From connectome to cognition: The search for mechanism in human functional brain networks

Ravi D. Mill[*], **Takuya Ito**, and **Michael W. Cole**

Center for Molecular and Behavioral Neuroscience, Rutgers University, USA

Abstract

Recent developments in functional connectivity research have expanded the scope of human neuroimaging, from identifying changes in regional activation amplitudes to detailed mapping of large-scale brain networks. However, linking network processes to a clear role in cognition demands advances in the theoretical frameworks, algorithms, and experimental approaches applied. This would help evolve the field from a descriptive to an explanatory state, by targeting network interactions that can mechanistically account for cognitive effects. In the present review, we provide an explicit framework to aid this search for "network mechanisms", which anchors recent methodological advances in functional connectivity estimation to a renewed emphasis on careful experimental design. We emphasize how this framework can address specific questions in network neuroscience. These span ambiguity over the cognitive relevance of resting-state networks, how to characterize task-evoked and spontaneous network dynamics, how to identify directed or "effective" connections, and how to apply multivariate pattern analysis at the network level. In parallel, we apply the framework to highlight the mechanistic interaction of network components that remain "stable" across task domains and more "flexible" components associated with on-task reconfiguration. By emphasizing the need to structure the use of diverse analytic approaches with sound experimentation, our framework promotes an explanatory mapping between the workings of the cognitive mind and the large-scale network mechanisms of the human brain.

Key terms

Functional connectivity; dynamic connectivity; directed connectivity; MVPA; multi-modal neuroimaging; computational modeling

---

[*]Corresponding Author: Ravi D. Mill, Center for Molecular and Behavioral Neuroscience, Rutgers University. 197 University Ave, Newark, NJ 07120, USA. rdm146@rutgers.edu.

1. A framework for mechanistic discovery in network neuroscience

Since the emergence of human functional neuroimaging, researchers have sought the optimal analytic framework to link non-invasively acquired brain data to cognitive function. An initial focus on changes in regional activation amplitudes (Friston et al., 1994; Kanwisher, 2010) has given way to examination of functional connectivity (FC) between regions and large-scale networks of regions (Biswal, Yetkin, Haughton, & Hyde, 1995; Medaglia, Lynall, & Bassett, 2015; Petersen & Sporns, 2015; Raichle, 2010; Sporns, 2014). This trend provides a macroscopic parallel to the search for the "neural code" in animal neurophysiology, which has also transitioned from analysis of spiking in individual neurons to deciphering patterns of spatiotemporal synchronization in neuronal populations (Fries, 2005; Goldman-Rakic, 1988; Kumar, Rotter, & Aertsen, 2010; Laughlin & Sejnowski, 2003; M. Siegel, Buschman, & Miller, 2015; Vaadia et al., 1995). Both human and animal literatures have drawn inspiration from earlier connectionist models linking cognition to interactions within simulated networks (Fodor & Pylyshyn, 1988; Rumelhart et al., 1986). Rather than being confined to abstract computational models or animal neurophysiology, recent improvements in technology (e.g. the spatial and temporal resolution of functional magnetic resonance imaging, fMRI; Duyn, 2012; Feinberg et al., 2010; Lewis, Setsompop, Rosen, & Polimeni, 2016), methodology (e.g. source modeling of magneto-/electro-encephalography data, MEG/EEG; Brookes et al., 2011; Hipp, Hawellek, Corbetta, Siegel, & Engel, 2012), and research strategy (e.g. "big data" initiatives such as the Human Connectome Project, HCP; Van Essen et al., 2013) have rendered questions over the network architecture of the human brain more amenable to direct empirical investigation than ever before.

However, whilst current methods allow for the *descriptive* mapping of networks in unprecedented detail and complexity, there remains scope for a clearer *explanatory* understanding of what and how these networks compute. In this review, we outline a framework that targets the discovery of "network mechanisms" to enable this deeper understanding. We operationally define a network mechanism as a set of interactions amongst large-scale neural populations (e.g. cortical regions) that take part in an explanation of a cognitive phenomenon. Concretely, we postulate that the explanatory relevance of a network mechanism can be clarified via two key research practices: i) careful experimental manipulations of cognitive and neural states, in combination with ii) recently developed FC estimation methods that characterize the operation of a network mechanism via its spatial, temporal and directional "components" (see Figure 1 and 2). Such network components encompass spatial topology (the spatial configuration of network interactions; see Figure 1a), temporal properties (the onset, duration and spectral features of these interactions; see Figure 1b), and directional method of communication (whether interacting regions have consistent asymmetries or lags in how they communicate; see Figure 1c). To clarify, "components" in our framework serve as critical details that collectively characterize the operation of a network mechanism (see Figure 2). Thus, throughout we emphasize the need to improve how hypothesized network mechanisms are experimentally tested, as well as how different components that characterize the function of a network mechanism are characterized.

The above concrete features address specific limitations in the FC literature focusing on the task-free "resting state", which we argue has provided primarily descriptive rather than explanatory insight. Our formalization of a clear framework for experimentally testing network-level mechanisms also distinguishes the present report from previous informative reviews of the human FC literature (Bullmore & Sporns, 2009; Medaglia et al., 2015; Sporns, 2014). Additionally, whilst these cited reviews share an emphasis on defining graph theory and summarizing its applications to human imaging, we opt for breadth in summarizing multiple FC analysis methods beyond graph theory. Critically, this method summary is structured by how the various analytic approaches might advance mechanistic inquiry in network neuroscience.

To clarify the importance of a mechanistic approach, we conceptualize a continuum representing the overall level of understanding of a system of interest (e.g. the brain). This continuum ranges from merely describing the system at one extreme, to explaining its full set of causal relations at the other extreme. Whilst a description of the system might be incredibly detailed and complex, this provides little insight into the causal relevance of these details without demonstrating how manipulating or perturbing the system affects its hypothesized function. Hence, testing a network mechanism via experimental manipulation of neurocognitive states offers a concrete means to clarify its function. This approach helps advance neuroscience along the explanatory continuum towards a causal and empirically verified understanding of the brain. In this conceptualization, we acknowledge that identifying a network mechanism does not by itself achieve a full causal understanding of its relations. This follows from suggestion that causal inference in complex systems such as the brain is a "messy and iterative" process (Illari & Williamson, 2012), requiring convergence across multiple analytic approaches. This convergence is necessary to provide the breadth to better characterize a mechanism's components along with empirical verification via replication of those core components across complementary methods. We emphasize that identifying network mechanisms via experimental manipulation of neural and cognitive states, in combination with methods that characterize its spatial, temporal and directional components, is key in advancing towards this idealized "causal" understanding of the brain[1].

Indeed, there are several ways in which the identification of a network mechanism can be leveraged in a broader research program that aims for causal understanding. First, network mechanisms can generate predictions for how to manipulate the system of interest via brain stimulation. Such "neural" manipulations are becoming increasingly viable given recent developments in transcranial magnetic stimulation (TMS) and transcranial direct/alternating current stimulation (TDCS/TACS). These methods offer the potential to non-invasively disrupt or enhance the connectivity between regions that are proposed to interact in a network mechanism. Linking neural manipulation of a network mechanism to accompanying cognitive and behavioral changes in experimental tasks will likely strengthen causal statements made about that mechanism.

---

[1]Note that we limit use of the term "causal" to prevent associations with the causal connectivity literature, which we consider a class of methods clarifying activation asymmetries or lags between regional time series (termed "directed functional connectivity", see section 5). Rather than highlighting these methods as preferentially important to the discovery of network mechanisms, we treat them in our framework as one of three functional components comprising a given mechanism (i.e. directional components, see Figure 1c and 2).

A second way in which network mechanisms can advance causal understanding is by informing computational models. Empirical work that provides insight into network mechanisms is necessary to constrain the set of causal factors contributing to a given cognitive phenomenon. This facilitates instantiation of that phenomenon within a simulation. In contrast, purely descriptive accounts provide insufficient insight into these causal factors and therefore fail to adequately inform simulations. Accurate neural simulations of cognitive phenomena are useful for producing artificial intelligence systems (with various practical applications; Kumaran, Hassabis, & McClelland, 2016; Piekniewski et al., 2016), as well as developing interventions in clinical samples. Importantly, a research program can utilize the same approaches that achieve these more practical outcomes to test mechanistic hypotheses that extend beyond particular applications; namely, experimental manipulations of the system (via neural and cognitive manipulations) and testing the plausibility of mechanisms in a causal sense via computational modeling.

To illustrate the "causal continuum" we advocate, consider a Pearson's correlation between fMRI BOLD time series of two regions (A-B) changing in a statistically reliable fashion across two task conditions. This would serve as a spatial component in a network mechanism, in that it suggests an explanatory link to a manipulation of cognitive state, and therefore provides greater functional insight than observing that these same regions tend to correlate during task-free rest. Nevertheless, this mechanistic component would lie at the low end of the continuum ranging from "descriptive" to "causal" insight into brain function. An advance along this continuum would result from using additional FC estimation methods to characterize more of this mechanism's temporal and directional components. For instance, the FC between the two regions might begin to reliably differ across task conditions around 300ms after stimulus onset. Furthermore, this pattern of FC change might be captured by a stable directionality from A→B that extends from 300–500ms (as revealed by a directed connectivity method applied over sliding time windows with MEG/EEG, see section 5). Further advances along the causal continuum might be achieved by replicating the recovery of this mechanism across variations in FC estimation methods and experimental paradigms, which would demonstrate robustness in explaining the cognitive state of interest. Brain stimulation methods could then strengthen the case that this network mechanism exerts a causal influence via perturbing the connectivity between these two regions, and observing concomitant disruptions in the cognitive state to which that mechanism is linked. Thus, pursuing a mechanistic program of research can readily produce progress in advancing basic cognitive neuroscience theory, although achieving a satisfactory neural explanation for a given cognitive phenomenon will undoubtedly require a long-term process.

To provide some historical context, neuroscientific mechanisms have been most commonly investigated at the cellular or microscopic level, leading to the identification of molecular mechanisms explaining the function of individual neurons (Hodgkin & Huxley, 1952) and local circuits involving small numbers of neurons (Catterall, Raman, Robinson, Sejnowski, & Paulsen, 2012; Douglas & Martin, 2004). This empirical focus builds on the philosophical concept of "supervenience" (Kim, 1993), which states that the properties at macroscopic levels of a complex system are effectively "fixed" by (or are "supervenient" on) mechanisms operating at lower microscopic levels. By extension, it is often assumed that processes operating at the micro level are most mechanistically relevant (Kim, 1993; Klee, 1984).

However, mounting theoretical and empirical evidence suggests that applying this view to complex deterministic systems such as the human brain is at best overly reductive, and at worst false. For example, recent animal neurophysiology findings highlight the adoption of neural codes by coordinated populations rather than single cells as a crucial mechanism of cognitive transfer (Fries, 2005; Gjorgjieva, Mease, Moody, & Fairhall, 2014; Kumar et al., 2010; Lopes da Silva, 1991; Montijn, Meijer, Lansink, & Pennartz, 2016). Brain simulations have gone further in linking mechanisms operating at the macro rather than micro level preferentially to the "emergence" of cognitive states (Hoel, Albantakis, & Tononi, 2013; Tononi, Sporns, & Edelman, 1999). These findings question the utility of focusing only on the activation of isolated neurons and local circuits.

Rather than dismissing the importance of micro level mechanisms, we endorse a complementary form of "weak" emergence in substantiating network-level mechanistic inquiry in the human brain (consistent with Bedau, 1997; Seth, 2008). This states that whilst macro-level properties are underpinned by microlevel interactions (consistent with supervenience), cognition emerges holistically from the collective operation of mechanisms across all spatial scales (see also Bassett & Gazzaniga, 2011), such that reducing mechanisms from the macro to the micro scale is a non-trivial operation. In other words, reducing macro-level mechanisms to their micro-level interactions might be ontologically possible, but epistemologically complex (Bedau, 1997). Hence, our network mechanisms framework both refutes reductive supervenience (prioritizing the micro scale) and a strong irreducible form of emergence (prioritizing the macro scale), in favor of a weakly emergent holistic hierarchy. This position states that macro levels emerge from micro interactions, with both micro and macro being ontologically important, but their mapping epistemologically complex. This view acknowledges the complexity of the human brain, whilst validating the need to empirically study mechanisms at all levels or organization – not just at the micro cellular scale historically prioritized in invasive animal research, but also the macro network scale now available to human neuroimaging.

We hence formalize a definition of "network mechanisms" as higher-level network interactions that build on lower-level processes conducted within discrete neural populations (themselves comprising perhaps hundreds of thousands of neurons), together producing cognition. This definition combines our above position on the brain as operating on weak emergence within a holistic hierarchy, with Craver's more abstract "compositional" framework (Craver, 2007; see also its extension by Illari & Williamson, 2012). It also accommodates the spatial resolution of current fMRI and MEG/EEG imaging modalities, which is limited to the large-scale population level. Following from our acknowledgement of the complexity in mapping micro to macro level mechanisms, we assume that micro levels can be reasonably summarized at higher spatial scales without *majorly* detracting from the explanatory potential of network mechanisms (e.g. coordinated local neuronal populations are summarized by regional fMRI BOLD amplitudes and evoked MEG/EEG fields, Hoel et al., 2013; Lopes da Silva, 1991). Hence, a mechanism serves as an explanation of a phenomenon (cognition) based on a characterization of its interacting elements (networks of brain regions; Craver, 2007). Consistent with Craver and Illari & Williamson, and many of the mechanistic frameworks proposed in the philosophy of science (e.g. Bechtel & Abrahamsen, 2005; Imai, Tingley, & Yamamoto, 2013; Pearl, 2001), we emphasize the role

of careful task manipulations in linking network interactions to statistically reliable explanations of cognitive states. We also advocate combining task manipulations with FC estimation methods that characterize different functional "components" operating within each mechanism (see Figure 1 and Figure 2). This feature of the framework is consistent with prior emphasis on elucidating "component parts and operations" of mechanisms (Bechtel & Abrahamsen, 2005; Illari & Williamson, 2012).

As a concrete extension from Craver's (and Illari & Williamson's) framework, we endorse attempts to embed network mechanisms within a broader neuroscientific hierarchy. Figure 3a depicts this hierarchy in the form of levels of mechanistic interactions occurring over increasing spatial scales – from lower cellular (micro) to higher network (macro) levels. Structuring by spatial scales provides a common platform to represent network mechanisms underlying diverse cognitive phenomena, whilst also encouraging mutual contact between micro and macro levels of inquiry. This is to help clarify the nontrivial issue of how micro and macro mechanisms relate in the brain's weakly emergent system (as highlighted above), as well as more generally harnessing important insights from decades of research conducted below the network level (in animals and via human imaging of regional activations[2]).

An application of this hierarchy of mechanisms is provided in Figure 3b, which charts empirically supported levels of mechanisms underlying working memory maintenance (Braver, Cole, & Yarkoni, 2010; Darki & Klingberg, 2015; Dumontheil et al., 2011; Miller & Cohen, 2001)[3]. The figure should highlight that the historical focus on identifying cellular mechanisms via animal neurophysiology may primarily permit insight into lower levels of this mechanistic hierarchy. This follows from methodological complications in studying large-scale networks in animals (Spira & Hai, 2013). Relative to invasive methods, human functional neuroimaging methods are optimally positioned to target network mechanisms that operate with lower-level neurophysiological mechanisms to produce cognition. Furthermore, the theoretical practice of considering the position of network mechanisms within the broader neural hierarchy can help identify neurophysiological experiments that target how lower-level mechanisms extend upwards. Similarly, linking network mechanisms to biophysically plausible neurophysiological processes would help validate the former. Hence, embedding the search for network mechanisms in a broader neural hierarchy could be beneficial to mechanistic inquiry at all spatial scales.

---

[2]Indeed, despite our focus on the network level, it could be argued that some cognitive phenomena are adequately explained by mechanisms operating at lower levels. For example, basic visual perceptual functions have been linked mechanistically to lateral inhibition of receptive fields in visual cortex neurons (Hubel & Wiesel, 1979), and spatial encoding of place has been linked mechanistically to activity in hippocampal neurons (O'Keefe & Nadel, 1978). The question of which level of mechanism is appropriate or optimal in explaining a cognitive process is not easy to answer. Indeed, these queries might distract from advancing our understanding of cognition. Rather, we endorse the more holistic view that all levels of mechanism operate collectively in the neural system, and argue that formalizing a hierarchy of mechanisms from micro to macro levels will help clarify inquiry at all levels.
[3]Whilst the depicted example centers prominently on prefrontal cortex mechanisms, conceptually similar levels of mechanism could be provided for cognitive phenomena that do not prominently engage PFC. For example, fear expression to a visual stimulus (the cognitive phenomenon to explain) can be linked to lower level encoding of visual features by neurons in visual cortex, which leads to object recognition at the population/regional-level, which in turn leads to network-level interactions between visual occipital regions, thalamus and amygdala that drive the final expression of fear (Gilmartin, Balderston, & Helmstetter, 2014; LeDoux, 2000).

2. Implications of network mechanisms in advancing human FC research

To provide an interim summary of the core tenets of our framework: i) network mechanisms are identified as interactions amongst brain regions that explain cognitive states; ii) experimental manipulations (both cognitive and neural) are key in reliably linking network interactions to a clear explanatory role in cognition; iii) task manipulations can be combined with recently developed FC estimation algorithms to characterize distinct functional components of each mechanism (spatial, temporal and directional); iv) attempts should be made to link higher-level network mechanisms to neurophysiological mechanisms operating at lower-levels (and vice versa), so as to mutually validate all levels. This framework is consistent with prior formulations in the philosophy of science literature characterizing complex systems with properties of weak emergence, albeit with concrete modifications that accommodate some of the nuances of neuroscientific research.

The study of network mechanisms can advance the field of human FC research in a number of pragmatic ways. Firstly, encouraging the analysis of networks across careful task manipulations can bridge the gap between the largely segregated task activation and FC literatures (with the latter largely focusing on task-free rest; Biswal et al., 1995). Secondly, elucidating network mechanisms underlying "healthy" cognitive function will by extension clarify how dysfunction of these mechanisms contributes to clinical conditions (Cole, Repov, & Anticevic, 2014; Craddock et al., 2013; Geerligs, Rubinov, Cam-CAN, & Henson, 2015). Finally, testing precise mechanistic hypotheses would help constrain the sizeable methodological "model space" in brain network analysis, spanning choices between multiple network node definitions, clustering algorithms and FC estimation methods. Large model spaces are increasingly the norm given the (undoubtedly fruitful) integration of methods and algorithms from fields such as computer science and engineering. Nevertheless, the increasing array of methods raises the need for principled ways to adjudicate between them. As a basis for this adjudication, we encourage consideration of the capabilities of a given network analysis method to advance an explanatory understanding of cognition *above and beyond* that which is provided by previous methods.

Critically, this latter point argues against using an algorithm purely on the basis that it has not been previously applied to human imaging data. Rather, a given method is advocated if it has clear provisions to make it more suitable to discovering network mechanisms over previous approaches. For example, a novel FC estimation method might be endorsed based on its better treatment of non-Gaussian information that is known to be functionally relevant in human imaging data (compared to methods that make Gaussian assumptions; Mumford & Ramsey, 2014). Relatedly, this mechanistic approach would ensure that the use of advanced FC methods is structured by theory-driven experimentation, rather than the inverse case of cognitive theory being tailored to the output of different methods. This advocates the selection of a method *a priori* based on its capacity for explanatory insight, rather than purely its novelty. The latter scenario increases the likelihood of *a posteriori* rationalization of an algorithm's output, which can slow the advancement of explanatory insight.

The remainder of this review outlines key methodological guidelines to constrain the search for network mechanisms. We begin by providing an overview of attempts to clarify the

cognitive relevance of functional networks mapped in the resting state (termed "resting-state networks"). Resting-state networks have become the primary research focus in the FC field, and hence for consistency we begin by considering them as potential "interacting elements" (Craver, 2007) of network mechanisms. Critically, however, we also highlight that the resting-state literature has largely provided descriptive rather than mechanistic insight into functional networks, given its primary emphasis on mapping networks during task-free rest. This arises primarily from the lack of experimental manipulation of cognitive states, leading to poor characterization of the phenomena to be explained and preventing insight into explanatory mechanisms. Nevertheless, the fruitful method development in this field can be harnessed to provide mechanistic insight. Hence, in subsequent sections, we detail how task manipulations can be combined with recently developed FC estimation methods to characterize a network mechanism's functional components – its spatial topology, temporal properties and directional asymmetries (see Figure 1 and 2)[4]. In particular, we highlight methods that can correct the present overreliance on outlining spatial components, at the expense of temporal and directional ones. In the final section, we detail recent developments in the field of multivariate pattern analysis (MVPA), which represents a broad class of algorithms that hold potential in clarifying many network components over the "univariate" methods currently adopted.

To provide an empirical application of our framework, we concurrently describe an example cognitive mechanism centering on interactions between a majority of network connections that remain spatially "stable" across task domains, and a minority of "flexible" ones (see Figure 6 for a hierarchical depiction of this mechanism). We also suggest temporal and directional components operating within this general "stable/flexible" mechanism. For example, we propose a temporal component that splits the spatially "flexible" connections into functionally distinct "slow" timescale connections linked with learning, and "fast" timescale connections linked with stimulus-evoked responding (see section 4 for further details). In keeping with Figure 3 and the accompanying text in section 1, we also make concerted efforts throughout to link this "stable/flexible" network mechanism to biophysically plausible mechanisms operating at lower spatial scales. Whilst we admit that this example network mechanism is basic, it nevertheless can serve as a useful template that can be refined to represent more functionally specific mechanisms. This furthers our main primary aim of emphasizing the utility of "network mechanisms" as a conceptual framework to guide future FC research.

3. Clarifying the cognitive relevance of resting-state networks

Thus far, the dominant focus in human FC research has been on mapping spatial patterns of fMRI BOLD synchronization in the resting state. Computing the Pearson's correlation coefficient between pairwise regional (or voxel-wise) BOLD time series, in the absence of a controlled task, has yielded a highly reproducible set of large-scale networks spanning many domains of cognitive function. These canonical resting-state networks include low-level sensory and motor networks (Biswal et al., 1995; Cordes et al., 2000), as well as higher-

---

[4]Whilst components other than spatial, temporal and directional are also likely involved in characterizing network mechanisms (e.g. neuromodulatory components), these are the three that can be most reliably measured via current non-invasive imaging modalities.

order default mode (Greicius, Krasnow, Reiss, & Menon, 2003), dorsal attention (Corbetta & Shulman, 2002) and frontoparietal control networks (Cole & Schneider, 2007; Fox et al., 2005). Subsequent refinements to this seed-based correlation approach have enabled more principled identification of functional brain regions from which to extract BOLD time series (Glasser et al., 2016; Power et al., 2011), and use of data-driven network definition approaches such as community detection (Power et al., 2011; Yeo et al., 2011). Graph theoretical methods have improved the quantification of large-scale properties of these resting-state networks, leading to identification of highly inter-connected network "hubs" (e.g. in the frontoparietal control network; Cole, Pathak, & Schneider, 2010) and functionally segregated "modules" (e.g. sensory and motor networks; Bullmore & Sporns, 2009). However, despite evident refinement in how resting-state networks are described, there remains much ambiguity as to precisely how cognitive function emerges from them. This ambiguity bears critically on the search for network mechanism, and its resolution calls for more targeted attempts at studying the function and reconfiguration patterns of resting-state networks during experimentally controlled tasks.

Initial attempts to clarify the function of resting-state networks focused on their visible spatial overlap with task-evoked activation patterns (e.g. as identified by general linear model contrasts, Biswal et al., 1995). Studies that have used meta-analytic methods to identify "canonical" task activations (via cross-experiment correlations) across a number of cognitive domains have yielded high spatial similarity with their functionally linked resting-state networks (Laird et al., 2013; Smith et al., 2009). Such findings have informed a view that the spatial topology of resting-state networks reflects a "history" of common activation during cognitive processing, likely arising from long-term synaptic potentiation (Wig, Schlaggar, & Petersen, 2011). The implied longitudinal relationship between the two measures was supported by the increase in resting-state correlation observed amongst regions commonly activated during a task after it had been repeatedly performed (Lewis, Baldassarre, Committeri, Romani, & Corbetta, 2009). However, other studies have highlighted divergences, in that regions commonly activated in a task activation contrast often correlate with separate resting-state networks (Barredo, Oztekin, & Badre, 2015; Mill, Cavin, & O'Connor, 2015), with this lack of correspondence enhanced by more functionally-specific contrasts (Mennes, Kelly, Colcombe, Castellanos, & Milham, 2013). These latter findings suggest that the relationship between resting-state networks and task activations is not one-to-one. Furthermore, the cited meta-analytic approach is poorly disposed to provide insight into network mechanisms with clear links to cognitive function. Such comparison of the spatial overlap between resting-state networks and task activation patterns between separate subject groups provides evidence only of a coarse associative (rather than explanatory) relationship between resting-state networks and on-task cognition.

With a view to more rigorous methods, recent work has compared the similarity of networks estimated across rest and task states in the same sample of subjects. A pair of independent studies converged in demonstrating a high degree of spatial overlap between functional networks estimated at rest and across a number of distinct tasks (Cole et al., 2014; Krienen et al., 2014). The shared variance in FC across rest and multiple task states reached as high as 80%, suggesting the existence of a "stable" (or "intrinsic") resting-state network architecture (see Figure 4a). This stable FC architecture might be the component driving the

noted high correspondence between resting-state networks and task activations (i.e. those identified by "general" rather than functionally-specific GLM contrasts, Mennes et al., 2013). However, the residual 20% of task FC variance that is unexplained by rest FC might still be highly relevant to more domain-specific cognitive processing. Indeed, both studies observed reliable differences in pairwise FC across different tasks, suggesting more "flexible" network components that underpin network reconfiguration via variation of a small proportion of connections. With reference to our earlier definition of network mechanisms (see Section 1), we theorize that cognition emerges from this general interaction between "stable" and "flexible" spatial network components.

Whilst this stable/flexible distinction is likely of broad importance to diverse forms of cognition, the challenge now is to better understand the function of each spatial component and how they interact. To this end, evidence from more targeted task manipulations suggests that "flexible" components might be preferentially localized to the frontoparietal control resting-state network. Prior research has shown that this network has the highest variability in its global connectivity patterns across multiple tasks in comparison to other resting-state networks (Cole et al., 2013). This insight emerged from combining a generalized psychophysiological interaction (gPPI) approach to task FC estimation (McLaren, Ries, Xu, & Johnson, 2012) with a graph theoretical measure of "global variable connectivity" (Cole et al., 2013). These methods were applied in a paradigm that instantiated multiple task states (the permuted rule operations paradigm, PRO; Cole, Bagic, Kass, & Schneider, 2010). Hence, as per our mechanistic framework, combining advanced FC methods with careful task manipulation served to extend prior graph theoretical description of "hubs" in the resting state (e.g. Cole et al., 2010) to reveal a more explanatory cognitive role. Cognitive states might emerge in a general sense from interactions between domain-general "flexible hub" regions in the frontoparietal control network and domain-specific content regions in sensory-motor networks.

The extent to which the "stable" component of resting-state networks is driven by underlying structural connectivity has also been questioned. Comparisons of structural networks identified via diffusion MRI (dMRI) demonstrate overlap with the spatial topology of resting-state networks, both in terms of their pairwise connections (Hagmann et al., 2008) and their graph theoretical features (e.g. overlap between structural and functional hubs, van den Heuvel & Sporns, 2013). More formalized computational models have also explained a reasonable proportion of the variance in resting-state network topology from structural connectivity (Goni et al., 2014; Mišić et al., 2015). These findings suggest that structure might indeed contribute to resting-state network architecture, possibly providing the "skeleton" that constrains paths of communication within them (Petersen & Sporns, 2015). However, computational modeling has also revealed the relationship between structural and functional connectivity to be only partially explanatory, and often mediated by other neurophysiological factors (Deco, Jirsa, McIntosh, Sporns, & Kotter, 2009; Honey, Kotter, Breakspear, & Sporns, 2007). In support of the outlined stable/flexible mechanism, it might be that structural connectivity alone fails to predict more flexible functional network components that vary across different task domains[5].

To more clearly characterize task-evoked network flexibility, recent computational models have directly predicted task activation patterns (identified via general linear model contrasts across a number of distinct tasks) on the basis of resting-state FC. A set of regression models trained on a large number of FC predictors and a small number of structural predictors was able to reliably capture individual differences in task activations (Tavor et al., 2016; see Figure 4b). Notably, exclusion of the structural features did not appreciably worsen model performance, whereas exclusion of the functional features did, supporting the idea that structural connectivity fails to capture more task-specific (and in this case individualized) aspects of cognitive function.

In search of a formalized mechanism linking resting-state FC to task activations (essentially explaining how the Tavor et al result was possible), Cole and colleagues (Cole, Ito, Bassett, & Schultz, 2016) developed a computational model centering on the concept of "activity flow" – an estimate of the propagation of task-evoked regional amplitudes across resting-state networks. Formally, activity flow for a given region reduces to the sum of activity in other regions weighted by its resting-state FC with those regions (see Figure 4c). A model based on this simple estimate could explain as much as 90% of the variance in activation patterns across a number of task domains (see Figure 4d), and even predict individual differences in held-out subjects. Whilst the activity flow model is simplified in its biophysical plausibility compared to other network models (Deco, Jirsa, & McIntosh, 2013; Timme et al., 2016), it nonetheless provides an explanatory network mechanism that clarifies longstanding ambiguity over the relevance of resting-state networks to task-related cognitive processing. Specifically, the success of the activity flow mapping approach provides evidence that activity flow is a network mechanism linking functionally relevant brain region activations to stable aspects of resting-state networks.

In subsequent sections, we detail advances in analysis methods that facilitate further insight into network mechanisms. In detailing these methods, we suggest that the search for network mechanisms will require going beyond the present overreliance on one analytic approach derived from the resting-state literature, i.e., undirected FC computed (typically via Pearson's correlation but also via other methods e.g. partial correlation, regression) on averaged (univariate) timeseries over large epochs and in one isolated imaging modality (typically fMRI). Referring back to our definition in section 1 (see Figure 1 and Figure 2), such methodological advances will be necessary to elucidate network components other than spatial topology, and hence enrich the characterization of each network mechanism. We begin by highlighting the need to capture network dynamics.

4. Capturing functionally relevant network dynamics

In keeping with the dominant approach in FC research, the previous section characterized the cognitive relevance of resting-state networks in terms of a single functional component – spatial topology i.e. the spatial pattern of connections between brain regions. However, it is

---

[5]Note that the failure to predict structural from functional connectivity could also arise from methodological issues, such as the use of network estimation methods (applied to both functional and structural connectivity) that do not clearly establish "direct" versus "indirect"/mediated connections. Problems with conclusively identifying structural connections in humans using dMRI have also been highlighted (Jones, Knösche, & Turner, 2013).

likely that temporal components of resting-state networks also provide critical insight into their function i.e. the spatial pattern of connections *and when that pattern emerges during cognition* (see Figure 1b). The importance of capturing functional dynamics is emphasized by the increased interest in deciphering the "temporal code" from recordings of neuronal populations made in animals (Fries, 2005; Kumar et al., 2010; Siegel et al., 2015). Harnessing recent developments in dynamic FC estimation therefore permits localization of network mechanisms in *both* space and time. To demonstrate the utility of these methods, below we highlight the segregation of the spatially "flexible" component of the stable/flexible network mechanism into functionally distinct learning and stimulus-evoked temporal components operating on distinct "slow" and "fast" timescales respectively (see Figure 1b).

Mathematically diverse FC estimation methods applied in task-free rest have provided basic evidence that resting-state networks undergo dynamic fluctuations. Such methods include the use of sliding windows (e.g. 30–60 seconds) to conduct time-varying correlation (Hutchison, Womelsdorf, Gati, Everling, & Menon, 2013) and spatial independent components analyses (spatial ICA; Allen et al., 2014), as well as more data-driven methods that do not require selection of a sliding window (e.g. temporal ICA; Smith et al., 2012). All methods converge in demonstrating that resting-state networks typically estimated over long durations (~5–10 minutes) undergo dynamic fluctuations over shorter periods. These "temporal network states" tend to recur over the entire rest scan (Allen et al., 2014; Karahan lu & Van De Ville, 2015; Liu & Duyn, 2013) and across subjects (Zalesky, Fornito, Cocchi, Gollo, & Breakspear, 2014). Hence, standard static approaches to FC estimation run the risk of obscuring dynamic transitions between temporal network states, which might be mechanistically relevant to cognition.

However, whilst dynamic transitions described amongst resting-state networks are assumed to reflect transitions in cognitive states, the task-free nature of rest FC leaves scope for alternative interpretations. This is because mere description of temporal network states during rest fails to provide a clear link to emergent cognitive states. The failure to elucidate this explanatory relation opens the possibility that network dynamics might be driven by artifactual rather than functionally relevant sources. For example, a recent study highlighted the contamination of resting-state network dynamics by head motion. This was quantified by the reduction in a multivariate kurtosis measure of dynamic non-stationarity after removal of high-motion time points (Laumann et al., 2016; see Figure 5a). Reassuringly, multivariate kurtosis also increased in task relative to rest states (after removal of high-motion time points), suggesting that functionally relevant network dynamics are obtainable (see Figure 5b). These findings make clear how merely describing dynamic network states at rest provides imperfect insight into their cognitive relevance. Consistent with the previous section detailing spatial components of network mechanisms, we argue that the identification of mechanistically relevant temporal components calls for combining dynamic FC estimation methods with careful task manipulations. This would link prior description of temporal state transitions in rest FC more clearly to accompanying transitions in cognitive states, as instantiated in controlled task settings.

Studies combining dynamic FC estimation with task manipulations remain nascent, but preliminary findings suggest refinement of the previously introduced stable/flexible mechanism. Specifically, whereas the spatially "stable" network component might reflect structural connectivity without overlaid dynamics, the "flexible" component might arise from structural connectivity in addition to functionally distinct dynamic reconfiguration processes occurring over distinct timescales (see Figure 1b and Figure 6). To clarify, "slow" dynamic fluctuations amongst resting-state networks have been shown to track the trajectory of learning in sensorimotor tasks (Bassett, Yang, Wymbs, & Grafton, 2015; see Figure 5c). Dynamic estimates of network modularity derived from a temporal clustering algorithm revealed a longitudinal increase (over days and weeks) in the autonomy of sensory and motor networks with learning, coupled with reduced recruitment of cognitive control networks (Bassett et al., 2015; see also Chein & Schneider, 2005; Lewis et al., 2009). Beyond adding a temporal component to the stable/flexible mechanism, this finding supports the specialized role of higher-order cognitive control networks in actively driving learning-induced flexibility. This extends its role as a spatial "flexible hub" introduced in the previous section (Cole et al., 2013).

In contrast, evidence of functionally relevant network dynamics emerging over relatively "fast" timescales was provided by Sadaghiani et al (Sadaghiani, Poline, Kleinschmidt, & D'Esposito, 2015). This was enabled by trial-wise FC estimation over pre-stimulus windows in an auditory detection task. Reduced modularity in default mode and visual networks in this baseline period (spanning ~20–40 seconds) was associated with impaired performance, thereby suggesting a mechanistic link between fast trial-evoked FC reconfiguration and overt behavior (see also Thompson et al., 2013). It is tempting to link these outlined "slow" learning and "fast" stimulus-evoked dynamic FC components to respective long-term synaptic potentiation and spike propagation neurophysiological substrates, given that these also differ in their slow versus fast timescales (Dobrunz & Stevens, 1997; Sejnowski, 1999; see Figure 6). Indeed, further scrutiny of this network mechanism might reveal a cyclic relationship linking these two temporal components, such that repetitions of "fast" stimulus-evoked network reconfigurations lead to learning-induced reconfigurations emerging over a "slower" timescale, with these "slow" reconfigurations then exerting reciprocal constraints on how "fast" reconfigurations emerge across resting-state networks. Whilst such links remain speculative at this stage, it is evident that capturing dynamic FC processes over multiple timescales will be key in characterizing network mechanisms that clarify multiple levels of the neural hierarchy (see Figure 3).

However, it is also worth highlighting that the present primary reliance on fMRI complicates the study of dynamic FC. This is due to well-known limitations in fMRI temporal resolution, arising from both the relatively low sampling rate of the BOLD signal (that only permit a maximum frequency resolution in the ~0.75 Hz slow delta range; Lewis et al., 2016), as well as its vascular underpinnings rendering it an indirect measure of neural activity. Despite advances in sub-second TR acquisition protocols (Lewis et al., 2016; Smith et al., 2013) and blind deconvolution methods that aim to separate the underlying neural activation from the HRF (detailed in the next section; Havlicek, Friston, Jan, Brazdil, & Calhoun, 2011), analysis of dynamic FC would undoubtedly benefit from greater involvement of MEG and

EEG. These modalities provide millisecond resolution estimates of neural activity that are not confounded by hemodynamics.

However, this improved temporal resolution comes at the expense of worse spatial resolution and the related issue of "field spread" – the spreading of activity from a single neural source across proximal MEG/EEG sensors. This has been shown to yield artifactual, locally dominant FC patterns (Schoffelen & Gross, 2009). However, simulations have shown that spatial filtering approaches to MEG/EEG source modeling (e.g. linear beamforming; Van Veen et al., 1997) can attenuate field spread (Schoffelen & Gross, 2009). To further counteract the problem, beamformer source modeling can be combined with FC estimation methods that exclude instantaneous or "0-lag" phase relationships between input time series. This follows from the fact that field spread is entirely carried at 0-lag, albeit at the potential cost of removing 0-lag signal of neural origin (Nolte et al., 2004). Indeed, pairwise correlations between beamformer-modeled MEG source time series have recovered reasonable analogues of fMRI resting-state networks, both with exclusion of 0-lag phase relationships (Hipp et al., 2012) as well as without (Brookes et al., 2011). Note that similar results might be expected with EEG, given recent use of anatomical MRI images and sophisticated models of skull tissue conductivity to create "individualized" head models. Such head models have been observed to yield equivalent EEG source modeling accuracy as with MEG (Klamer et al., 2015).

Beyond demonstrating the overlap in spatial network topology across fMRI and MEG, studies have begun to capitalize on the above advances in MEG source modeling to target dynamics. For example, estimation of dynamic FC over sliding windows yielded a decomposition of MEG resting-state networks into temporal network states, albeit over much shorter timescales than anticipated by fMRI studies (i.e. at millisecond-to-second resolution; Baker et al., 2014; de Pasquale et al., 2010). Another intriguing application of MEG/EEG source connectivity is in clarifying whether different network mechanisms communicate via different frequency bands. Thus far, MEG homologues of resting-state networks have emerged primarily in a moderately low frequency range (theta to beta; Brookes et al., 2011; de Pasquale et al., 2010; Hipp et al., 2012). However, a recent study that corrected for the lower signal-to-noise ratio at higher frequencies also recovered resting-state analogues in the gamma band (Hipp & Siegel, 2015). Probing the functional relevance of such band-limited or "multiplexed" (Siegel, Donner, & Engel, 2012) temporal components holds much promise in formalizing the previously raised distinction between "slow flexible" and "fast flexible" components of stable/flexible mechanism. The former might be linked to FC changes in lower frequencies, whereas as the latter might emerge at higher frequencies.

Overall, dynamic FC analyses have revealed that the topology of resting-state networks masks a diversity of temporal processing over multiple timescales. Whilst this sub-field remains at an early stage, emerging findings support a distinction between "stable" components, "slow flexible" components (linked with learning and underlying synaptic potentiation) and "fast flexible" components (linked with stimulus-evoked reconfiguration and underlying spike propagation; see Figure 6). Although this distinction serves to integrate temporal and spatial components as part of a unified network mechanism (consistent with

Figure 2), further research in this vein is undoubtedly required. For example, future work might probe whether dynamic FC fluctuations are dependent on or arise as a natural consequence of the propagation of regional activations across resting-state networks, as suggested by the activity flow mapping approach (Cole et al., 2016).

5. Revealing asymmetries in activity propagation via directed functional connectivity

The majority of research into human brain networks has focused on one FC estimation algorithm – pairwise Pearson's correlation computed between regional time series – which conveys whether two regions A and B communicate in a general "undirected" sense (connectivity A-B). This is especially true for fMRI connectivity studies, whereas MEG/EEG connectivity has been commonly computed via both correlation and undirected coherence approaches. In contrast, a class of "directed" or "effective" FC algorithms provides additional insight into directions of activity propagation – whether region A communicates downstream to region B (connectivity A→B) or vice versa (connectivity B→A). Adding directional information entails a conceptual advance over networks described via undirected FC methods only, as reflected by the emphasis on parameters capturing directional communication in a number of computational network models (for review see Woolrich & Stephan, 2013). However, directed FC methods have not been widely adopted in empirical research, which is in part due to the majority of anatomical cortico-cortical connections being bidirectional (Felleman & Van Essen, 1991; Markov et al., 2011). However, we have already highlighted the imperfect correspondence between functional and structural connectivity, and there is neurophysiological evidence of asymmetries in directed activity propagation arising from experience-dependent synaptic plasticity processes independent of anatomy (Zheng, Quan, Yang, & Zhang, 2011). Even if directed connections are fewer than undirected ones, they might still have particular influences on network mechanisms underlying cognitive function. Hence, clarifying whether networks communicate via directed or undirected forms of activity propagation adds a distinct functional component to each network mechanism (see Figure 1c and 2).

The primary impediment to more widespread use of directed connectivity has been the numerous methodological uncertainties plaguing these algorithms in comparison to the mathematically simpler undirected methods. This includes uncertainty over the choice of directed FC algorithm from multiple mathematically diverse options, such as Granger causality (Roebroeck, Formisano, & Goebel, 2005), directed coherence (Nolte et al., 2008), structural equation modeling (SEM; Gates & Molenaar, 2012), dynamic causal modeling (DCM; Chen, Kiebel, & Friston, 2008; Friston, Harrison, & Penny, 2003) and Bayesian approaches (Mumford & Ramsey, 2014; Patel, Bowman, & Rilling, 2006). Uncertainty also persists over appropriate pre-processing steps (i.e. how to suppress noise sources whilst preserving relevant fine-grained temporal information), algorithmic parameters (e.g. the number of lagged observations included in Granger causality models) and how to maintain computational tractability with increasing number of input variables (e.g. the search space in model-based methods such as DCM becomes prohibitively large for >10 input variables). This uncertainty extends to the choice of imaging modality, especially as concerns fMRI

given temporal limitations arising from its low sampling rate and hemodynamic-induced confounds. The latter feature is particularly problematic given established regional and inter-subject variability in the hemodynamic response function (Handwerker, Ollinger, & D'Esposito, 2004), which can give rise to "baseline" differences in regional BOLD morphology that complicate directional inference (especially in between-subject designs; Schippers, Renken, & Keysers, 2011). Whilst MEG/EEG modalities provide a more "direct" and higher temporal resolution signal, the higher sampling rate permits contamination by a wider range of artifacts that can reduce overall signal-to-noise.

These methodological concerns call for concerted attempts to validate directed FC across algorithms and modalities, as a necessary precursor to applying them in the search for network mechanisms. However, studies aiming to validate in simulated data have yielded equivocal results. One fMRI simulation study (Smith et al., 2011) that compared the efficacy of multiple directed FC algorithms across variation in input parameters and pre-processing steps found only modest recovery of the embedded directional ground truth (~65% maximum detection accuracy). In contrast, a later study employing a similar multi-algorithmic validation, as well as explicitly modeled lags in communication at the neural level (a key difference from the generative model used in the Smith et al., 2011 study) found much higher detection accuracies across a number of algorithms, in both simulated fMRI and MEG/EEG data (Wang et al., 2014).

Recently, we devised a principled approach to validating directed FC in real fMRI and source-modeled MEG data acquired in the same subjects (Mill, Bagic, Bostan, Schneider, & Cole, 2016). This empirical validation sidesteps debate over the validity of various simplifying assumptions made in simulations (Deshpande, Sathian, & Hu, 2010; Ramsey, Hanson, & Glymour, 2011). Our approach centered on a clearly explicated "empirical ground truth" directed connectivity pattern. This took the form of an experimentally-induced reversal in activity propagation between auditory and visual regions that was predicated on the widely replicated "sensory reactivation" effect in episodic memory (e.g. Wheeler et al., 2006). This ground truth was successfully recovered by a number of directed FC methods, such as Granger causality and IMAGES Bayes network methods (Ramsey et al., 2011), and across fMRI and source-modeled MEG. These broadly positive results confirm the base validity of directional algorithms applied to human imaging data. The findings also provide more practical insight into the broad virtue of analyzing the experimental modulation of directionality estimates across task conditions (as highlighted previously; Roebroeck et al., 2005) and the improvement in fMRI results after applying "blind deconvolution" methods that remove the influence of HRF variability (Havlicek et al., 2011). Future empirical validations might test the recovery of directed connectivity ground truths spanning multiple regions, thereby testing for the removal of "indirect" connections that can arise via pairwise methods. Nevertheless, results from the Mill et al study support the use of available algorithms in combination with task manipulations to recover directional components of network mechanisms.

Methodological uncertainties have also meant that estimations of directionality at the large-scale network level (beyond a subset of regions) remain nascent, yet preliminary results have been promising. For example, analysis of lagged temporal dependencies between cortical

regions yielded a decomposition of the established fMRI resting-state networks into 8 recurring large-scale directed processing sequences (termed "lag motifs"; Mitra, Snyder, Blazey, & Raichle, 2015). This lag-based result suggests a critical role for directed activity propagation in mediating the emergence of these networks. Analyses of multivariate Granger causality during experimentally controlled task settings have identified two large-scale directional influences amongst resting-state networks (increased dorsal attention→ventral attention connectivity, and increased frontoparietal control→default mode connectivity), which were linked with "top-down" cognitive control processes that improved behavioral performance (Wen, Liu, Yao, & Ding, 2013; Wen, Yao, Liu, & Ding, 2012). These studies highlight the potential for directed FC analyses to both refine descriptions of network properties in the resting state, as well as clarify the on-task mechanisms linking them to cognition.

Contingent on further validation work, future research might examine the role of directed activity propagation in our example stable/flexible mechanism. For example, it might be that spatially flexible components linked preferentially to the frontoparietal control network enforce downstream (i.e. "top-down") directional relationships with content-specific networks to coordinate processing across multiple domains (i.e. frontoparietal→content networks; see Figure 1c). More generally, future research might strive to identify whether explanatory network interactions are better characterized as *either* undirected/bidirectional or directional. This argues against the idea that more widespread application of directed FC methods will reveal previously undirected relationships to be exclusively directional. Rather, evidence supports the presence of real FC relationships lacking in any stable directionality, in the form of synchronized oscillations observed at multiple spatial scales (Rajagovindan & Ding, 2008; Siegel et al., 2012; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Approaches to analyzing directed connectivity therefore need to better provide for instances of real undirected connectivity. Such approaches might also accommodate a role for dynamics, given observation of transitions between directed and undirected FC relationships in animals (Gregoriou, Gotts, Zhou, & Desimone, 2009) and humans (Womelsdorf et al., 2007). Indeed, linking dynamic directionality fluctuations to a clearer explanatory role in cognition would serve as an idealized recovery of a network mechanism's full complement of spatial, temporal and directional components (as exemplified in Figure 2).

6. Increasing the sensitivity of network components via multivariate pattern analysis

Another feature of the standard FC estimation pipeline is the extraction of averaged time series from isolated brain voxels or as the average across neighboring voxels within putative brain regions. Extracting such "univariate" estimates of brain activation might occlude FC mechanisms encoded by "multivariate" representational patterns amongst multiple voxels or areas of cortex. Indeed, the application of multivariate pattern analysis (MVPA) methods to multi-unit recordings in animals has refined relationships between spatiotemporal patterns in neuronal populations and cognitive function (Montijn et al., 2016; Siegel et al., 2015; Stokes et al., 2013). Application of MVPA to human imaging data has provided further insight into the functional relevance of regional activation amplitudes by characterizing multivariate

patterns amongst voxels. In comparison to univariate methods, MVPA algorithms can increase the sensitivity of detecting neural representations underlying a wide array of cognitive states. Whilst the application of MVPA at the network level remains in a fledgling state, integrating this powerful class of algorithms holds much promise in the search for network mechanisms. These tools can be leveraged to better characterize a network mechanism, both via its spatial components (by decoding cognitive representations in topological network patterns) and temporal components (by dynamic decoding of temporal network states; see Figure 7).

To provide a brief overview (for further details see Haxby, Connolly, & Guntupalli, 2014 and Haynes, 2015), a fundamental concept linking different MVPA algorithms is that of a high-dimensional space spanning multiple brain regions/voxels ("features") that are classified over a set of input samples ("observations") as representing a particular experimental condition. In essence, MVPA identifies patterns of brain regions/voxels that contain "information" relevant to decoding a given cognitive state (Kriegeskorte, Goebel, & Bandettini, 2006). MVPA methods include linear and non-linear classifiers that estimate the accuracy in decoding a cognitive state from a multivariate pattern (Haxby, 2001). Representational similarity analysis (RSA; Kriegeskorte, Mur, & Bandettini, 2008) can provide more detailed information as to the similarity between multivariate patterns for two or more stimulus classes. This helps distinguish between multivariate representations that have equally high classification accuracies but might be classifying via very different functions. These MVPA algorithms have been observed to be more sensitive to fine-grained cognitive information than alternative univariate approaches (Haxby, 2001; Kriegeskorte et al., 2006). Rather than being confined to lower-level sensory classifications, MVPA has also been successful in decoding higher-order functions (Cole, Etzel, Zacks, Schneider, & Braver, 2011; Soon, Brass, Heinze, & Haynes, 2008). Application of MVPA in a sliding window approach has also highlighted its utility in decoding dynamics, with above-chance classifications observed earlier than equivalent univariate approaches in trial-locked analyses applied to both fMRI (Kohler et al., 2013) and MEG (Borst, Ghuman, & Anderson, 2016; Sudre et al., 2012). Overall, findings from the application of MVPA algorithms to task activations emphasize their considerable utility in capturing richer and more specific cognitive representations in space and time.

These regional activation-based findings can be extended to the search for mechanisms at the network level in three broad ways. Firstly, MVPA can be applied to identify more sensitive (i.e. higher signal-to-noise) spatial components via extraction of multivariate timecourses from the pattern amongst multiple regions or voxels prior to FC estimation (see Figure 7a). For example, Coutanche and colleagues (Coutanche & Thompson-Schill, 2013) devised a method of "informational connectivity", wherein a searchlight MVPA analysis decodes the timepoint-by-timepoint classification accuracy across a set of voxels with reference to a multivariate "prototype" (the mean response within that multivariate searchlight to a given task condition). The resulting informational timecourses were submitted to standard pairwise correlations to provide a connectivity map of multivariate representations. A related multivariate connectivity method was found to yield more reliable rest FC maps (across sessions and subjects) than those obtained from univariate approaches (Linda Geerligs, Cam-CAN, & Henson, 2016). With a view to linking these findings to explanatory network

mechanisms, future studies might seek to apply multivariate connectivity methods during experimentally controlled tasks, so as to capitalize on the likely higher signal-to-noise ratio of the input time series.

The second approach to MVPA network analysis also targets more refined spatial components, by decoding cognitive states directly from topological FC patterns (see Figure 7b). Rather than basing an MVPA classification on activation patterns amongst input voxels, this approach performs a classification based on connectivity estimates between multiple regions or large-scale networks. Cole et al (2013) demonstrated the utility of this approach by identifying more specialized spatial components linked to flexible network reconfiguration. Specifically, a network representational similarity analysis linked the increasing similarity of FC patterns involving the frontoparietal control network with increasing similarity in cognitive processing across 64 unique task states. Each of these task states comprised a permutation of shared low-level rules (4 motor x 4 logic x 4 sensory rules). This approach therefore related "flexible" frontoparietal network engagement to the "flexible" integration of these shared low-level rules in novel task contexts, consistent with the involvement of cognitive control in rapid instructed task learning, (Cole, Laurent, & Stocco, 2013). A follow-up analysis trained a linear classifier to successfully decode the current task state on the basis of frontoparietal FC patterns with domain-specific networks (e.g. frontoparietal to motor network FC successfully decoded the motor task rules). These findings highlight the use of network MVPA methods both to identify more specialized "flexible" components of the stable/flexible network mechanism, as well as in directly decoding cognitive states that emerge from this network mechanism. Whilst these results support a specific role for the frontoparietal network in mediating "flexible" cognitive control, it would be useful to extend this classification approach to other networks linked with control functions (e.g. the dorsal attention network and the cingulo-opercular network, Corbetta & Shulman, 2002; Dosenbach et al., 2007).

The third application of network MVPA methods is in identifying temporal components of network mechanisms via dynamic decoding of temporal states (King & Dehaene, 2014; Stokes et al., 2013). By evaluating moment-to-moment variability of multivariate representations, insight into the timescale of task-related information in specific networks can be gained. An extension of this dynamic MVPA approach was recently employed in the classification of temporal network states during task-free rest (Chen, Ito, Kulkarni, & Cole, 2016; see Figure 7c and previous discussion of temporal network states in section 4). Specifically, the similarity in multi-region spatial patterns across individual time points was used to identify periods during resting-state scans in which the global activation state was similar. A clustering algorithm applied to the resulting "temporal similarity matrix" (left panel, Figure 7c) identified temporal network states and sub-states (middle panel, Figure 7c), which critically were associated with distinct functional roles via a meta-analysis database of task activation patterns (right panel, Figure 7c; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). This study provides one example of the profitable combination of dynamic MVPA methods in elucidating clearer explanatory linkages between spontaneous temporal network states and cognitive states. Following from similar suggestion in section 4, future extensions of this dynamic network MVPA approach might instantiate careful task manipulations to link each temporal network state directly to a cognitive state. Future studies

might also probe the correspondence of these fMRI states with those classified in higher temporal resolution MEG/EEG modalities (King & Dehaene, 2014)

To summarize, MVPA represents a novel class of methods that opens up a host of possibilities for research into network mechanisms. Whilst the full scope of combining MVPA with other advanced FC methods has yet to be realized, preliminary findings highlight the particular utility of MVPA in elucidating more refined spatial and temporal components within network mechanisms, and linking these components more directly to cognitive function. In this context, we have alluded to the potential for MVPA methods to refine our stable/flexible network mechanism – from comprising interactions between fairly general "stable" and "flexible" spatial components, to a more specific role for the frontoparietal control network in underpinning the latter component.

7. Summary of key challenges and future directions

Whilst the preceding sections have detailed numerous advances in the study of human brain networks, a number of key challenges are posed to the search for network mechanisms. Firstly, there remains a broad need for more principled validations of FC estimation strategies. Confidence in available methodologies is a necessary precursor to generating meaningful mechanistic insight from them, and there remains much ambiguity over optimal preprocessing steps (e.g. minimization of artifacts), choice of FC algorithm, and approaches to significance testing. It will be important for such validations to be applied in empirical as well as synthetic datasets, so as to avoid debate over the assumptions made by different simulation models. As we have repeatedly emphasized, validation of FC methods should also be accompanied by a renewed focus on sound experimental design. This would increase the odds of providing explanatory rather than merely descriptive insight into functional networks. Combining advanced FC estimation methods with targeted task manipulations will also prevent over-emphasizing data-driven methods to the detriment of cognitive theory.

We have also highlighted the need to address the present overreliance on fMRI, given the relatively low temporal resolution of this modality and parallel improvements in the spatial resolution of source-modeled MEG/EEG. Whilst the utility of MEG/EEG source connectivity in clarifying network dynamics is fairly obvious, the use of these modalities might also be beneficial for MVPA, given that fMRI's low sampling rate often means that the number of observations is not much greater than the number of voxel features (which increases the risk of "overfitting" noise in the data; Haxby et al., 2014). Network analyses applied to multi-modal imaging datasets, such as simultaneously acquired EEG-fMRI, represents a largely unexplored but potentially remunerative avenue, given that both imaging modalities enable complementary improvements in spatial and temporal resolution. Multi-modal network analyses might also involve physiological measurements (e.g. pupillometry), which could clarify interactions between cortical networks and autonomic neuromodulatory systems that are increasingly viewed as integral to cognitive function (Aston-Jones & Cohen, 2005; Mill, O'Connor, & Dobbins, 2016; Shine et al., 2016). Furthermore, recently released large-sample electrocorticography datasets (ECoG; e.g. the Restoring Active Memory project, RAM) could be invaluable in clarifying the micro scale underpinnings of network mechanisms studied in human imaging.

As alluded to in section 1, recent advances in combining brain stimulation technology, such as TMS and TDCS/TACS, with fMRI and EEG raise the potential to combine "neural" manipulations of brain networks with experimental manipulations of cognitive states. This approach can strengthen causal links between a network mechanism and its cognitive function. Combining stimulation of network nodes and connections with simultaneous fMRI or EEG can clarify their causal influence both over other network nodes and connections (Eldaief, Halko, Buckner, & Pascual-Leone, 2011; O'Shea, Johansen-Berg, Trief, Göbel, & Rushworth, 2007), as well as cognitive outcomes through links with on-task behavior (Polanía, Nitsche, Korman, Batsikadze, & Paulus, 2012; Zanto, Rubens, Thangavel, & Gazzaley, 2011). An important extension of this general approach would be to devise network stimulation protocols to treat mental disease (Fox, Halko, Eldaief, & Pascual-Leone, 2012). Although technological development is ongoing to enable stimulation of multiple cortical areas (greater than the 1–3 cortical sites currently feasible, see Dayan, Censor, Buch, Sandrini, & Cohen, 2013 for review), this class of methods offer the potential to directly assess the causal influence of proposed network mechanisms.

We have also highlighted the potential to combine advanced FC methods in novel and innovative ways. An associated goal is to probe relationships between FC and behavioral measures to capture individual differences in network processing and how these relate to task performance (Schultz & Cole, 2016). Such individual difference approaches might capitalize on recently developed MVPA algorithms that enable "hyperalignment" of representational spaces across subjects (Guntupalli et al., 2016). Relationships between FC and behavior can help verify that network changes across task conditions arise from functionally relevant rather than artifactual sources. However, this approach is not infallible when applied in isolation, as supported by recent evidence that head motion negatively correlates with a number of individual difference measures (Siegel et al., 2016). Rather, truly explanatory network mechanisms will likely be revealed through a convergence across multiple methods, combining FC-behavioral relationships with careful task manipulations, as well as multiple FC estimation algorithms and imaging modalities. Critically, despite our suggestion to use more advanced FC methods, it should again be noted that we encourage use of only those methods that provide additional mechanistic insight into cognitive function, rather than applying a given method simply because it has never been done before.

8. Conclusion

The field of human functional connectivity research is tantalizingly poised. Recent technical and methodological advances have opened new avenues of inquiry, and the challenge now is to develop optimal strategies to navigate these avenues without getting lost in the labyrinth of "big data". Our goal in this review was to demonstrate that the ongoing feedback loop between FC method development and insight into cognitive function would be fine-tuned by a specific focus on identifying network mechanisms. To exemplify this, our initial summary of attempts to clarify the functional relevance of resting-state networks highlighted the interaction between spatially "stable" and "flexible" network components as a mechanism linked to the general emergence of cognitive states. Findings from the application of dynamic FC methods revealed fluctuations in network processing occurring over multiple timescales, supporting further segregation of the flexible spatial component into "slow

flexible" and "fast flexible" temporal components, each associated with memory encoding and stimulus-evoked functions respectively (see Figure 6). Studies employing directed FC methods and/or capitalizing on burgeoning advances in MVPA would likely further refine this general mechanism. The search for network mechanisms is embedded in the broader goal of neuroscience to add functionally relevant components to each level of the neural hierarchy (depicted in Figure 3). Hence, it remains to be seen exactly how "stable", "slow flexible" learning and "fast flexible" stimulus-evoked network components emerge from lower-level neurophysiological processes related to structural connectivity, synaptic potentiation and spike propagation respectively. Insights into this general cognitive mechanism will likely clarify more "specialized" mechanisms involving specific functional networks. To this end, mounting evidence suggests that the frontoparietal control network serves as a "flexible hub", interacting with content-specific networks to drive processing across multiple cognitive domains. Overall, the conceptual framework and advanced tools outlined above have the potential to further our understanding of network mechanisms in both health and disease.

References

Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD. Tracking whole-brain connectivity dynamics in the resting state. Cerebral Cortex (New York, NY: 1991). 2014; 24(3): 663–676. https://doi.org/10.1093/cercor/bhs352.

Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. Annual Review of Neuroscience. 2005; 28(1):403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709.

Baker, AP., Brookes, MJ., Rezek, IA., Smith, SM., Behrens, T., Probert Smith, PJ., Woolrich, M. Fast transient networks in spontaneous human brain activity; eLife. 2014. p. 3https://doi.org/10.7554/eLife.01867

Barredo J, Oztekin I, Badre D. Ventral Fronto-Temporal Pathway Supporting Cognitive Control of Episodic Memory Retrieval. Cerebral Cortex. 2015; 25(4):1004–1019. https://doi.org/10.1093/cercor/bht291. [PubMed: 24177990]

Bassett DS, Gazzaniga MS. Understanding complexity in the human brain. Trends in Cognitive Sciences. 2011; 15(5):200–209. https://doi.org/10.1016/j.tics.2011.03.006. [PubMed: 21497128]

Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, Grafton ST. Dynamic reconfiguration of human brain networks during learning. Proceedings of the National Academy of Sciences. 2011; 108(18):7641–7646. https://doi.org/10.1073/pnas.1018985108.

Bassett DS, Yang M, Wymbs NF, Grafton ST. Learning-induced autonomy of sensorimotor systems. Nature Neuroscience. 2015; 18(5):744–751. https://doi.org/10.1038/nn.3993. [PubMed: 25849989]

Bechtel W, Abrahamsen A. Explanation: a mechanist alternative. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences. 2005; 36(2):421–441. https://doi.org/10.1016/j.shpsc.2005.03.010.

Bedau, M. Weak emergence. In: Tomberlin, J., editor. Philosophical Perspectives: Mind, Causation, and World. Vol. 11. Malden, MA: Blackwell; 1997. p. 375-399.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magnetic Resonance in Medicine. 1995; 34(4):537–541. [PubMed: 8524021]

Borst JP, Ghuman AS, Anderson JR. Tracking cognitive processing stages with MEG: A spatio-temporal model of associative recognition in the brain. NeuroImage. 2016; 141:416–430. https://doi.org/10.1016/j.neuroimage.2016.08.002. [PubMed: 27498135]

Braver TS, Cole MW, Yarkoni T. Vive les differences! Individual variation in neural mechanisms of executive control. Current Opinion in Neurobiology. 2010; 20(2):242–250. https://doi.org/10.1016/j.conb.2010.03.002. [PubMed: 20381337]

Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, … Morris PG. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proceedings of the National Academy of Sciences. 2011; 108(40):16783–16788. https://doi.org/10.1073/pnas.1112685108.

Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience. 2009; 10(3):186–198. https://doi.org/10.1038/nrn2575. [PubMed: 19190637]

Catterall WA, Raman IM, Robinson HPC, Sejnowski TJ, Paulsen O. The Hodgkin-Huxley Heritage: From Channels to Circuits. Journal of Neuroscience. 2012; 32(41):14064–14073. https://doi.org/10.1523/JNEUROSCI.3403-12.2012. [PubMed: 23055474]

Chein JM, Schneider W. Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. Cognitive Brain Research. 2005; 25(3):607–623. https://doi.org/10.1016/j.cogbrainres.2005.08.013. [PubMed: 16242923]

Chen CC, Kiebel SJ, Friston KJ. Dynamic causal modelling of induced responses. NeuroImage. 2008; 41(4):1293–1312. https://doi.org/10.1016/j.neuroimage.2008.03.026. [PubMed: 18485744]

Chen, RH., Ito, T., Kulkarni, KR., Cole, MW. Large-scale multivariate activation states of the human brain. 2016. (No. biorxiv;068221v1). Retrieved from http://biorxiv.org/lookup/doi/10.1101/068221

Cole MW, Bagic A, Kass R, Schneider W. Prefrontal Dynamics Underlying Rapid Instructed Task Learning Reverse with Practice. Journal of Neuroscience. 2010; 30(42):14245–14254. https://doi.org/10.1523/JNEUROSCI.1662-10.2010. [PubMed: 20962245]

Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. Intrinsic and Task-Evoked Network Architectures of the Human Brain. Neuron. 2014; 83(1):238–251. https://doi.org/10.1016/j.neuron.2014.05.014. [PubMed: 24991964]

Cole MW, Etzel JA, Zacks JM, Schneider W, Braver TS. Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. Frontiers in Human Neuroscience. 2011; 5:142. https://doi.org/10.3389/fnhum.2011.00142. [PubMed: 22125519]

Cole, MW., Ito, T., Bassett, DS., Schultz, DH. Activity flow over resting-state networks shapes cognitive task activations. 2016. (No. biorxiv;055194v1). Retrieved from http://biorxiv.org/lookup/doi/10.1101/055194

Cole MW, Laurent P, Stocco A. Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. Cognitive, Affective, & Behavioral Neuroscience. 2013; 13(1):1–22. https://doi.org/10.3758/s13415-012-0125-7.

Cole MW, Pathak S, Schneider W. Identifying the brain's most globally connected regions. NeuroImage. 2010; 49(4):3132–3148. https://doi.org/10.1016/j.neuroimage.2009.11.001. [PubMed: 19909818]

Cole MW, Repov G, Anticevic A. The Frontoparietal Control System: A Central Role in Mental Health. The Neuroscientist. 2014; 20(6):652–664. https://doi.org/10.1177/1073858414525995. [PubMed: 24622818]

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. Multi-task connectivity reveals flexible hubs for adaptive task control. Nature Neuroscience. 2013; 16(9):1348–1355. https://doi.org/10.1038/nn.3470. [PubMed: 23892552]

Cole MW, Schneider W. The cognitive control network: Integrated cortical regions with dissociable functions. NeuroImage. 2007; 37(1):343–360. https://doi.org/10.1016/j.neuroimage.2007.03.071. [PubMed: 17553704]

Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. Neuroscience. 2002; 3(3):201–215. https://doi.org/10.1038/nrn755. [PubMed: 11994752]

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, … Meyerand ME. Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR. American Journal of Neuroradiology. 2000; 21(9):1636–1644. [PubMed: 11039342]

Coutanche, MN., Thompson-Schill, SL. Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain; Frontiers in Human Neuroscience. 2013. p. 7https://doi.org/10.3389/fnhum.2013.00015

Craddock RC, Jbabdi S, Yan C-G, Vogelstein JT, Castellanos FX, Di Martino A, … Milham MP. Imaging human connectomes at the macroscale. Nature Methods. 2013; 10(6):524–539. https://doi.org/10.1038/nmeth.2482. [PubMed: 23722212]

Craver, CF. Explaining the brain: mechanisms and the mosaic unity of neuroscience. Oxford: New York: Oxford University Press: Clarendon Press; 2007.

Darki F, Klingberg T. The Role of Fronto-Parietal and Fronto-Striatal Networks in the Development of Working Memory: A Longitudinal Study. Cerebral Cortex. 2015; 25(6):1587–1595. https://doi.org/10.1093/cercor/bht352. [PubMed: 24414278]

Dayan E, Censor N, Buch ER, Sandrini M, Cohen LG. Noninvasive brain stimulation: from physiology to network dynamics and back. Nature Neuroscience. 2013; 16(7):838–844. https://doi.org/10.1038/nn.3422. [PubMed: 23799477]

Deco G, Jirsa VK, McIntosh AR. Resting brains never rest: computational insights into potential cognitive architectures. Trends in Neurosciences. 2013; 36(5):268–274. https://doi.org/10.1016/j.tins.2013.03.001. [PubMed: 23561718]

Deco G, Jirsa V, McIntosh AR, Sporns O, Kotter R. Key role of coupling, delay, and noise in resting brain fluctuations. Proceedings of the National Academy of Sciences. 2009; 106(25):10302–10307. https://doi.org/10.1073/pnas.0901831106.

de Pasquale F, Della Penna S, Snyder AZ, Lewis C, Mantini D, Marzetti L, … Corbetta M. Temporal dynamics of spontaneous MEG activity in brain networks. Proceedings of the National Academy of Sciences. 2010; 107(13):6040–6045. https://doi.org/10.1073/pnas.0913863107.

Deshpande G, Sathian K, Hu X. Effect of hemodynamic variability on Granger causality analysis of fMRI. NeuroImage. 2010; 52(3):884–896. https://doi.org/10.1016/j.neuroimage.2009.11.060. [PubMed: 20004248]

Dobrunz LE, Stevens CF. Heterogeneity of release probability, facilitation, and depletion at central synapses. Neuron. 1997; 18(6):995–1008. [PubMed: 9208866]

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, … Petersen SE. Distinct brain networks for adaptive and stable task control in humans. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(26):11073–11078. https://doi.org/10.1073/pnas.0704320104. [PubMed: 17576922]

Douglas RJ, Martin KAC. Neuronal circuits of the neocortex. Annual Review of Neuroscience. 2004; 27(1):419–451. https://doi.org/10.1146/annurev.neuro.27.070203.144152.

Dumontheil I, Roggeman C, Ziermans T, Peyrard-Janvid M, Matsson H, Kere J, Klingberg T. Influence of the COMT Genotype on Working Memory and Brain Activity Changes During Development. Biological Psychiatry. 2011; 70(3):222–229. https://doi.org/10.1016/j.biopsych.2011.02.027. [PubMed: 21514925]

Duyn JH. The future of ultra-high field MRI and fMRI for study of the human brain. NeuroImage. 2012; 62(2):1241–1248. https://doi.org/10.1016/j.neuroimage.2011.10.065. [PubMed: 22063093]

Eldaief MC, Halko MA, Buckner RL, Pascual-Leone A. Transcranial magnetic stimulation modulates the brain's intrinsic activity in a frequency-dependent manner. Proceedings of the National Academy of Sciences. 2011; 108(52):21229–21234. https://doi.org/10.1073/pnas.1113103109.

Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Glasser MF, … Yacoub E. Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging. PLoS ONE. 2010; 5(12):e15710. https://doi.org/10.1371/journal.pone.0015710. [PubMed: 21187930]

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex (New York, NY: 1991). 1991; 1(1):1–47.

Fodor JA, Pylyshyn ZW. Connectionism and cognitive architecture: a critical analysis. Cognition. 1988; 28(1–2):3–71. [PubMed: 2450716]

Fox MD, Halko MA, Eldaief MC, Pascual-Leone A. Measuring and manipulating brain connectivity with resting state functional connectivity magnetic resonance imaging (fcMRI) and transcranial magnetic stimulation (TMS). NeuroImage. 2012; 62(4):2232–2243. https://doi.org/10.1016/j.neuroimage.2012.03.035. [PubMed: 22465297]

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks.

Proceedings of the National Academy of Sciences. 2005; 102(27):9673–9678. https://doi.org/10.1073/pnas.0504136102.

Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends in Cognitive Sciences. 2005; 9(10):474–480. https://doi.org/10.1016/j.tics.2005.08.011. [PubMed: 16150631]

Friston KJ, Harrison L, Penny W. Dynamic causal modelling. NeuroImage. 2003; 19(4):1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7. [PubMed: 12948688]

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping. 1994; 2(4):189–210. https://doi.org/10.1002/hbm.460020402.

Gates KM, Molenaar PCM. Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. NeuroImage. 2012; 63(1):310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026. [PubMed: 22732562]

Geerligs L, Cam-CAN &, Henson RN. Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. NeuroImage. 2016; 135:16–31. https://doi.org/10.1016/j.neuroimage.2016.04.047. [PubMed: 27114055]

Geerligs L, Rubinov M, Cam-CAN &, Henson RN. State and Trait Components of Functional Connectivity: Individual Differences Vary with Mental State. Journal of Neuroscience. 2015; 35(41):13949–13961. https://doi.org/10.1523/JNEUROSCI.1324-15.2015. [PubMed: 26468196]

Gilmartin MR, Balderston NL, Helmstetter FJ. Prefrontal cortical regulation of fear learning. Trends in Neurosciences. 2014; 37(8):455–464. https://doi.org/10.1016/j.tins.2014.05.004. [PubMed: 24929864]

Gjorgjieva J, Mease RA, Moody WJ, Fairhall AL. Intrinsic Neuronal Properties Switch the Mode of Information Transmission in Networks. PLoS Computational Biology. 2014; 10(12):e1003962. https://doi.org/10.1371/journal.pcbi.1003962. [PubMed: 25474701]

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, … Van Essen DC. A multi-modal parcellation of human cerebral cortex. Nature. 2016; 536(7615):171–178. https://doi.org/10.1038/nature18933. [PubMed: 27437579]

Goldman-Rakic PS. Topography of cognition: parallel distributed networks in primate association cortex. Annual Review of Neuroscience. 1988; 11:137–156. https://doi.org/10.1146/annurev.ne.11.030188.001033.

Goni J, van den Heuvel MP, Avena-Koenigsberger A, Velez de Mendizabal N, Betzel RF, Griffa A, … Sporns O. Resting-brain functional connectivity predicted by analytic measures of network communication. Proceedings of the National Academy of Sciences. 2014; 111(2):833–838. https://doi.org/10.1073/pnas.1315529111.

Gregoriou GG, Gotts SJ, Zhou H, Desimone R. High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. Science. 2009; 324(5931):1207–1210. https://doi.org/10.1126/science.1171402. [PubMed: 19478185]

Greicius MD, Krasnow B, Reiss AL, Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100(1):253–258. https://doi.org/10.1073/pnas.0135058100. [PubMed: 12506194]

Guntupalli JS, Hanke M, Halchenko YO, Connolly AC, Ramadge PJ, Haxby JV. A Model of Representational Spaces in Human Cortex. Cerebral Cortex. 2016; 26(6):2919–2934. https://doi.org/10.1093/cercor/bhw068. [PubMed: 26980615]

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the Structural Core of Human Cerebral Cortex. PLoS Biology. 2008; 6(7):e159. https://doi.org/10.1371/journal.pbio.0060159. [PubMed: 18597554]

Handwerker DA, Ollinger JM, D'Esposito M. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. NeuroImage. 2004; 21(4):1639–1651. https://doi.org/10.1016/j.neuroimage.2003.11.029. [PubMed: 15050587]

Havlicek M, Friston KJ, Jan J, Brazdil M, Calhoun VD. Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering. NeuroImage. 2011; 56(4):2109–2128. https://doi.org/10.1016/j.neuroimage.2011.03.005. [PubMed: 21396454]

Haxby JV. Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. Science. 2001; 293(5539):2425–2430. https://doi.org/10.1126/science.1063736. [PubMed: 11577229]

Haxby JV, Connolly AC, Guntupalli JS. Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. Annual Review of Neuroscience. 2014; 37(1):435–456. https://doi.org/10.1146/annurev-neuro-062012-170325.

Haynes JD. A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. Neuron. 2015; 87(2):257–270. https://doi.org/10.1016/j.neuron.2015.05.025. [PubMed: 26182413]

Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nature Neuroscience. 2012; 15(6):884–890. https://doi.org/10.1038/nn.3101. [PubMed: 22561454]

Hipp JF, Siegel M. BOLD fMRI Correlation Reflects Frequency-Specific Neuronal Correlation. Current Biology. 2015; 25(10):1368–1374. https://doi.org/10.1016/j.cub.2015.03.049. [PubMed: 25936551]

Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology. 1952; 117(4):500–544. [PubMed: 12991237]

Hoel EP, Albantakis L, Tononi G. Quantifying causal emergence shows that macro can beat micro. Proceedings of the National Academy of Sciences. 2013; 110(49):19790–19795. https://doi.org/10.1073/pnas.1314922110.

Honey CJ, Kotter R, Breakspear M, Sporns O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proceedings of the National Academy of Sciences. 2007; 104(24):10240–10245. https://doi.org/10.1073/pnas.0701519104.

Hubel DH, Wiesel TN. Brain mechanisms of vision. Scientific American. 1979; 241(3):150–162. [PubMed: 91195]

Hutchison RM, Womelsdorf T, Gati JS, Everling S, Menon RS. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques: Dynamic Functional Connectivity. Human Brain Mapping. 2013; 34(9):2154–2177. https://doi.org/10.1002/hbm.22058. [PubMed: 22438275]

Illari PM, Williamson J. What is a mechanism? Thinking about mechanisms across the sciences. European Journal for Philosophy of Science. 2012; 2(1):119–135. https://doi.org/10.1007/s13194-011-0038-2.

Imai K, Tingley D, Yamamoto T. Experimental designs for identifying causal mechanisms: *Experimental Designs for Identifying Causal Mechanisms*. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2013; 176(1):5–51. https://doi.org/10.1111/j.1467-985X.2012.01032.x.

Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. NeuroImage. 2013; 73:239–254. https://doi.org/10.1016/j.neuroimage.2012.06.081. [PubMed: 22846632]

Kanwisher N. Functional specificity in the human brain: A window into the functional architecture of the mind. Proceedings of the National Academy of Sciences. 2010; 107(25):11163–11170. https://doi.org/10.1073/pnas.1005062107.

Karahano lu FI, Van De Ville D. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. Nature Communications. 2015; 6:7751. https://doi.org/10.1038/ncomms8751.

Kim, J. Supervenience and mind: selected philosophical essays. New York, NY, USA: Cambridge University Press; 1993.

King JR, Dehaene S. Characterizing the dynamics of mental representations: the temporal generalization method. Trends in Cognitive Sciences. 2014; 18(4):203–210. https://doi.org/10.1016/j.tics.2014.01.002. [PubMed: 24593982]

Klamer S, Elshahabi A, Lerche H, Braun C, Erb M, Scheffler K, Focke NK. Differences Between MEG and High-Density EEG Source Localizations Using a Distributed Source Model in Comparison to fMRI. Brain Topography. 2015; 28(1):87–94. https://doi.org/10.1007/s10548-014-0405-3. [PubMed: 25296614]

Klee RL. Micro-determinism and concepts of emergence. Philosophy of Science. 1984; 51(1):44–63.

Kohler PJ, Fogelson SV, Reavis EA, Meng M, Guntupalli JS, Hanke M, … Tse PU. Pattern classification precedes region-average hemodynamic response in early visual cortex. NeuroImage. 2013; 78:249–260. https://doi.org/10.1016/j.neuroimage.2013.04.019. [PubMed: 23587693]

Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proceedings of the National Academy of Sciences. 2006; 103(10):3863–3868. https://doi.org/10.1073/pnas.0600244103.

Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience. 2008; 2:4. https://doi.org/10.3389/neuro.06.004.2008. [PubMed: 19104670]

Krienen FM, Yeo BTT, Buckner RL. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. Philosophical Transactions of the Royal Society B: Biological Sciences. 2014; 369(1653):20130526–20130526. https://doi.org/10.1098/rstb.2013.0526.

Kumaran D, Hassabis D, McClelland JL. What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. Trends in Cognitive Sciences. 2016; 20(7):512–534. https://doi.org/10.1016/j.tics.2016.05.004. [PubMed: 27315762]

Kumar A, Rotter S, Aertsen A. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nature Reviews Neuroscience. 2010; 11(9):615–627. https://doi.org/10.1038/nrn2886. [PubMed: 20725095]

Laird AR, Eickhoff SB, Rottschy C, Bzdok D, Ray KL, Fox PT. Networks of task co-activations. NeuroImage. 2013; 80:505–514. https://doi.org/10.1016/j.neuroimage.2013.04.073. [PubMed: 23631994]

Laughlin SB, Sejnowski TJ. Communication in neuronal networks. Science (New York, NY). 2003; 301(5641):1870–1874. https://doi.org/10.1126/science.1089662.

LeDoux JE. Emotion Circuits in the Brain. Annual Review of Neuroscience. 2000; 23(1):155–184. https://doi.org/10.1146/annurev.neuro.23.1.155.

Lewis CM, Baldassarre A, Committeri G, Romani GL, Corbetta M. Learning sculpts the spontaneous activity of the resting human brain. Proceedings of the National Academy of Sciences. 2009; 106(41):17558–17563. https://doi.org/10.1073/pnas.0902455106.

Lewis LD, Setsompop K, Rosen BR, Polimeni JR. Fast fMRI can detect oscillatory neural activity in humans. Proceedings of the National Academy of Sciences. 2016; 113(43):E6679–E6685. https://doi.org/10.1073/pnas.1608117113.

Liu X, Duyn JH. Time-varying functional network information extracted from brief instances of spontaneous brain activity. Proceedings of the National Academy of Sciences. 2013; 110(11):4392–4397. https://doi.org/10.1073/pnas.1216856110.

Lopes da Silva F. Neural mechanisms underlying brain waves: from neural membranes to networks. Electroencephalography and Clinical Neurophysiology. 1991; 79(2):81–93. [PubMed: 1713832]

Markov NT, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, … Knoblauch K. Weight Consistency Specifies Regularities of Macaque Cortical Networks. Cerebral Cortex. 2011; 21(6):1254–1272. https://doi.org/10.1093/cercor/bhq201. [PubMed: 21045004]

McLaren DG, Ries ML, Xu G, Johnson SC. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. NeuroImage. 2012; 61(4):1277–1286. https://doi.org/10.1016/j.neuroimage.2012.03.068. [PubMed: 22484411]

Medaglia JD, Lynall ME, Bassett DS. Cognitive Network Neuroscience. Journal of Cognitive Neuroscience. 2015; 27(8):1471–1491. https://doi.org/10.1162/jocn_a_00810. [PubMed: 25803596]

Mennes M, Kelly C, Colcombe S, Castellanos FX, Milham MP. The Extrinsic and Intrinsic Functional Architectures of the Human Brain Are Not Equivalent. Cerebral Cortex. 2013; 23(1):223–229. https://doi.org/10.1093/cercor/bhs010. [PubMed: 22298730]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annual Review of Neuroscience. 2001; 24:167–202. https://doi.org/10.1146/annurev.neuro.24.1.167.

Mill RD, Bagic A, Bostan A, Schneider W, Cole MW. Empirical validation of directed functional connectivity. NeuroImage. 2016; 146:275–287. https://doi.org/10.1016/j.neuroimage.2016.11.037. [PubMed: 27856312]

Mill RD, Cavin I, O'Connor AR. Differentiating the functional contributions of resting connectivity networks to memory decision-making: fMRI support for multistage control processes. Journal of Cognitive Neuroscience. 2015; 27(8):1617–1632. https://doi.org/10.1162/jocn_a_00808. [PubMed: 25803597]

Mill RD, O'Connor AR, Dobbins IG. Pupil dilation during recognition memory:Isolating unexpected recognition from judgment uncertainty. Cognition. 2016; 154:81–94. https://doi.org/10.1016/j.cognition.2016.05.018. [PubMed: 27253862]

Miši B, Betzel RF, Nematzadeh A, Goñi J, Griffa A, Hagmann P, … Sporns O. Cooperative and Competitive Spreading Dynamics on the Human Connectome. Neuron. 2015; 86(6):1518–1529. https://doi.org/10.1016/j.neuron.2015.05.035. [PubMed: 26087168]

Mitra A, Snyder AZ, Blazey T, Raichle ME. Lag threads organize the brain'sintrinsic activity. Proceedings of the National Academy of Sciences. 2015; 112(17):E2235–E2244. https://doi.org/10.1073/pnas.1503960112.

Montijn JS, Meijer GT, Lansink CS, Pennartz CMA. Population-Level Neural Codes Are Robust to Single-Neuron Variability from a Multidimensional Coding Perspective. Cell Reports. 2016; 16(9):2486–2498. https://doi.org/10.1016/j.celrep.2016.07.065. [PubMed: 27545876]

Mumford JA, Ramsey JD. Bayesian networks for fMRI: A primer. NeuroImage. 2014; 86:573–582. https://doi.org/10.1016/j.neuroimage.2013.10.020. [PubMed: 24140939]

Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M. Identifying true brain interaction from EEG data using the imaginary part of coherency. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology. 2004; 115(10):2292–2307. https://doi.org/10.1016/j.clinph.2004.04.029. [PubMed: 15351371]

Nolte G, Ziehe A, Nikulin VV, Schlögl A, Krämer N, Brismar T, Müller K-R. Robustly Estimating the Flow Direction of Information in Complex Physical Systems. Physical Review Letters. 2008; 100(23) https://doi.org/10.1103/PhysRevLett.100.234101.

O'Keefe, J., Nadel, L. The hippocampus as a cognitive map. Oxford: New York: Clarendon Press ; Oxford University Press; 1978.

O'Shea J, Johansen-Berg H, Trief D, Göbel S, Rushworth MFS. Functionally Specific Reorganization in Human Premotor Cortex. Neuron. 2007; 54(3):479–490. https://doi.org/10.1016/j.neuron.2007.04.021. [PubMed: 17481399]

Patel RS, Bowman FD, Rilling JK. A Bayesian approach to determining connectivity of the human brain. Human Brain Mapping. 2006; 27(3):267–276. https://doi.org/10.1002/hbm.20182. [PubMed: 16092131]

Pearl, J. Direct and indirect effects. Proc. 17th Conf. Uncertainty in Artificial Intelligence; San Francisco: Morgan Kaufmann; 2001.

Petersen SE, Sporns O. Brain Networks and Cognitive Architectures. Neuron. 2015; 88(1):207–219. https://doi.org/10.1016/j.neuron.2015.09.027. [PubMed: 26447582]

Piekniewski F, Laurent P, Petre C, Richert M, Fisher D, Hylton T. Unsupervised Learning from Continuous Video in a Scalable Predictive Recurrent Network. 2016 arXiv, (1607.06854).

Polanía R, Nitsche MA, Korman C, Batsikadze G, Paulus W. The Importance of Timing in Segregated Theta Phase-Coupling for Cognitive Performance. Current Biology. 2012; 22(14):1314–1318. https://doi.org/10.1016/j.cub.2012.05.021. [PubMed: 22683259]

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, … Petersen SE. Functional Network Organization of the Human Brain. Neuron. 2011; 72(4):665–678. https://doi.org/10.1016/j.neuron.2011.09.006. [PubMed: 22099467]

Raichle ME. Two views of brain function. Trends in Cognitive Sciences. 2010; 14(4):180–190. https://doi.org/10.1016/j.tics.2010.01.008. [PubMed: 20206576]

Rajagovindan R, Ding M. Decomposing Neural Synchrony: Toward an Explanation for Near-Zero Phase-Lag in Cortical Oscillatory Networks. PLoS ONE. 2008; 3(11):e3649. https://doi.org/10.1371/journal.pone.0003649. [PubMed: 18987745]

Ramsey JD, Hanson SJ, Glymour C. Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study. NeuroImage. 2011; 58(3):838–848. https://doi.org/10.1016/j.neuroimage.2011.06.068. [PubMed: 21745580]

Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. NeuroImage. 2005; 25(1):230–242. https://doi.org/10.1016/j.neuroimage.2004.11.017. [PubMed: 15734358]

Rumelhart, DE., Rumelhart, DE. Group, JLM. the PR. Foundations. Cambridge u.a: 1986. (12. Aufl)

Sadaghiani S, Poline JB, Kleinschmidt A, D'Esposito M. Ongoing dynamics in large-scale functional connectivity predict perception. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112(27):8463–8468. https://doi.org/10.1073/pnas.1420687112. [PubMed: 26106164]

Schippers MB, Renken R, Keysers C. The effect of intra- and inter-subject variability of hemodynamic responses on group level Granger causality analyses. NeuroImage. 2011; 57(1):22–36. https://doi.org/10.1016/j.neuroimage.2011.02.008. [PubMed: 21316469]

Schoffelen JM, Gross J. Source connectivity analysis with MEG and EEG. Human Brain Mapping. 2009; 30(6):1857–1865. https://doi.org/10.1002/hbm.20745. [PubMed: 19235884]

Schultz DH, Cole MW. Higher Intelligence Is Associated with Less Task-Related Brain Network Reconfiguration. Journal of Neuroscience. 2016; 36(33):8551–8561. https://doi.org/10.1523/JNEUROSCI.0358-16.2016. [PubMed: 27535904]

Sejnowski TJ. The Book of Hebb. Neuron. 1999; 24(4):773–776. https://doi.org/10.1016/S0896-6273(00)81025-9. [PubMed: 10624941]

Seth, AK. Measuring emergence via nonlinear Granger causality. In: Bullock, S.Watson, R.Noble, J., Bedau, M., editors. Artificial life XI: proceedings of the 11th international conference on the simulation and synthesis of living systems. Cambridge: MIT Press; 2008. p. 41-49.

Shine, JM., Bissett, PG., Bell, PT., Koyejo, O., Balsters, JH., Gorgolewski, KJ., … Poldrack, RA. The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. Neuron. 2016. https://doi.org/10.1016/j.neuron.2016.09.018

Siegel, JS., Mitra, A., Laumann, TO., Seitzman, BA., Raichle, M., Corbetta, M., Snyder, AZ. Data Quality Influences Observed Links Between Functional Connectivity and Behavior. Cerebral Cortex. 2016. https://doi.org/10.1093/cercor/bhw253

Siegel M, Buschman TJ, Miller EK. Cortical information flow during flexible sensorimotor decisions. Science. 2015; 348(6241):1352–1355. https://doi.org/10.1126/science.aab0551. [PubMed: 26089513]

Siegel, M., Donner, TH., Engel, AK. Spectral fingerprints of large-scale neuronal interactions. Nature Reviews Neuroscience. 2012. https://doi.org/10.1038/nrn3137

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, … Glasser MF. Resting-state fMRI in the Human Connectome Project. NeuroImage. 2013; 80:144–168. https://doi.org/10.1016/j.neuroimage.2013.05.039. [PubMed: 23702415]

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, … Beckmann CF. Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(31):13040–13045. https://doi.org/10.1073/pnas.0905267106. [PubMed: 19620724]

Smith SM, Miller KL, Moeller S, Xu J, Auerbach EJ, Woolrich MW, … Ugurbil K. Temporally-independent functional modes of spontaneous brain activity. Proceedings of the National Academy of Sciences. 2012; 109(8):3131–3136. https://doi.org/10.1073/pnas.1121329109.

Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, … Woolrich MW. Network modelling methods for FMRI. NeuroImage. 2011; 54(2):875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063. [PubMed: 20817103]

Soon CS, Brass M, Heinze HJ, Haynes JD. Unconscious determinants of free decisions in the human brain. Nature Neuroscience. 2008; 11(5):543–545. https://doi.org/10.1038/nn.2112. [PubMed: 18408715]

Spira ME, Hai A. Multi-electrode array technologies for neuroscience and cardiology. Nature Nanotechnology. 2013; 8(2):83–94. https://doi.org/10.1038/nnano.2012.265.

Sporns O. Contributions and challenges for network models in cognitive neuroscience. Nature Neuroscience. 2014; 17(5):652–660. https://doi.org/10.1038/nn.3690. [PubMed: 24686784]

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J. Dynamic Coding for Cognitive Control in Prefrontal Cortex. Neuron. 2013; 78(2):364–375. https://doi.org/10.1016/j.neuron.2013.01.039. [PubMed: 23562541]

Sudre G, Pomerleau D, Palatucci M, Wehbe L, Fyshe A, Salmelin R, Mitchell T. Tracking neural coding of perceptual and semantic features of concrete nouns. NeuroImage. 2012; 62(1):451–463. https://doi.org/10.1016/j.neuroimage.2012.04.048. [PubMed: 22565201]

Tavor I, Jones OP, Mars RB, Smith SM, Behrens TE, Jbabdi S. Task-free MRI predicts individual differences in brain activity during task performance. Science. 2016; 352(6282):216–220. https://doi.org/10.1126/science.aad8127. [PubMed: 27124457]

Thompson GJ, Magnuson ME, Merritt MD, Schwarb H, Pan WJ, McKinley A, … Keilholz SD. Short-time windows of correlation between large-scale functional brain networks predict vigilance intraindividually and interindividually: Short-Time Window Activity Predicts Vigilance. Human Brain Mapping. 2013; 34(12):3280–3298. https://doi.org/10.1002/hbm.22140. [PubMed: 22736565]

Timme NM, Ito S, Myroshnychenko M, Nigam S, Shimono M, Yeh FC, … Beggs JM. High-Degree Neurons Feed Cortical Computations. PLOS Computational Biology. 2016; 12(5):e1004858. https://doi.org/10.1371/journal.pcbi.1004858. [PubMed: 27159884]

Tononi G, Sporns O, Edelman GM. Measures of degeneracy and redundancy in biological networks. Proceedings of the National Academy of Sciences. 1999; 96(6):3257–3262. https://doi.org/10.1073/pnas.96.6.3257.

Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature. 1995; 373(6514):515–518. https://doi.org/10.1038/373515a0. [PubMed: 7845462]

van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends in Cognitive Sciences. 2013; 17(12):683–696. https://doi.org/10.1016/j.tics.2013.09.012. [PubMed: 24231140]

Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. WU-Minn HCP Consortium. The WU-Minn Human Connectome Project: an overview. NeuroImage. 2013; 80:62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041. [PubMed: 23684880]

Van Veen BD, Van Drongelen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Transactions on Biomedical Engineering. 1997; 44(9):867–880. https://doi.org/10.1109/10.623056. [PubMed: 9282479]

Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nature Reviews. Neuroscience. 2001; 2(4):229–239. https://doi.org/10.1038/35067550. [PubMed: 11283746]

Wang, HE., Bénar, CG., Quilichini, PP., Friston, KJ., Jirsa, VK., Bernard, C. A systematic framework for functional connectivity measures; Frontiers in Neuroscience. 2014. p. 8https://doi.org/10.3389/fnins.2014.00405

Wen X, Liu Y, Yao L, Ding M. Top-Down Regulation of Default Mode Activity in Spatial Visual Attention. Journal of Neuroscience. 2013; 33(15):6444–6453. https://doi.org/10.1523/JNEUROSCI.4939-12.2013. [PubMed: 23575842]

Wen X, Yao L, Liu Y, Ding M. Causal Interactions in Attention Networks Predict Behavioral Performance. Journal of Neuroscience. 2012; 32(4):1284–1292. https://doi.org/10.1523/JNEUROSCI.2817-11.2012. [PubMed: 22279213]

Wheeler ME, Shulman GL, Buckner RL, Miezin FM, Velanova K, Petersen SE. Evidence for separate perceptual reactivation and search processes during remembering. Cerebral Cortex (New York, NY: 1991). 2006; 16(7):949–959. https://doi.org/10.1093/cercor/bhj037.

Wig GS, Schlaggar BL, Petersen SE. Concepts and principles in the analysis of brain networks: Brain networks. Annals of the New York Academy of Sciences. 2011; 1224(1):126–146. https://doi.org/10.1111/j.1749-6632.2010.05947.x. [PubMed: 21486299]

Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P. Modulation of Neuronal Interactions Through Neuronal Synchronization. Science. 2007; 316(5831):1609–1612. https://doi.org/10.1126/science.1139597. [PubMed: 17569862]

Woolrich MW, Stephan KE. Biophysical network models and the human connectome. NeuroImage. 2013; 80:330–338. https://doi.org/10.1016/j.neuroimage.2013.03.059. [PubMed: 23571421]

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nature Methods. 2011; 8(8):665–670. https://doi.org/10.1038/nmeth.1635. [PubMed: 21706013]

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, … Buckner RL. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology. 2011; 106(3):1125–1165. https://doi.org/10.1152/jn.00338.2011. [PubMed: 21653723]

Zalesky A, Fornito A, Cocchi L, Gollo LL, Breakspear M. Time-resolved resting-state brain networks. Proceedings of the National Academy of Sciences. 2014; 111(28):10341–10346. https://doi.org/10.1073/pnas.1400181111.

Zanto TP, Rubens MT, Thangavel A, Gazzaley A. Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nature Neuroscience. 2011; 14(5):656–661. https://doi.org/10.1038/nn.2773. [PubMed: 21441920]

Zheng C, Quan M, Yang Z, Zhang T. Directionality index of neural information flow as a measure of synaptic plasticity in chronic unpredictable stress rats. Neuroscience Letters. 2011; 490(1):52–56. https://doi.org/10.1016/j.neulet.2010.12.024. [PubMed: 21168473]

Highlights

- Mechanisms are needed to link descriptive network maps to cognitive function

- Network mechanisms involve explanatory interactions amongst brain regions

- Task manipulations and advanced analytic methods can reveal network mechanisms

- Each mechanism is characterized by its spatial, temporal and directional components

- Cognition emerges from mechanisms at macro network as well as micro cellular levels
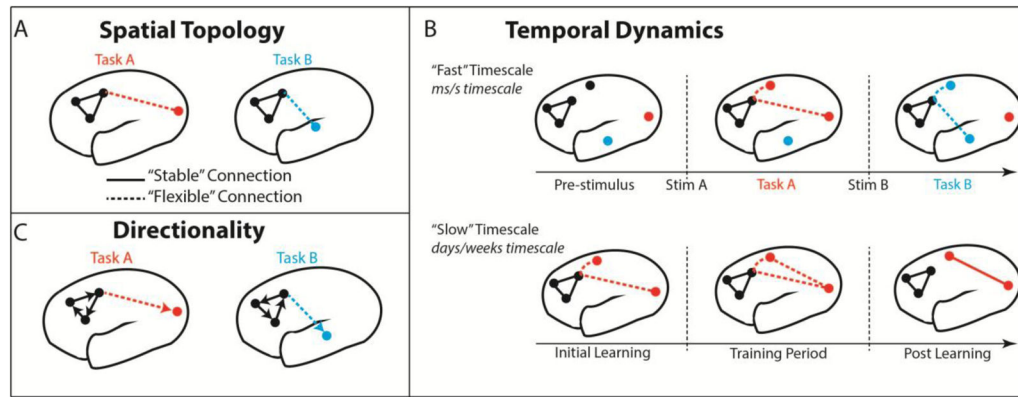
Figure 1. FC methods provide insight into different functional components to improve the characterization of each network mechanism

All panels depict hypothetical network components for a visuomotor task (Task A) and an audiomotor task (Task B). **A)** Comparing the spatial topology of network connections across different task conditions reveals spatial components of each network mechanism. Solid FC lines here (and in subsequent panels) represent the spatial component that remains "stable" across multiple task domains, whereas dashed FC lines represent components that are more "flexible" across task domains. **B)** Combining well-designed tasks with dynamic FC analyses allows for the temporal components of network organization to be investigated at different timescales. Example network reconfigurations operating on the "fast" scale (associated with stimulus-evoked responding; upper panel) are shown for both visuomotor and audiomotor tasks. Network dynamics operating at the "slow" scale are associated with learning (lower panel), and are visualized for the visuomotor task (Task A) over the course of training (i.e. practice). **C)** Establishing whether communication amongst networks occurs in a directed (i.e. lagged, represented by arrows) or undirected fashion serves as another key component, enabling a fuller characterization of activity propagation in the brain.
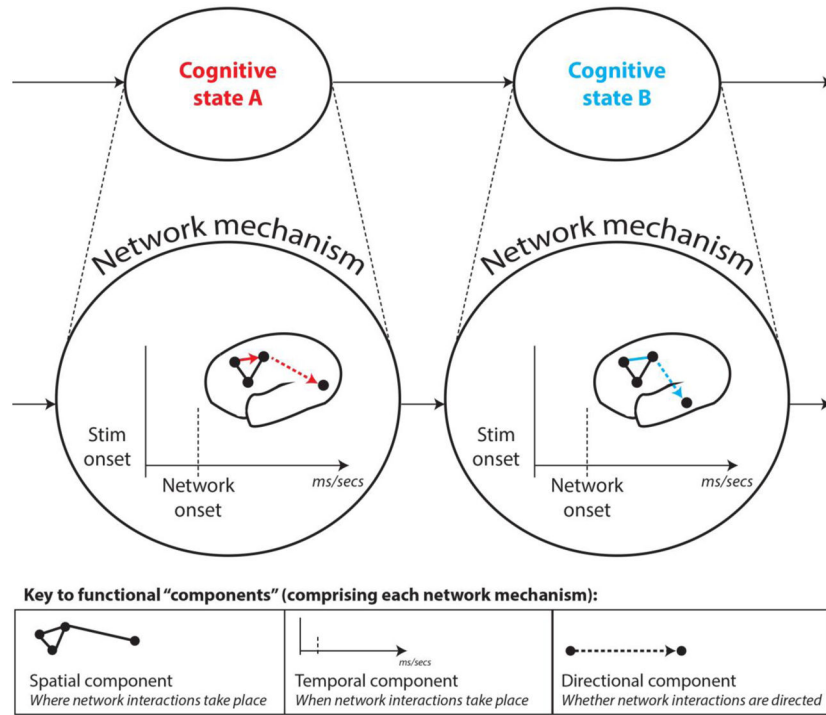
Figure 2. Depiction of a hypothetical network mechanism comprised of spatial, temporal and directional components

Following from the previous figure, we depict a network mechanism underlying the emergence of two example cognitive states: A (visuomotor task state) and B (audiomotor task state). Colored lines denote connections that are flexible across these cognitive states, whereas black lines denote stable connections. The key in the lower panel distinguishes between each functional component that is superimposed in the full mechanism in the upper panel. This should demonstrate how different functional components combine to enrich the characterization of a given network mechanism. Although the field has primarily focused on describing the spatial topology of brain networks (typically in task-free rest), we argue that combining careful task manipulations with more advanced FC estimation methods can help characterize how spatial, temporal and directional components operate collectively within network mechanisms to drive cognition.
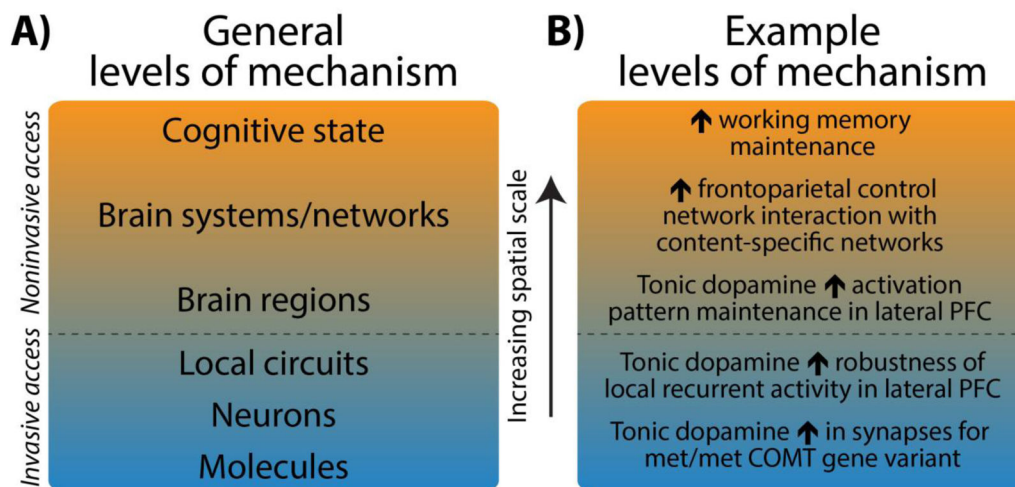
Figure 3. Neuroscientific levels of mechanism are compositional, and organized in a weakly emergent holistic hierarchy

A) We conceptualize neuroscientific levels of mechanism as particular neural elements interacting at successively higher spatial scales. Each mechanistic level is comprised of interactions occurring at the lower levels, such as brain networks being composed of a particular set of interacting brain regions. Rather than formalize a separate hierarchy for each cognitive state being explained (as in Craver, 2007), we formalize a common neural hierarchy to provide a platform to represent diverse cognitive states. This common hierarchy is structured so that each level marks an increase in spatial scale (i.e. from micro cellular scales accessible to animal neurophysiology, to higher network scales accessible to human imaging). Different cognitive states are theorized to emerge from this common hierarchy via different sets of mechanistic interactions operating at each scale (as characterized by functional network "components", see Figure 1 and 2). **B)** We provide an empirically supported example of a hierarchy of neural mechanisms underpinning working memory maintenance (the cognitive state under explanation). The schematic illustrates that cognitive states emerge from mechanisms operating at multiple spatial scales, thereby correcting the historical over-emphasis on micro cellular mechanisms and instead validating the need for mechanistic inquiry at the network level also. Note that network interactions underlying working memory maintenance are characterized using network definitions taken from the resting-state literature (e.g. "frontoparietal control network"). Note that the effects in intermediate levels may be less straightforward in reality than this simple illustration (Braver et al., 2010).
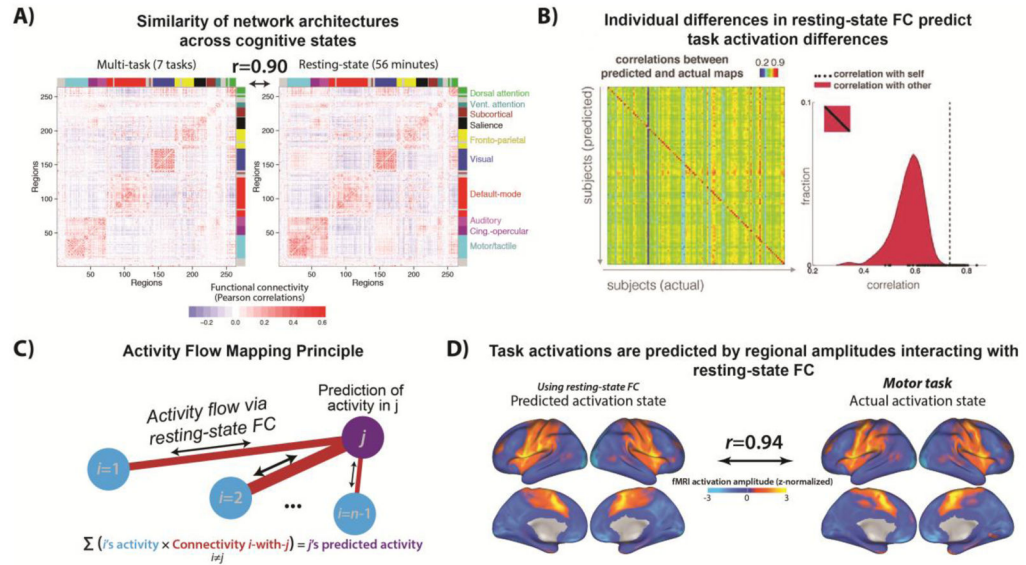
Figure 4. Clarifying the cognitive relevance of resting-state networks via analyses of spatial topology

A) Analyzing the spatial topology of resting-state networks across multiple cognitive tasks highlights majority "stable" (accounting for ~80% shared variance) and minority "flexible" (accounting for ~20% unexplained variance) spatial components. Adapted from Cole et al (2014). **B)** Using a regression model that includes resting-state FC measures as predictors, Tavor et al (2016) were able to predict individual differences in task activation patterns from individual differences in resting-state FC. Specifically, an individual's own resting-state connectivity map was a better predictor for that individual's task activation map than any other individual's resting-state connectivity map (as demonstrated by the diagonal-dominant correlation matrix). Adapted from Tavor et al (2016). **C)** Activity flow mapping principle. Cole et al (2016) developed a mechanistic model that explains how task activations emerge from the propagation of regional amplitude changes across routes constrained by resting-state FC. Adapted from Cole et al (2016). **D)** An example of applying the activity flow mapping procedure in panel C to predict activation patterns during a motor task. In essence, these predictions are made by multiplying the task activation map excluding the held-out region/voxel (i.e. the to-be-predicted "j" term in Panel C) with the resting-state FC map of that held-out region with the rest of the brain. This simple model could explain ~90% of the variance in the actual activation patterns observed across a number of distinct tasks (beyond the motor task depicted here). Adapted from Cole et al (2016).
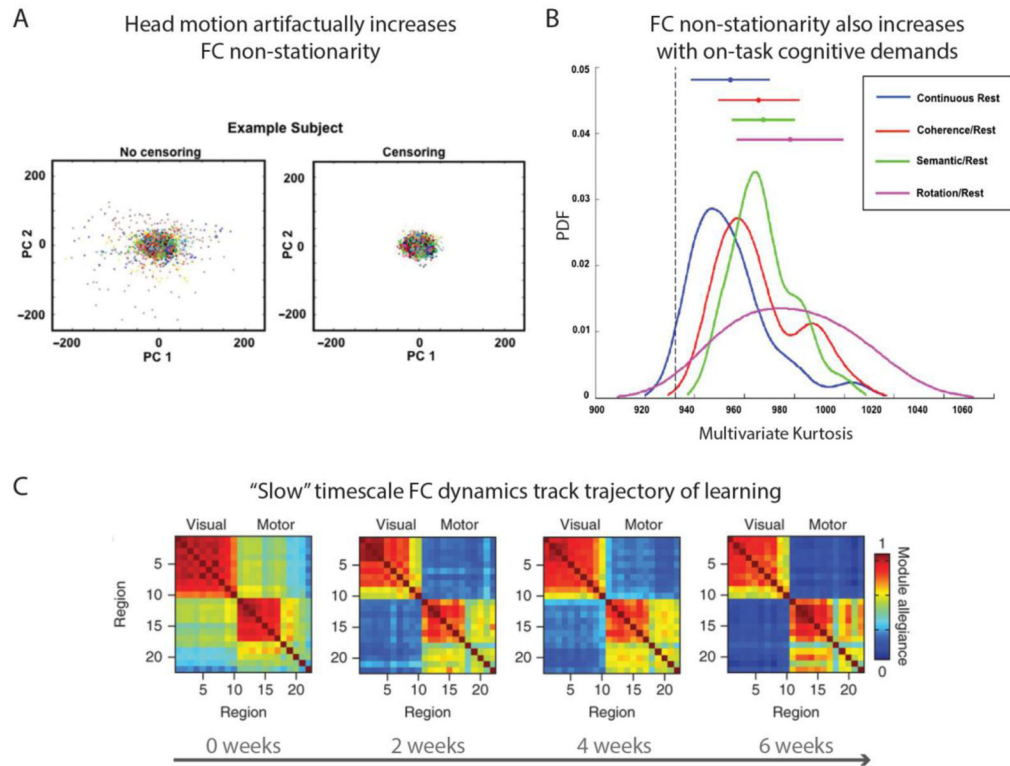
Figure 5. FC dynamics emerge from both artifactual and cognitively relevant sources
A) Plot of the artifactual increase in a multivariate kurtosis measure of FC non-stationarity with increased head motion. The left panel plots multivariate kurtosis for the first two principal components of functional networks estimated across 10 resting-state sessions, with each point denoting a single timeframe. The right panel removes timeframes with greater than 0.2 mm framewise displacement (a measure of head motion) which reduces the overall kurtosis. Adapted from Laumann et al (2016). **B)** Multivariate kurtosis also increases in task versus rest, highlighting the existence of cognitively relevant network dynamics. Note that censoring of motion with framewise displacement > 0.2 mm was performed prior to running these comparisons. Adapted from Laumann et al (2016). **C)** Changes in a graph theoretical measure of dynamic network modularity occurring over a "slow" timescale (multiple weeks) track the learning of a motor task. Specifically, modularity in content-specific visual and motor networks increased over the course of learning, whereas FC between these content networks and cognitive control networks was reduced (not depicted). This furthers demonstration in panel B that cognitively relevant network dynamics exist, by linking these dynamics to a clear mechanistic role in learning. Adapted from Bassett et al (2015).
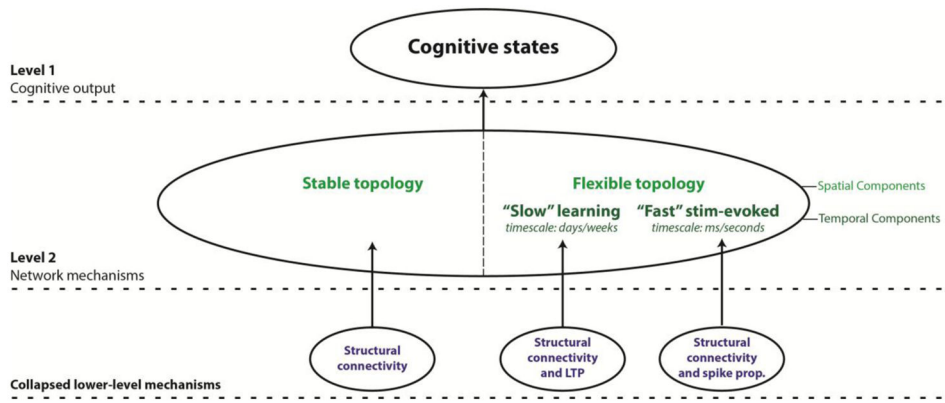
Figure 6. Summary of the stable/flexible mechanism linked to the general emergence of cognitive states

This provides a tentative formalization of the example network mechanism detailed progressively in the main text. The network-level mechanism links interactions between "stable" and "flexible" spatial components to the emergence of cognitive states. Crucially, we highlight how the "flexible" spatial component of the network mechanism might be split by functionally distinct temporal components operating on "slow" and "fast" timescales (associated with learning and stimulus-evoked responding respectively). We also provide tentative links between these stable, "slow" flexible and "fast" flexible network components and different underlying neurophysiological mechanisms (lower panel). Network mechanisms formalized in this way represent a significant explanatory advance over descriptive network mappings of the resting state. Note that in the above, we assume that all direct (rather than indirect) functional connections are underpinned by direct structural ones, with the segregation into spatially stable and flexible components arising from the absence and presence of dynamics respectively. Note that we make this assumption despite prior discrepancies in generally linking structural to functional connectivity (see footnote 5 and section 3 for details).
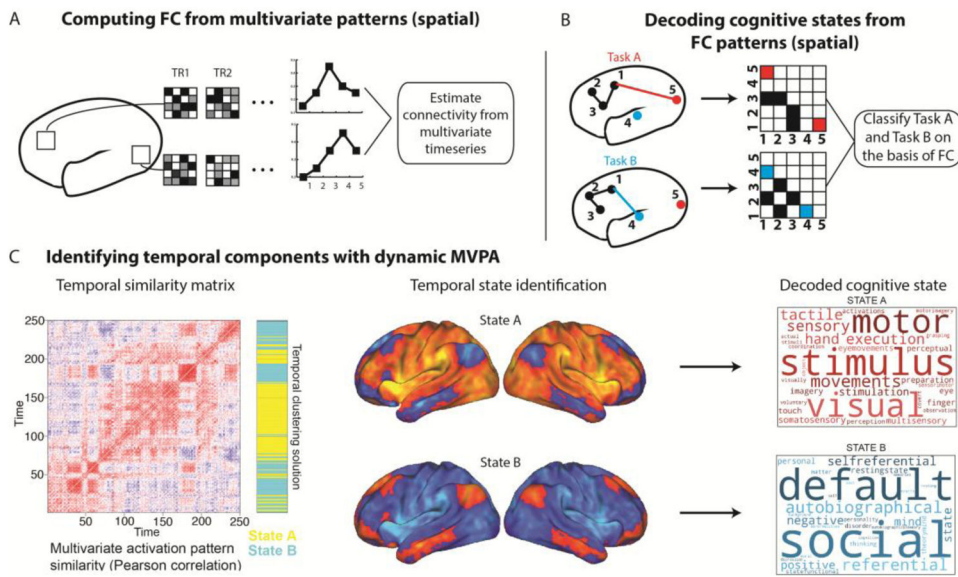
Figure 7. MVPA methods can identify spatial and temporal components of network mechanisms
A) Extraction of informational time series from multivariate voxel patterns within brain regions prior to FC estimation. This approach yields network FC maps comprised of multivariate representational patterns (treating voxels as features) linked to particular cognitive states, which can yield more sensitive spatial components of network mechanisms. **B)** MVPA can directly decode cognitive states from already estimated multivariate FC patterns (treating network connections as features). This approach also recovers spatial network components that are more directly linked to cognitive states. **C)** Empirical example of an application of dynamic MVPA to elucidate temporal network states. The left panel is a "temporal similarity matrix" illustrating the timepoint-by-timepoint similarity of multivariate activation patterns. This temporal similarity matrix is clustered to identify two temporal network states that dynamically fluctuate over the course of a rest scan (State A and State B, rendered at the voxel scale in the middle panel). These temporal states are then linked to respective "task positive" and "task negative" cognitive functions via a meta-analysis of task activations (right panel; Yarkoni et al., 2011). Whilst direct manipulation of these states in task contexts is required to substantiate their functional roles, this dynamic MVPA approach represents an improvement in linking temporal states decomposed in the resting-state to clearer cognitive functions. Adapted from Chen et al (2016).