

What Is the Nature of the Hierarchical Organization of Lateral Prefrontal Cortex?

David Badre

Abstract

An influential view of lateral prefrontal cortex (IPFC) is that it is organized hierarchically to support cognitive control function. Specifically, regions more rostrally are hypothesized to engage in more abstract control processing than those caudally. Further, rostral regions are proposed to asymmetrically influence those caudal to them. This chapter provides an updated background on this view of IPFC organization and reviews evidence for two theoretical commitments of IPFC hierarchy: (a) functional differentiation along the rostro-caudal dimension of the IPFC and (b) super-to-subordinate hierarchical interactions within the IPFC. It will be seen that the standard view has undergone important revisions. In particular, what makes control more or less abstract along the rostro-caudal axis has been defined and redefined. The original assumption of a rostro-caudal gradient has been revised in favor of a hierarchy of interacting networks, which include association cortex outside of IPFC and subcortical structures. In addition, the apex of the hierarchy has shifted from rostro-lateral prefrontal cortex at the most anterior extent of the PFC to the mid-dorsolateral prefrontal cortex (mid-dIPFC) that lies just caudal to it. This discussion speaks directly to the topic of the functional organization of the PFC.

Introduction

The lateral prefrontal cortex (IPFC) has an established association with higher cognitive function, including cognitive control or executive function (Badre 2020; Devinsky and D'Esposito 2004; Duncan 2013; Miller and Cohen 2001; Stuss and Benson 1987). Broadly speaking, these functions control and organize our behavior, flexibly mapping sensory input to action outputs based on internal representations of goals, plans, and our behavioral context. They allow us to perform a wide range of different behaviors in the open-ended complexity of our everyday world.

However, beyond this broad functional association, there is little agreement regarding the functional organization of IPFC. We do not understand whether or how the various systems and networks that are encompassed within IPFC are distinguished from one another in terms of their computational nature or how they interact to support the complex cognitive control functions we attribute to them. Indeed, some theorists have proposed that IPFC is functionally homogenous and without a systematic organization, at least in the portion of IPFC that supports cognitive control (Assem et al. 2020; Duncan 2010, 2013). The strong version of this perspective proposes that IPFC is part of a multiple demand network that supports performance of demanding tasks in a domain general way, but that no particular part of PFC is devoted to a particular task. Thus, what functional differentiation IPFC might exhibit from task to task is not governed by consistent organizing principles that would generalize across task-independent cognitive demands or other computational-level factors.

In this chapter, I review the state of an alternative, influential class of theory regarding IPFC organization that does not ascribe particular tasks to localized regions of the IPFC, but which does assume functional organizing principles at a computational level that are systematic and generalizable. Specifically, I will consider the proposal that the IPFC is organized as a functional hierarchy along its rostral-caudal axis in the service of control function (Badre and Nee 2018; Christoff et al. 2009; Fuster 2001; Koechlin et al. 2003; Nee 2021; Soltani and Koechlin 2022).

It is important to clarify what is meant by an organizing hierarchy with regard to IPFC, as there are at least two ways hierarchy is used when discussing IPFC. The first is mostly uncontroversial. Most theorists accept the proposal that IPFC holds a hierarchical relationship to the rest of the brain. Theories of cognitive control propose that IPFC broadcasts top-down signals to influence processing in other areas or networks of the brain that support basic cognitive processes like perception and memory (Badre and D'Esposito 2009; Badre and Nee 2018; Cole et al. 2015b; Duncan 2013; Hazy et al. 2006; Miller and Cohen 2001). These control signals modulate ongoing processing in these networks so that they are coordinated toward particular behavioral goals. This is a hierarchical organization of the brain, and theories of IPFC organization are situated within a commitment to this larger architecture.

Nonetheless, the hierarchy we will primarily be concerned with in this chapter—and the one that remains controversial in the cognitive neuroscience literature—describes the intrinsic functional organization of IPFC itself. Specifically, this refers to the proposal that functionally distinct regions, networks, or gradients within the IPFC form a representational and/or processing hierarchy, with higher-order areas controlling and influencing the activity in lower-order areas.

Though specifics differ, this architecture is generally hypothesized to support the control of behavior at multiple levels of abstraction and/or over multiple timescales that array along the rostral-caudal axis of the frontal lobe. In

particular, the most caudal areas of frontal cortex—those participating in motor, premotor, and attention networks—are associated with control over specific, concrete movements and externally directed attention. Moving rostrally in IPFC, progressively complex control functions are proposed to be supported, serving goals or rule structures that are more abstract, multiply contingent, internally generated, counterfactual, and/or that prevail over longer timescales.

Though this axis of organization has sometimes been used to describe both dorsal and ventral aspects of lateral frontal lobe organization, debate about the hierarchical organization of IPFC has mostly concerned the networks related to cognitive control and adaptive task performance. These include the frontoparietal control networks (Fedorenko et al. 2013; Gordon et al. 2017; Gratton et al. 2018b; Ji et al. 2019; Power et al. 2011; Yeo et al. 2011) as well as sensorimotor and attention networks. As such, our discussion of hierarchy in IPFC will primarily concern the motor and dorsal premotor areas in caudal frontal cortex, pre-premotor (prePM) and mid-dorsolateral prefrontal cortex (mid-dIPFC around the inferior frontal sulcus) more rostrally, and the lateral frontal pole or rostro-lateral prefrontal cortex (rIPFC) at the most rostral extent.

Over the last two decades, this hierarchical view of IPFC organization has been supported in some ways, and in others, has undergone important revisions. Here, I will provide a brief update and background on the current state of the literature on the hierarchical organization of the IPFC.

This discussion is of direct relevance to the central topic of this Forum on the organization of PFC. Not only does it provide one putative answer to the question of how part of the frontal lobe is organized, it also has implications for the theory of the organization of PFC more generally. First, hierarchical theories are examples of theories of functional organization that do not localize particular task-related functions or executive skills, like task switching, to particular areas of the frontal cortex, per se. Rather, a hierarchical organization defines a processing architecture that describes how general functions like cognitive control emerge from the interactions among regional or network computations. Second, as the IPFC has direct interactions with other parts of the frontal lobe, the organization of IPFC will have implications for the organization of those other areas that interact with it. Thus, the debates, challenges, and discoveries arising from investigation of hierarchy in IPFC are of relevance to understanding frontal lobe organization more generally.

Functional Hierarchy in Rostro-Caudal Lateral PFC

Functional Differentiation Along the Rostro-Caudal Dimension of IPFC

A hierarchy along the rostro-caudal axis of the IPFC takes as its premise that there exists functional differentiation along that axis that can be related to

generalizable task demands. A persistent obstacle to finding such evidence is that experiments testing particular task contrasts, while being well-controlled and thus amenable to mechanistic theory, will also feature many idiosyncratic choices, both at the level of the task and study implementation. Any of these choices could drive observed effects in a region like LPFC that is adapted to shape task performance. Thus, inconsistencies in the literature might arise from overgeneralizing these idiosyncratic effects.

Meta-analyses are one way of testing the premise of a rostro-caudal organizing axis in LPFC that can overcome the limitations of particular idiosyncratic tasks (see also de la Vega et al. 2018). Recently, Abdallah et al. (2022) studied meta-analytic connectivity in fMRI activation across 14,371 studies from the NeuroSynth database, spanning a wide range of tasks and contrasts. Meta-analytic connectivity refers to a pattern of co-activation across study contrasts. More concretely, for any point in LPFC, one can compute the probabilities that (a) activity is reported in other brain regions given activity in that LPFC location and (b) activity is reported in other brain regions when there is no activity in that LPFC location. Meta-analytic connectivity, then, is the odds ratio computed from these two probabilities. At the scale of thousands of task contrasts, it provides an estimate of systematic co-activity across diverse differences in tasks and study implementations.

Importantly, Abdallah et al. (2022), tested whether the high-dimensional variance in these meta-analytic connectivity values across LPFC locations could be reduced to lower-dimensional components. They found that a rostro-caudal dimension of organization emerged from this analysis to account for the most variance in meta-analytic connectivity of LPFC regions (around 40%), followed next by the dorsoventral dimension of organization (around 20%) (Figure 7.1a). To clarify, these dimensions of organization are in reference to variance in meta-analytic connectivity. They do not speak to organization of function along these gradients. In other words, across many different tasks and studies, the variance in co-activation between regions of LPFC and other regions in the brain is systematically related to its position along a rostro-caudal axis. This observation was robust across several controls and ways of doing the analysis; it was also evident at the single subject level using an independent dataset in which people were scanned doing many different tasks.

Two further insights were evident from this study. First, this rostro-caudal LPFC gradient was situated within a broader hierarchy of brain networks defined by their distance from unimodal sensorimotor regions (Huntenburg et al. 2018). Breaking the LPFC into quintiles from caudal to rostral, the meta-analytic connectivity of the most caudal LPFC quintile was found to overlap with visual networks and external attention networks more so than the rostral portions of the gradient. The connectivity of the most rostral LPFC quintile overlapped with cognitive control and default mode networks, with a gradual transition from visual and attention networks to cognitive control and default

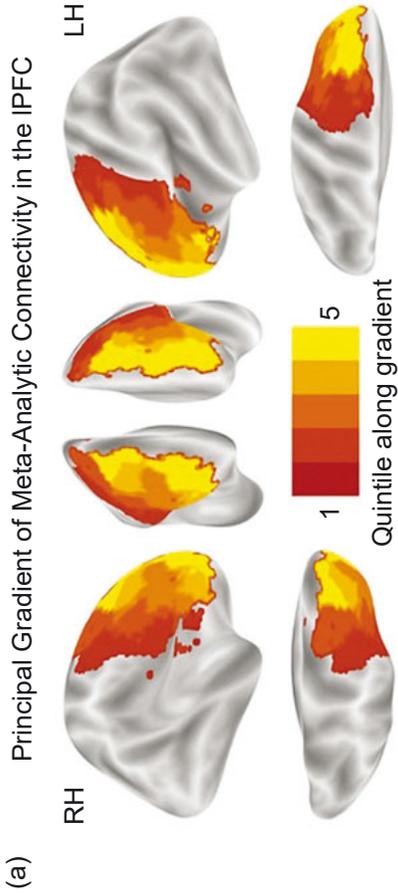
networks across the quintiles. Notably, this pattern of network overlap is also consistent with observations by Choi et al. (2018), who directly compared activations from a task of hierarchical control (Badre and D'Esposito 2007) to the Yeo et al. (2011) parcellation.

Importantly, this network connectivity pattern situates the local IPFC rostro-caudal gradient within a global hierarchical organizing principle in the brain. Specifically, Margulies et al. (2016) analyzed whole brain connectivity data in both human and macaque monkey and found a principal gradient spanning networks in terms of distance from sensorimotor regions, with cognitive control and default mode networks at the furthest extreme. Directly comparing their local gradient of meta-analytic connectivity in IPFC with that macroscale hierarchy of networks identified by Margulies et al. (2016) in resting state connectivity, Abdallah et al. found that the network profile of the local IPFC gradient of meta-analytic connectivity correlated with this macroscale hierarchical organization.

The second insight to emerge from this meta-analysis is that the gradient is differentially related to distinct task demands. Abdallah et al. used a topics analysis (Poldrack et al. 2012) of terms, describing studies in the NeuroSynth database to cluster experiments into 38 topics, ranging from attention and decision making to lexical semantics and memory encoding. They then tested how these topics were associated with activation along the gradient.

This analysis found a systematic pattern of associations across tasks that are consistent with an abstraction gradient (Figure 7.1b). Roughly, the most caudal zone was primarily associated with tasks involving basic sensorimotor functions such as eye movements and attention. The middle caudal zone was most associated with more complex controlled tasks, such as those involving response selection or task switching. The middle rostral zone was chiefly associated with mentalizing and decision making. The furthest rostral zone was most associated with topics related to subjective experience, empathy, declarative memory, and feedback/error processing. Notably, this set of task associations is roughly in line with the gradient of function inferred from the network connectivity profile, ranging from more concrete, external, sensorimotor function to more internal, abstract cognitive function.

Thus, while the IPFC is clearly associated with multiple demands, it is also not uniformly or arbitrarily associated with all tasks without an organization that can be described at a cognitive or computational level of analysis. Rather, a rostro-caudal abstraction gradient is supported across thousands of contrast measurements from fMRI studies of the human brain. Another recent meta-analysis (de la Vega et al. 2018), while coarser in resolution, reached similar conclusions to Abdallah et al. regarding functional differences in IPFC. Thus, whether hierarchical in nature or not, a rostro-caudal functional gradient in IPFC appears evident and robust.



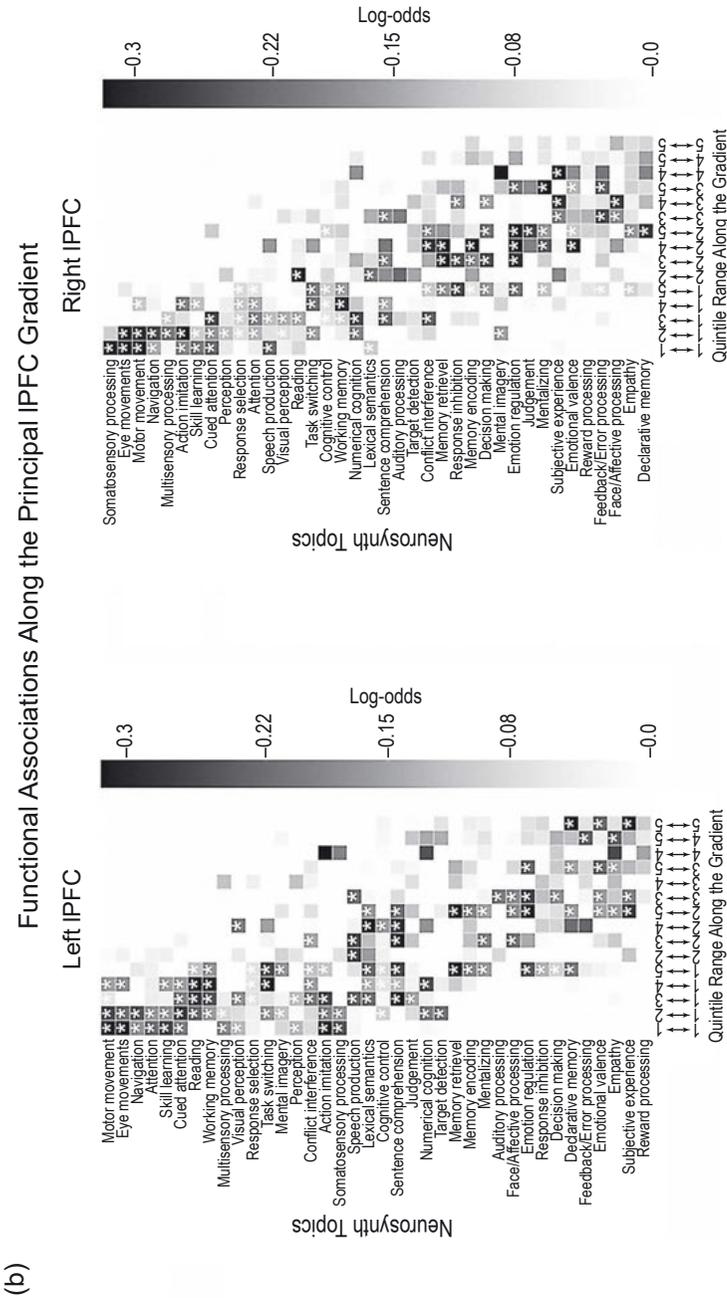


Figure 7.1 Meta-analysis results from Abdallah et al. (2022). (a) Surfaces showing the rostro-caudal gradient in meta-analytic connectivity. (b) Results from the topics analysis associating experiment topics with quintiles of the meta-analytic connectivity gradient going from 1 caudally to 5 rostrally. Box shading shows log-odds and an asterisk indicates statistical significance. Adapted from Abdallah et al. (2022, Figures 1 and 4).

What Defines the Functional Hierarchy in IPFC?

Meta-analysis provides evidence that IPFC has a rostro-caudal organization based on cognitive demands that generalize across tasks. Nonetheless, the specifics of these demands, as can be described at a computational or mechanistic level, remain vague when using this large-scale approach. Even in the Abdallah et al. analysis, for example, it is not clear why particular topics, such as “cognitive control,” are distributed more caudally than others, such as “response inhibition.” More specific task analyses and computational accounts are needed for this level of interpretation.

Carefully controlled laboratory experiments have been useful for testing hypotheses at this more specific computational and functional level. However, several proposals and tests (reviewed in Badre 2008; Badre and Nee 2018) have failed to produce a consistent interpretation. Here, we will consider three that have received recent attention: policy abstraction, internal versus external attention, and present versus future focus.

Policy Abstraction

The earliest establishing fMRI experiments that tested rostro-caudal differences in IPFC manipulated rule complexity in terms of the number of contingencies needed to make a response (Badre and D’Esposito 2007; Koechlin et al. 2003; Yao and Hsieh 2022). These studies found that as choice conflict was manipulated among responses, stimulus-to-response mappings, context-to-stimulus-to-response mappings, and episode-to-context-to-stimulus-to-response mappings, the focus of activation related to that choice demand moved rostrally from premotor cortex to prePM to mid-dIPFC to rIPFC.

One interpretation of this observation is that it reflects a change in policy abstraction needed for hierarchical control. Cognitive control generally refers to cases where a higher-order context is needed in order to distinguish which response to make (Badre and Nee 2018; Botvinick 2008). Hierarchical control refers to cases in which these contextual signals are themselves conditioned on superordinate contexts, which can be conceptualized as a hierarchical or branching rule tree.

The concept of policy abstraction is closely tied to this definition of hierarchical control. Policy, as a concept, comes from machine learning and reinforcement learning, where it refers to a relationship between a particular context, an action, and the expected outcome this will produce (Botvinick et al. 2009). Policy abstraction refers to learning abstractions over these relationships, wherein classes of lower-order policy are captured within a particular higher-order policy. For example, a sequence of specific policies that enact individual movements of an artificial agent around a grid could be abstracted over in terms of a policy that defines the start and end position. Abstraction of not only states and actions, but also their combination, is increasingly influential

in AI and robotics, where these agents face similar dilemmas in planning and decision making in complex settings as humans do (Konidaris 2019).

In the context of hierarchical control, policy takes a similar definition and its abstraction is defined in terms of contingency. Thus, a simple S-R mapping, such as pressing a key when a particular color is presented in order to succeed at a trial, represents a concrete form of policy; it determines what action to take based on the stimulus context that yields a desired outcome. Policy can be more abstract by adding contingencies. Thus, task switching involves grouping sets of S-R mappings to follow based on a task cue (e.g., following all the color response rules when a stimulus appears in the upper half of the screen). In this sense, task-set-level policy is more abstract because the task context does not specify what specific response to make (as occurs in simple S-R mapping), but rather specifies how to interpret a set of stimulus contexts with regard to selecting a response. Further contingencies could be added, for example, specifying how to interpret screen position with regard to the task to perform. As contingencies are added, the policy becomes higher order.

People's performance suggests that they break complex tasks down and represent them efficiently as hierarchically structured, abstract policies. In task-switching manipulations, switches of higher-order policy, as defined above, show larger switch costs than lower-order switches, consistent with traversing different choice points on a policy tree (Kleinsorge and Heuer 1999; Rac-Lubashevsky and Frank 2021; Ranti et al. 2015). An advantage to hierarchies is that decisions can be made at each level separately, and the status of lower-level decisions need not affect upper-level ones. Consistent with this, people will solve hierarchical rules in parallel, in line with decisions being made partly independently at each level (Rac-Lubashevsky and Frank 2021; Ranti et al. 2015). Hierarchies also permit generalization, such as over lower-order S-R mappings and transfer to new tasks with the same contingency structure. Indeed, people learn hierarchical rules more rapidly through reinforcement than rule sets that are not structured hierarchically (Badre and Frank 2012; Badre et al. 2010; Eichenbaum et al. 2020; Frank and Badre 2015). Further, when possible, they impose a hierarchical structure on tasks which builds a more abstract structure that can support transfer to new tasks (Collins and Frank 2013; Collins et al. 2014). Indeed, people structure tasks hierarchically, even when doing so conveys no immediate behavioral advantage and potentially comes with a cost in mental effort (Sayali et al. 2023).

It is evident, then, that people control their behavior during complex tasks in hierarchically structured ways based on more abstract policy. Thus, one account of abstraction differences along the rostro-caudal axis is that they reflect the neural processing needed to behave according to increasing levels of policy abstraction. The results from the fMRI experiments described above are consistent with this interpretation, in that manipulating selection demands at higher levels of policy abstraction resulted in more rostral IPFC activation. More recent experiments using different tasks (Nee 2021; Nee and D'Esposito

2016), but which can again be analyzed in terms of progressive increases in policy abstraction (Sayali et al. 2023), show a similar rostro-caudal pattern of activation differences. Further, complementary support for the hierarchical assumptions made in these experiments has come from studies in patients with lesions in LPFC (Azuar et al. 2014; Badre et al. 2009) and transcranial magnetic stimulation (Nee and D’Esposito 2017), and it supports differences in the necessity of rostral versus caudal PFC for following complex versus simpler rules, respectively.

Observations from other experiments, however, are difficult to explain within the simplest policy-abstraction account. Hierarchical manipulations of the 12AX-CPT task—in which a higher-order number context (1 or 2) determines which context (A or B) determines whether lower-order items (Xs or Ys) are targets—have produced inconsistent results with regard to the rostro-caudal gradient, despite the clear policy manipulation. One fMRI study using a blocked design failed to locate differences along this axis when comparing these different levels of contingency (Reynolds et al. 2012). However, other designs using this task did locate rostro-caudal differences in fMRI activity associated with higher- and lower-order contexts, though in different locations along the rostro-caudal axis than would be predicted by prior studies (Nee and Brown 2012, 2013).

It is conceivable that these differences might relate to specific aspects of the experimental protocols, such as the serial versus simultaneous nature of presentation (see Badre and Nee 2018). Nonetheless, if that is the case, it also indicates that factors beyond policy abstraction are relevant to the LPFC organization. Similarly, a recent study of the multiple demand network found that while there was consistent functional differentiation rostro-caudally in this network, it was due to factors like reward and time pressure that were not clearly attributable to a policy-abstraction hierarchy (Crittenden and Duncan 2014; Shashidhara et al. 2019). Thus, while policy abstraction may be important for the rostro-caudal organization of LPFC, it is evidently not the only relevant factor.

The function of rLPFC presents another problem for a single policy-abstraction gradient. While there is ample evidence from neurophysiological recording of abstract rule processing in mid-dLPFC (e.g., Mansouri et al. 2020; Wallis and Miller 2003a; Wallis et al. 2001), the few recording studies of the anterior frontal pole in the monkey have not found more abstract rule coding or, indeed, rule coding at all (Tsujimoto et al. 2010). Neuroimaging studies of hierarchical sequence control in humans have found a strong association of rLPFC with superordinate or sequence-level effects that could be interpreted as higher-order policy (Desrochers et al. 2015a, b, 2019). The specifics, though, do not fit with a simple policy gradient. The activity in rLPFC ramped toward the end of sequences and brain stimulation with TMS also had disruptive effects at the end of the sequence. How this ramping dynamic relates to simple policy-abstraction demands is unclear. Further, in terms of its functional relationship to other regions of PFC, rLPFC is unlikely to be the apex of

the control hierarchy and so should not necessarily be expected to represent the highest levels of policy abstraction.

Thus, while our ability to behave according to higher-order policy is closely related to IPFC and its rostro-caudal gradient, it is unlikely that policy abstraction ranks the entire functional gradient in IPFC. Other ideas have continued to be pursued in recent years regarding the organization of IPFC along its rostro-caudal axis that relate to policy abstraction, but which also differ in important ways.

External- versus Internal-Oriented Cognition

Returning to the ordering of tasks in Figure 7.1b, another hypothesized organizing dimension of the lateral gradient can be recognized. Specifically, one can informally describe the tasks as shifting from those primarily requiring externally oriented processing, such as tasks of multisensory perception and cued attention, to those requiring an internal orientation, such as tasks of declarative memory or emotion regulation.

In this view, going from sensorimotor networks to internal processing of the default mode network, the gradient in IPFC is not different qualitatively from the larger hierarchical network organization of the brain (Margulies et al. 2016). According to this hypothesis, mid-dIPFC has a multi-demand and integrative nature due to its interposition between areas or networks involved primarily in internal (rIPFC and default mode network) versus externally oriented (sensorimotor and dorsal/ventral attention networks) control. The mid-dIPFC may, therefore, play a crucial role in linking the internal control of thought with its externalization in behavior.

As with policy abstraction, the first neuroimaging experiments that provided evidence of a rostro-caudal organization of IPFC could be interpreted in terms of a progression from internally to externally oriented control (Badre and D'Esposito 2007; Koechlin et al. 2003). In particular, the highest levels of both tasks—those associated with the most rostral IPFC activation in each study, though differing in their specific location of activation—placed a demand on “episodic control,” which meant that the particular task episode acted as a context for selecting the appropriate set of S-R mappings. There was no external cue for this episode, so it had to be tracked internally.

A similar observation was made in a series of studies by Nee and D'Esposito (2016, 2017) using the Comprehensive Control Task (CCT), which manipulated simple S-R selection (sensorimotor control), contextual S-R selection (contextual control), and temporally extended selection based on items held in memory (temporal control). These manipulations and their network associations fit with a shift from external to internal processing and have consistently associated these three demands with progressively rostral areas of IPFC. Hence, there is some empirical support for defining the rostro-caudal organization of IPFC in terms of an external to internal processing dimension.

Nevertheless, there are also limitations. First, at a more mechanistic level, outside of proximity to sensory versus default networks, it is not fully clear what defines internal versus external processing. For example, most models of cognitive control assume that information must be held in working memory by PFC for it to affect current processing, even of contextual information available in the external world. This requirement is important as not all information we sense should serve as a control signal. So, deciding to represent something in working memory constitutes a decision about whether to allow it to influence behavior. As such, this information must be internally represented, even for simple sensorimotor control. Indeed, studies of selection from within working memory, based on an internally maintained context, have located activation in prePM, caudal to mid-dIPFC (Chatham et al. 2014). This type of selection from within working memory, also termed output gating in the context of hierarchical control, is an internally oriented cognitive control demand, yet, it is not selectively associated with rostral PFC.

Second, the direct experimental support distinguishing episodic control from other types of control demands is, at present, weak. Pitts and Nee (2022) modified the CCT in a way that manipulated episodic control demands factorially, relative to other demands. The contrast of low versus high episodic control resulted in a pattern of overlapping activation, which for contextual control was based on a stimulus cue and caudal to the original “temporal control” manipulation. Thus, at present, conceptualizing the rostro-caudal hierarchy along a strict gradient of external to internal processing dimension has mixed theoretical and empirical support.

Present- versus Future-Oriented Control

A related alternative to the internal-external gradient distinguishes between present- versus future-oriented control (Badre and Nee 2018; Nee 2021; Soltani and Koechlin 2022). Most experimental manipulations of cognitive control are in the moment. Whether one is naming the ink color of a word or switching between tasks, one is selecting a particular response or task set to perform right now as well as in expectation of whatever outcomes follow from that behavior. However, we are also able to monitor cues and encode information that is relevant to our future or hypothetical behavior, rather than what we are doing right now. Pitt and Nee (2022) have pointed out that the external versus internal focus of control is often confounded in experiments with this temporal focus on present versus future.

The hypothesis that the rostro-caudal axis of IPFC is differentiated by a present to future orientation is broadly consistent with several observations from the neuroimaging literature regarding, in particular, rIPFC. FMRI studies have associated with rIPFC cognitive branching (Koechlin et al. 1999; Koechlin and Hyafil 2007), with monitored conditions driving exploration of future alternative behaviors over exploitation of current behaviors that are leading to present

rewards (Badre et al. 2012; Boorman et al. 2009; Culbreth et al. 2023; Daw et al. 2006), and with counterfactual task set control, as in what task was performed versus what could have been performed (Donoso et al. 2014a; Soltani and Koehlin 2022). In common, these tasks require some monitoring of which alternative paths to take. This form of monitoring may not affect behavior on the present trial, but to the degree that one successfully tracks future or possible rewards, this tracked information could affect behavior on future trials. Consistent with the present/future characterization of IPFC organization, Nee (2021) found that activation in more caudal areas of IPFC correlated with response times on current trials of the CCT, whereas activation in the more rostral areas was associated with response times on future trials in these experiments. In sum, there is some convergence across different tasks and experiments that a shift of focus from present to future orientation might characterize processing along rostro-caudal IPFC.

The distinction between future- and present-oriented control may also fit with the connectivity of these regions of IPFC and at a computational level, a distinction between representations suitable for action planning versus action execution. Vaidya and Badre (2022) recently observed that independent lines of research have separately implicated medial temporal lobe and orbitomedial PFC (MTL-OMPFC) versus frontoparietal networks in the same function, the representation of abstract task set information. Why do distinct networks represent the same abstract information? Vaidya and Badre (2022) proposed that this might reflect differences in the format or use of these representations rather than their content (Lovett and Anderson 2005), with the MTL-OMPFC network representing abstract tasks to plan what to do versus frontoparietal networks that format abstract task information for efficient task performance (Figure 7.2). Specifically, MTL-OMPFC representations may organize task information in a map-like format that allows multiple relationships among task states to be represented, and new relations to be inferred. By contrast, frontoparietal networks may represent task information as productions that specify what action to take in a given a state. Productions are unidirectional, and so while not as useful for planning, they can yield efficient and controlled action selection given a set of states.

Several convergent lines of research support this hypothesis. Results across species implicate a network involving ventromedial and orbitofrontal cortex along with the medial temporal lobe (MTL-OMPFC network) with the representation of abstract task information, such as latent contextual states, and with drawing inferences based on these representations (e.g., Bradfield et al. 2015; Chan et al. 2016; Coutureau et al. 2002; Iordanova et al. 2007; Jones et al. 2012; Schuck et al. 2016; Wilson et al. 2014; Zhou et al. 2021a, b). These results have specified that this MTL-OMPFC network encodes a cognitive map of task-space as a way of efficiently representing structured task relationships useful for planning and inference.

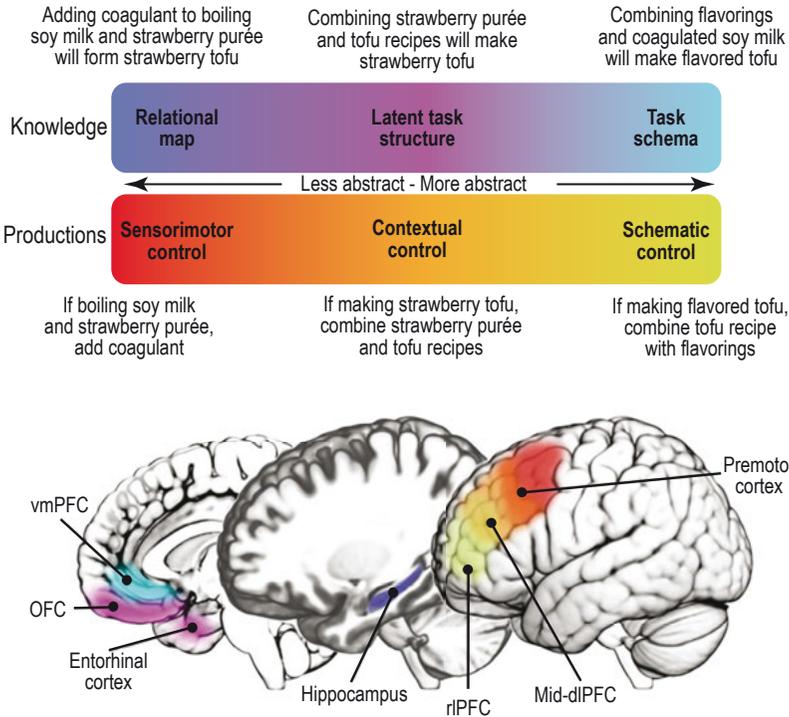


Figure 7.2 Schematic from Vaidya and Badre (2022) summarizing the hypothesized relationship between the IPFC control network and the MTL-OMPFC network in representing abstract task knowledge. It shows two abstraction gradients of organization task planning and inference (blue to cyan) and task production (red to yellow). The gradient in IPFC follows the distinction from sensory to cognitive to schematic control proposed by Badre and Nee (2018). The examples in black are intended to illustrate the shared content at each level, but differences in format. Reproduced with permission from Vaidya and Badre (2022).

On the other hand, there is evidence that the lateral frontoparietal network represents abstract task sets and can leverage inferred relationships in order to behave. For example, we recently observed that the frontoparietal network, particularly the mid- and anterior portions, represented inferred abstract task relationships while performing an acquired equivalence task. (Vaidya et al. 2021). Specifically, across learning phases, participants learned a latent, abstract task set that they could use to generalize behavior to new cases they had not previously encountered through inference. People were able to do this generalization and perform the task. However, possibly because learning occurred during performance of the task rather than the period of inference, decoding results found the latent task set information to be maintained in mid and anterior IPFC, with limited activity in MTL-OMPFC.

From “The Frontal Cortex: Organization, Networks, and Function,” edited by Marie T. Banich, Suzanne N. Haber, and Trevor W. Robbins. 2024. Strüngmann Forum Reports, vol. 34, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262549530

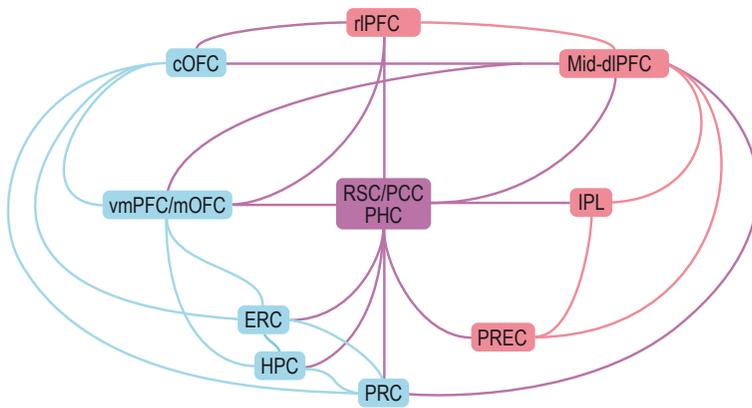


Figure 7.3 Schematic of connections between regions along the rostro-caudal axis of IPFC and those in the MTL-OMPFC network, based on a review of anatomical studies in the nonhuman primate. Light red shows regions and connections in the frontoparietal network. Light blue shows regions and connections in the MTL-OMPFC network. Purple marks the connections between the networks. Notably, rIPFC and mid-dIPFC, in the rostral IPFC, hold direct connections with regions in the MTL-OMPFC network. Reproduced with permission from Vaidya and Badre (2022).

Importantly, to connect our plans with their realization in our behavior, there must be an interface between these systems for planning and for execution of tasks. Figure 7.3 summarizes connections between these networks, based on a review of anatomical studies in nonhuman primates (Vaidya and Badre 2022). In the IPFC, it is notable that rIPFC and mid-dIPFC in the IPFC share connections with orbitofrontal cortex and the ventromedial PFC, perhaps positioning them as the interface between the MTL-OMPFC network and the frontoparietal control network. These ideas elaborate what Badre and Nee (2018) distinguished as “schematic control” (as distinct from sensory and cognitive control; see Figure 7.2) in their review of the literature around hierarchical control. Nevertheless, direct testing of these ideas in experiments designed to distinguish planning from execution is needed.

Hierarchical Interactions Within the Rostro-Caudal Organization of IPFC

Our discussion to this point has concerned the functional attributes that might characterize processing or representations along the rostro-caudal gradient of IPFC. An important implication of hierarchy in IPFC is, however, that it is not merely a description of function going from more concrete sensorimotor control to more abstract cognitive and then schematic control, but that it reflects a processing architecture. Specifically, a hierarchy of processing within the IPFC

would mean that superordinate regions of IPFC asymmetrically influence the processing of subordinate regions of IPFC.

Several anatomical gradients have been noted across species in IPFC that are consistent with a hierarchical organization, including the transitions from granular to agranular cortex and changes in connectivity across areas from caudal to rostral (Badre and D'Esposito 2009; Jacobs et al. 2001; Phillips et al. 2021; Sanides and Sanides 1972; Thiebaut de Schotten et al. 2016; Yeterian et al. 2012). However, over the last several years, more directed anatomical and functional investigations have found evidence for a processing hierarchy in IPFC, as defined above, that emphasizes two main points. First, the rostral mid-dIPFC is the top of the hierarchy, in that it exerts an asymmetric influence over other IPFC regions, including rIPFC which is rostral to it. Second, cortico-striatal interactions appear integral to hierarchical processing within IPFC.

Mid-dIPFC As the Apex of the Hierarchy

One way to index position within a processing hierarchy is in terms of input versus output connectivity. As influence should be asymmetrically super- to subordinate, regions higher in the hierarchy should exert broader influence than those which are lower. It follows that superordinate regions would have a higher ratio of output to input than subordinate regions (Badre and D'Esposito 2009).

Goulas et al. (2014) used the CoCoMac database of anatomy (Kotter 2004) in the macaque prefrontal cortex to test this ratio of efferent to afferent connections across regions of IPFC. A strictly rostro-caudal hierarchy predicts that the rIPFC should show the highest ratio. However, Goulas et al. did not find evidence that area 10 in the monkey was at the top of the hierarchy. Rather, they observed that mid-dIPFC showed the highest output asymmetry and so should be the apex of the hierarchy by this definition.

Evidence from effective connectivity studies in humans provided converging support. Across multiple studies using dynamic causal modeling of fMRI data from the CCT (Nee 2021; Nee and D'Esposito 2016; Pitts and Nee 2022; Wood and Nee 2023), a consistent pattern of effective connectivity within IPFC has emerged that places mid-dIPFC on top (Figure 7.4). In particular, mid-dIPFC exerts influence over both caudal sensorimotor control areas of IPFC, as well as rIPFC and schematic control areas. Further, stimulation of regions in this network using TMS produced hierarchical effects on behavior that were consistent with this pattern of super/subordinate relationships (Nee and D'Esposito 2017).

Taken together with the functional attributions we have discussed along the rostro-caudal axis of IPFC, the picture emerging from these studies is one of multiple zones of integration within IPFC, hierarchically ordered with respect to each other. In particular, caudal regions of IPFC sit at the interface of sensory input and movement. The most rostral regions are at the interface between planning and inference networks. In the middle, at the apex, the mid-dIPFC

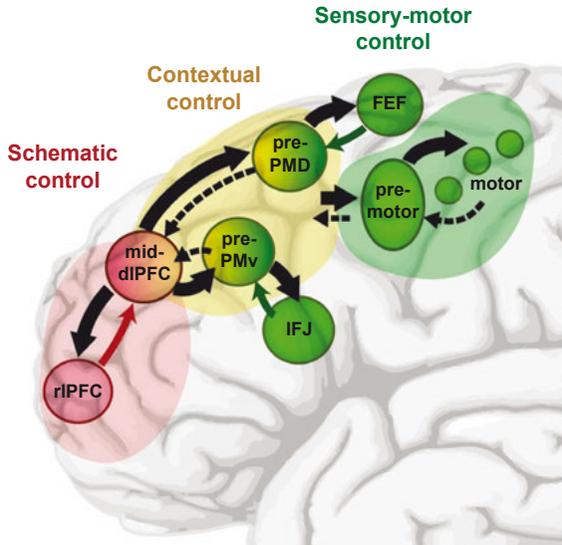


Figure 7.4 Schematic showing summarizing interactions between regions along the rostro-caudal axis of IPFC. The three zones defined by Badre and Nee (2018) of schematic, contextual, and sensorimotor control are distinguished by colored shading with regions labeled in each. Large arrows show the consistent primary direction of influence. Broken or colored arrows are weak or task-dependent influences. Mid-dIPFC shows the strongest outgoing influences both rostrally to schematic control regions, and caudally, to cognitive and sensory control regions. Reproduced with permission from Badre and Nee (2018).

is the link between these two systems, allowing control to be influenced by multiple forms of information from both the world around us, as well as our internal planning, memory and affective systems.

Cortico-Striatal Circuits and Hierarchical Control

Evidence from effective connectivity analysis of fMRI data in humans is consistent with the asymmetric anatomical connections found in nonhuman primates. However, effective connectivity measured in fMRI is not necessarily due to direct cortico-cortical interactions. Functional and effective connectivity can also reflect complex, polysynaptic network interactions. Indeed, one such contributor to hierarchical interactions among regions of IPFC may be the basal ganglia and its interactions with IPFC through the thalamus.

A recent study using Granger causality analysis of fNIRS data at rest (Schumacher et al. 2019) reproduced the asymmetric cortical pattern of functional connectivity observed using fMRI, in that there was an overall rostro-caudal pattern of influence, but with mid-dIPFC showing the strongest influence on other regions. Interestingly, a second study (Schumacher et al.

2021) of patients diagnosed with Parkinson disease undergoing deep brain stimulation of the subthalamic nucleus replicated this observation, but then found that turning on the stimulator enhanced this pattern of rostro-to-caudal IPFC asymmetry in patients early in the progression of the disease. Though limited by sample size and sample heterogeneity, and not linked directly to behavior, these observations provide evidence that interactions with basal ganglia structures, like the subthalamic nucleus, might play a causal role in the asymmetry of influence among IPFC regions.

What is the functional role of the basal ganglia with regard to the rostro-caudal hierarchy? One hypothesis is that the basal ganglia supports working-memory gating in the service of hierarchical cognitive control (Frank and Badre 2012; Frank and O'Reilly 2006). As already introduced, working memory plays a central role by maintaining information that can serve as a control signal. The IPFC, in particular, is hypothesized to maintain the control representations needed for this function (Miller and Cohen 2001). Computational models of this mechanism have demonstrated, however, that working-memory gating is required to select what information to hold as a control signal in working memory, "input gating" and when to allow it to influence behavioral choice, "output gating" (Frank et al. 2001; Hochreiter and Schmidhuber 1997; O'Reilly and Frank 2006). Further, tasks often require selectively input gating and/or output gating one representation, while holding others in mind. This is particularly important for hierarchical control tasks in which superordinate contexts are held steady, while subordinate goals are frequently updated.

Frank and O'Reilly (2006) proposed the prefrontal cortex basal ganglia working memory (PBWM) model, which implements selective working-memory gating using a cortico-striatal-thalamic mechanism that is assumed to perform the same core computations as more established mechanisms of motor control. In motor control, candidate actions are represented in premotor cortex but not executed because cortico-thalamic drive is under inhibition from the globus pallidus. However, topographically organized loops through striatum can disinhibit or gate particular responses. Moreover, through dopaminergic signaling, reinforcement learning can modulate synaptic weights in the striatum so that its gating responses reflect a cost-benefit decision about the candidate actions.

PBWM proposes that parallel cortico-striatal-thalamic loops perform analogous computations on cortical representations in IPFC, thereby acting as an adaptive gate on working memory. Further, because of its topographic organization, subcircuits can modulate specific cortical populations thereby allowing selective gating. The relationship between cortico-striatal interactions, working memory, and dopamine signaling predicted by the model has been supported by multiple lines of evidence in humans and animals (Chatham et al. 2014; Cools et al. 2006; Dagher and Robbins 2009; Frank et al. 2004; Jin et al. 2014; Lee et al. 2015; McNab and Klingberg 2008; Schmitt et al. 2017; Schonberg et al. 2010; Stollstorff et al. 2010; Tai et al. 2012; Voon et al. 2010).

PBWM also proposes dynamics among multiple cortico-striatal loops that serve as the basis of hierarchical cognitive control (Badre and Frank 2012; Frank and Badre 2012). There is a well-established organization of anatomical connections rostro-caudally between striatum and IPFC. This organization has been observed in detailed tracing studies in animals (Haber 2003; Haber et al. 2020), as well as in human studies using diffusion-weighted tractography (Verstynen et al. 2012) and functional connectivity (Choi et al. 2012, 2018). PBWM suggests that different loops can update contextual representations in working memory at different levels of a task hierarchy. For example, one loop might gate working-memory superordinate goals (e.g., making a sandwich), while another loop is gating working memory for the subordinate goal (e.g., slicing bread). Importantly, asymmetric diagonal connections from the superordinate loops higher in the rostro-caudal organization, to the striatal region gating the subordinate loops, allow these higher-order contexts to influence differentially the gating decisions made at the lower level. Note, this hierarchical gating system could operate over whatever factors are functionally differentiating regions of IPFC along its rostro-caudal axis. They would allow control to be made compositional and executable through productions to match the complex tasks we confront (Bhandari and Badre 2018).

Some initial evidence supports this hypothesized nested looping architecture in the context of hierarchical control. Evidence from model-based fMRI, which correlates parametric functions estimated from a computational model of learning with fMRI BOLD signal change, indicates that specific cortical and striatal sites, aligned along the rostro-caudal dimension, were sensitive to reward prediction errors at specific levels of policy abstraction (Badre and Frank 2012). A study of artificial grammar learning observed three separate pairs of IPFC-striatal foci associated with different levels of task complexity (Jeon et al. 2014). These sites were connected based on diffusion tractography. Further, it has been observed in high-fidelity diffusion tractography that IPFC-striatum connections are not only ordered rostro-caudally; when connections deviate from this pattern, they are more likely to do so from rostral IPFC to caudal striatum, than vice versa, consistent with asymmetric diagonal connections that imply a hierarchy (Verstynen et al. 2012). Nonetheless, more evidence is needed to connect selective gating at multiple levels during cognitive control to interacting cortico-striatal loops.

In sum, there is evidence that both cortico-cortical and cortico-striatal connections in IPFC may support a hierarchical architecture with mid-dIPFC at its apex. Taken together with the functional divisions described in the preceding section, these interactions may describe how information from the sensory and planning systems are not only integrated but used as contextual signals for control.

Conclusions and Future Directions

Over the last two decades, our understanding of the hierarchical organization of the IPFC has progressed in a number of ways. There is now strong evidence that an axis of functional differentiation exists from caudal to rostral. Further, it seems clear that this IPFC organization is situated within a larger organization of brain networks that bridge from concrete sensorimotor function at one end, to more abstract cognitive function at the other. There is also consistent evidence that this functional organization is hierarchical in its processing character, with the mid-dIPFC at the apex of this hierarchy and cortico-striatal interactions playing an important role.

In addition to continuing efforts to test and revise these hypotheses, there are a number of future directions and open questions to be addressed. For example, mechanistic investigation of the ways that this IPFC hierarchical organization interacts with other areas of PFC—or the brain more generally—such as with medial PFC (Shenhav et al. 2018; Venkatraman et al. 2009b; Wood and Nee 2023) or cerebellum (D’Mello et al. 2020), is needed to add specificity regarding the IPFC role as a controller. This line of investigation should be combined with lesion or other manipulations that allow us to understand the causal influence that PFC has on other frontal lobe regions and processing in the brain more generally. Understanding how this architecture changes and contributes to the development of cognitive control will provide important insights (e.g., Amso et al. 2019; Freier et al. 2021; Unger et al. 2016).

Further, while considerable evidence for a functional hierarchy in PFC has been reported from humans studies using fMRI, lesion, and TMS approaches, this organization has not been thoroughly examined and tested in animals models. What physiological evidence we do have suggests that neural populations in IPFC encode most task information, and while there are some gradients (see Rich and Averbeck, this volume), there are not large qualitative differences in cell coding along the rostral-caudal axis. Thus, reconciling these literatures will require direct study of hierarchical control and tasks thought to engage this axis in animal models, as well as complementary approaches in humans, such as those using intracranial recordings (e.g., Johnson et al. 2023).

A related open question concerns the organization of the neural representations themselves that occur in different regions along the rostral-caudal axis. There is a growing focus in the broader field on the geometry and dynamics of neural representations in terms of how similarly neural populations represent their inputs during a task (e.g., stimuli, contexts, responses, task sets) in their patterns of neural activity, and how these patterns of similarity change over short and long timescales (Badre et al. 2021). The geometry of neural population coding is known to affect computation (Fusi et al. 2016). For example, whether neural populations encode their inputs as a small set of abstract low-dimensional components or as high-dimensional mixtures balances a trade-off between generalizability versus separability. That trade-off might affect

behavior, if one is engaged in learning versus interference resolution (Badre et al. 2021; Fusi et al. 2016). Thus, how these features of neural representational geometry interact with the rostro-caudal organizing gradient of IPFC will be important to understand. For example, it may be that lower-dimensional representations that comprise the components of a hierarchical task structure are represented separately from more integrated conjunctive representations that represent specific instances of a task. From studies with human EEG, we now know that such high-dimensional, conjunctive representations are important for determining performance on a trial-to-trial basis (Kikumoto and Mayr 2020) as well as for maintaining and prioritizing action plans in working memory (Kikumoto et al. 2022). These neural representations are involved when performing hierarchically structured tasks; thus, it is important to understand their relationship to the rostro-caudal organization of the IPFC.

Just as repeated experimentation and testing of ideas around the rostro-caudal organization of IPFC has changed our view of this organization over the last several years, experiments in these domains promise to continue to do so. Study of this problem across levels of analysis and using a range of approaches will give us a clearer picture of the functional significance of this dimension of PFC organization.

Acknowledgments

I am grateful to current and former Badre Lab members, as well as to Mark D’Esposito, Derek Nee, John Duncan, Theresa Desrochers, Michael Frank, and many other colleagues for helpful discussions on these topics over the years.



From “The Frontal Cortex: Organization, Networks, and Function,” edited by Marie T. Banich, Suzanne N. Haber, and Trevor W. Robbins. 2024. Strüngmann Forum Reports, vol. 34, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262549530