

# Brain Networks for Cognitive Control

## Four Unresolved Questions

David Badre

### Abstract

The last decade has witnessed a marked shift of emphasis in cognitive neuroscience away from simple localization of function and toward the organization, coding, and dynamics of brain networks. This is surely a healthy evolution of our science, and the study of cognitive control has benefited from this shift, as much as any domain. However, the emphasis on brain-wide networks for cognitive control has reopened some older debates, once thought resolved, while also introducing some new ones. This chapter focuses on four questions viewed as unresolved and fundamental because one's particular answer to them commits to some basic theoretical differences regarding cognitive control function: Are there one, many, or any networks whose primary function is best described as cognitive control? Are the networks supporting cognitive control in the brain “hub-like” or “hierarchical” in their intrinsic and extrinsic organization? Are the networks for cognitive control modulatory or transmissive in the pathway from thought to action? Does controllability apply at the level of cognitive function or brain state? Each question is defined and relevant background is presented that could inform a resolution.

### Introduction

A longstanding problem in cognitive science and neuroscience concerns how the brain supports cognitive control. In broad terms, cognitive control refers to the set of mechanisms needed to organize our thoughts or actions to achieve a goal, particularly when the behaviors involved are not well learned or habitual (Stuss and Benson 1987; Logan and Gordon 2001; Miller and Cohen 2001; Badre and Nee 2018; Badre 2020). Cognitive control allows us to strategically select responses appropriate to our circumstances, to adjust our behavior on

the fly, and to adapt to open-ended problems and novel situations. It allows us to sustain goal-directed behavior over multiple timescales and to withhold inappropriate responses, even when those responses are prepotent, habitual, or stem from the prevailing urges of the moment. Cognitive control function lies close to the heart of human intelligence and ingenuity. It is also vulnerable to deficits across many, if not most, psychiatric and neurological disorders, being at the base of many of the behavioral problems arising in those conditions. Thus, understanding the mechanisms by which the brain supports cognitive control is a problem of fundamental importance.

Understanding cognitive control is of direct importance for intrusive thinking, the definition and scope of which is addressed in detail in other chapters of this volume. Most definitions, however, require that intrusions are unwanted and are unrelated to our goals or the task at hand. Thus, control mechanisms are an important means by which we both avoid intrusive thoughts and manage their impact. It follows that understanding the brain systems that support cognitive control function will have important implications for intrusive thinking, both in identifying its sources and seeking its potential remediation. In this chapter I review the brain networks that support cognitive control as a general background for more direct consideration of intrusive thinking.

As with most domains of cognitive neuroscience, the last decade of research into cognitive control in the brain has witnessed a shift away from a paradigm of functional localization toward one of functional networks. Among the most robust and important observations to emerge from the overall network approach has been that sets of brain areas tend to covary mostly with each other and not with other areas (Power et al. 2011; Yeo et al. 2011; Buckner et al. 2013). Further, the structure of this covariation is not entirely due to spatial proximity. Rather, affiliated areas can be distributed in each lobe of the brain, whereas other areas that are spatially contiguous may not affiliate. These basic properties have allowed for definition of brain networks or clusters of areas that covary with each other at different scales (Power et al. 2011; Yeo et al. 2011).

Here, I focus on four big questions that are provoked when one takes a network view of cognitive control seriously:

- Are there one, many, or any networks whose primary function is best described as cognitive control?
- Are the networks supporting cognitive control in the brain “hub-like” or “hierarchical” in their intrinsic and extrinsic organization?
- Are the networks for cognitive control modulatory or transmissive in the pathway from thought to action?
- Does controllability apply at the level of cognitive function or brain state?

Obviously, this is not intended as an exhaustive list of questions about control networks. Rather, these are the kinds of questions that I find myself asking routinely, whether in my own work or in reading about others'. No one has definitive answers, and so these questions also remain contentious or unresolved.

My goal is not to provide answers in this essay, though I will express my own view. Rather, I will define each question and present some relevant background in the hope of provoking further discussion.

### **Are There One, Many, or Any Networks Whose Function Is Cognitive Control?**

One of the oldest questions in the study of cognitive control or executive function is whether there exists one or many executive controllers in the mind and brain, or if there are executive controllers at all. The majority view has mostly been that, while there exists cognitive control function, it is not simply one thing. Rather, what we call executive function or cognitive control actually refers to a variety of specific functions and capacities.

Two camps reject this basic view. First, there are those who contend that there is one central system for cognitive control or executive function and that little to no decisive evidence exists for strong dissociations among subtypes of cognitive control functioning. The second camp argues that control is emergent from network processing in the brain, but that no particular area or network of areas is best characterized as primarily supporting “control.” Finally, even among those who agree that cognitive control exists and has many facets, there has been little agreement about the exact type and number of these facets.

This core debate has unfolded in almost every domain in which cognitive control has been studied: from behavior to individual differences to neuropsychology to neuroimaging. Currently, it is playing out again in network neuroscience. I will devote some more space to this first question than the other questions as it also provides an opportunity to summarize some background on the networks relevant to cognitive control.

#### **The Multiple Demand System: One Network to Control Them All**

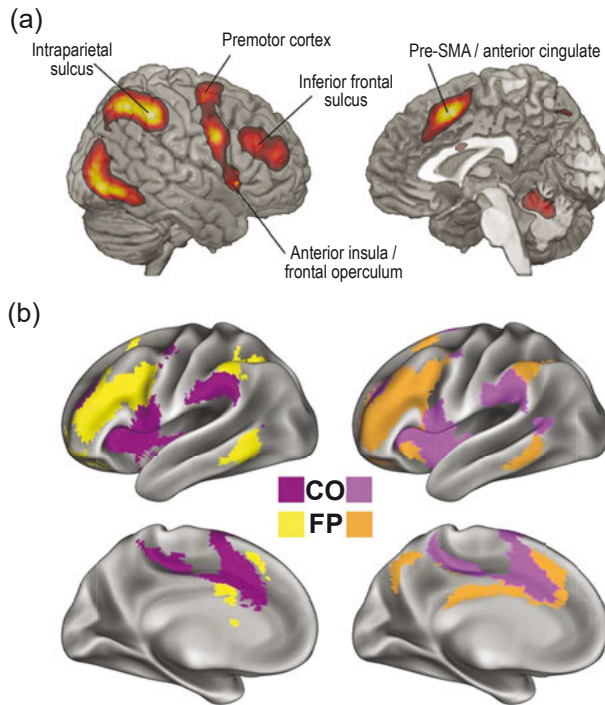
One reason that the unitary hypothesis has been so hard to falsify conclusively is that it is often the null hypothesis (Aron et al. 2015). It predicts that in any setting in which one attempts to locate a difference based on a type of cognitive control, there will be no difference. Thus, any imprecision in design, logic, or measurement has the potential to find evidence consistent with this unitary view by virtue of being inconsistent with the alternative. As a consequence, the unitary view has been something of a “zombie hypothesis” over the years: falsified in experiments that show dissociations in the brain or behavior, only to rise again a few years later when the same distinction is not found to generalize to a new task or the methodology changes. However, it is important to acknowledge that a failure to locate a difference, even in direct replication, is not itself positive evidence for a unitary controller. Rather, unitary controllers need positive predictions and evidence of their own.

In this light, the definition of the multiple demand system put forth by John Duncan and colleagues is appealing as a unitary controller view of brain organization, because it is based on a positive prediction: the multiple demand system is engaged when you perform any challenging or difficult task (Duncan 2013). Under these difficult circumstances, one needs to sequence the set of attentional states required to perform the task. It is also in these “hard” settings where one should expect the unitary cognitive control system to be engaged. Importantly, however, the specifics of the task in question or the demands that made the task difficult are not important. This system should be fundamental and domain general, so that it participates across these different task settings.

To test this hypothesis, Fedorenko et al. (2013) conducted an fMRI experiment in which they contrasted difficult versus easy conditions in a wide range of tasks. Difficulty was simply defined as a condition that took longer and induced more errors behaviorally. The tasks differed in their specific demands and the domain of input, such as between verbal or spatial. Nonetheless, when one contrasted the hard with the easy conditions of these tasks, a consistent set of areas was activated in each participant, as shown in Figure 11.1a. Given its definition, this network was dubbed the “multiple demand system” or MD system (Fedorenko et al. 2013).

The MD system has been studied extensively. It includes premotor cortex, lateral prefrontal cortex (PFC) around the inferior frontal sulcus, the intraparietal sulcus, the anterior cingulate cortex (ACC), the frontal operculum, and subregions of the basal ganglia (Fedorenko et al. 2013). This network has been associated with a variety of measures of flexible behavior, including general intelligence (Woolgar et al. 2010) and novel rule following (Tschemtscher et al. 2017). In addition, most recently, it has been found to line up with the Human Connectome Project parcellation that is defined based on a range of structural and functional anatomical features (Assem et al. 2020).

As a unitary system, the MD system is proposed to serve a very general control function needed across multiple complex tasks; namely, the assembly of attentional episodes that are the smallest unit chunks of a complex problem (Duncan 2013). When people seek to solve a new or difficult task, it has long been thought that they must break the problem into parts (Newell 1990). From the MD theory, each part is defined by a set of input-output relations that are coordinated by attentional systems. The MD system is proposed to manage these attentional episodes and the transitions from one to the next. Thus, neural coding within this network is thought to be highly dynamic, changing from moment to moment in a trajectory determined by the flow of attentional episodes. The consistent and widespread observation of flexible and dynamic coding of prefrontal neurons from electrophysiological recording in the nonhuman primate shares a qualitative correspondence to this view of multiple demand coding (Rainer et al. 1998; Stokes et al. 2013).



**Figure 11.1** Networks activated across multiple task demands. (a) Activated regions of multiple demand systems: contrast of hard versus easy conditions in all tasks run (after Fedorenko et al. 2013). (b) Frontoparietal (FP) and cingulo-opercular (CO) networks defined through functional connectivity: different methods of network definition find convergent network definitions (after Gratton et al. 2018).

### Frontoparietal Control System and the Cingulo-Opercular, Other Control System

As already noted, the functional definition of the MD system encompasses a wide and consistent set of frontal, parietal, and subcortical regions. However, evidence from analysis of functional connectivity in large resting-state data sets indicates that these areas are separable into at least two networks: a frontoparietal (FP) network and a cingulo-opercular (CO) network (Power et al. 2011; Gratton et al. 2018). Whole brain parcellations repeatedly locate differences in connectivity between these two networks across multiple methods, in large samples, and repeatedly in “deep sampled” fMRI subjects (Power et al. 2011; Yeo et al. 2011; Gordon et al. 2017; Gratton et al. 2018; Ji et al. 2019). Data from patients with brain damage to either regions of the FP or CO networks exhibit reduced functional connectivity at rest within that network but not across the networks, amounting to a double dissociation of

functionally connected networks (Nomura et al. 2010). Further, this network distinction does not depend on studying connectivity at rest. A recent fMRI study reproduced this network difference in connectivity within the functionally defined MD system while participants performed a cognitive control task (Crittenden et al. 2016). Thus, the evidence is quite strong for a network distinction in functional connectivity among two major networks that make up the MD system.

Importantly, however, while the evidence for two networks within the MD system suggests distinct functions are served by these two networks, the evidence for what those functions might be is neither strong nor clear. A meta-analysis of executive function tasks proposed a functional distinction between the FP and CO networks based loosely on timescale of control (Dosenbach et al. 2006, 2007, 2008). This analysis noted that the FP network was activated in tasks involving task cueing or adjustments of a task from feedback. The CO network, by contrast, was activated for these features in addition to demands to sustain control over time. Based on these observations and follow-up work, Dosenbach, Petersen, and colleagues proposed a distinction between “control implementation” by the FP network and “task set maintenance” by the CO system (see Gratton et al. 2018). These functional designations are intuitive, but they are not specified in a concrete mechanistic or process-specific way. To date, no study has cleanly operationalized these processes and pitted them against each other. Thus, no evidence of a functional double dissociation between control implementation and task set maintenance presently exists for the FP and CO networks.

It is notable in this context that other prominent frameworks have attributed more mechanistic functional differences to the lateral PFC and dorsal ACC areas that overlap with the FP and CO networks, respectively. For example, Botvinick proposed that the dACC may be important for detecting conditions that require control, such as response conflict, and thereby signaling upregulation of control signals by lateral PFC (Botvinick et al. 2004). Recently, Shenhav et al. (2013) updated the conflict detection model to suggest that ACC computes the expected value of control, a signal that specifies the type and intensity of control carried out by lateral PFC. Others have proposed that dACC has access to stimulus-response policies which allows it to make predictions and detect errors in response outcomes (Alexander and Brown 2011). The predicted response-outcome model captures this mechanism and can account for a wide range of results from both electrophysiology and neuroimaging. Still other models have suggested that dACC plays a role in computing value of counterfactual plans to be executed in the future (Fouragnan et al. 2019). Thus, several models propose a functional distinction between dACC and lateral PFC, which might extend to the FP and CO networks, though there is presently little agreement about what these differences might be or consistent empirical evidence for these distinctions.

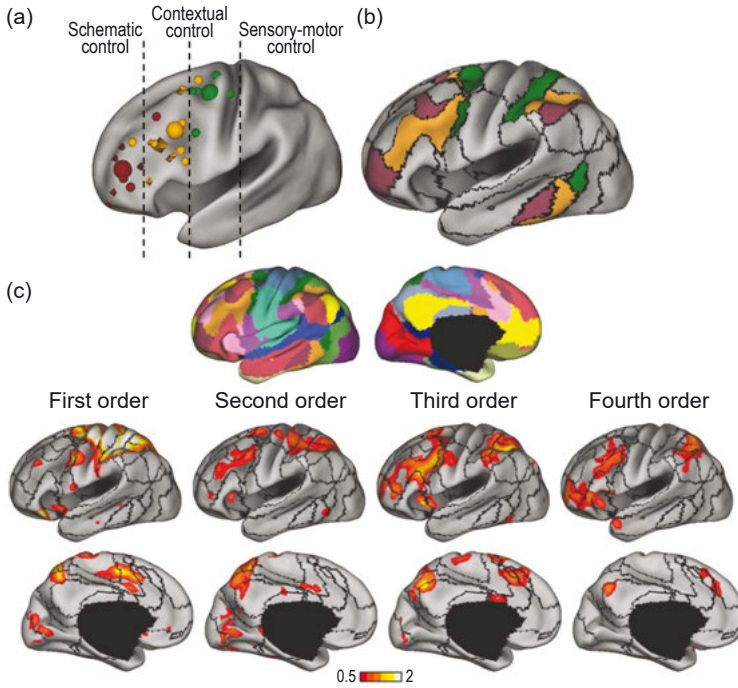
## Hierarchical Control and Distinctions within the Frontoparietal System

Within the broadly defined FP network there is evidence for further functional distinctions and subnetworks (e.g., Dixon et al. 2018). These distinctions have been most consistently observed in the context of complex tasks that are designed to test hierarchical cognitive control (Badre and Nee 2018). Hierarchical cognitive control refers specifically to cases where we must control actions based on immediate contextual signals, while also being influenced by higher-order superordinate control signals that are either more abstract policy or extended in time.

In general, if a task requires tracking multiple contextual signals to keep overlapping behavioral policies separate, demands on hierarchical control grow. For example, in a recent experiment, children and adults were instructed with a set of mappings between cartoon characters and left or right button presses, the “Go” task (Verbruggen et al. 2018). Prior to performing the Go task, however, participants were asked to view all of the cartoon characters, pressing the right button to advance to the next character (the “Next” task). This meant that while performing the Next task, participants would occasionally press the rightward arrow to a character that required a left response later on during the Go task. Such an overlap of responses can cause conflict, evidenced in slowed response time during the Next task. However, this conflict, is reduced if one can impose a latent context that separates the episode of the Next task from the later episode of the Go task and their respective response sets.

Interestingly, when doing the Next task, children exhibited more conflict than adults; children had a harder time imposing this context episode on the task. Notably, this conflict was evident even though they had never performed the Go task and were only instructed on the response rules for this task. So, it was not rule following that was a problem for the children, perhaps counter to the widely held view. The conflict indicates they immediately implemented the rules just from the instruction. Rather, their slow response was a symptom of diminished hierarchical control capacity: they could not keep the latent task contexts separate.

Studies testing hierarchical control have consistently exposed differences within the FP control system (Figure 11.2a). Across a range of studies using fMRI, transcranial magnetic stimulation (TMS), and testing of patients with frontal lobe lesions, differences in policy abstraction (defined in terms of the number of conditions or branches in a decision tree between stimulus and response) yield differences along the caudal to rostral PFC, with the highest levels of abstraction associating with the rostral mid-dorsolateral PFC (Koechlin et al. 2003; Badre and D’Esposito 2007, 2009; Nee and D’Esposito 2016, 2017; Badre and Nee 2018). Further, manipulations of temporal abstraction, which refers to the degree to which a goal or task must be held pending over time, have found fMRI activation in the most rostral portion of the frontal cortex, the rostralateral PFC (Koechlin and Hyafil



**Figure 11.2** Relationship of functional studies of hierarchical control and brain networks defined from functional connectivity. (a) Results from a meta-analysis of hierarchical control studies (after Badre and Nee 2018). The colors distinguish three functional zones related to different hierarchical control demands related to using simple (sensorimotor control) or complex (contextual control) contexts to control responses. Schematic control refers to studies that manipulated temporal abstraction or subgoaling and branching demands. Small shapes are individual studies. Large shapes are average coordinates. Within the contextual control zone, spheres refer to second-order control and diamonds to third-order control and show a further separation rostral to caudal of these studies. (b) The 17-network parcellation from Yeo et al. (2011) with the three networks most overlapping the three zones highlighted (after Badre and Nee 2018). (c) The direct comparison of the Yeo et al. (2011) network parcellation with activation across four levels of hierarchical control, from Badre and D’Esposito (2007), shows the consistent network overlap in multiple lobes of the brain (after Choi et al. 2018).

2007; Desrochers et al. 2015; Nee and D’Esposito 2016; Badre and Nee 2018). Perhaps relatedly, the rostrolateral PFC has also been implicated in tasks requiring information from memory, future directed thought, counterfactual or alternative courses of action, or pending actions to act as control signals (reviewed in Badre and Nee 2018). For this reason, Nee and Badre gave this zone a general label of “schematic control” to emphasize its relationship with these types of computations.

Important to the present discussion, these distinctions along the lateral PFC are mostly encompassed within the broadly defined FP network. However,



brain networks can be decomposed at multiple scales. Yeo et al. (2011) applied a clustering procedure to a functional connectome collected at rest in a large sample of participants. While one clustering solution, termed the 7-network parcellation, agreed with the coarse FP- and CO-network distinction, they also identified a finer-grained 17-network parcellation that broke up the FP network into more than one network (Figure 11.2b). This included, for example, a separate network for the rostralateral PFC as distinct from the mid-dorsolateral PFC and from the premotor cortex.

By comparing this 17-network structure with the functional delineations identified by Badre and D'Esposito (2007) in the fMRI study of hierarchical control (Figure 11.2c), Choi et al. (2018) found that there was a correspondence between the functional bounds associated with task-based differences in levels of hierarchical control and distinctions within the 17-network structure. Further, there were also effects of hierarchical control in distinct regions of the parietal cortex and medial frontal cortex in accord with the network structure (Choi et al. 2018). In a direct comparison, it was found that network membership, rather than rostrocaudal location, best predicted the hierarchical level of a particular voxel (Badre and D'Esposito 2007). Thus, rather than a set of areas or a gradient going from front to back along the lateral frontal cortex, ranked by a factor like policy abstraction, Choi et al. (2018) found that there are a set of subnetworks within the FP network (or MD network) that are differentially activated, based on complex control demands such as policy or temporal abstraction.

### Frontostriatal Circuits and Gating Interactions

A further network property of control that has been highlighted by the study of hierarchical control is the potential importance of corticostriatal loops in controlling interactions between separate frontal circuits (O'Reilly and Frank 2006; Collins and Frank 2014; Chatham and Badre 2015). It is well established that the basal ganglia form a series of loops with the frontal cortex via the thalamus (Alexander et al. 1986; Haber 2003). In motor control, these loops are thought to support a feedback-based gating function (Mink 1996). Specifically, candidate actions represented by cell populations in premotor cortex are initially too weak to fire, because thalamic drive is under tonic inhibition by the globus pallidus. However, these candidate actions in premotor cortex also send descending inputs to the striatum. The striatum, including putamen and caudate, receive broad inputs, not just from this premotor region but from cortex more broadly. Cells in the striatum are modulated by the presence of dopamine, which also induces plasticity so that these cells can learn which combinations of actions and context have been adaptive or not. Thus, the value of the actions considered in premotor cortex is computed as a function of what is being processed in cortex more broadly. If these actions have a history of being adaptive in this context,

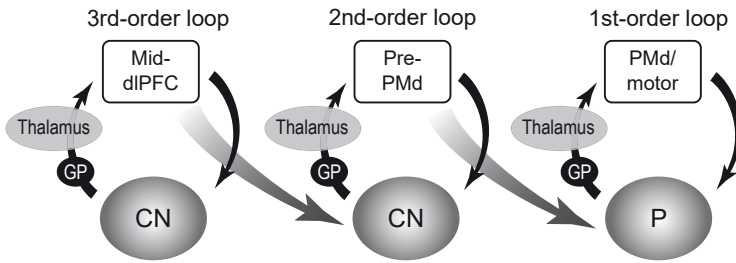
“Go” cells in the striatum will elicit a cascade that ultimately disinhibits the thalamus and allows the action to be output (Mink 1996; Wickens, 1993; O’Reilly and Frank 2006).

One influential hypothesis is that these same corticostriatal feedback loops can operate over goal and context representations that are needed for cognitive control and that are maintained in working memory by the lateral PFC (O’Reilly and Frank 2006). This computation is described using the metaphor of a gate on working memory. When the gate is closed, information does not pass in or out of working memory. When it is open, working memory can be updated and top-down control signals deployed. The feedback loops of the basal ganglia could operate as these gates by controlling transmission from one cortical network to another through their disinhibitory action on the thalamus.

Consistent with this hypothesis, there is evidence from fMRI, patient, and pharmacology studies for these corticostriatal interactions during tasks that specifically manipulate input and output gating of working memory (Frank and O’Reilly 2006; McNab and Klingberg 2008; Baier et al. 2010; Chatham et al. 2014; Chatham and Badre 2015). Furthermore, the loops between the lateral PFC and the basal ganglia are ordered and topographic, such that there are both macro- and microlevel loops between cortex and striatum that are arrayed in an orderly fashion along the rostrocaudal dimension of the frontal lobes (Verstynen et al. 2012). Choi et al. (2018) reported convergent evidence of hierarchical ordering within the striatum in resting-state functional connectivity. Further, some evidence from fMRI and TMS provides functional support for separate loops that control context- and motor-level processing during rule learning and execution (Badre and Frank 2012; Jeon et al. 2014; Korb et al. 2017).

Interaction among multiple corticostriatal loops is a candidate mechanism for hierarchical control (see Figure 11.3; Frank and Badre 2012). Specifically, the gated output of superordinate contexts maintained in working memory by one network can act as a top-down influence on the corticostriatal gating loop controlling subordinate networks. In this way, multiple contingent contexts can interact hierarchically to control action.

Models of these multiple corticostriatal loop interactions have shown that they can efficiently learn abstract hierarchical rules, transfer these structures to new tasks, and exhibit the same quasi-parallel decision dynamics that humans employ when they perform hierarchical control tasks (Frank and Badre 2012; Collins and Frank 2013; Ranti et al. 2015). Further, gating of contextual representations is a means of controlling input and output through lateral PFC, thus breaking down hard problems into more manageable chunks. In this sense, these gating computations resemble Duncan’s conception of an attentional episode (Duncan 2013). These computations emerge, however, from an interaction among separate, hierarchically ordered subnetworks.



**Figure 11.3** Schematic of a nested, interacting corticostriatal loop network for hierarchical control. The details of the corticobasal ganglia loops have been simplified in this diagram. Each loop is a feedback loop for one cortical network. However, the output of each network can act as a top-down influence on a lower-order loop. This nesting can provide a mechanism for multiple-contingent gating needed for complex, hierarchical control of behavior. Labeled areas are motor cortex (motor), dorsal premotor cortex (PMd), anterior dorsal premotor cortex (pre-PMd), mid-dorsolateral prefrontal cortex (mid-dlPFC), globus pallidus (GP), putamen (P), and caudate nucleus (CN). Reprinted with permission from Badre and Nee (2018).

### The Stopping Network

A rigorous and compelling line of research has associated a separate corticobasal ganglia network with a distinct form of cognitive control from rule following and hierarchical control, namely stopping. Inhibitory control has long been a mainstay of cognitive control function. However, not all inhibitory behavior (e.g., slowing, stopping, or withholding what we are doing) is the consequence of an inhibitory process.

The distinction between inhibition as an outcome and inhibition as a countermanding or stopping process has caused considerable confusion in the literature (Macleod et al. 2003). For instance, the Go/No-Go task commonly used to study inhibitory control might tap into an inhibition mechanism that prevents an urge to respond on No-Go trials. Not responding to a No-Go cue, however, could simply reflect a decision not to go rather than an actual suppression of a Go response. This ambiguity clearly poses a challenge to the study of the systems underlying inhibitory control. Thus, to understand inhibitory control in the brain, it is important to test cases where an inhibitory process is required to stop an ongoing or initiated action or thought.

To test inhibitory control, the Stop-Signal task (SST) is the closest to a gold-standard paradigm that we have (Logan and Cowan 1984). An action must be selected on every trial of the task in response to a “Go” stimulus. However, these initiated responses must be occasionally stopped when a “Stop” stimulus onsets at a delay after the Go stimulus. Success on the SST will thus depend on the deployment and intensity of an inhibitory process, measured behaviorally as the Stop-Signal response time and correlated with individual differences in inhibitory control, including relating to real-world impulsive behaviors such

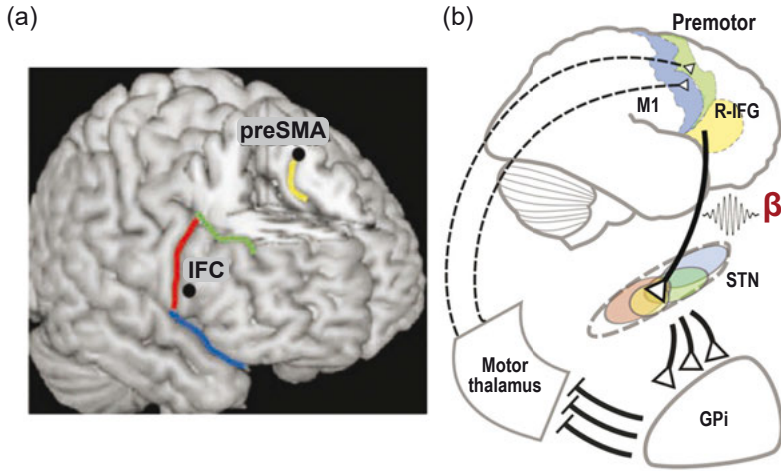
drug addiction (Dalley and Robbins 2017). Not all impulsive behavior, however, is linked to inhibition tested by the SST.

Strong evidence from multiple sources has associated stopping in the SST with a brain network that includes the right inferior frontal cortex, the pre-supplementary motor area (preSMA), and the subthalamic nucleus (STN) (see Figure 11.4; Aron et al. 2004, 2007; Aron and Poldrack 2006). These regions are consistently activated in fMRI studies that employ the SST. Damage to the right ventrolateral PFC and preSMA causes deficits in stopping that are dissociable from other frontal regions, such as the dorsolateral PFC. Importantly, the right ventrolateral PFC, preSMA, and STN interact as a dynamic network to inhibit behavior (Aron et al. 2016; Wessel and Aron 2017). These regions are connected by direct white matter connections, the integrity of which correlates with the speed of stopping (Forstmann et al. 2012).

STN is a key node in this stopping network (Isoda and Hikosaka 2008; Li et al. 2008; Schmidt et al. 2013). It projects an excitatory influence onto the globus pallidus, thereby enhancing its inhibitory influence over the thalamus. This pathway can rapidly bypass the gating computations occurring in the corticostriatal loops and put the brakes on behavior. Recent evidence from an elegant optogenetic study in the mouse confirms these basic features in the context of the stopping that occurs during surprise (Fife et al. 2017). Specifically, excitatory stimulation of the STN cells that project to the globus pallidus caused cessation of licking responses in a mouse. Then, inhibition of the STN eliminated stopping due to a surprising stimulus.

The stopping network lies clearly distinct from the FP network involved in contextual control that was discussed above (Aron et al. 2015). Even subcortically, it appears most related to the distinct hyperdirect (rather than direct/indirect) pathways through the basal ganglia. Thus, motor inhibition may be another example of a dissociable form of control.

Further, there is growing evidence for a broader inhibitory role for this network beyond countermanding motor actions. For example, we observed increased theta band oscillations between preSMA and STN under conditions of greater uncertainty, and this coupling correlated with slowing of responses during the decision (Frank et al. 2015). Ostensibly through motor inhibition, the impact of control was functionally at the level of decision making. By stopping the output of a response, more evidence was allowed to accumulate before committing to a response; this is formally equivalent to setting a higher evidence threshold and making a more conservative decision. Finally, there is evidence that components of the stopping network, including the right inferior frontal cortex, may also inhibit cognitive actions, specifically the act of retrieval from long-term memory (Guo et al. 2018; Castiglione et al. 2019). In sum, there is a separate brain network for fast stopping, and there may also be further subnetwork distinctions within this domain.



**Figure 11.4** Networks critical for stopping. (a) Cortically, the right inferior frontal cortex (IFC), sometimes termed ventrolateral prefrontal cortex, and the presupplementary motor area (preSMA) have been consistently implicated as playing a causal role in stopping during the Stop-Signal task (after Aron et al. 2007). (b) Schematic of the pathways between cortex, the subthalamic nucleus (STN), internal globus pallidum (GPi), and thalamus that are thought to support fast stopping (after Aron et al. 2016). R-IFG: right inferior frontal gyrus.

### Control without Controllers

The evidence presented above supports either one or several networks involved in control. However, a third perspective, most recently argued by Eisenreich et al. (2017), holds that none of these networks truly supports cognitive control as a unique function. Rather, since neurons are systems of distributed computation, they have emergent features of control that arise naturally in such systems. From this perspective, cognitive control is an emergent property of network computation, and there is no specific system devoted to cognitive control in the brain (Eisenreich et al. 2017).

There are many examples of distributed systems in nature that display controlled behavior without the presence of a central controller. Eisenreich et al. (2017) gives the example of a bee swarm searching for a good site to build a hive. Bees use dances to communicate to other bees that they have found, for instance, a good hive site. More bees will come to the site and do the dance if they agree with the location. Once enough bees are dancing at the site, the dance changes to a “build here” decision. At that point, a decision threshold has been passed, and the bees start to build. However, if there are multiple sites, there is conflict. The bee swarms at each location grow more slowly, and so more time is required to reach a decision. Importantly, this control adjustment is carried out at a “swarm level,” not at the level of any individual bee, as

no bee is aware of both locations. Thus, distributed systems exhibit dynamics that can be characterized as control.

To what degree is cognitive control similarly emergent? The strong version of this perspective proposes that there is no population of neurons in the brain that is devoted to representing a goal or that is directing actions toward it. Rather, goal-directed behavior emerges naturally from the systems devoted to action and perception and their local control dynamics. Control is distributed throughout the brain rather than being a function with a locus, whether that locus is a brain region or a network.

There is little positive evidence for the strong version of this viewpoint. Rather, the primary evidence is negative, questioning evidence for cognitive control networks in the brain and evidence of loss of function from brain lesions. For example, Eisenreich et al. (2017) focus on the complex and nonlinear mixed selectivity of neurons within the frontal and parietal association cortices in the networks discussed above to question whether these represent a goal or context as a top-down influence. In addition, Eisenreich et al. (2017) take a relatively dim view of human fMRI and neuropsychology, noting that there are debates in these literatures, or mixed results, regarding most proposed networks. They take these debates to suggest that the evidence in favor of regions or networks that are devoted to cognitive control function is inconclusive at best.

However, in my view, the evidence in favor of networks for control is not as ambivalent as Eisenreich et al. suggest. The mere presence of complex and nonlinear mixed selective cell coding constrains the mechanisms for control, but it is not, in and of itself, at odds with these neural representations serving a control function. Likewise, Eisenreich et al. overstate the inconsistencies found in the neuroimaging and neuropsychological literatures to some degree. As described in the preceding sections, many findings are now quite consistent and highly robust. The debates have boiled down to disagreements over functional interpretations of fairly consistent distinctions. Nonetheless, from their skeptical position, Eisenreich et al. make a crucial point not to underestimate the inherent controllability of any distributed system. In theorizing, it is important to distinguish this kind of distributed but local control from the centrally organized, goal-directed control we associate with cognitive control function.

### **Are Networks Supporting Control “Hub-Like” or “Hierarchical”?**

To some degree, merely recognizing that a network rather than a specific brain area is important for a function like cognitive control commits the same shallow theoretical error as “blobology” did in the early days of cognitive neuroscience. Labeling a network merely assigns it a location without providing mechanistic insight or constraint on theory. However, the focus on networks for control, rather than individual areas, does offer an opportunity to consider new questions about the macro-level processing dynamics and functional

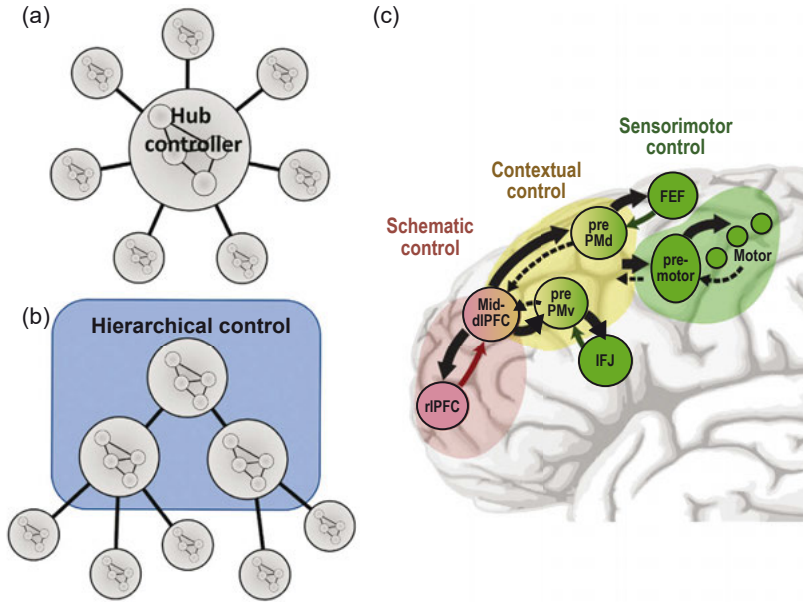
organization of those networks. Note that this is a distinct question from that considered above, which is concerned simply with whether there are multiple control-related networks and what their function might be. The question considered here is what the nature of the interaction among these networks, and others, might be. In this respect, one question to emerge out of the study of networks for cognitive control is whether these networks are *hub-like* or *hierarchical* in organization.

Viewing cognitive control systems as hubs is an intuitive and appealing theoretical idea. In essence, cognitive control systems manage or modulate routing between other systems of perception and action to carry out tasks. Thus, these networks are central to the network dynamics of the brain, will be active across most tasks, and will exert broad influence. In other words, cognitive control networks are flexible hubs, with near proximity to all other networks, and with the ability to change their connectivity with multiple other systems as needed to coordinate their dynamics during a task (Figure 11.5a).

Evidence from fMRI functional connectivity has provided some support for the hub hypothesis. Cole et al. (2013) scanned participants while they performed sixty-four different mini-tasks in the scanner. This procedure allowed changes in connectivity to be assessed while people were shifting the rules and domains over which they performed the tasks. Cole et al. observed that the FP network showed the greatest variability in its connections with other networks across all the tasks relative to any other network, including the CO network. Furthermore, rather than just reflecting random variability in a small set of connections, FP also had the highest participation coefficient, which derives from how uniform its connections are across all networks. From these observations, Cole et al. (2013) concluded the FP network was acting as a flexible hub, changing its connectivity based on the task and thereby modulating the relevant network for a particular task. In subsequent work, this global cross-task connectivity of the FP network has been associated with fluid intelligence, a further clue to its potential importance for cognitive control, particularly during rapid task instruction and execution (Cole et al. 2015).

To some degree, the flexible hub model resembles a unitary central controller that is required to modulate all other dynamics in the brain. As already noted, however, there are likely distinctions among networks for control, even within the FP system itself. Indeed, a recent analysis of functional connectivity patterns of the FP network across multiple task conditions found that this hub-like network was decomposable into at least two networks with different patterns of connectivity, and that these patterns were similar to those identified by Yeo et al.'s (2011) 17-network parcellation (Dixon et al. 2018).

An important alternative hypothesis to a global hub is that the subnetworks for control relate to each other hierarchically, such that some networks exert higher-order influence over other subnetworks, which in turn exert control over more restricted domains (Figure 11.5b). The cascade model (Koechlin et al. 2003) essentially proposed such a dynamic along the rostrocaudal axis



**Figure 11.5** Hub versus hierarchical network organizations. (a) A hub network organization places a control network, like the frontoparietal network, at the center of coordinating other networks where it serves a general and fundamental role in organizing all other networks. (b) A hierarchical network organization allows for multiple controlling networks to share asymmetric influences with each other and to have differences in their domains of control and proximity to other networks. (c) A schematic summary of the results from Nee and D’Esposito (2016, 2017) showing hierarchical interactions among frontal lobe networks (after Badre and Nee 2018). Regions along lateral prefrontal cortex are shown within the three control zones referenced in Figure 11.2a. Heavy, unbroken arrows show strong directions of influence. Broken arrows depict weak influences. Colored arrows are domain- or task-specific influences. Abbreviations: mid-dorsolateral prefrontal cortex (mid-dIPFC), rostral lateral prefrontal cortex (rIPFC), ventral premotor cortex (prePMv), inferior frontal junction (IFJ) area, anterior dorsal premotor cortex (pre-PMd), frontal eye field (FEF).

of the frontal lobe, such that abstract temporally extended control signals in rostral frontal cortex influence more temporally proximate contextual signals in lateral PFC, which in turn influence action control by premotor and motor cortex. Other models of hierarchical control have shared similar dynamics, including among nested corticostriatal loops and through medial PFC (Frank and Badre 2012; Alexander and Brown 2015).

In a set of two fMRI experiments, Nee and D’Esposito (2016, 2017) provided evidence for a hierarchical structure within lateral PFC. These studies used estimates of effective or directional connectivity from dynamic causal modeling while subjects performed a set of complex tasks that engaged varying degrees of hierarchical control in verbal versus spatial input domains.



Hierarchical strength was defined in terms of greater outward than inward connectivity (i.e., a region has broader outputs than inputs, as defined in Badre et al. 2009).

The basic results from these experiments are summarized in Figure 11.5c: Mid-dorsolateral PFC was active in higher- (more abstract) but not lower-order tasks across both input domains. It exerted an influence on the more caudal dorsal premotor cortex and ventral premotor cortex regions that were active across both the simpler and more complex tasks, but only in the spatial or verbal domain, respectively. These caudal contextual control regions also received domain-specific input from sensorimotor regions. The mid-dorsolateral PFC received greater input from the rostromedial PFC during conditions where temporal abstraction was required. In a follow-up TMS study, Nee and D'Esposito replicated these findings and showed that stimulation of nodes in this network produced behavioral effects that were broadly consistent with this information flow.

These findings from fMRI in humans converge with earlier anatomical studies in the macaque monkey. Goulas et al. (2014) performed an extensive meta-analysis of monkey anatomical projections using the CoCoMac database and focused on the connective asymmetry that might drive hierarchy. They coded multiple sites in the PFC based on the same definition of hierarchy as above: any area higher in the hierarchy would have broader efferent connections to lower-order areas than the reverse. Consistent with Nee and D'Esposito, anterior mid-dorsolateral PFC (areas 45 and 46) showed the greatest asymmetry on this metric, relative to regions caudal to the mid-dorsolateral PFC or to the rostromedial PFC which is anterior to it (Goulas et al. 2014). Notably, although Goulas et al. (2014) did find evidence that the mid-dorsolateral PFC was higher in terms of this network definition of hierarchy, it was not the most hub-like, based on a measure of betweenness centrality. This appears consistent with structural connectivity metrics in humans as well (van den Heuvel and Sporns 2013).

It remains open how one should characterize the dynamics among networks supporting cognitive control. The broadly defined FP system exhibits a hub-like character, with high participation and flexibility in connectivity across multiple tasks. There is also evidence that subnetworks within this overall system relate to each other hierarchically. In that system, there is no central domain general hub. Rather, the rostral mid-dorsolateral PFC is not active or necessary across all tasks; it is necessary during those complex tasks that require higher-order contextual control. Lower-order areas within the FP system are activated across more tasks, but they are domain specific. Thus, a hierarchical control architecture assumes that global control of the whole system emerges from limited, local, and hierarchical interactions among control networks. This contrasts with a hub network that manages interactions broadly and globally.

Finally, it should be noted that there are other hierarchical models we have not discussed that yield different organizations. For example, Barbas and Rempel-Clover (1997) proposed a laminar definition of hierarchy which

distinguishes regions based on their output versus input layers of cortex. This laminar definition of hierarchy also predicts hierarchical interactions within the FP system, but places the more rostral areas, like rostromedial PFC, higher (Goulas et al. 2014). Thus, the architecture and organization of networks for control remains a mostly open question at present, but at least one core distinction is between those proposing hub-based interactions and those proposing hierarchical ones.

### **Are the Networks for Control Modulatory or Transmissive?**

A core assumption in most brain and network theories of cognitive control is that their function is modulatory rather than transmissive. This dichotomy was highlighted by Miller and Cohen (2001) in their seminal review on the PFC and cognitive control. Their claim was that the PFC does not lie along the pathway from stimulus to action. Rather, a series of pathways from input to output exist in the brain that differ in their various strengths of connection. Collectively, these pathways represent the full action repertoire of the system. What PFC contributes is a system set apart for maintaining contexts in working memory and deploying them as control signals that can bias competition among these pathways for behavior. From this perspective, then, PFC is modulatory, not transmissive. As such, one could remove the PFC and this would not prohibit actions from occurring in response to inputs. However, as PFC maintains high-level goals and contexts, its loss would prevent the system from selecting action pathways based on abstract, temporally remote, or task set information that is not available in the immediate stimulus. The result is primarily automatic behavior based on the strongest stimulus-to-response mapping.

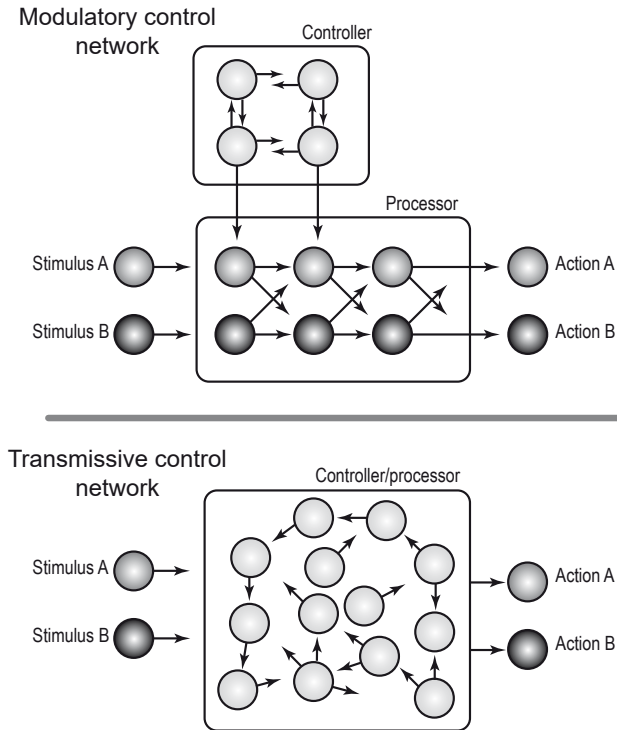
As already noted, most current perspectives on the brain networks for control take this modulatory view. These networks serve control by maintaining contexts; then through gated hierarchical interactions or flexible hubs, they bias the right organization of the system to carry out the task that will achieve the desired goal or fits with that context.

There is renewed reason, however, to reconsider this accepted view, or at least the strong version that the PFC selectively maintains a context representation required for modulating other systems that route inputs to outputs. To see why, consider the problem of control as a route driven between two locations in a town. A good control system is set up such that any start point can reach any end point. This is often done by building some main roads through town that everybody uses. This is a generalizable system because the right combination of these roads can assemble any route. But, it also causes problems. As they are general and everyone uses them, such roads are susceptible to traffic. Thus, we have to add gates (traffic lights) and monitor where we are going. This is analogous to the interference or competition among stimulus-response (SR) pathways that we experience in a task that overlaps with other tasks because

we are using general rules. Now consider that you have a particular route that is being used a lot. You could build an express road between those two locations in town: at some cost of time and asphalt, you then add some dimensions to your road system and gain a low traffic route. Increasing your dimensionality is costly but highly efficient, if you know you need a particular set of routes.

From this analogy, one way to think about the transition from controlled to automatic behavior is to view it as a transition from a reliance on generalizable, low-dimensional neural representations that are subject to interference to high-dimensional neural representations that take time to build but which directly map a combination of inputs to a response. These transformations could occur as transformations within frontal systems themselves. Early on, coordination among more networks, using gates and so forth, is necessary because a new task has to be assembled from low-dimensional components. However, over time, it is efficiently supported by a high-dimensional representation that allows a more direct route from input to output. It is still routed through the PFC, where multiple contexts and goals can affect it, but just differently in terms of the format of the routing (e.g., from low to high dimensions). This is different from the modulatory view which requires that there are always the same separate tracks from input to output: control acts like a switch operator deciding which track gets to run and when. This is among the distinctions that Eisenreich et al. (2017) made in their argument about emergent control systems and is captured by their schematic representations of different control architectures (Figure 11.6).

The evidence for this transmissive rather than modulatory model of control is limited at present, but there are intriguing clues. First, the computational trade-off described above between generalizable low-dimensional representations versus parallel high-dimensional representations has been shown in theoretical work using neural networks (Fusi et al. 2016; Musslick et al. 2017). Second, there is evidence from physiology in the nonhuman primate and multivoxel pattern analyses of human fMRI data that the FP network does not encode single contextual features of tasks but rather large conjunctions of multiple task features (Woolgar et al. 2011, 2016; Rigotti et al. 2013; Pischedda et al. 2017). Presently, we lack clear evidence that separate areas or networks represent separate contexts or elements of a task. Third, maintenance in working memory may not be a fixed-point system, wherein information is maintained in a single stable form to be accessed at any point as an external control signal. Rather, evidence from electrophysiology in monkeys and EEG in humans suggests that neural ensembles undergo dynamic change over time (Stokes et al. 2013). Thus, these representations are themselves expressed in trajectories toward an end point. Finally, evidence from nonhuman primates has shown that the nonlinear mixed selectivity of PFC neural representations supports high-dimensional capacity during task performance (Rigotti et al. 2013). This is what allows these populations to encode multiple mixtures of their inputs in unique patterns that can be read out by downstream cells. Rigotti et al. also



**Figure 11.6** Schematic of the difference between modulatory versus transmissive control networks (after Eisenreich et al. 2017). In the modulatory control network (top panel), the contextual controller lies outside the pathways from stimulus to response. Its influence is like a switch operator, choosing which path from stimulus to response is enacted. Its removal removes controlled behavior, leaving behind only automatic behavior. In the transmissive control network (bottom), control networks are a part of the pathway from stimulus to response. The nature of these representations, however, changes over the course of experience with a task. Thus, transitions from controlled to automatic behavior are supported by features in the geometry of the population coding, such as high- versus low-dimensional coding.

provided evidence that this high-dimensional coding is behaviorally relevant, as trials in which a monkey committed an error were associated with a reduction in dimensionality.

It is important to emphasize that these results could be interpreted in several ways. It could be that high-dimensional representations are why the PFC can be flexible. In other words, high dimensionality allows multiple input states to be mapped to multiple output states. Alternatively, as monkeys in these experiments have been extensively trained on these tasks, they have automated the task and formed direct mappings from multiple input states to output states in a compact way. Regardless, both perspectives are largely transmissive in their view of the PFC rather than modulatory. PFC is part of the routing, but the

*nature* of the routing is constrained by the format of its representation, either high or low dimensional. These observations reopen the question of whether control networks are best characterized as modulatory, sitting outside the basic flow of perception to action, versus being a direct part of it but changing their format and coding as a function of automaticity.

### **Does Controllability Apply at the Level of Cognitive Function or Brain State?**

The conception of the mind and brain as a control system is one of executive function's most animating theoretical ideas, dating back at least to Norbert Wiener's mental-servo notions in *Cybernetics* (Wiener 1948) and the seminal studies on the human operator in motor control by Kenneth Craik (1948). These ideas and their descendants rely on the engineering formalism of optimal control theory. In a control system, there is a set point, which is the desired system state, and mechanisms of feedback or prediction that lead the system to adjust toward the set point either in response to or in anticipation of changes to the system state.

In the classic example of a thermostat, the state of the system that matters is the temperature in your home. The set point is the desired temperature. Feedback to the system in the form of temperature measurements can result in heating or cooling actions that will change the temperature of the environment until the set point is reached. This is feedback control. Fancier modern thermostat systems may also anticipate or learn about how ambient temperatures change over the course of a day. Such a system can engage proactive cooling or heating to maintain a stable set point, and so implement feedforward control. Regardless of its specifics, however, the efficiency with which a control system can reach its set point and the range of set points it can reach are a means of assessing its quality. Control systems can be evaluated, compared, and optimized on the basis of their ability to reach any particular desired state from any initial state (termed controllability) and the efficiency with which they do so.

In cognitive control, control theory concepts have been historically posed at the cognitive-functional level. The set point is defined with reference to some real-world defined goal or target, such as drinking coffee or making it to your connecting flight or naming the ink color in a Stroop task. An effective control system is one that allows us to reach the widest range of such goal states efficiently, either in the world or our cognitive system, given a similarly wide range of initial contexts and situations. In this conception, maximal controllability (i.e., being able to get to any output state given any input state) is presumably what cognitive neuroscientists intuitively mean when they use the term *flexible behavior*.

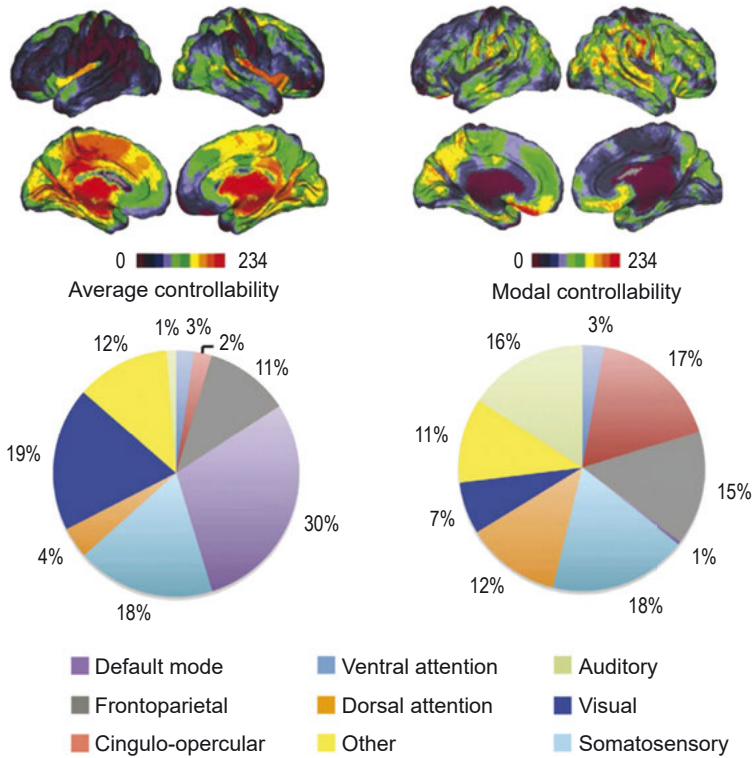
From a control theory perspective, psychological or neural mechanisms must gather feedback or make predictions about the distance to desired set

points at this cognitive-functional level. Then, some mechanisms or processes are proposed that select and implement mental or physical actions to reach them (e.g., moving, remembering, thinking, naming, inhibiting). Learning is similarly based on feedback from the world about one's current functional-level state and the actions that were taken. In elaborating these models, neuroscience focuses on the neural mechanisms that implement these functional-level control operations.

The expected value of control model introduced above is an example of a theory of cognitive control that emphasizes optimal control theory (Shenhav et al. 2013). From this perspective, a control problem occurs when there is a disparity between a goal state and the current state, such as response conflict arising during a Stroop task. The control system, in this case the dACC, is able to compute not only this distance but the mental effort needed to resolve it, in terms of the type and intensity of the control signal needed. This is the expected value of control; namely, the value of achieving the goal discounted by the effort required to reach it. Decisions about what and when to engage control, then, can be made optimally by the brain as a function of these computations.

As with expected value of control, most models of cognitive control organize the control problem at this functional level and then draw links to neural mechanisms at different degrees of specificity. Recently, however, a new set of perspectives on control have emerged within a network connectivity framework and these emphasize a subtly different level of controllability. For example, a line of sophisticated work has applied advanced network analysis techniques of white matter connections to characterize the controllability of the brain (Laurent et al. 2015a; Betzel et al. 2016; Gu et al. 2017; Khambhati et al. 2018). In brief, these analyses have emphasized how the density and organization of connections affect transitions within the space of possible brain states. Analyses within this paradigm have emphasized two kinds of transitions (Laurent et al. 2015a): transitions to common, "easy to reach" states are associated with densely connected networks, like the default mode network, whereas transitions toward rare, "hard to reach" states are facilitated by networks with weak connections, such as the FP and CO networks (Figure 11.7). This suggests that these control networks are well positioned for maximal controllability, shunting the system into any desired brain state.

These exciting new ideas offer a powerful approach to understanding control in brain networks, and the search for translational, developmental, and clinical correlates of these metrics is ongoing (Cornblath et al. 2019). However, as described, the underlying model of control in these cases differs fundamentally from the traditional functional-level control systems described above. These systems define the control problem at the level of *brain state* rather than *cognitive function* or *real-world goal state*. In other words, the control problem is not how do you get that particular drink you want, but rather how do you get the brain into a particular state that corresponds to having that drink. The



**Figure 11.7** Analysis of network controllability in the structural connectome (after Gu et al. 2015). Left: Networks show the highest average controllability, which reflects how efficiently they can move to “easy to reach” brain states. The default network is the highest on this metric. Right: Networks show the highest modal controllability, which reflects the efficiency with which they can move into “hard to reach” states. Highlighted in this analysis are the frontoparietal and cingulo-opercular networks.

set point, then, is a target pattern of brain activity, not a real-world objective, and the control system must plot the distance between your starting pattern of brain activity and that goal pattern of brain activity through a functional connectome. Finally, controllability is not defined in terms of how you behave but rather how readily you can shift from the brain state you are in to any desired brain state.

So, why aren’t these levels of theorizing about control the same? Isn’t this network conception just a reductionist reframing of the original functional-level control problem in terms of brain states? This is certainly the way it is often posed and interpreted. However, a complication arises because of the classic philosophical problem of multiple realizability. This is exactly the case

where this abstract philosopher's thought problem has some real implications for scientific theory.

Multiple realizability was famously raised by Hilary Putnam as an argument against identity theories in which each psychological state had one and only one implementation (Putnam 1967; Block and Fodor 1972). The argument is that multiple species or individuals within a species can realize the same psychological process, like pain or vision, with very different brains. Thus, there is a many-to-one mapping of brain states to psychological states. David Marr's idea of distinct levels of analysis relies on similar arguments about the asymmetry as one goes from computational to algorithmic to implementational levels (Marr 1982). Essentially, the functional or psychological level is abstraction over multiple possible realizations of algorithms and implementations within the brain.

The implications of this concept for cognitive control are important. For the reductionist reframing to correspond directly to the functional-level models, there must be a one-to-one correspondence between any one functional-level state (e.g., drinking) and a corresponding brain state (or a highly correlated class of such states). However, as the problem of multiple realizability highlights, this assumption is hardly guaranteed. Rather, the goal of having that drink may entail a wide range of activities to get there and a wide set of possible realizations of actually quenching one's thirst. Each of these is associated with a set of brain states. Some may not be more strongly correlated with each other than with other states, and so comprise a disjunct set. Thus, what ultimately connects this disjunct class of brain states is the functional-level outcome, drinking. As such, conducting control at the functional level would be the best way to ensure success, rather than making specific brain states a set point.

A further issue concerns feedback in a brain state control system. To work at the brain state level, the control system needs a means of detecting its distance from its set point. But, we don't have explicit access through the senses into our actual brain state, in terms of what neurons are firing and when. We do, however, have access to a functional-level description, like whether we are drinking or not. This feedback is also essential for the control system to learn and know what actions to take to reach a goal in the future. Without assuming a direct correspondence between the functional and brain state level that violates multiple realizability, feedback about the specific brain state target is not available to the control system in an obvious way. This is a problem if one's control system is operating primarily at a brain state level. We do, however, have perceptual systems that can assess the real-world outcome of our actions and can use these to assess our state.

What this discussion highlights is that computing the distance to a goal and seeking inputs that minimize that distance is different, depending on whether one is mapping the distance to a particular brain state or the distance to a disjoint set of such states defined by a functional outcome. Even if one



allows for some correlation among that set of states, it is easy to see how the problem gets quite complicated in the latter case, if the control system has no access to the functional-level description and is instead optimizing control over brain states.

These complications notwithstanding, the issue of level of controllability remains unresolved. It is not clear to me that the conventional functional view is necessarily always the correct one or whether both options might not be true and influence matters under different circumstances. At the level of planning, awareness, and explicit control, the functional-level description of control might be the appropriate level at which to understand cognitive control for the reasons discussed above. As such, learning and feedback mechanisms must ultimately reference this level of analysis. Neural accounts must explain how the brain supports this functional-level control system, deploying neural mechanisms that interpret the state of the world with respect to goals, compute distances to real-world hypothetical and counterfactual outcomes, define the means to cross that space, and monitor progress as it goes. However, for other kinds of control, such as switching among well-learned tasks or adjusting on the fly to maintain a stable trajectory of behavior, things may be different. In these cases, the principles and constraints of brain-level network control may be the most relevant feature determining individual variability in success, even for functional-level outcomes.

In sum, whether control plays out at a primarily functional versus brain level is an important open question, particularly as we take more sophisticated approaches to understanding dynamics within the brain's connectome. These levels of controllability are not mutually exclusive, as both may influence controlled behavior. This discussion highlights the importance of being explicit about the level at which we assume control is occurring in our theorizing, and that it is not trivial to assume that brain state control is isomorphic with functional-level control.

### **Concluding Thoughts**

The emergence of network neuroscience has brought with it an opportunity to reevaluate some older questions about control and to raise some new and exciting ones. Answers to these questions frame most of the major theories of cognitive control in the brain, with each of them staking out a position implicitly or explicitly along these lines. As I noted at the beginning, however, this is hardly an exhaustive list of the major questions facing control theorists. For example, neural dynamics clearly constitute a very important aspect of brain processing, and how dynamics among networks relate to control is only starting to be understood. One can easily think of other such questions. This essay is merely a starting point for considering the implications that network neuroscience holds for our understanding of cognitive control function and the level

of neural and psychological mechanism. With relevance to this volume, these concerns will also constrain any conception of control of intrusive thought as well as the ways one might intervene.

### **Acknowledgments**

This work was supported by the National Institute of Mental Health (MH111737) and National Institute of Neurological Disease and Stroke (NS108380) of the NIH, a MURI award from the Office of Naval Research (N00014-16-1-2832), and an award from the James S. McDonnell Foundation. I am grateful to Apoorva Bhandari, Olga Lositsky, and other Badre Lab members for helpful comments and discussions on these topics, and to Emily Chicklis for her help with figures and illustrations. Figure 11.7, from Gu et al. (2015), and Figure 11.4b, from Aron et al. (2016), are used under the terms of the Creative Commons Attribution 4.0 International License (CC-BY). Copyright for Figure 11.4a, from Aron et al. (2007), is held by the Society for Neuroscience