-state fMRI to define "functional" connectivity. These methods are far from perfect, likely creating many spurious edges. One simple solution for this is to threshold the graph at a certain "cost". For example, at a cost of 0.15, only the strongest 15 percent of the edge are kept in the graph. A lower cost that retains more edges usually results in more communities(56). This is useful in practice, because all community detection algorithms that depend on maximizing modularity suffer from a resolution limit(279), in that many smaller communities are not discovered, even when they are clearly defined. While some approaches have attempted to fix this resolution limit, they rely on global parameters that have been shown to be inaccurate if the network contains communities that are of different sizes, as the larger communities will be divided into two communities prematurely(279). One way around this is to simply threshold the graph to be sparser. However, this results in disconnected nodes and throws away data, likely ignoring important biological information present in the weaker connections. Thus, here, instead of thresholding the connections, we exponentially weight the connections, which increases the difference between strong and weak connection values. This obtains a larger number of communities. This allows for an accurate definition of the communities without ignoring weaker connections or removing nodes with weak connections from the graph. Like a resolution parameter, it can be finely tuned to estimate a particular number of communities.

Methods

rs-fcMRI data and preprocessing. Publically available resting-state fMRI data were used from 62 healthy adults (mean age = 22.59, SD = 2.45, age range =19-27, 26 males) that were acquired as part of the NKI-Rockland sample. For each subject, 9 minutes and 35 seconds of rs-fMRI data were acquired using a multiband gradient-echo echo-planar imaging sequence (TR = 1400 ms, echo time = 30 ms, multiband factor = 4, flip angle = 65 degrees, 2 mm^3 isotropic voxels with 64 axial slices). Structural data were acquired using a MPRAGE (TR = 1900 ms, TE= 2.51 ms, flip angle = 9 degree, 1 mm^3 isotropic

voxel). For both datasets, subjects were instructed to stay awake and keep their eyes open. Image preprocessing was performed the software Configurable Pipeline for the Analysis of Connectomics. First, brain images were segmented into white matter (WM), gray matter, and cerebral spinal fluid (CSF). Rigid body motion correction was then performed to align each volume to a temporally averaged volume, and a boundary-based registration algorithm was used to register the EPI volumes to the anatomical image. Advanced Normalization Tools (ANTS) was used to register the images to MNI152 template using a nonlinear normalization procedure. We then performed nuisance regression to further reduce non-neural noise and artifacts. To reduce motion-related artifacts, we used the Friston-24 regressors model during nuisance regression. WM and CSF signals were regressed using the CompCor approach. Linear and quadratic drifts were also removed. Data were bandpass filtered from 0.009–0.08 Hz. Finally, signal intensity was scaled to a whole-brain mode value of 1000. No spatial smoothing was performed.

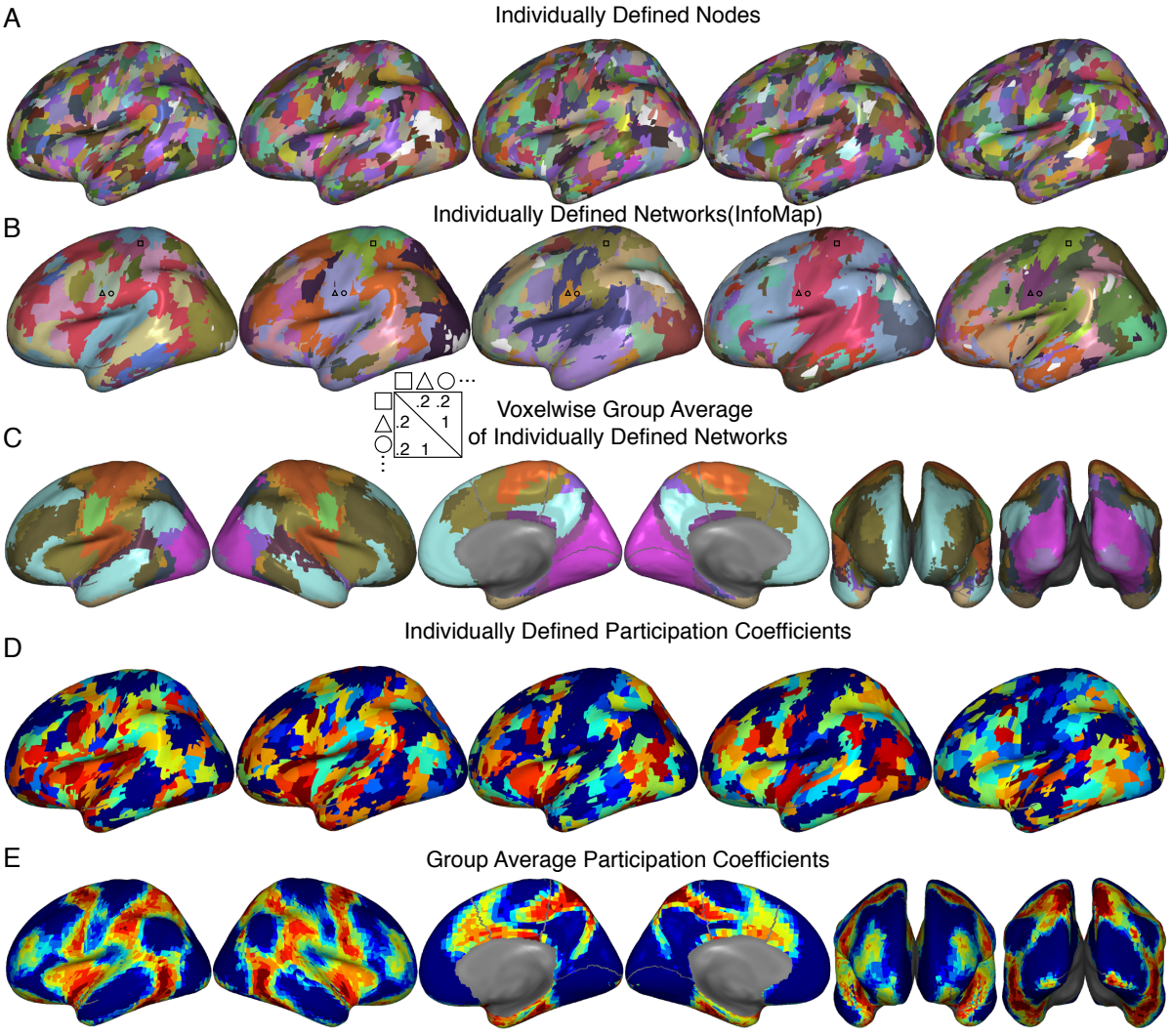


Figure 1. *Nodeless Networks Algorithm and Results*. A, nodes are defined in each subject. B, communities are defined in individual subjects. C, a voxel-wise consensuses

matrix is formed, representing the individual subject voxel level communities results, which are then averaged. D, in each subject, each node is assigned to the community it overlaps with the most, and the participation coefficient is calculated. E, participation coefficients are averaged across subjects.
Our method, which we call the nodeless network algorithm, relies on each individual having a unique parcellation. Any parcellation method can be used with our algorithm. Thus, we focus our attention on the algorithm to handle unique parcellations, not the unique parcellations themselves, as methods for generating unique parcellations are constantly evolving and improving. We use a method that has been shown to work well for group average data and for which the code is publically available(92). This method can either group voxels based on their similarity of activity or their similarity of spatial correlation; we chose the latter, as this metric leverages information about the voxels' connectivity pattern. This method allows for a choice of the number of nodes. To test our algorithm across a range of resolutions (i.e., number of nodes), we ran our analyses with 100, 200, 300, 600, 900, and 1500 nodes in the graph.

For each subject, we define the nodes based on their fMRI data at a certain number of nodes (Figure 1A). Next, we take the mean time series of each node. We create a correlation matrix of the Pearson r between every pair of nodes in the graph. Next, we exponentially weight the correlation matrix, starting at an exponent of 1, increasing in 0.1 steps. Each iteration uses the InfoMap algorithm(56, 129), but other graph theory community detection methods could be employed (e.g., Louvain). Exponents are increased until the number of communities requested by the user is achieved. If the number of communities requested is not found, we use the smallest exponent for which the highest number of communities was found. We chose to detect 18 communities. This was chosen to be in line with the resolution of other clustering techniques(56, 85, 114, 141, 227). Given that the InfoMap algorithm is probabilistic, once the correct exponent is found, 1000 iterations are run, with uniformly distributed exponents in the range from the original exponent minus 0.1 to the original exponent plus 0.1. A consensus matrix is formed for each run, where nodes in the same community have a 1 and nodes in different communities have a 0. Finally, we run the InfoMap algorithm on the average of those matrices 10,000 times (choosing the partition with the shortest map equation length, similar to the highest Q run for algorithms that maximize modularity) to obtain the final division of nodes into communities (Figure 1B). Supplemental Figure 1 shows the number of communities detected across all exponents and graphs.

Next, our algorithm relies on consensus clustering to obtain a projection from nodes to voxels. Traditionally, consensus clustering generates a matrix, where nodes in the same community have a 1 and nodes in different communities have a 0. This is how the matrix was utilized above. However, at this step, we project that consensus matrix from the node level to the voxel level (Figure 1C, inset). We make a voxel-level consensus matrix for each subject; voxels in the same community have a value of 1, and voxels in different communities have a value of 0. We then average across these matrices. For

example, if two voxels were in the same community in 34 of 65 subjects, it would have an edge weight of 0.523. We also ignored nodes that are not in a community with at least 5 nodes in this calculation. We then apply InfoMap community detection, exponentially weighting the matrix until the number of communities the user requests its achieved (Figure 1C).

Given this group voxel-level community assignment, we now can assign individual parcels to communities. For each node, we assign the node to the mode of the communities present at voxels in the node. For example, if a node has 18 voxels, and 10 were assigned to a community, it is assigned to that community. We then calculated the participation coefficient, which measures the diversity of a node's connections across communities. The participation coefficient *i* is defined as:

$PC = 1 - \sum_{s=1}^{N_M} \left(\frac{K_{is}}{K_i}\right)^2$,

the sum across communities M, where $K_i$is the sum of connectivity weights of *i* and K_{is} is the sum of connectivity weight of *I to* community M. If a region has connections uniformly distributed to all communities, then its participation coefficient value will be close to 1; on the other hand, if its connectivity is purely within a community, its participation coefficient value will be 0.

While exponentially weighting the matrix obtains a larger number of communities, it is not necessary to do this to accurate capture the participation coefficient. Thus, we averaged these results across a range of costs, from 0.05 to 0.15. This was done for each subject (Figure 1D). We then average each voxels participation coefficient scores across subjects (Figure 1F).

For comparison, we also ran more traditional graph theory analyses. We used the 900 node atlas that was derived using the same method we applied to individual subjects here; however, in this parcellation the results are then averaged at the group level, creating a single group-level parcellation. For each subject, we built a correlation matrix representing the Pearson r between all pairs of nodes' time series. Each matrix is fisher z transformed, and then the average is taken. We then applied the same community detection that we did for an individual subject's matrix to this average matrix. We also calculated participation coefficient across the same range of costs and took the average. All results were visualized on the cortical surface with pycortex(280).

Results

As our algorithm is automated and requires no parameter fitting; thus, analyses and comparisons across different graphs with different numbers of nodes are trivial. Figures 2-4 show the frontal, lateral, and medial views of the community detection across graphs with different numbers of nodes (100, 200, 300, 600, 900, 1500). Many subdivisions are not achieved in graphs with lower numbers of nodes, such that networks are not as distributed across cortex and interdigitated. For example, the dorsolateral pre-frontal cortex, at 100 nodes, is grouped into the fronto-parietal network;

however, above 100 nodes, the dorsolateral pre-frontal cortex becomes part of the cingulate-opercular network (Figure 3). This suggests that a distinct signal that is more correlated with the cingulate-opercular network exists, but, when too many voxels in that region are grouped together, the original signal is lost via blurring that occurs by averaging, and the blurry signal is now more correlated with the fronto-parietal network, likely given that the original signal was blurred with voxels that are actually part of this network. Moreover, motor cortex forms a topographic map with distinct hand and mouth networks(56). With under 300 nodes, this distinction was not detected (Figure 3). Finally, on the medial wall (Figure 4), in all graphs, the motor network or networks and the cingulate-opercular networks are present. However, in the graphs with more nodes (e.g., 900), the anterior and posterior default mode network is consistently fractioned, and the secondary visual network, the dorsal attention, fronto-parietal are present on the medial wall. Given these results and that these nodes were defined to maximize homogeneity in individual subjects, it is highly unlikely that nodes that are roughly the same size or larger and are based on gross anatomy (e.g., AAL or Harvard Oxford Atlas) are not blurring distinct signals and are able to accurately capture the community structure of the network.

The participation coefficient values are shown in Figure 5 for all graphs. Interestingly, 100 nodes appeared to capture participation coefficient values at a higher granularity than 200, 300, or even 600 nodes. However, the most consistent results were at 900 and 1500 nodes. The graphs with more nodes detected areas of high participation coefficient in the fronto-parietal network, premotor, dorsolateral prefrontal cortex, temporal-parietal-occiptial junction, ventral and dorsal attention networks, the precuneous, the cingulate, parietal occipital fissure, and the hippocampus and fusiform gyri.

We compared these participation coefficients and community results to two other techniques. First, as described in the methods, we ran a traditional graph theory analysis. Second, a highly principled ontological model of cognitive functions was recently developed. This method jointly estimates cognitive functions via a hierarchical Bayesian model (i.e., Author-Topic model) of the task categories (e.g., n-back, flanker, visual pursuit) and task activation in the BrainMap. It allows for each task to engage multiple cognitive functions, and each voxel to be associated with multiple cognitive functions. Areas that are highly associated with many cognitive functions are likely areas of integration across functions, similar to arears of high participation coefficient being areas of integration across communities. This allows us to compare the community and participation coefficient detection of our nodeless network algorithm and the more traditional graph theory community and participation coefficient detection to the cognitive functions and areas where many cognitive functions are active (Figure 6).

To compare the similar of communities to cognitive functions we used Normalized Mutual Information (MI). If two divisions of voxels into communities are identical, this value is 1. If no voxels fall into the same community across divisions, MI value is 0. The

MI of the nodeless network algorithm has a MI with the cognitive functions that 38% higher than the MI between traditional graph theory method and the cognitive functions (Figure 6). Moreover, the spatial correlation between participation coefficient and the number of cognitive functions active at a voxel was 39% percent higher for the nodeless network algorithm compared to the traditional graph theory method (Figure 6).

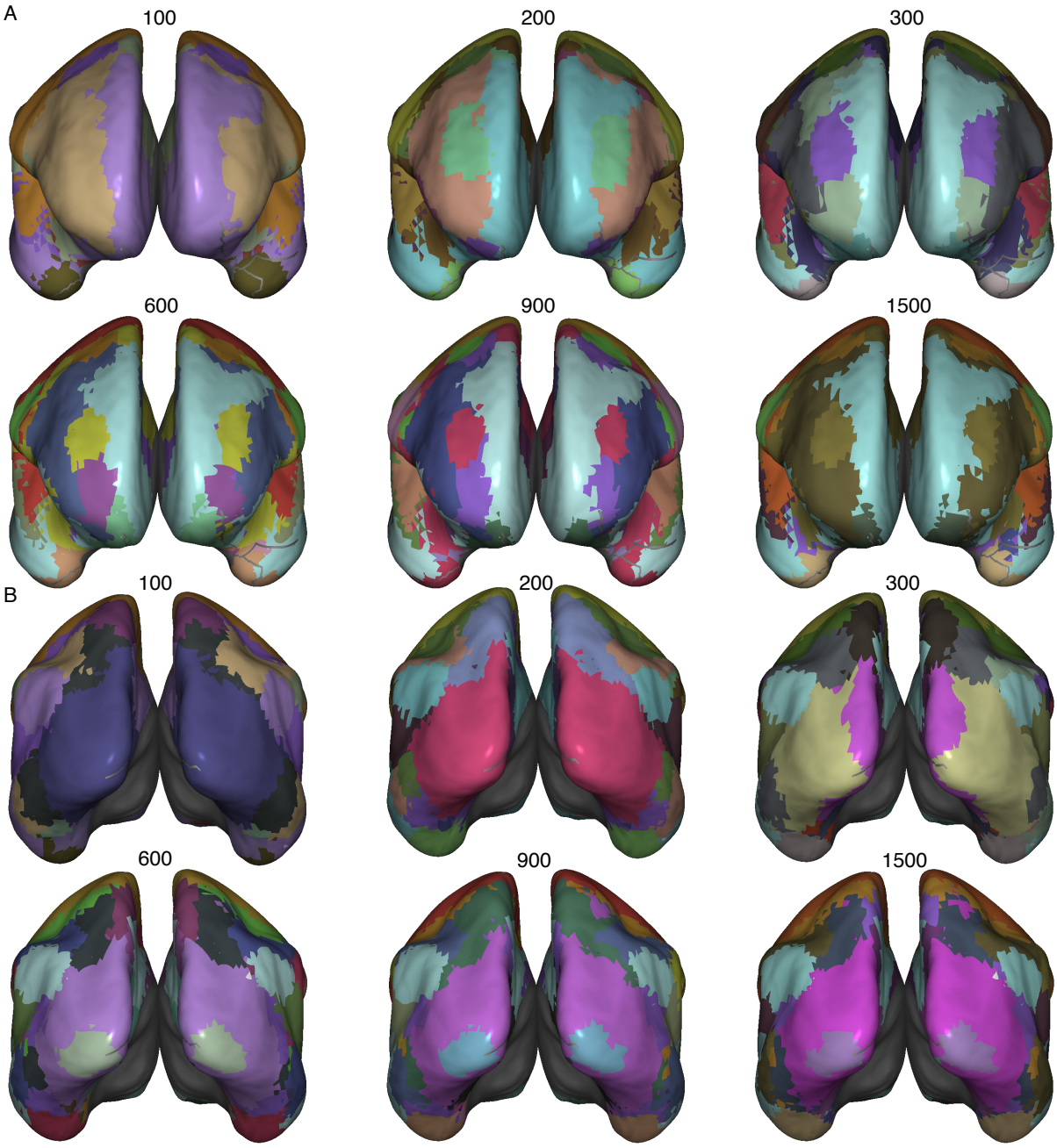


Figure 2. Anterior (a) and Posterior (b) views of group-average community detection in graphs with different numbers of nodes. Each color represents a unique community. Higher numbers of nodes result in a more interdigitated community structure. For

example, prefrontal cortex splits into three different communities, and early and late visual cortex split into two different communities.

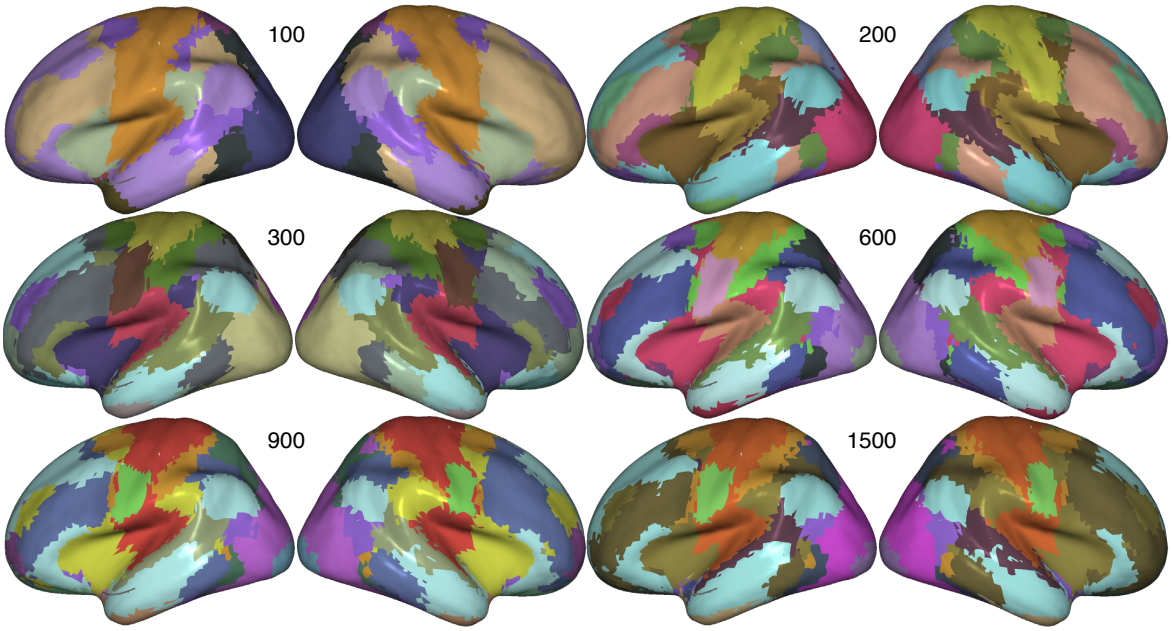


Figure 3. Lateral views of group-average community detection in graphs with different numbers of nodes. Each color represents a unique community. Higher numbers of nodes result in a more interdigitated community structure. For example, the motor cortex separates into two distinct networks that fit the homotopic organization of motor cortex.

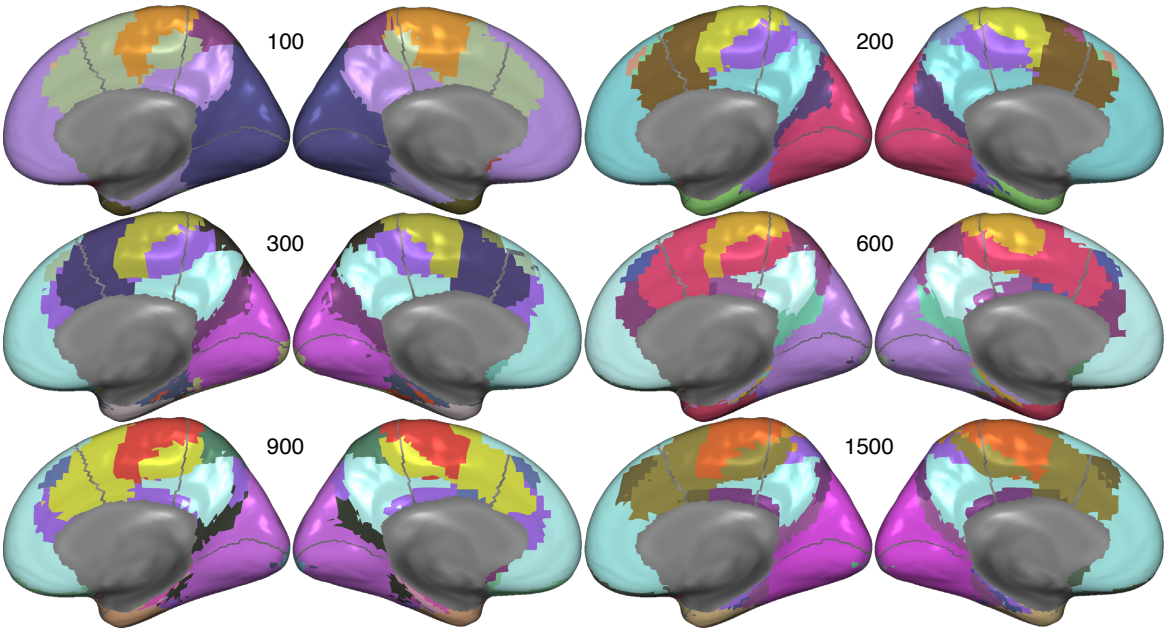


Figure 4. Medial views of group-average community detection in graphs with different numbers of nodes. Each color represents a unique community. Higher numbers of nodes result in a more interdigitated community structure. For example, the posterior

default mode network splits from the cingulate in the networks with a higher number of nodes.

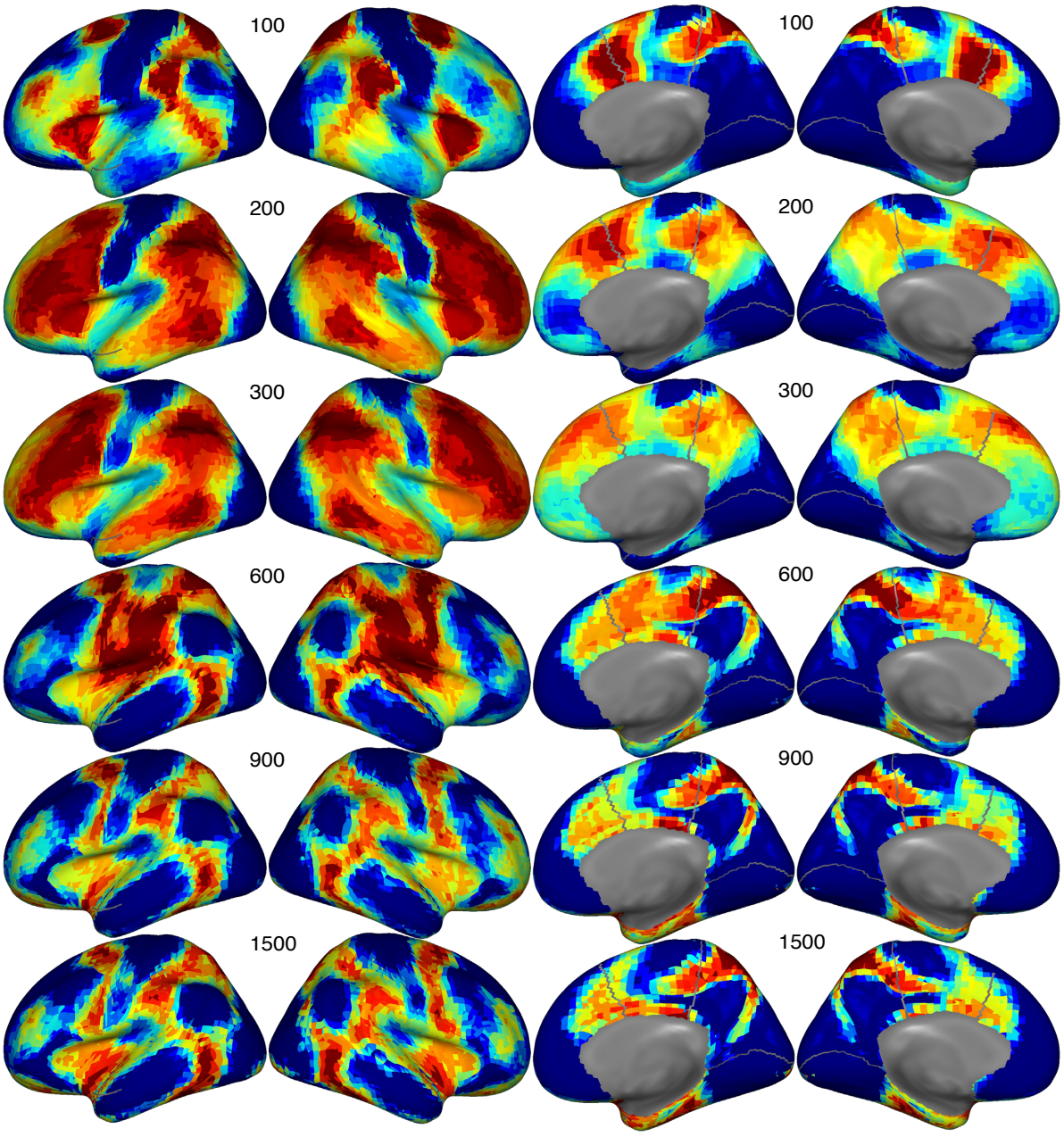


Figure 5. Group average participation coefficients across graphs with different numbers of nodes. In general, the specificity of the metric increases in graphs with greater numbers of nodes. However, the specificity of the metric in a graph with 100 nodes is comparable to the graphs with 600, 900, and 1500 nodes.

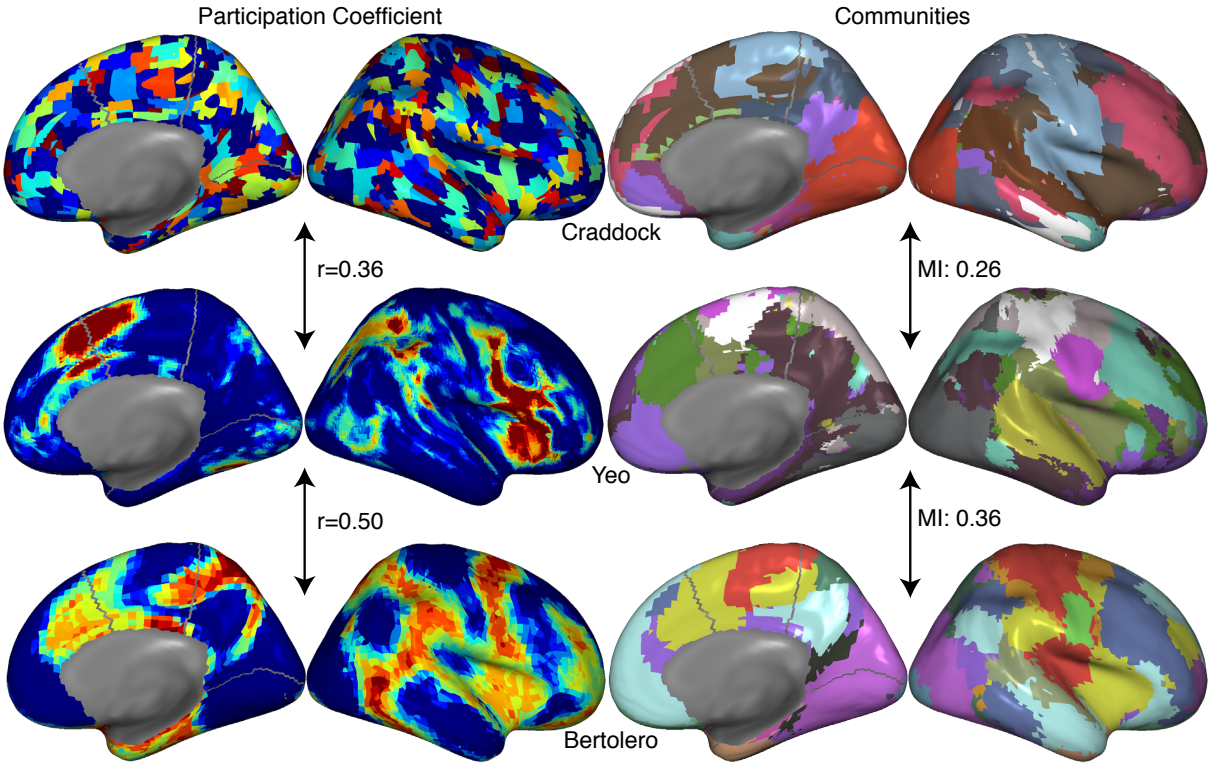


Figure 6. Comparison of clustering and integrative hub methods. Top, the communities and participation coefficient is calculated under a more traditional graph theory method. Middle, “flexible” hub regions and “cognitive components” from an author-topic model of BrainMap activation data. Bottom, the current method for community detection and measuring the participation coefficient. The current method detects participation coefficients that are more highly correlated with the flexible regions and communities that are more similar to the cognitive components.

Discussion

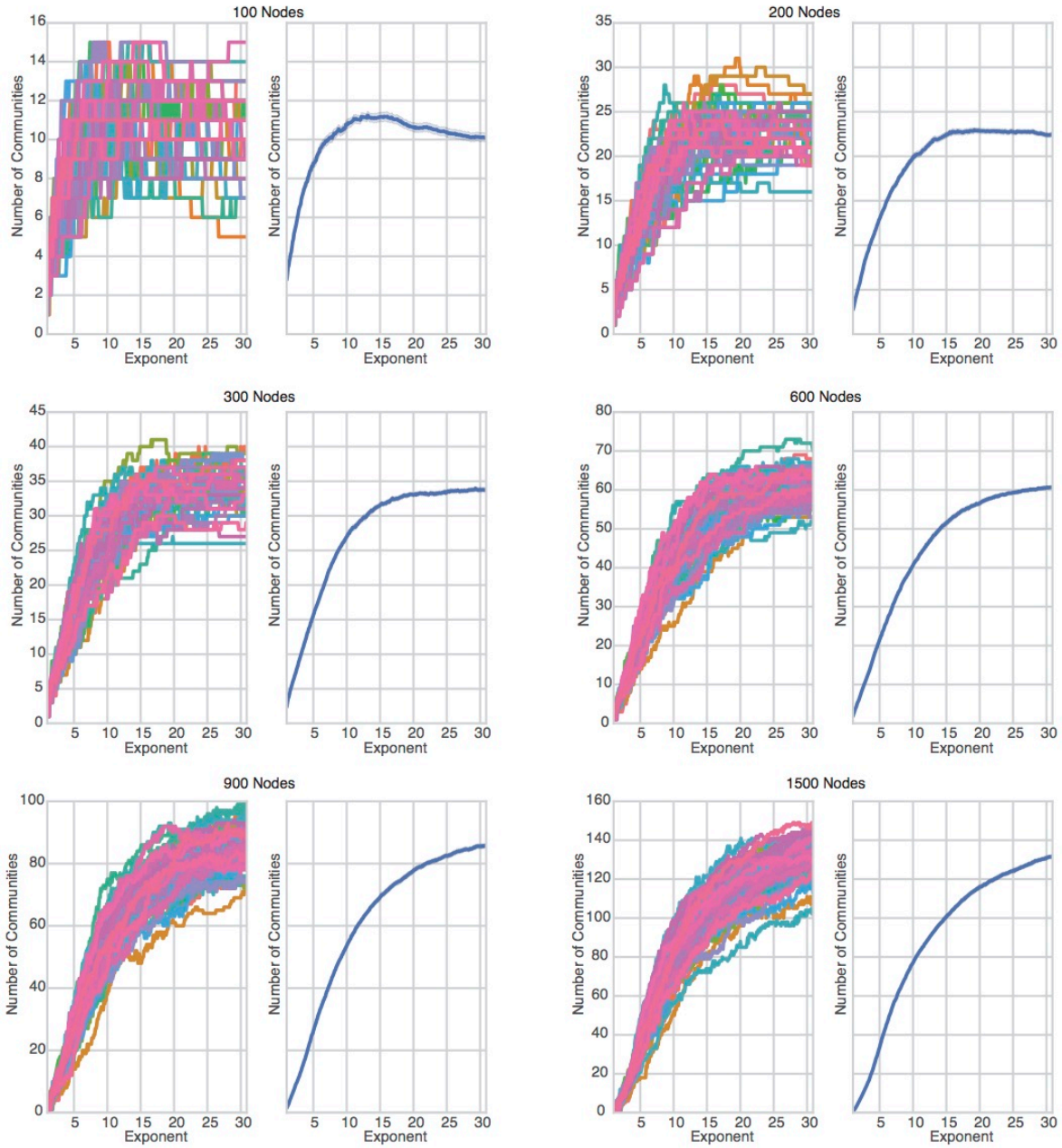
Modeling the human brain as a complex network is one of the most powerful methods to capture fMRI data. Here, we propose a method that allows for group average graph theory analyses while simultaneously allowing for a mathematically optimal definition of nodes—parcellations that are fit ideally to each individual subject. Moreover, our method allows for the highest resolution possible—the voxel level. We also propose exponentially weighting the correlation matrix during community detection to resolve a particular number of communities.

Compared to a more traditional parcellation method and graph theory analysis, our proposed method detected communities and participation coefficient values that are more in line with the cognitive functions and areas that are associated with multiple cognitive functions detected by a large meta-analysis of task evoked activity. While the meta-analysis is certainly not a “ground truth”, it is a very principled model of cognitive functions and was applied to a very large set of data (over 10,000 experiments).

Moreover, our results suggest that maintaining the functional boundaries in individual subject data is critical, as this method resulted in a model that is more similar to a model of task evoked data than using a single parcellation.

The maps from theses analyses are publically available (NeuroVault). These maps can be used by researchers as an a priori and independent identification of communities and each voxel's participation coefficient. However, the code that implements this method is in python (available on github.com/mb3152/nodelessnetworks), highly automated, memory efficient, and fast. The complete set of analyses (post pre-processing of the fMRI images) presented here complete in under a day (20 cores, 512 GB RAM), with most of the time being dedicated to the final community detection. Thus, this approach is easy to implement and scales easily to larger sets of data with more subjects.

Our method could be applied to other datasets to generate communities and a map of participation coefficients. This method is perhaps most useful while comparing groups. While many studies use a single parcellation (which was usually defined on a healthy and young population) for every group, our method allows for each individual to be parcellated, and then each group to have a community and participation map based on those individual parcellations. Thus, the parcellation does not fit one group better than the other, as each individual has a unique parcellation fit to that individual. This solves a major potential flaw in many network science comparisons—if the parcellation fits one group better than the other, any comparison across groups is invalid. Moreover, this method resolves the communities and the participation coefficients at the voxel level, allowing for a high level of spatial specificity when comparing groups. Finally, this method evades the issue of a researcher having to test a single hypothesis with many different parcellations.



Supplemental Figure 1. Exponentially weighting the network's weights increases the number of communities detected. Left panel, each line is an individual subject, showing the relationship between the exponent used and the number of communities detected. Right panel, the group average relationship between the exponent used and the number of communities detected. Regardless of the number of nodes in the network, exponentially weighting the network increases the number of communities detected at a logarithmic rate.

Reverse Network Inference

Max Bertolero
Thomas Yeo
Mark D'Esposito

Abstract

The practice of reverse inference–inferring that a subject is engaged in a cognitive process given a particular pattern of brain activity–is widely viewed as unprincipled, inaccurate, and ineffective. Here, we show it can be principled, accurate, and effective. Given an principled model of brain activity and an ontology of cognitive functions, reverse inference can be accurate and can progress our understanding of the brain. Here, we develop a method for reverse inference–jointly and probabilistically–based on the intuition that task evoked brain activity is the result of the activity of multiple cognitive functions, each of which is subserved by a network of brain regions, some of which overlap across cognitive functions. We show that individual frames of brain activity during tasks can be captured with this model, and the task the subject is in can be decoded very accurately. Moreover, we generated two new insights into the brain by using this reverse inference method. First, we replicate a previous meta-data finding at a high temporal and spatial resolution in individual subjects. Second, we find that subjects' task evoked activity can be clustered into states, each of which correspond to a single resting-state network or a combination of resting-state networks. While states are rather similar across tasks, each task has one to two unique states.

Introduction

Two influential articles by Mark Desposito (131) and Russ Poldrack (133) cautioned against the use of reverse inference — the practice of inferring that a cognitive function is engaged based on observing an activated brain region previously evoked by tasks that presumably probe that cognitive function. This inference is only deductively valid if the brain region is involved in only that cognitive function, and the original task (contrast) engaged only that cognitive function. For example, consider the language example from Russ Poldrack's influential paper (133). If we observe activation in the inferior frontal gyrus, does this indicate language processing? In other words, we want to know the probability of language processing given activation in the inferior frontal gyrus. To compute this probability, Poldrack used Bayes rule (Figure 1) and analyzed brain activity maps in the BrainMap database. There are 166 language studies that activate the inferior frontal gyrus and 703 language studies that do not activate the inferior frontal gyrus. Therefore, under this method, the probability that language studies activate the gyrus is 0.19 (166 divided by 166 plus 703. However, there are 199 non-language studies that activate the inferior frontal gyrus and 2154 studies that do not activate the gyrus. Therefore the probability of non-language studies activating the gyrus is 0.08 (199 divided by 199 plus 2154). Considering these non-language studies is critical. If non-language studies consistently activate the inferior frontal gyrus, then the gyrus is not specific to language processing, but perhaps some other cognitive function. Finally, we also have to set the prior probabilist of langue processing being engaged. If we assume a neutral probability of 0.5 (281), this gives us a 22% percent probability of language processing given inferior frontal gyrus activity.

$$P(\text{lang}|+\text{act}) = \frac{P(+\text{act}|\text{lang})P(\text{lang})}{P(+\text{act}|\text{lang})P(\text{lang}) + P(+\text{act}|\text{-lang})P(\text{-lang})}$$

Figure 1. Bayes Rule applied to language processing and inferior frontal gyrus activity

However, there are various hidden assumptions in this method. First, each probability is computed independently, considering only language and all non-language studies. However, by lumping all non-language studies together, the probability of non-language studies activating the inferior frontal gyrus is diluted. For example, of the 199 non-language studies, let us assume that 90 are studying the cognitive function A and find activity at the inferior frontal gyrus. Let us also assume that, of these roughly 2154 non-language studies that don't find activity at the inferior frontal gyrus, 10 are studying cognitive function A. Under these assumptions, the probability of cognitive function A activating the inferior frontal gyrus is 0.9. In fact, inferior frontal gyrus activation would indicate the presence of cognitive function A, rather than language processing. Thus, the inferior frontal gyrus is specific for function A. Language studies recruit function A 17% of the time (a probability of 0.19 times a probability of 0.9), which is why language studies activate the inferior frontal gryus 19% of the time. Second, the probabilities are computed for each voxel or brain region independently. But observing a single region may not be very informative for two reasons. First, cognitive functions potentially arises from interacting and overlapping brain networks. We assume this model here, instead of the inaccurate or at least incomplete model in which single regions each execute a distinct cognitive function. Second, a single region's activation, relative to a collection of regions' activation, does not provide as much information when inferring whether one of two cognitive functions is engaged.

Thus, reverse inference must utilize a model in which cognitive tasks engage common and distinct cognitive functions, which are in turn supported by multiple, possibly overlapping, brain regions. However, this model requires an ontology of cognitive functions. In the above example, we need to know what cognitive function A is, and where it is most likely to be activated, along with all the other cognitive functions, and where they are most likely to be activated. To achieve this, we recently instantiated a mathematical model that encodes this notion (151), allowing us to estimate a nested cognitive ontology based on the BrainMap database. The ontology consists of a collection of cognitive components, the probability that a task would recruit a component and the probability a component would activate a voxel.

Here we consider how our mathematical framework and resulting ontology provide conceptual and computational advances over the original reverse inference framework described above. Our mathematical framework explicitly models how interactions between cognitive functions give rise to brain activation. Therefore, it is easy to perform reverse inference across the whole brain within our framework. Here, we perform whole brain reverse inference during task states on individual frames (i.e., single timepoints

during a scan). We characterize the accuracy of the reverse inference, both in terms of how well the model fits the data and how well the model can be used to decode what task the subject is in. We also provide two novel applications. First, we extend a previous meta-analysis finding—when more cognitive functions are engaged during a task, activity at connector nodes in the fronto-parietal and dorsal attention networks increases. While the original analysis looked across tasks in the BrainMap database, we look across timepoints in a single scanning session during multiple task states. Our new method also allows us to test the extent to which each voxel is dependent on the number of cognitive functions engaged. Second, we clustered individual timepoints into "states", and find that each state corresponds remarkably well to the activation of one or two traditional resting-state networks, and that connector node regions are activated in the most states. Thus, we argue that reverse inference, when computed jointly and probabilistically across the brain with a principled model, is valid and can be applied to deepen our understanding of global brain function.

Method

Joint Probabilistic Reverse Inference

To estimate cognitive functions, we used an author-topic model. This application has been described previously(176), but we describe it here in the context of reverse inference. An author-topic model defines an exact probabilistic model relating documents, authors, topics, and words(282). Consider a collection of scientific documents or papers. Across this collection, we have authors and their words. The author-topic model lets us discover abstract topics, which are made concrete by their association with certain authors and their association with certain words. An important point is that the topics are never observed because topics are abstract. The model generates what the actual topics are, in that each topic has a probability of being written about by a particular author, and each word has a probability of appearing given a certain topic. We analyzed the BrainMap database with the author-topic model. The BrainMap database to which the model is applied contains 10,000 imaging experiments. Each experiment is tagged with one of 83 tasks. Each experiment is also tagged with the activated voxels in MNI space. If we think of documents as experiments, authors as behavioral tasks, topics as cognitive components, and voxels as words, there is a one-to-one mapping between the author-topic model in text mining and our problem. We applied the author-topic model to the BrainMap database. We don't assume functions like "emotion", "language" or "vision". The model estimates them. One parameter we have to set is the number of cognitive components. Here, we consider the 12 component model (see (176) for a complete discussion of choosing the number of components).

Essentially, we formulate an optimization problem where we want the parameters θ and β that maximizes the posterior probability $\Pr(\theta, \beta \mid BrainMap\ \mathcal{E}ata)$, where θ is the $Pr(Component|Task)$ and β is the $Pr(Voxel|Component)$. This is achieved with an expectation maximization algorithm. Given the estimate of β, it is actually quite simple to perform reverse inference, as we have the probability of a voxel being active given a component. More specifically, for a reverse inference, we want to find the new $\theta *$ that

maximizes the posterior probability of θ given the BrainMap estimate of β , which was estimated with the author-topic model, and activation data from a new "experiment's" whole brain activation; here, the experiment is a single time point (i.e., frame or volume) of an fMRI time series. Essentially, we find the $\theta *$ that maximizes $\Pr(\theta|\beta, Whole\ Brain\ \mathcal{E}ata|)$. This is also done via expectation maximization. This gives us, for each component, $P(component \mid Whole\ Brain\ Data)$. Thus, for each time point in an fMRI time series, we have a probability of each component being engaged. This distribution sums to 1.

fMRI data and preprocessing

We used the tfMRI minimally processed data from the Human Connectome Project. For each task, we used the 100 subjects with the lowest mean frame wise displacement. We used AFNI to preprocess the images, matching traditional resting-state functional connectivity studies. The AFNI command *3dTproject* was used, passing the mean signal from the cerebral spinal fluid mask, the mean signal from the white matter mask, the mean whole brain signal, and the motion parameters to the "-ort" options, which remove the signals via linear regression. The options "-automask", which generates the mask automatically was used. The "-passband 0.009 0.08" option, which removes frequencies outside 0.009 and 0.08, was used. Finally, the "-blur 6", was used, which smooths the images (inside the mask only) with a filter that has a width (FWHM) of 6mm after the time series filtering. Because of the short length of the Emotion task, it was not included in our analyses. Given limitations of the author-topic model in handling negative values, all values below zero were set to zero, allowing us to convert brain activity into "word count".

Fitting the model to the data

To analyze how well the joint probabilistic reverse inference was able to accurately model the original data, we reconstructed the original data using the model estimates of each component's engagement. For each component, we multiplied the component's probability across voxels by the reverse inference's probability for that component being engaged. For example, if the joint probabilistic reverse inference estimated that component 1 has a .1 probability, and particular voxel has a probability of .01 for that component, that voxel is scored as 0.001. We do this for each voxel, for each component, giving us 12 whole brain maps. We then take the sum across the 12 maps (Figure 2). We then calculate the spatial correlation between the original data and this map. We call this value the "fit". To measure if the fit of the model to the data was better than random, we compared the spatial correlation value of the fit of the model to the data to the spatial correlation value of the fit of the model to all other frames (i.e., the data the model was not fit to).

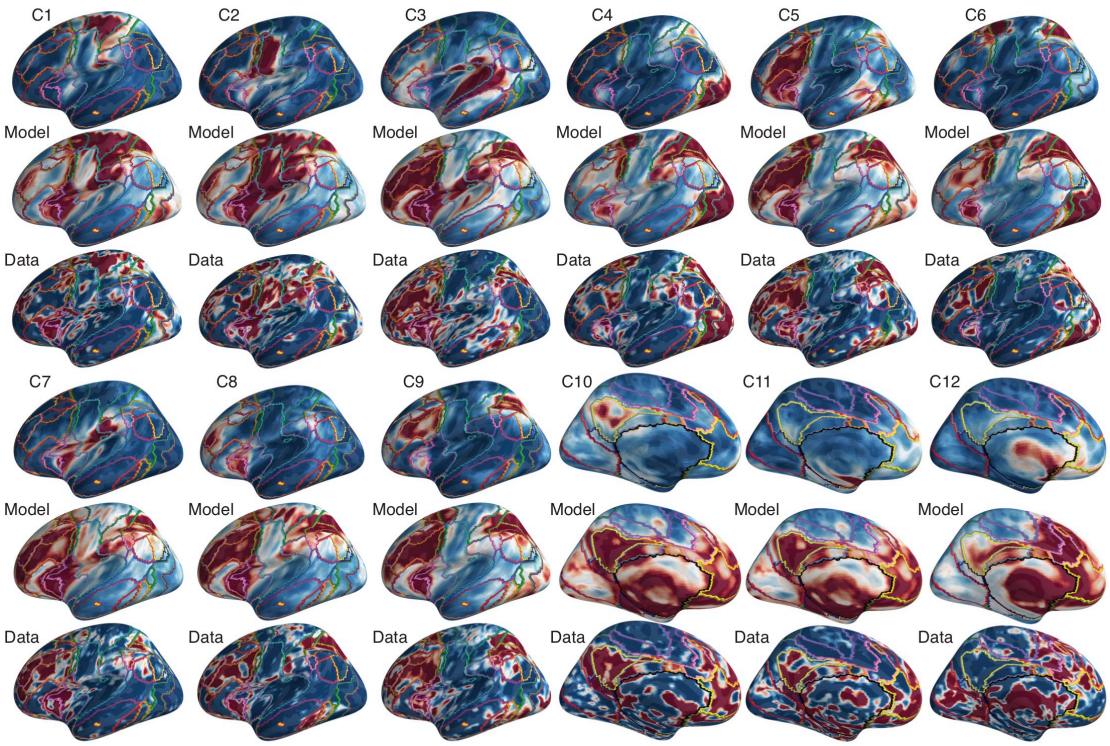


Figure 2. Cognitive Components, models of the data, and original data. For each component, a frame was chosen for which that component was highly likely to be engaged. Next, a model of the data using the reverse inference probabilities and the cognitive components was built.

Prediction of task

To ensure that our method of reverse inference not only fit the original data, but also picked up on task specific differences, we used the correlation matrix of the cognitive components' probabilities across time, and each component's mean activity in the task (relative to the mean of that component in that subject in other tasks), to predict which task the subject is executing. Thus the model contains 78 unique features. We have 600 of these feature sets; 6 per subject. We iteratively split the data into 450 feature sets and 150 feature sets, fitting the model to the 450 feature sets and then testing the model on the 150 feature sets. We did this 10,000 times, saving the accuracy of the model to predict the task of the unseen matrices. We measured accuracy with precision. The precision is the ratio tp / (tp + fp) where tp is the number of true positives and fp the number of false positives. Intuitively, the precision is the ability of the classifier not to label as positive a sample that is negative. The best value is 1 and the worst value is 0(283).

Analysis of components engaged

One of our previous analyses (140) using the author-topic model of BrainMap showed that, when more components are engaged in a task, activity at connector nodes–regions mostly in the fronto-parietal network–is increased. This analysis was across

BrainMap tasks, which are themselves the mean across multiple experiments, each of which averages across individuals, and is a contrast of that task, relative to another task. Here, we were able to do a more thorough analysis at a higher temporal resolution in individual subjects. Moreover, instead of taking the mean activity across a certain set of regions (as the meta-data contains a large amount of 0 values) we were able to simply correlate each voxel's activity with the number of components engaged at each time point, which we measured with 1 minus the standard deviation of the distribution of component probability values. We also executed this analysis where we set voxels in the original data to zero if 3 or more components have a high probability of being engaged at the voxel (> 1e-5). This ensures that the estimate of the number of components engaged is driven by voxels that are relatively specialized to a single component, and not driven by voxels where many components have a high probability, as these areas overlap highly with connector nodes (140), which is precisely where we expect to see increased activity as more components are engaged.

States

We sought to reduce individual time-points into states. Given our joint probabilistic reverse inference model, for each task, for each subject, for each time point, we have the distribution of component probabilities. For each time point *i* and *j* in a particular subject and task, we calculate the correlation between their distributions. This is done for each pair of time points, resulting in a square correlation matrix. This matrix is then used to form a graph, where each node is a time point and each edge is the similarity of the time points' component probability distributions. Community detection is then applied to cluster time points into communities. The edges weights are raised to a particular exponent to more heavily weight the strong connections; higher exponents lead to a greater number of communities while using InfoMap. The exponent is raised until the number of communities preferred is reached. We chose to analyze 8 communities or states. For each community, we then take the mean of all time points in that community. For each of these images (i.e., all the mean images from all the subjects), we compute the joint probabilistic reverse inference. The same graph construction and community detection procedure above is then applied, treating each mean image from a subject as a node, giving us a final clustering of original brain data into states.

Results

Joint Probabilistic Reverse Inference can accurately capture which components are engaged and model the original data

For each task, for each subject, for each frame, we measured the fit of the model to the data. The fit, on average, across tasks, was 0.29. Essentially, this value is the spatial correlation between the model's reduction of the data, and the original data. In all tasks, the model fits the data it was fit significantly better than data it was not fit to. Thus, the model was able to fit the data with accuracy across tasks (Figure 3).

Figure 3. How well the model fits the data. For each task, a violin plot that shows the spatial correlations between the model and the frame the model was fit to (blue) and a violin plot that shows the spatial correlations between the model and the frames the model was not fit to (green).

Joint Probabilistic Reverse Inference can accurately predict the subject's task state

We were able to predict the task the subject was engaged in with between .7 precision to 1 precision (0.83 on average; Figure 4). The Working Memory and Language tasks exhibited the highest precision. The Motor and Social tasks exhibited the median precision. The Gambling and Relational tasks exhibited the lowest precision.

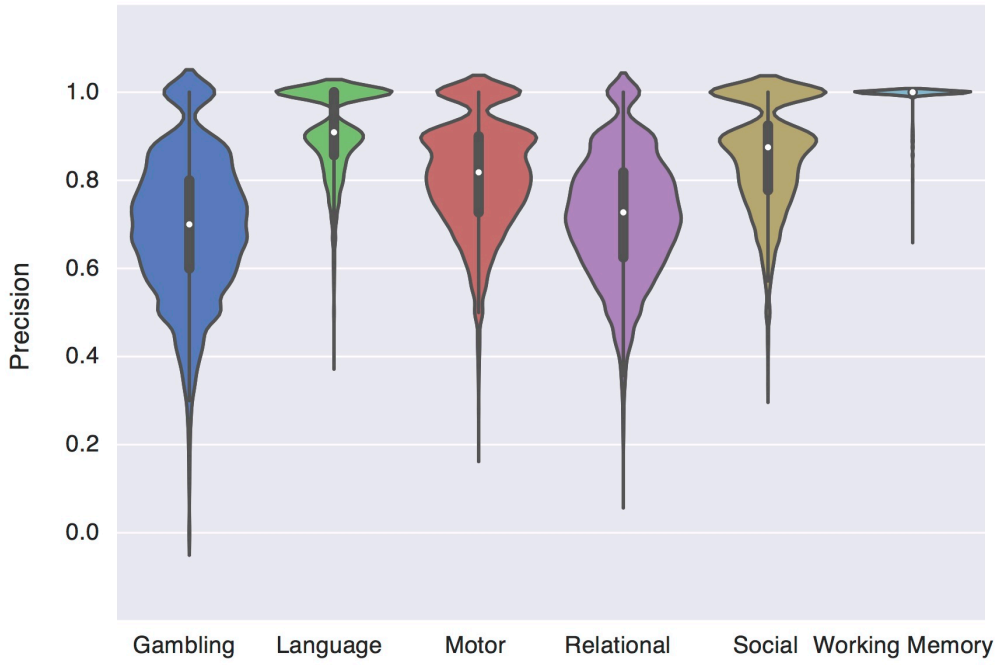


Figure 4. Precision accuracy while predicting which task the subject is engaged in. The distribution across 10,000 trials is shown for each task.

Activity in fronto-parietal control and dorsal attention networks correlates with number of components engaged

Next, we measured which voxels correlated with the number of cognitive components engaged at each time point. We calculated this correlation for each subject in each task, and then took the mean for each task. In every task, we saw positive correlation between the number of cognitive components engaged and activity at voxels in the fronto-parietal network and dorsal attention network. We observed negative correlations between the number of cognitive components engaged and activity at voxels in the default mode network. These results are shown in Figure 5 for every task. Moreover, results were practically identical when ignoring voxels where more than 3 cognitive components had high probabilities during the joint probabilistic reverse inference.

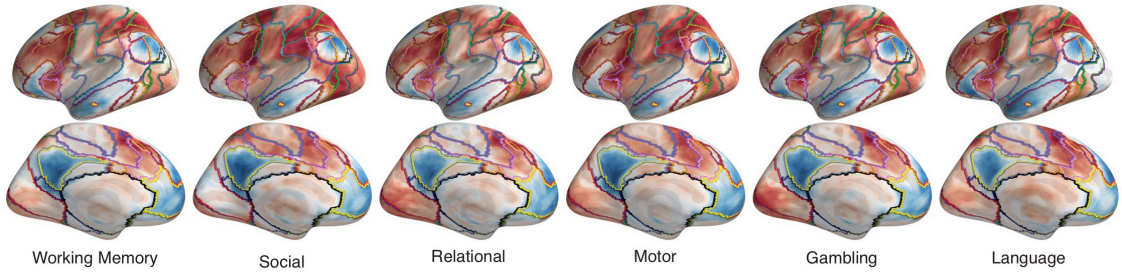


Figure 5. Correlation between brain activity and the number of cognitive components engaged in each task. The average across subjects is shown.

Brain states reflect functional connectivity networks and are unique across tasks
We clustered subjects individual time points into group-wise task states. In each task, states appears to be combinations of, or single, resting-state networks (Figure 5). Moreover, across tasks, similar states appeared (Figure 6). For example, in the Motor and Working memory tasks, 7 of the 8 states were highly similar, with the Motor task having a state that was highly specific to the ventral sensory motor network, likely because, in the Motor task, subjects are presented with visual cues that ask them to move their tongue, while subject presumedly are not moving their tongue in the Working Memory task. Interestingly, our method is able to pick up the tongue movement state, and, while all the other states were similar to the working memory task state, the tongue state did not exist in the working memory task. The same phenomena was observed in the Language task; while states similar to the overlapping states in the Working Memory and Motor tasks were observed, two unique tasks were observed (Figure 7).

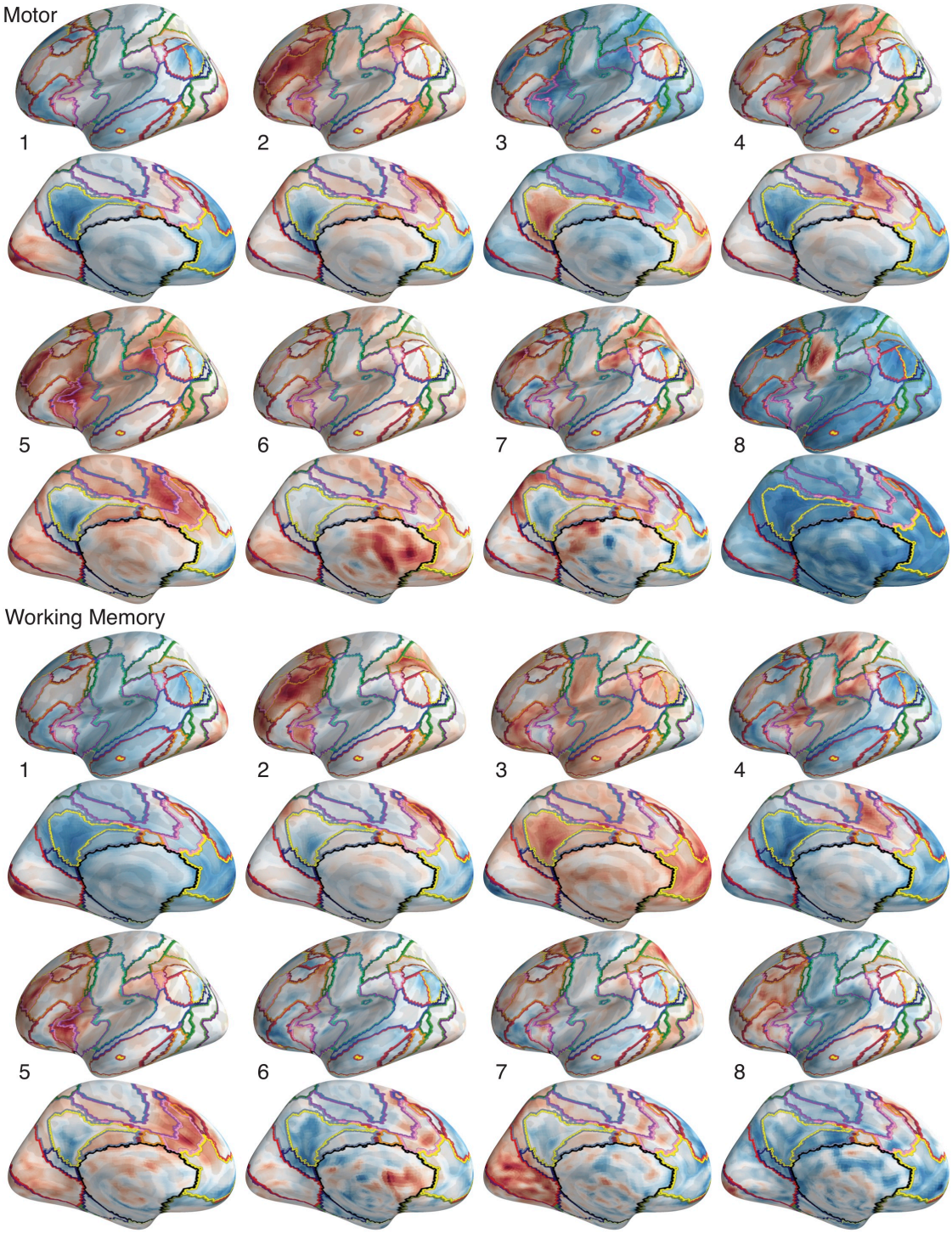


Figure 6. Eight states from the Motor and Working Memory tasks

The same phenomena was observed in the Language task; while states similar to the overlapping states in the Working Memory and Motor tasks were observed, two unique task states were observed (Figure 6). The first state involved activation at the frontal aspect of the fronto-parietal network, the inferior frontal gyrus, and the lateral temporal

lobe, an area known for semantic processing. It also involved deactivation of the default mode network. The second state involved deactivation of the fronto-parietal network, and activation of the inferior frontal gyrus and the same lateral temporal lobe region as the first state. Finally, in the Gambling task, we observed a unique state, where the fronto-parietal network and the medial aspect of the default mode, two networks that are typically anti-correlated (284), are active. However, the orbital-frontal cortex aspect of the default mode network has been associated with representing value (285), while the fronto-parietal network has been associated with decision making (286) and potentially controls the default mode network (287). Thus, these two network likely interact during gambling tasks.

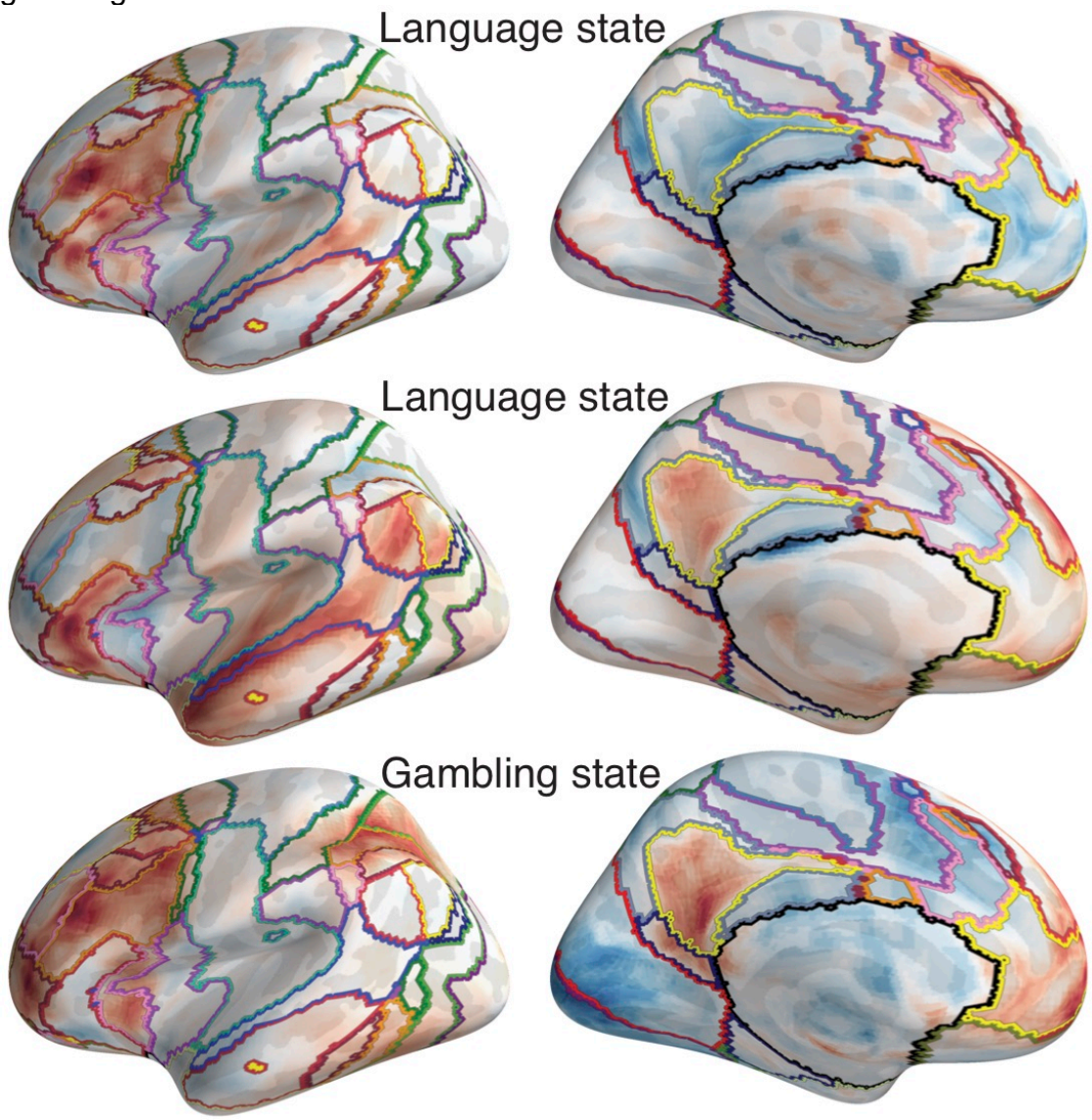


Figure 7. Unique States

Discussion

The ability to observe brain activity and, with reasonable certainty, know the subject's mental state, is highly sought after. However, the accuracy of a reverse inference is typically low. Given that a single region can be activated by many different cognitive

tasks, in general, observing that a single region is engaged provides limited additional evidence for the engagement of a cognitive function(288). Moreover, the applications of reverse inference has, so far, been rather limited to practical business applications (289) or outright abused(290). While the selectivity of brain regions to particular cognitive processes has been studied (291), reverse inference has not been widely applied to progress our understanding of the brain. The inaccuracy of reverse inference and the dearth of fruitful scientific applications are both due to the false assumptions that typically underlie reverse inferences. If the reverse inference is not based on how the brain actually functions, it likely won't be accurate, and its utility certainly cannot progress our understanding of the brain.

While many cognitive processes are certainly localized, there is no simple one-to-one mapping between every cognitive process and a brain region. Many cognitive processes studied in psychology are likely not natural kinds (9) and others are highly distributed (16). An emerging idea is that most cognitive processes (besides, for example, edge detection in visual cortex) are the result of the combinatorial interactions of relatively stable functional networks that each have a specific function(140, 194, 196, 227, 292). Most notably, previous work termed this idea network co-occurrence, and, similar to the current work, modeled task activity by combinations of canonical brain networks(292). The upshot of this work was the ability to model task activity with combinations of networks.

The previous network co-occurance work, however, was focused on modeling the activity differences between two tasks states, relying on traditional group-level task contrasts (e.g., math versus language processing). Here, we presented a reverse inference method that is based on network co-occurance, but we focused on modeling the raw individual frames of individual's brain activity during cognitive tasks. We lend further credence to the network co-occurance model of brain activity–our method fits the data and was an accurate method of detecting which task the subject was engaged in. Thus, a network co-occurance model can accurately capture individual frames of data as well as differences in activity across different cognitive tasks.

The ability to, given a principled model of the brain and method of reverse inference, gain further insight into the brain was also demonstrated. We presented two novel applications. First, we replicated a previous finding at a much higher temporal resolution in individual subjects, rendering more detailed information about which regions of the brain are sensitive to the number of unique cognitive functions engaged. As the previous analysis looked across group-level activation maps, one could only conclude that, in general across tasks, when more cognitive functions are required, activity in the fronto-parietal network is increased. However, we demonstrated here, with increased spatial resolution, that this phenomenon exists on the sub-second temporal resolution. Second, we used the cognitive component probabilities to detect discrete states that occur during cognitive tasks. We found states that were present across tasks, as well as states that were unique to that task and reflected the processes unique

to that task. While much research has been done to study the dynamic interactions between brain networks(293-295) the temporal resolution of functional connectivity does not allow for analysis of these interactions at high temporal resolutions (i.e., what happens in each time point and changes from timepoint to timepoint), and there are many methodological issues(296). Our method to detect states found unique interactions between networks that are likely not captured by dynamic functional connectivity methods. Moreover, out method depends on univariate activity, which eludes many of the methodological issues of dynamic functional connectivity. For example, in the Gambling tasks, we observed a state that was unique to the Gambling task, and involved two networks–the fronto-parietal network and the default mode–being active together, despite these networks traditionally exhibiting negative functional connectivity 8. Moreover, each states boundaries (where activity transitions from high to low) overlap highly with functional connectivity networks. This suggests that we observe functional connectivity networks because there is repeated transient activation of each network (or combinations of networks).

In conclusion, we presented a principled model and method for reverse inference. This method was able to accurately capture individual timepoints of data and predict which task the subject was engaged in. Moreover, we demonstrated that this method can be used to process our understanding of the brain.

Summary

Modularity

Different parts of the brain do different things. But what boundaries can we carve the brain up with to define these parts? The current work answered this question in further detail than previous work (56, 72, 73) defining the boundaries around the parts of the brain that are dedicated to specific functions. We used functional connectivity to define the “parts”. Functional connectivity between two regions simply measures how strongly correlated the two regions’ activity levels are across time (e.g., the Dow Jones Industrial Average and the Standard & Poor’s 500 are strongly functionally connected). We build a network model of the brain, where brain regions are nodes and the strength of functional connectivity between those two nodes is the edge that connects those two nodes. We then group nodes into sub-networks by maximizing the extent to which each sub-network coherently activates and deactivates together (i.e., strong functional connectivity or edges between nodes in the same sub-network) and the extent to which the pattern of each sub-network is unique to that sub- network (i.e., weak functional connectivity or edges between nodes in the different sub-networks). In the human brain, we typical find 12-14 sub-networks. The sub-networks’ coherent and unique activity is always present, regardless of what a person is doing, only retreating upon death(297). Because of the sustained, coherent, and unique activity of each sub-network, we hypothesized that each sub-network executes a discrete cognitive function mostly autonomously. Simply put, if a sub-network of regions is firing at different times and in a different pattern than other sub-networks of regions, that sub-network is probably up to something unique.

To test this hypothesis, we needed a model of which cognitive functions exist and which parts of the brain execute them. Thus, we leveraged the application of the author-topic model to the brain (151). The author topic model was originally designed to find literary topics (282) Imagine a library. What are the topics that exist in all the books? What words are associated with which topics? What authors write about which topics? Of course, you could make a reasonable guess. For example, Hemingway writes about war, and the word “death” usually appears in war related works. This is what librarians used to do. However, the author-topic model can discover topics based on the data. The author-topic model defines an exact probabilistic model relating authors, topics, and words. Each topic has a probability of being written about by a particular author, and each word has a probability of appearing given a certain topic.

We have the same problem in cognitive neuroscience. Similar to how the first librarians organized libraries, we have been guessing as to which cognitive functions exists, which tasks tap them, and which parts of the brain execute them. Our library is a large amount of cognitive tasks with a large amount of brain activity evoked by those tasks. Just as each author writes about multiple topics, each task likely utilizes multiple cognitive functions. Just as the model estimates the topics, as well as which authors write about

them and which words are associated with them, the model can estimate the cognitive functions, as well as which tasks tap each cognitive function and which brain regions execute that cognitive function. Thus, we do not have to make the inaccurate assumption that each cognitive tasks taps a single cognitive function, nor do we have to guess which cognitive function each task taps. Best of all, we do not even have to guess about which cognitive functions (e.g., working memory, vision) exist. The model discovers them. Thus, the mental kinds as natural kinds problem goes away, because we do not posit mental kinds. We discover them based on the data. The critical insight into why this model works is that it defines the brain's components in the context of those components functioning together as a large scale network. This is how ones carves the brain at its joints.

The model estimates 12-14 cognitive functions and which parts of the brain execute each cognitive function, essentially dividing the brain up into sub-networks (151). When we look at these sub-networks, and the functional connectivity sub-networks described above, they are very similar; if two brain regions are in the same functional connectivity sub-network, they are also likely to be in the same cognitive function sub-network. Thus, when we claim that different parts of the brain do different things, we now know how to divide the brain up into those parts–functional connectivity sub-networks. This knowledge is essentially the beauty and utility of the author topic-model combined with the explanatory power of knowing the neural mechanism–coherent and unique activity–that defines the boundaries of the cognitive functions that the author-topic model discovered.

While it is certainly interesting that each sub-network, which coherently and uniquely activates and deac- tivates together, executes a discrete function, how dependent is the processing of one sub-network on the other sub-networks(26)? One computational way to test this is to ask if adding more nodes or sub-networks to a network increases the computational complexity of the network. Thus, we grew (i.e., added in nodes and edges) a network of concepts and relationships between concepts (i.e., 'dog' and 'cat' are connected in the network, as they are both animals), and measured the computational complexity of an inference in the network as the network grew. The relationship between the size of the network and computational complexity was logarithmic. This suggests that, at a certain point, even in a network with many edges and nodes, computationally complexity does not increase as more nodes are added. However, we wanted to test this in the brain. We used a very similar logic: when more sub-networks are engaged in a task, if their processing is dependent, then activity should increase in all the sub-networks as more sub-networks become engaged. For example, if the motor sub-network is engaged, and then the memory sub-network becomes engaged, does the motor network have to process any of the additional information that is computed by the memory network? Does any part or parts of the brain have to? We used network science to quantify which regions do and do not increase their activity as more sub-networks are engaged in a task. We found that

nodes that are only connected to a single sub-network (provincial hubs or regular nodes) do not. This suggests that most of the brain's processing is modular, in that it is mostly informationally encapsulated or autonomous.

However, nodes that are connected to many different sub-networks (connector nodes) do increase their activity as more sub-networks are engaged in a task. Thus, connector nodes are sensitive to the amount of global information being processed. Integration is a computation that is necessarily more complex as more sub-networks are active, because there is simply more information that must be integrated. Thus, the increased activity of connector nodes as more sub-networks are engaged is highly consistent with their putative integrative role. Another role that fits with this evidence is that connector nodes coordinate the interactions between sub-networks. When there are more networks to coordinate, the computation is necessarily more complex.

Integration in Modular Networks

We found evidence that suggests that connector nodes integrate information or coordinate communication across the brain. An additional and more speculative notion that these analyses generated is that connector nodes are actually necessary to maintain a modular network. This notion also has support from lesion evidence, which shows that the extent to which the network's connectivity is modular (i.e., how unique or different the activity profiles of each sub-network are from one another) decreases when connector nodes are damaged (152). Essentially, if you do not have connector nodes to integrate across the modular sub-networks, the sub-networks interact directly in a more holistic and less modular fashion. Thus, we wanted to gain more insight into the role of connector nodes in maintaining the modular nature of the network. We did this in three different ways.

First (Chapter 2.1), we investigated the interconnectivity of connector nodes, along with high degree nodes (the nodes with the most edges or connections). Both connector nodes and high degree nodes have been proposed to perform integrative and coordinative functions. One property that has been claimed to be evidence for high degree node's integrative and coordinative properties is the fact that they form a club of highly interconnected nodes termed the rich club (142). We measured the interconnectivity of both connector nodes and high degree nodes. We looked at a variety of networks–the c. elegans, the macaque brain, the human brain, the United States power grid, and global air traffic. In all networks, connector nodes were more strongly interconnected than high degree nodes. Moreover, we found that in silico damage to connector nodes in each club decreases the network's efficiency more than damage to the high degree nodes. Critically, we found that, if we simulate natural selection in networks for a balance between modularity and efficient integration, a highly interconnected club of connector nodes, but not high degree nodes, emerges. Thus, the connector nodes and their interconnectivity potentially evolved via selective pressures

that favored both modularity and efficient integration.

Next, we dug deeper into the notion that connector nodes both integrate information across sub-networks and coordinate connectivity between sub-networks, and that these two functions serve to maintain the modular structure of the brain's functional network. The logic was that, when connector nodes are connected to many sub-networks, they can integrate information and coordinate connectivity optimally, allowing other regions to perform more modular local processing. Across subjects, we found that the increased strength of connector nodes' connectivity to many sub-networks is correlated with increases in the modularity of the global network. Moreover, we found that the enhanced modular architecture afforded by strong connector nodes is accompanied by enhanced cognitive performance (i.e., fitness). These analyses further demonstrated that connector nodes integrate information across sub-networks and coordinate the connectivity between sub-networks, which maintains the modular structure of the network. Additionally, our analyses showed that connector nodes function is critical to optimal cognitive performance.

Finally, in order to further understand the connectivity of connector nodes, we studied the functional connectivity of the fronto-parietal sub-network, which contains the highest concentration of connector nodes, and the thalamus, a sub-cortical structure that has been shown to exhibit connector node properties(197) We found that every sub-network in the brain has a distinct region in the fronto-parietal sub-network for which it is strongly functionally connected. These distinct regions become interdigitated in the fronto-parietal sub-network, forming a gradient towards convergent regions where the fronto-parietal sub-network is functionally connected to many different sub-networks. Moreover, for each sub-network's region in the fronto-parietal network, there is a corresponding region in the thalamus that that sub-network is strongly functionally connected to. For each sub-network, these two regions (fronto-parietal and thalamic) are functionally connected. Finally, functional connectivity between the fronto-parietal network and cortical sub-networks and the functional connectivity between the thalamus and cortical sub-networks similarly predict the functional connectivity between sub-networks across individuals.

This connectivity pattern is a more granular look into what allows connector nodes to be integrative and coordinative. The combination of discrete and convergent connectivity in the fronto-parietal sub-network allows information flow to be bi-directional. Information from all sub-networks can be transferred to the fronto- parietal sub-network via distinct pathways that exists between sub-networks and the fronto-parietal network. Information is then likely integrated in the convergent regions of the fronto-parietal sub-network, and can then be sent back to specific sub-networks via the distinct pathways that exist between the fronto-parietal sub-network and specific sub-networks. The thalamus likely plays a critical role, potentially mediating connectivity between the fronto-parietal network and other networks, aiding in information integration and signal propagation

across cortex.

More speculatively, it has been suggested that thalamocortical interactions provide synchronized signals across cortex that coordinate the development of cortical sub-networks, and these signals aid in organizing connectivity across the massively expanded human cortex (220). The fronto-parietal- thalamic connectivity, and their similar relationship to cortical-cortical connectivity, found here supports this idea. This connectivity pattern also generates the idea that an evolutionary older integrative structure—the thalamus—might have been "copied" (i.e., it performs similar functions and has similar connectivity patterns) and expanded in the cortex as the fronto-parietal sub-network, and these two integrative structures now function together. Thus, the fronto-parietal network might have been naturally selected because it allows for greater control over the connectivity between sub-networks both during development and ongoing cognition, allowing for more complex computations to occur. It is then no surprise that this network is often referred to as a "control" network and its activity is related to "top down" regulation of other networks. It is also not surprising that similar global network changes (decreased modularity) occur when regions of the thalamus and the fronto-parietal network that are functionally connected to many sub-networks are damaged(152, 197).

Network Discovery and Models

Given the focus of many of the analyses on sub-networks and connector nodes, we developed a new method for capturing sub-networks and connector nodes. Typically, a single set of nodes is used for a large group of subjects. However, every brain is slightly differently; thus, for each subject, we developed a set of nodes that ideally work for that subject's brain. Just as a custom orthotic fits better than the one off the shelf, customs sets of nodes that fit each subject are better than a single set of nodes for all subjects. This increased fit allows us to best measure the network structure of different brains. Moreover, in order to still observe group-level network properties, we created an algorithm that averages the information across these different sets of nodes, giving us sub-networks and the location of connector nodes at the highest resolution possible. Under this method, functional connectivity sub-network's boundaries align even better with cognitive functions' (from the author-topic model) boundaries than we had previously found.

We also developed a new method for reference inference–predicting which cognitive functions a subject is executing based on that subject's brain activity. Typically, reverse inference is executed by observing that a brain region that is associated with a particular cognitive function is active, and thus that that subject is executing that cognitive function. However, this is simply looking at the flawed look- up dictionary I presented in the Introduction. Cognition arises from interacting and potentially overlapping sub-networks. Thus, we utilized the author-topic model of cognitive functions described above, which models brain activity as the joint combination of sub-

networks, to do reverse network inference–what sub-networks are engaged together? Thus, the reverse inference method reflects how the brain actually works.
We were able to accurately model individual frames of data, suggesting that this network model of the brain captures how the brain activates at any given time. Moreover, we were able to predict which task the subject was in almost 100 percent of the time. Finally, we used this model to replicate our previous findings; when more sub-networks are engaged, activity is greater at connector node regions. As the previous analysis looked across group-level activation maps, one could only conclude that, in general across tasks, when more cognitive functions are required, activity at connector nodes (essentially, the fronto-parietal sub-network) is increased. However, we demonstrated here, with increased spatial resolution, that this phenomenon exists on the sub-second temporal resolution in individual subjects in different tasks.
Finally, we used the method to detect brain states–patterns of brain activity that repeat during cognition. Each state's boundaries (where activity transitions from high to low) overlap highly with functional connectivity sub-networks. All states were either a single functional connectivity sub-network or a combination of functional connectivity sub-networks. While most states were present in all tasks, each task typically had 1-2 states that were unique. This suggests that we observe functional connectivity sub-networks because there is repeated transient activations of each network or unique combinations of networks. Thus, sub-networks do not only emerge as units because of their coherent and unique activity across time. Sub-networks also activate as potentially overlapping units at any given time.

Conclusion

Brain regions cluster into sub-networks based on their unique and coherent fluctuating activity. Each of these sub-networks executes a discrete cognitive function. Each sub-network's processing is modular, in that the computational complexity of processing in that sub-network is not dependent on processing in other sub-networks. Somewhat counterintuitively, what appears to allow for this modular processing is a select set of nodes that have connectivity to many sub-networks in the brain and form a tightly interconnected club. These nodes are connector nodes. This mix of interconnectivity and wide connectivity across the brain allows connector nodes to integrate across the sub-networks and coordinate information transfer between sub-networks, maintaining the modular structure of the brain. Why does the brain function like this? It appears that a club of highly interconnected connector nodes is nature's solution to achieve modularity–which is computationally attractive–and integration–which is required for more complex computations. We observe this solution in many different species, and even in man-made networks like the power grid. Strong connector nodes actually increase the modularity of the network and the cognitive fitness of the subject. At a more granular level, the combination of discrete and convergent connectivity of connector nodes to the rest of the brain allows for integration and for information flow to be bi-directional between connector nodes and other sub-networks. Cortical connector

nodes also function with, and might be an evolutionary “copy” of, an evolutionary older brain structure–the thalamus, which also exhibits connector node properties. Cognition appears to arise from the integrated and coordinated activity of multiple functional sub-networks. The brain is not like a network. It is a network.

References

1. Lloyd G (2007) Pneuma between body and soul. *Journal of the Royal Anthropological Institute* 13(s1):S135–S146.

2. Debru A (2015) Galen's Approach to Anatomy and the Soul. *History of Medicine/ru* 2(2). doi:10.17720/2409-5834.v2.2.2015.12e.

3. Finger S (2005) Andreas Vesalius: The New “Human” Neuroanatomy. *Minds Behind the Brain* (Oxford University Press), pp 53–68.

4. Singer C (1943) To Vesalius on the fourth centenary of his De Humani Corporis Fabrica. *J Anat* 77(Pt 4):261–265.

5. Lyons SL (2010) *Species, Serpents, Spirits, and Skulls* (SUNY Press).

6. Lichteim L (1885) On Aphasia. *Brain* 7(4):433–484.

7. Schomers MR, Garagnani M, Pulvermüller F (2017) Neurocomputational Consequences of Evolutionary Connectivity Changes in Perisylvian Language Cortex. *Journal of Neuroscience* 37(11):3045–3055.

8. Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532(7600):453–458.

9. Quine WV (1969) Natural Kinds. *Essays in Honor of Carl G. Hempel*, ed Rescher N (Springer Netherlands, Dordrecht), pp 5–23.

10. Fodor J (1974) Special sciences (or: The disunity of science as a working hypothesis). *Synthese* 28(2):97–115.

11. Guest O, Love BC (2017) What the success of brain imaging implies about the neural code. *eLife* 6:49.

12. O'Keefe J, Conway DH (1978) Hippocampal place units in the freely moving rat: why they fire where they fire. *Exp Brain Res* 31(4):573–590.

13. O'Keefe J (1976) Place units in the hippocampus of the freely moving rat. *Exp Neurol* 51(1):78–109.

14. Tambini A, Davachi L (2013) Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences* 110(48):19591–19596.

15. Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during

working memory. *Trends in Cognitive Sciences* 7(9):415–423.

16. Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes J-D (2017) The Distributed Nature of Working Memory. *Trends in Cognitive Sciences* 21(2):111–124.

17. Coull JT, Nobre AC (2008) Dissociating explicit timing from temporal expectation with fMRI. *Current Opinion in Neurobiology* 18(2):137–144.

18. Breska A, Ivry RB (2016) Taxonomies of timing: where does the cerebellum fit in? *Current Opinion in Behavioral Sciences* 8:282–288.

19. Canolty RT, et al. (2006) High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313(5793):1626–1628.

20. Cole MW, et al. (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience* 16(9):1348–1355.

21. Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE (2014) Intrinsic and Task-Evoked Network Architectures of the Human Brain. 83(1):238–251.

22. Bassett DS, Gazzaniga MS (2011) Understanding complexity in the human brain. *Trends in Cognitive Sciences* 15(5):200–209.

23. Nguyen JP, et al. (2016) Whole-brain calcium imaging with cellular resolution in freely behaving Caenorhabditis elegans. *Proc Natl Acad Sci USA* 113(8):E1074–81.

24. Russell SJ, Norvig P (2010) *Artificial Intelligence* (Prentice Hall).

25. Fodor J (1983) *The modularity of mind: An essay on faculty psychology.* (Cambridge) Available at: http://scholar.google.com/scholar?q=related:zNtQfHIOd3IJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

26. Fodor J (1983) *The Modularity of Mind* (MIT Press).

27. Hume D, Norton DF, Norton MJ (2000) *A treatise of human nature* (Oxford University Press, USA).

28. Koller D, Friedman N (2009) *Probabilistic Graphical Models* (MIT Press).

29. Albert R, Jeong H, Barabási A-L (1999) Internet: Diameter of the World-Wide Web. *Nature* 401(6749):130–131.

30. Pearl J (2014) *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann).

31. Koller D, Lerner U, Angelov D (1999) A General Algorithm for Approximate Inference and Its Application to Hybrid Bayes Nets. *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*:324–333.

32. Kevin Murphy YW (2000) The Factored Frontier Algorithm for Approximate Inference in DBNs. *Proceedings of the th Confer- ence on Uncertainty in AI*:378–385.

33. Speer R, Havasi C (2013) ConceptNet 5: A Large Semantic Network for Relational Knowledge. *The People's Web Meets NLP*, Theory and Applications of Natural Language Processing. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 161–176.

34. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL (2000) The large-scale organization of metabolic networks. *Nature* 407(6804):651–654.

35. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review* 29(4):251–262.

36. Meunier D, Bullmore ET (2010) Modular and hierarchically modular organization of brain networks. *Front Neurosci* 4:200.

37. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69(2 Pt 2):026113.

38. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008(10):P10008.

39. Steyvers M, Tenenbaum JB (2010) The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science* 29(1):41–78.

40. Sporns O, Honey CJ, Kötter R (2007) Identification and Classification of Hubs in Brain Networks. *PLOS ONE* 2(10):e1049.

41. Ellis AW, Lambon Ralph MA (2000) Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: insights from connectionist networks. *J Exp Psychol Learn Mem Cogn* 26(5):1103–1123.

42. Ihmels J, Friedlander G, Bergmann S, Sarig O (2002) Revealing modular organization in the yeast transcriptional network. *Nature*.

43. Newman MEJ (2006) Modularity and Community Structure in Networks. *Proc Natl Acad Sci USA* 103(23):8577–8582.

44. Wagner GP, Pavlicev M, Cheverud JM (2007) The road to modularity. *Nature Reviews Genetics* 8(12):921–931.

45. Chen ZJ, He Y, Rosa-Neto P, Germann J, Evans AC (2008) Revealing Modular Architecture of Human Brain Structural Networks by Using Cortical Thickness from MRI. *Cereb Cortex* 18(10):2374–2381.

46. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10(3):186–198.

47. Bassett DS, et al. (2010) Efficient Physical Embedding of Topologically Complex Information Processing Networks in Brains and Computer Circuits. *PLoS Comput Biol* 6(4):e1000748.

48. Oldham MC, et al. (2008) Functional organization of the transcriptome in human brain. *Nature Neuroscience* 11(11):1271–1282.

49. Hawrylycz MJ, et al. (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 489(7416):391–399.

50. Chen C-H, et al. (2012) Hierarchical Genetic Organization of Human Cortical Surface Area. *Science* 335(6076):1634–1636.

51. Richiardi J, et al. (2015) Long-term climate forcing by atmospheric oxygen concentrations. *Science* 348(6240):1238–1241.

52. Brodmann K (1909) *Vergleichende Lokalisationslehre der Groshirnrinde* (Leipzig: Barth) Available at: http://scholar.google.com/scholar?q=related:MBhmwZui5vsJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

53. Hagmann P, et al. (2008) Mapping the Structural Core of Human Cerebral Cortex. *PLOS Biology* 6(7):e159.

54. van den Heuvel MP, Sporns O (2013) An Anatomical Substrate for Integration among Functional Networks in Human Cortex. 33(36):14489–14500.

55. Bruno AM, Frost WN, Humphries MD (2015) Modular Deconstruction Reveals the Dynamical and Physical Building Blocks of a Locomotion Motor Program. 86(1):304–318.

56. Power JD, et al. (2011) Functional Network Organization of the Human Brain. *Neuron* 72(4):665–678.

57. Blank I, Kanwisher N, Fedorenko E (2014) A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal

fluctuations. *Journal of Neurophysiology* 112(5):1105–1118.

58. Bassett DS, Bullmore E (2006) Small-World Brain Networks. 12(6):512–523.

59. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* 52(3):1059–1069.

60. Sporns O (2014) Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience* 17(5):652–660.

61. Vincent JL, et al. (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447(7140):83–86.

62. Honey CJ, et al. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci USA* 106(6):2035–2040.

63. Honey CJ, Kötter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. 104(24):10240–10245.

64. Hermundstad AM, et al. (2013) Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proc Natl Acad Sci USA* 110(15):6169–6174.

65. Honey CJ, et al. (2012) Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. 76(2):423–434.

66. Wang L, Saalmann YB, Pinsk MA, Arcaro MJ, Kastner S (2012) Electrophysiological Low-Frequency Coherence and Cross-Frequency Coupling Contribute to BOLD Connectivity. 76(5):1010–1020.

67. Bassett DS, Yang M, Wymbs NF, Grafton ST (2015) Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience* 18(5):744–751.

68. van den Heuvel MP, Mandl RCW, Kahn RS, Hulshoff Pol HE (2009) Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. 30(10):3127–3141.

69. Greicius MD, Supekar K, Menon V (2009) Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex*.

70. Raichle ME (2011) The Restless Brain. *http://wwwliebertpubcom/brain* 1(1):3–12.

71. Spadone S, et al. (2015) Dynamic reorganization of human resting-state networks during visuospatial attention. *Proc Natl Acad Sci USA* 112(26):8112–8117.

72. Smith SM, et al. (2009) Correspondence of the Brain's Functional Architecture during Activation and Rest. *Proc Natl Acad Sci USA* 106(31):13040–13045.

73. Laird AR, et al. (2011) Behavioral Interpretations of Intrinsic Connectivity Networks. *http://dxdoiorg/101162/jocn_a_00077* 23(12):4022–4037.

74. Crossley NA, et al. (2013) Cognitive relevance of the community structure of the human brain functional coactivation network. *Proc Natl Acad Sci USA* 110(28):11583–11588.

75. Krienen FM, Yeo BTT, Buckner RL (2014) Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Phil Trans R Soc B* 369(1653):20130526–20130526.

76. Kiani R, et al. (2015) Natural Grouping of Neural Responses Reveals Spatially Segregated Clusters in Prearcuate Cortex. 85(6):1359–1373.

77. Romano SA, et al. (2015) Spontaneous Neuronal Network Dynamics Reveal Circuit's Functional Adaptations for Behavior. 85(5):1070–1085.

78. Power JD, Schlaggar BL, Lessov-Schlaggar CN, Petersen SE (2013) Evidence for Hubs in Human Functional Brain Networks. 79(4):798–813.

79. Guimerà R, Amaral LAN (2005) Functional cartography of complex metabolic networks. *Nature* 433(7028):895–900.

80. Guimera R, Mossa S, Turtschi A (2005) The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles.

81. Guimerà R, Sales-Pardo M, Amaral LAN (2007) Classes of complex networks defined by role-to-role connectivity profiles. *Nature Physics* 3(1):63–69.

82. Liska A, Galbusera A, Schwarz AJ, Gozzi A (2015) Functional connectivity hubs of the mouse brain. *NeuroImage* 115:281–291.

83. Scholtens LH, Schmidt R, de Reus MA, van den Heuvel MP (2014) Linking Macroscale Graph Analytical Organization to Microscale Neuroarchitectonics in the Macaque Connectome. 34(36):12192–12205.

84. Anderson ML, Kinnison J, Pessoa L (2013) Describing functional diversity of brain regions and brain networks. *NeuroImage* 73:50–58.

85. Yeo B, Krienen FM, Eickhoff SB, Yaakub SN (2014) Functional specialization and flexibility in human association cortex. *Cerebral ….*

86. Fodor J (2001) *The Mind Doesn't Work that Way* (MIT Press).

87. Raff EC, Raff RA (2000) Dissociability, modularity, evolvability. *Evolution & Development* 2(5):235–237.

88. Bolker JA (2000) Modularity in Development and Why It Matters to Evo-Devo. *American Zoologist* 40(5):770–776.

89. Fox PT, Lancaster JL (2002) Mapping context and content: the BrainMap model. *Nature Reviews Neuroscience* 3(4):319–321.

90. Shen X, Tokoglu F, Papademetris X, Constable RT (2013) Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage* 82:403–415.

91. Gordon EM, Laumann TO, Adeyemo B (2014) Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral ….*

92. Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. 33(8):1914–1928.

93. Ginestet CE, Nichols TE, Bullmore ET, Simmons A (2011) Brain Network Analysis: Separating Cost from Topology Using Cost-Integration. *PLOS ONE* 6(7):e21570.

94. Lancichinetti A, Fortunato S (2012) Consensus clustering in complex networks. 2:47.

95. Simpson SL, Moussa MN, Laurienti PJ (2012) An exponential random graph modeling approach to creating group-based representative whole-brain connectivity networks. *NeuroImage* 60(2):1117–1126.

96. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59(3):2142–2154.

97. Danon L, Díaz-Guilera A, Duch J, Arenas A (2005) Comparing community structure identification. 2005(09):P09008–P09008.

98. From sensation to cognition. (1998) From sensation to cognition. 121 (Pt 6):1013–1052.

99. Mesulam MM (1990) Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology* 28(5):597–613.

100. Posner MI, Petersen SE, Fox PT, Raichle ME (1988) Localization of cognitive operations in the human brain. *Science*.

101. Cosmides L, Tooby J (1992) Cognitive adaptations for social exchange. *The adapted mind*. Available at: http://books.google.com/books?hl=en&lr=&id=SxX4gRzOS6oC&oi=fnd&pg=PA163&dq=Cognitive+Adaptations+for+Social+Exchange&ots=Bh0m0G0NvK&sig=9p1sCMxRnbL2_vbL4tbJMFHbXYE.

102. Cajal SR *Histology of the Nervous System of Man and Vertebrates (History of Neuroscience, No 6)* (Oxford University Press) Available at: http://scholar.google.com/scholar?q=related:2HxgvniR6FkJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

103. Carruthers P (2006) *The architecture of the mind: Massive modularity and the flexibility of thought: Massive modularity and the flexibility of thought* Available at: http://scholar.google.com/scholar?q=related:Sotg9Yx2_TYJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

104. Pinker S *The Blank Slate: The Modern Denial of Human Nature (New York: Viking, 2002)* Available at: http://scholar.google.com/scholar?q=related:TIkWqGFbk3UJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

105. van den Heuvel MP, Sporns O (2013) Network hubs in the human brain. *Trends in Cognitive Sciences* 17(12):683–696.

106. Bullmore E, Sporns O (2012) The economy of brain network organization. 74:47–14.

107. Park HJ, Friston K (2013) Structural and Functional Brain Networks: From Connections to Cognition. 342(6158):1238411–1238411.

108. Deco G, Tononi G, Boly M, Kringelbach ML (2015) Rethinking segregation and integration: contributions of whole-brain modelling. 16(7):430–439.

109. Shen K, et al. (2012) Information Processing Architecture of Functionally Defined Clusters in the Macaque Cortex. 32(48):17465–17476.

110. Clune J, Mouret J-B, Lipson H (2013) The evolutionary origins of modularity. *Proc Biol Sci* 280(1755):1–9.

111. Kashtan N, Kashtan N, Alon U, Alon U (2005) Spontaneous evolution of modularity and network motifs. 102(39):13773–13778.

112. Tosh CR, McNally L (2015) The relative efficiency of modular and non-modular networks of different size. *Proc Biol Sci* 282(1802):20142568–20142568.

113. Nomura EM (2010) Double dissociation of two cognitive control networks in patients with focal brain lesions. 1–6.

114. Yeo BTT, Krienen FM, Chee MWL, Buckner RL (2014) Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex. 88(C):212–227.

115. Siegel M, et al. (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348(6241):1352–1355.

116. Shih C-T, et al. (2015) Connectomics-Based Analysis of Information Flow in the Drosophila Brain. *Current Biology* 25(10):1249–1258.

117. Mišić B, et al. (2015) Cooperative and Competitive Spreading Dynamics on the Human Connectome. 86(6):1518–1529.

118. Alavash M, Hilgetag CC, Thiel CM, Gießing C (2015) Persistency and flexibility of complex brain networks underlie dual-task interference. *Human Brain Mapping* 36(9):3542–3562.

119. Sporns O (2013) Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology* 23(2):162–171.

120. McIntosh AR (2000) Towards a network theory of cognition. *Neural networks : the official journal of the International Neural Network Society* 13(8-9):861–870.

121. Anderson ML (2014) After Phrenology. Available at: http://books.google.com/books?hl=en&lr=&id=rCjbBQAAQBAJ&oi=fnd&pg=PR7&dq=After+Phrenology+Neural+reuse+and+the+interactive+brain&ots=0-lxN5f24c&sig=WuNOz5VdmBrvNqRw_wO6s0oEZIE.

122. Bassett DS, et al. (2011) Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences* 108(18):7641–7646.

123. Gratton C, et al. (2000) On imputing function to structure from the behavioural effects of brain lesions. *Phil Trans R Soc B* 355(1393):147–161.

124. Gratton C (2013) The effect of theta-burst TMS on cognitive control networks measured with resting state fMRI. 1–14.

125. Warren DE, et al. (2014) Network measures predict neuropsychological outcome after brain injury. *Proc Natl Acad Sci USA* 111(39):14247–14252.

126. Vaishnavi SN, et al. (2010) Regional aerobic glycolysis in the human brain. 107(41):17757–17762.

127. Crossley NA, et al. (2014) The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* 137(8):2382–2395.

128. Dynamic reconfiguration of frontal brain networks during executive cognition in humans (2015) Dynamic reconfiguration of frontal brain networks during executive cognition in humans. 112(37):11678–11683.

129. Rosvall M, Rosvall M, Bergstrom CT, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. 105(4):1118–1123.

130. Lancichinetti A, Fortunato S (2009) Community detection algorithms: A comparative analysis. *Phys Rev E* 80(5):056117–11.

131. Desposito M, Ballard D, Aguirre GK, Zarahn E (1998) Human Prefrontal Cortex Is Not Specific for Working Memory: A Functional MRI Study. *NeuroImage* 8(3):274–282.

132. Henson R (2005) What can functional neuroimaging tell the experimental psychologist? *The Quarterly Journal of Experimental Psychology Section A* 58(2):193–233.

133. Poldrack RA (2006) Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences* 10(2):59–63.

134. Poldrack RA, Poldrack RA (2010) Mapping Mental Function to Brain Structure: How Can Cognitive Neuroimaging Succeed? *Perspectives on Psychological Science* 5(6):753–761.

135. Machery E (2005) Concepts Are Not a Natural Kind*. *Philosophy of science* 72(3):444–467.

136. Piccinini G, Scott S (2006) Splitting Concepts*. *Philosophy of science* 73(4):390–409.

137. Churchland PM (1981) Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*. Available at: http://www.jstor.org/stable/2025900.

138. Churchland PM (1988) Perceptual plasticity and theoretical neutrality: A reply to Jerry Fodor. *Philosophy of science*. Available at: http://www.jstor.org/stable/187956.

139. Poldrack RA, Yarkoni T (2016) From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. *Annu Rev Psychol* 67(1):587–612.

140. Bertolero MA, Yeo BTT, D'Esposito M (2015) The modular and integrative functional architecture of the human brain. *Proc Natl Acad Sci USA*

112(49):E6798–807.

141. Yeo BTT, et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 106(3):1125–1165.

142. van den Heuvel MP, Sporns O (2011) Rich-Club Organization of the Human Connectome. *Journal of Neuroscience* 31(44):15775–15786.

143. Schmidt R, LaFleur KJR, de Reus MA, van den Berg LH, van den Heuvel MP (2015) Kuramoto model simulation of neural hubs and dynamic synchrony in the human cerebral connectome. *BMC Neuroscience* 16(1):143.

144. Scholtens LH, Schmidt R, de Reus MA, van den Heuvel MP (2014) Linking Macroscale Graph Analytical Organization to Microscale Neuroarchitectonics in the Macaque Connectome. *Journal of Neuroscience* 34(36):12192–12205.

145. Towlson EK, Vértes PE, Ahnert SE, Schafer WR, Bullmore ET (2013) The Rich Club of the C. elegans Neuronal Connectome. *Journal of Neuroscience* 33(15):6380–6387.

146. Grayson DS, et al. (2014) Structural and Functional Rich Club Organization of the Brain in Children and Adults. *PLOS ONE* 9(2):e88297.

147. Bacik KA, Schaub MT, Beguerisse-Díaz M, Billeh YN, Barahona M (2016) Flow-Based Network Analysis of the Caenorhabditis elegans Connectome. *PLoS Comput Biol* 12(8):e1005055.

148. Alstott J, Panzarasa P, Rubinov M, Bullmore ET, Vértes PE (2014) A Unifying Framework for Measuring Weighted Rich Clubs. *Scientific Reports* 4:7258.

149. de Reus MA, van den Heuvel MP (2013) Rich Club Organization and Intermodule Communication in the Cat Connectome. *Journal of Neuroscience* 33(32):12929–12939.

150. Gratton C, Laumann TO, Gordon EM, Adeyemo B, Petersen SE (2016) Evidence for Two Independent Factors that Modify Brain Networks to Meet Task Goals. *Cell Reports* 17(5):1276–1288.

151. Yeo BTT, et al. (2015) Functional Specialization and Flexibility in Human Association Cortex. *Cereb Cortex* 26(1):465–465.

152. Gratton C, Nomura EM, Pérez F, D'Esposito M (2012) Focal Brain Lesions to Critical Locations Cause Widespread Disruption of the Modular Organization of the Brain. *Journal of Cognitive Neuroscience* 24(6):1275–1285.

153. Jacomy M, Venturini T, Heymann S, Bastian M (2014) ForceAtlas2, a

Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE* 9(6):e98679.

154. Cajal SRY (2000) *Texture of the Nervous System of Man and the Vertebrates* (Springer Vienna, Vienna) doi:10.1007/978-3-7091-6315-3.

155. Fornito A, et al. (2011) Genetic Influences on Cost-Efficient Organization of Human Cortical Functional Networks. *Journal of Neuroscience* 31(9):3261–3270.

156. Bassett DS, et al. (2009) Cognitive fitness of cost-efficient brain functional networks. *Proceedings of the National Academy of Sciences* 106(28):11747–11752.

157. Langer N, et al. (2011) Functional brain network efficiency predicts intelligence. *Human Brain Mapping* 33(6):1393–1406.

158. Khundrakpam BS, et al. (2017) Imaging structural covariance in the development of intelligence. *NeuroImage* 144(Pt A):227–240.

159. Li Y, et al. (2009) Brain Anatomical Network and Intelligence. *NeuroImage* 47:S105.

160. Watts DJ, Strogatz SH (1998) Collective dynamics of “small-world” networks. *Nature* 393(6684):440–442.

161. Zamora-López G, Chen Y, Deco G, Kringelbach ML, Zhou C (2016) Functional complexity emerging from anatomical constraints in the brain: the significance of network modularity and rich-clubs. *Scientific Reports* 6(1):38424.

162. Simon HA (1991) The Architecture of Complexity. *Facets of Systems Science* (Springer US, Boston, MA), pp 457–476.

163. Coltheart M (1999) Modularity and cognition. *Trends in Cognitive Sciences* 3(3):115–120.

164. Robinson PA, Henderson JA, Matar E, Riley P, Gray RT (2009) Dynamical Reconnection and Stability Constraints on Cortical Network Architecture. *Physical Review Letters* 103(10):108104.

165. Clune J, Mouret J-B, Lipson H (2013) The evolutionary origins of modularity. *Proc Biol Sci* 280(1755):20122863–20122863.

166. Stevens AA, Tappon SC, Garg A, Fair DA (2012) Functional brain network modularity captures inter- and intra-individual variation in working memory capacity. *PLOS ONE* 7(1):e30468.

167. Arnemann KL, et al. (2015) Functional brain network modularity predicts response to cognitive training after brain injury. *Neurology* 84(15):1568–1574.

168. Gollo LL, Zalesky A, Hutchison RM, van den Heuvel M, Breakspear M (2015) Dwelling quietly in the rich club: brain network determinants of slow cortical fluctuations. *Philos Trans R Soc Lond, B, Biol Sci* 370(1668):20140165–20140165.

169. Margulies DS, et al. (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci USA* 113(44):12574–12579.

170. White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode Caenorhabditis elegans. *Phil Trans R Soc B* 314(1165):1–340.

171. Van Essen DC, et al. (2013) The WU-Minn Human Connectome Project: an overview. *NeuroImage* 80:62–79.

172. Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29(3):162–173.

173. Young MP (1993) The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society B: Biological Sciences* 252(1333):13–18.

174. Conaco C, et al. (2012) Functionalization of a protosynaptic gene expression network. *Proc Natl Acad Sci USA* 109 Suppl 1(Supplement_1):10612–10618.

175. Henzler CM, et al. (2013) Staged miRNA re-regulation patterns during reprogramming. *Genome Biol* 14(12):R149.

176. Yeo BTT, et al. (2015) Functional Specialization and Flexibility in Human Association Cortex. *Cereb Cortex* 25(10):3654–3672.

177. Schultz DH, Cole MW (2016) Higher Intelligence Is Associated with Less Task-Related Brain Network Reconfiguration. *J Neurosci* 36(33):8551–8561.

178. Wagner GP, Zhang J (2011) The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics* 12(3):204–213.

179. Krienen FM, Yeo BTT, Ge T, Buckner RL, Sherwood CC (2016) Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proc Natl Acad Sci USA* 113(4):E469–78.

180. Hawrylycz M, et al. (2015) Canonical genetic signatures of the adult human

brain. *Nature Neuroscience* 18(12):1832–1844.

181. Kuzawa CW, et al. (2014) Metabolic costs and evolutionary implications of human brain development. *Proceedings of the National Academy of Sciences* 111(36):13010–13015.

182. Laughlin SB, de Ruyter van Steveninck RR, Anderson JC (1998) The metabolic cost of neural information. *Nature Neuroscience* 1(1):36–41.

183. Lord L-D, Expert P, Huckins JF, Turkheimer FE (2013) Cerebral Energy Metabolism and the Brain's Functional Network Architecture: An Integrative Review. *Journal of Cerebral Blood Flow & Metabolism* 33(9):1347–1354.

184. Raichle ME, Gusnard DA (2002) Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences* 99(16):10237–10239.

185. Harris JJ, Attwell D (2012) The Energetics of CNS White Matter. *Journal of Neuroscience* 32(1):356–371.

186. Rubinov M, Ypma RJF, Watson C, Bullmore ET (2015) Wiring cost and topological participation of the mouse brain connectome. *Proc Natl Acad Sci USA* 112(32):10032–10037.

187. Zalesky A, Fornito A, Bullmore E (2012) On the use of correlation as a measure of network connectivity. *NeuroImage* 60(4):2096–2106.

188. Brandes U, Delling D, Gaertler M (2008) On modularity clustering. *IEEE transactions on ….*

189. Betzel RF, Bassett DS (2016) Multi-scale brain networks. *NeuroImage*. doi:10.1016/j.neuroimage.2016.11.006.

190. Siegel JS, Mitra A, Laumann TO, Seitzman BA (2016) Data quality influences observed links between functional connectivity and behavior. *Cerebral ….*

191. Gu S, Satterthwaite TD, Medaglia JD (2015) Emergence of system roles in normative neurodevelopment.

192. Berenstein AJ, Piñero J, Furlong LI, Chernomoretz A (2015) Mining the Modular Structure of Protein Interaction Networks. *PLOS ONE* 10(4):e0122477.

193. Clauset A, Newman M, Moore C (2004) Finding community structure in very large networks. *Phys Rev E*.

194. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA (2011) Functional network organization of the human brain. *Neuron*.

195. Goldman-Rakic PS (1988) Topography of cognition: parallel distributed networks in primate association cortex. *Annu Rev Neurosci* 11(1):137–156.

196. Laird AR, Riedel MC, Okoe M, Jianu R, Ray KL (2017) Heterogeneous fractionation profiles of meta-analytic coactivation networks. *NeuroImage*.

197. Hwang K, Bertolero MA, Liu W, D'Esposito M (2016) The human thalamus is an integrative hub for functional brain networks. doi:10.1101/056630.

198. Chou SJ, Babot Z, Leingärtner A, Studer M (2013) Geniculocortical input drives genetic distinctions between primary and higher-order visual areas. ….

199. Jones EG, Powell T (2016) *An anatomical study of converging sensory pathways within the cerebral cortex of the monkey* (Brain).

200. Colby CL, Gattass R, Olson CR (1988) Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative ….*

201. Mesulam MM, Van Hoesen GW, Pandya DN (1977) Limbic and sensory connections of the inferior parietal lobule (area PG) in the rhesus monkey: a study with a new method for horseradish peroxidase histochemistry. *Brain research*.

202. Stanton GB, Cruce WLR, Goldberg ME, Robinson DL (1977) Some ipsilateral projections to areas PF and PG of the inferior parietal lobule in monkeys. *Neuroscience Letters* 6(2-3):243–250.

203. Bruce C, Desimone R (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*.

204. Bignall KE (1969) Bilateral temporofrontal projections in the squirrel monkey: origin, distribution and pathways. *Brain research*.

205. Pandya DN, Seltzer B (1982) Association areas of the cerebral cortex. *Trends Neurosci*.

206. Pandya DN, Kuypers H (1969) Cortico-cortical connections in the rhesus monkey. *Brain research*.

207. Nauta WJ (1971) The problem of the frontal lobe: a reinterpretation. *J Psychiatr Res* 8(3):167–187.

208. Siegel RM, Read HL (1997) Analysis of optic flow in the monkey parietal area 7a. *Cereb Cortex* 7(4):327–346.

209. Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the

parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*.

210. Selemon LD, Goldman-Rakic PS (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *Journal of Neuroscience*.

211. Selemon LD, Goldman-Rakic PS (1988) Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network …. *Journal of Neuroscience*.

212. Goldman-Rakic PS (1987) Circuitry of the frontal association cortex and its relevance to dementia. *Arch Gerontol Geriatr* 6(3):299–309.

213. Preuss TM, Rakic PG (1987) Crossed corticothalamic and thalamocortical connections of macaque prefrontal cortex. *Journal of Comparative …*.

214. Sharad S, et al. (2014) Towards Automated Analysis of Connectomes: The Configurable Pipeline for the Analysis of Connectomes (C-PAC). *Front Neuroinform* 8. doi:10.3389/conf.fninf.2014.08.00117.

215. Avants BB, et al. (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* 54(3):2033–2044.

216. Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*.

217. Hyvärinen J (1982) Functional Role of Parietal Cortex. *The Parietal Cortex of Monkey and Man*, Studies of Brain Function. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 154–177.

218. Hyvärinen J (1982) Neural Connections in the Posterior Parietal Lobe of Monkeys. *The Parietal Cortex of Monkey and Man*, Studies of Brain Function. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 43–47.

219. Hyvärinen J (1982) *Posterior parietal lobe of the primate brain.* (Physiological Reviews).

220. Buckner RL, Krienen FM (2013) The evolution of distributed association networks in the human brain. *Trends in Cognitive Sciences* 17(12):648–665.

221. Jones EG (2001) The thalamic matrix and thalamocortical synchrony. *Trends Neurosci* 24(10):595–601.

222. Sherman SM, Guillery RW (2013) *Functional Connections of Cortical Areas* (MIT Press).

223. Sherman SM (2007) The thalamus is more than just a relay. *Current Opinion in Neurobiology* 17(4):417–422.

224. Giguere M, Goldman-Rakic PS (1988) Mediodorsal nucleus: areal, laminar, and tangential distribution of afferents and efferents in the frontal lobe of rhesus monkeys. *J Comp Neurol* 277(2):195–213.

225. Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4–26.

226. Sherman SM (2016) Thalamus plays a central role in ongoing cortical functioning. *Nature Neuroscience* 19(4):533–541.

227. Smith SM, et al. (2009) Correspondence of the Brain's Functional Architecture during Activation and Rest. *Proc Natl Acad Sci USA* 106(31):13040–13045.

228. Cole MW, Pathak S, Schneider W (2010) Identifying the brain's most globally connected regions. *NeuroImage* 49(4):3132–3148.

229. Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9(1):357–381.

230. Laird AR, Lancaster JL, Fox PT (2005) BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics* 3(1):65–78.

231. Holmes AJ, et al. (2015) Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Sci Data* 2:150031.

232. Nooner KB, et al. (2012) The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. *Front Neurosci* 6. doi:10.3389/fnins.2012.00152.

233. Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 12(1):26–41.

234. Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R (1996) Movement-related effects in fMRI time-series. *Magn Reson Med* 35(3):346–355.

235. Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* 37(1):90–101.

236. Lancichinetti A, Fortunato S (2012) Consensus clustering in complex networks. *Scientific Reports*.

237. Yuan R, Di X, Taylor PA, Gohel S, Tsai YH (2016) Functional topography of the thalamocortical system in human. *Brain Structure and …*.

238. Arcaro MJ, Pinsk MA, Kastner S (2015) The Anatomical and Functional Organization of the Human Visual Pulvinar. *J Neurosci* 35(27):9848–9871.

239. Behrens TEJ, et al. (2003) Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* 6(7):750–757.

240. Kumar VJ, van Oort E, Scheffler K, Beckmann CF (2017) Functional anatomy of the human thalamus at rest. *NeuroImage*.

241. Krauth A, et al. (2010) A mean three-dimensional atlas of the human thalamus: generation from multiple histological data. *NeuroImage* 49(3):2053–2062.

242. Morel A, Magnin M, Jeanmonod D (1997) Multiarchitectonic and stereotactic atlas of the human thalamus. *Journal of Comparative …*.

243. Zhang D, et al. (2008) Intrinsic functional relations between human cerebral cortex and thalamus. *Journal of Neurophysiology* 100(4):1740–1748.

244. Newman MEJ (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 103(23):8577–8582.

245. Kumar V, Mang S, Grodd W (2015) Direct diffusion-based parcellation of the human thalamus. *Brain Structure and Function*.

246. Ji B, et al. (2016) Dynamic thalamus parcellation from resting-state fMRI data. *Human Brain Mapping* 37(3):954–967.

247. Kim D-J, Park B, Park H-J (2013) Functional connectivity-based identification of subdivisions of the basal ganglia and thalamus using multilevel independent component analysis of resting state fMRI. *Human Brain Mapping* 34(6):1371–1385.

248. Johansen-Berg H, et al. (2005) Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cereb Cortex* 15(1):31–39.

249. Sporns O, Betzel RF (2016) Modular brain networks. *Annu Rev Psychol*.

250. Seeley WW, Menon V, Schatzberg AF (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of …*.

251. Dosenbach NUF, et al. (2007) Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci USA* 104(26):11073–11078.

252. Saalmann YB, Pinsk MA, Wang L, Li X (2012) The pulvinar regulates information transmission between cortical areas based on attention demands. *Science*.

253. Theyel BB, Llano DA, Sherman SM (2010) The corticothalamocortical circuit drives higher-order cortex in the mouse. *Nature Neuroscience*.

254. Adams MM, Hof PR, Gattass R, Webster MJ, Ungerleider LG (2000) Visual cortical projections and chemoarchitecture of macaque monkey pulvinar. *J Comp Neurol* 419(3):377–393.

255. Saalmann YB, Kastner S (2011) Cognitive and perceptual functions of the visual thalamus. *Neuron*.

256. McFarland NR, Haber SN (2002) Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. *Journal of Neuroscience*.

257. Markowitsch HJ, Irle E, Emmans D (1987) Cortical and subcortical afferent connections of the squirrel monkey's (lateral) premotor cortex: evidence for visual cortical afferents. *Int J Neurosci* 37(3-4):127–148.

258. Zhou H, Schafer RJ, Desimone R (2016) Pulvinar-Cortex Interactions in Vision and Attention. *Neuron* 89(1):209–220.

259. de Bourbon-Teles J, et al. (2014) Thalamic control of human attention driven by memory and learning. *Curr Biol* 24(9):993–999.

260. Wimmer RD, Schmitt LI, Davidson TJ, Nakajima M (2015) Thalamic control of sensory selection in divided attention. *Nature*.

261. Gratton C, Lee TG, Nomura EM, D'Esposito M (2013) The effect of theta-burst TMS on cognitive control networks measured with resting state fMRI. *Front Syst Neurosci* 7:124.

262. Carrera E, Tononi G (2014) Diaschisis: past, present, future. *Brain* 137(Pt 9):2408–2422.

263. Buckner RL, Krienen FM, Yeo B (2013) Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience*.

264. Choi EY, Yeo B, Buckner RL (2012) The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*.

265. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT (2011) The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 106(5):2322–2345.

266. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15(1):273–289.

267. Desikan RS, et al. (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31(3):968–980.

268. Valencia M, et al. (2009) Complex modular structure of large-scale brain networks. *Chaos* 19(2):023119.

269. Meunier D, Lambiotte R, Fornito A, Ersche KD, Bullmore ET (2009) Hierarchical modularity in human brain functional networks. *Front Neuroinform* 3:37.

270. Shen X, Papademetris X, Constable RT (2010) Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *NeuroImage*.

271. Power JD, Schlaggar BL, Petersen SE (2015) Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage*.

272. Power JD, Schlaggar BL, Petersen SE (2014) Studying brain organization via spontaneous fMRI signal. *Neuron*.

273. Sohn WS, et al. (2015) Influence of ROI selection on resting state functional connectivity: an individualized approach for resting state fMRI analysis. *Front Neurosci* 9(37):550.

274. Mohr P, Nagel IE (2010) Variability in brain activity as an individual difference measure in neuroscience? *Journal of Neuroscience*.

275. Hahn A, Kranz GS, Sladky R, Ganger S (2015) Individual diversity of functional brain network economy. *Brain*.

276. Wang D, Liu H (2014) Functional connectivity architecture of the human brain not all the same. *The Neuroscientist*.

277. Mueller S, Wang D, Fox MD, Yeo B, Sepulcre J (2013) Individual variability in functional connectivity architecture of the human brain. *Neuron*.

278. Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ (2015) Parcellating cortical functional networks in individuals.

279. Lancichinetti A, Fortunato S (2011) Limits of modularity maximization in community detection. *Phys Rev E* 84(6):066122.

280. Gao JS, Huth AG, Lescroart MD, Gallant JL (2015) Pycortex: an interactive surface visualizer for fMRI. *Front Neuroinform* 9(22):162.

281. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods* 8(8):665–670.

282. Rosen-Zvi M, Chemudugunta C, Griffiths T, Smyth P, Steyvers M (2010) Learning author-topic models from text corpora. *ACM Transactions on Information Systems* 28(1):1–38.

283. Pedregosa F, et al. (2011) Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12(Oct):2825–2830.

284. Fox MD, Snyder AZ, Vincent JL (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks.

285. Schoenbaum G, Takahashi Y, Liu TL (2011) Does the orbitofrontal cortex signal value? *Annals of the New …*.

286. Andersen RA, Cui H (2009) Intention, action planning, and decision making in parietal-frontal circuits. *Neuron*.

287. Chen AC, Oathes DJ, Chang C (2013) Causal interactions between fronto-parietal central executive and default-mode networks in humans.

288. Poldrack RA (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*.

289. Ariely D, Berns GS (2010) Neuromarketing: the hope and hype of neuroimaging in business. *Nature reviews neuroscience*.

290. Kaplan JT, Freedman J, Iacoboni M (2007) Us versus them: Political attitudes and party affiliation influence neural response to faces of presidential candidates. *Neuropsychologia*.

291. Lieberman MD, Eisenberger NI (2015) The dorsal anterior cingulate cortex is selective for pain: Results from large-scale reverse inference.

292. Bzdok D, et al. (2016) Formal Models of the Network Co-occurrence Underlying Mental Operations. *PLoS Comput Biol* 12(6):e1004994.

293. Shine JM, Koyejo O (2016) Temporal metastates are associated with differential patterns of time-resolved connectivity, network topology, and attention.

294. Liao X, Cao M, Xia M, He Y (2017) Individual differences and time-varying features of modular brain architecture. *NeuroImage*.

295. Shine JM, Koyejo O, Bell PT, Gorgolewski KJ, Gilat M (2015) Estimation of dynamic functional connectivity using Multiplication of Temporal Derivatives.

NeuroImage.

296. Power JD, Plitt M, Laumann TO, Martin A (2017) Sources and implications of whole-brain fMRI signals in humans. *NeuroImage*.

297. Boly M, et al. (2009) Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. *Human Brain Mapping* 30(8):2393–2400.