

Pathways Forward Toward an Understanding of Frontal Lobe Function

Marie T. Banich, Suzanne N. Haber, and Trevor W. Robbins

In this concluding chapter we examine some of the cross-cutting themes that emerged from the Forum. Here, we consider issues that transcend the individual working groups, which we believe are ripe for further discussion and investigation, notably translation from animal to human models, the role of connectivity in frontal function, and unique aspects of human cognition that are supported by the frontal lobes.

Animal Models: Utility, Limitations, and Future Potential

Comparisons across species were a recurrent theme at the Forum, figuring to variable extents in each of the four working groups. There are essentially two major reasons for this. First is the intriguing issue of how prefrontal cortex and its associated functions evolved, which is inevitably bound up with consideration of what were the main drivers of human evolution. Second, is the more pragmatic issue of how studies on infra-human animals can inform the understanding and possible treatment of clinical neurological and psychiatric disorders associated with the frontal lobe, through “animal models.” Although human brain imaging methods are constantly being refined, in terms of resolution, modality, and analytic sensitivity, to provide sophisticated regional and functional network maps of the human cortex, they still cannot provide detailed information at the cellular, molecular, and circuit levels. This information can only be provided by animal models, which are essential for understanding the underlying the causal mechanisms that lead to effective treatments of human disorders.

Comparing possible behavioral functions across species can also be problematic given the complexity of human cognition. However, it is a useful exercise to identify test procedures, such as the stop signal reaction time task, that procedurally appear to utilize comparable requirements across species, for

example, in terms of contingencies and inferred requirements for perception, memory, and behavioral control. Thus, classical tests of prefrontal function in humans were reported at the Forum to have several parallels in nonhuman primates (NHPs) and rodents. However, operational parallels in performance may not necessarily be matched by the quite the same psychological processes in humans and other animals, given the additional capabilities of humans in language, insight, and rapid learning. Nevertheless, it was deemed reasonable to assume that comparable performance in test paradigms in experimental animals may at least identify some of the “cognitive building blocks” of more advanced functions in humans. This notion of “building blocks” is supported by the evident hierarchical nature and rostral-caudal gradients of organization of the prefrontal cortex across species, an important theme of the Forum.

This comparison can, of course, be strengthened by finding factors that appear to affect behavioral performance in the same manner across species, including importantly neural mediation. If the same brain regions cross-species can be shown to be necessary for mediating behavioral performance, that then heightens the likelihood of effective translation of findings trans-species. This translational approach runs into special difficulties, however, in the case of the prefrontal cortex, in terms of the homology of its component regions. Homology (i.e., shared origins of structure among species) was a major theme for the first discussion group (Weiner et al., Chapter 4). This group considered several criteria for establishing the principle of homology, which entails conservation of structures during evolution, and agreed that the two most important were (a) the detailed histological (cytoarchitectonic) composition of different (pre)frontal regions and (b) their neural connectivity, not only within themselves but also with other brain regions. The latter highlighted a major recurrent theme; namely, networks within prefrontal regions as well as between prefrontal regions and other areas of the brain are likely very important in influencing and determining prefrontal function.

Overall, the classical position was generally supported: only some parts of the rodent prefrontal cortex—mainly posterior orbitofrontal cortex, prelimbic and infralimbic (medial prefrontal cortex) and some parts of dorsal anterior cingulate—have obvious counterparts in the primate brain, whereas counterparts to the dorsolateral (dlPFC) and ventrolateral prefrontal cortex and frontopolar cortex are not so evident. Vertes et al. (Chapter 3) provide an intriguing alternative proposition, positing “two prefrontal streams”: one based on orbitofrontal cortex and the other based on anatomical observations made on the tree shrew prefrontal cortex. This may provide some supporting evidence of dlPFC precursors existing in rodent medial prefrontal cortex, but this position is clearly controversial and it would need considerably more evidence to overturn the classical view that rodents do not have a dlPFC. Overall, infra-human primates, such as the macaque rhesus monkey, undoubtedly provide the best structural model of the human brain and may even indicate precursors of specialized human regions, such as Broca’s area, whereas New World Monkeys,

such as the marmoset, may provide an economically effective, compromise option for translational research involving NHPs (see Rowe et al., Chapter 16).

Neuroimaging and neural network analysis in humans can be paralleled quite readily in NHPs. Murray and Constantinidis (Chapter 6) and Rich and Averbek (Chapter 5) illustrate the potential of these animal models in providing a more refined analysis of prefrontal organization and function at cellular levels through multiunit recordings. Ideally, these types of recordings should now be performed concurrently in several prefrontal regions to elucidate the interactions among different (pre)frontal regions as well as the connectivity of these regions to noncortical areas (e.g., as in the frontostriatal pathways). In a complementary manner, investigations using sensitive anatomical tracing methods that help to define prefrontal connectivity relevant for human prefrontal networks (see Weiner et al., Chapter 4) will continue to be essential. However, NHPs were considered not to be entirely optimal as pragmatic animal models of human mental health disorders because NHPs cannot be used in the large numbers required to match those employed in large-scale human clinical or neuroimaging studies nor for drug discovery (see Roberts and Liston, Chapter 13), especially as such animal models frequently require an intervention such as a stressor or genetic manipulation which are ethically, as well as technically, difficult to employ in NHPs.

The utility of rodent models is that they allow for more invasive perturbations, such as genetic manipulations (e.g., genetic knock-out preparations), lesions, or stress as well as richer and more sophisticated procedures for tracking the activity of functional neural circuitry, including optogenetics, chemogenetics, or fiber photometry to measure calcium or neurotransmitter fluxes (see Izquierdo, Chapter 2). Although promising examples of the use of NHPs in some these techniques were described by Weiner et al. (Chapter 4) and Rowe et al. (Chapter 16), the general observation was that some studies are pursued more effectively and economically in rodents, especially mice.

Limitations

This partial dependence on rodents for understanding aspects of prefrontal organization and function leads to two major problems. First, an obvious one, is that only some “primate” prefrontal regions may be homologous in the rodent brain; hence rodents cannot effectively model the roles of such regions in primates. Second, there are instances where homology did appear to apply, but the underlying functions studied in both rodents and primates did not align, either because they were distinct or did not appear to operate in the same manner (for an example, see Weiner et al., Chapter 4). Of course, it is conceivable that some (pre)frontal structures that appear to be homologous based on cellular organization and connectivity may nonetheless have evolved to perform different functions, but this possibility is not helpful for the triangulation approach to translation. It is then often necessary to focus on behavioral similarities across

species and the common effects of other variables such as drug treatments (for examples, see Rowe et al., Chapter 16), genetic expression and environmental challenges such as stress (including during early life and forms of social deprivation) to achieve translational validity.

Future Potential: Linkage to Clinical Issues

Several interfaces have to be negotiated when translating findings from rodents to NHPs to humans for clinical use. Moreover, this translation should be bidirectional. Clinical observations and issues should not only be inspirational, they should also influence the precise questions that are posed and used to design preclinical studies. For example, detailed understanding of the neurobehavioral basis of distinctive symptoms should be pursued using current theoretical psychological or cognitive conceptions as this will also help to test the utility of those theories. We suggest that this goes beyond the Research Domain Criteria (RDoC) approach (Kozak and Cuthbert 2016) which identifies a matrix of relevant constructs (e.g., inhibition/suppression) applied to specific psychiatric disorders to identify commonalities. This modified RDoC approach is potentially useful when considering transdiagnostic and “dimensional” approaches to neuropsychiatric nosology; for example, the existence of impulsive-compulsive symptoms in many diagnoses (from ADHD and OCD to neurodegenerative disorders such as progressive supranuclear palsy), in which dysfunction of frontal regions or frontal connectivity is implicated. Commonality of symptoms across disorders may indicate some commonalities in approaches to their treatment, based on the fact that they may have overlapping impairments in neural circuitry (see Rowe et al., Chapter 16).

The availability of some modern techniques, such as optogenetic or chemogenetic stimulation, may enable, at the level of cellular resolution, the simulation of particular deficits in neural network function with much greater specificity than hitherto possible (see Izquierdo, Chapter 2). Such manipulations may be important, as a gross malfunction of a neural network could result from a variety of different deficits at the level of its contributory nodes or from impairments in distinct molecular or cellular components. Hence, superficially similar behavioral phenotypes may result from deficits in different mechanisms. In addition, the moot question of how many treatments—including pharmacological (Roberts and Liston, Chapter 11), neuromodulatory (e.g., deep brain stimulation, and rTMS (Rasmussen, Chapter 15), and psychological interventions (Jaeggi et al., Chapter 14)—actually work in neural network terms is currently not well understood. Better understanding of their mode of operation and underlying mechanisms may help lead to more refined versions of those treatments. One futuristic projection from the first working group (Weiner et al., Chapter 4) was that combined pharmacological/surgical procedures involving chemogenetic interventions may hold promise in the treatment of human psychiatric and neurological disorders, given improved knowledge and precision of their neural

correlates. Reaching this goal in the future will depend upon a phylogenetic, anatomical, and functional “vertical integration” of work on prefrontal cortex extending across species.

How Does Connectivity of Frontal Regions Enable Its Functions?

Throughout the Forum, we confronted this question in different ways and consider it to be a major gap that is ripe for future investigation. As alluded to above, connectivity can be examined at multiple levels: between cells, between subdivisions of frontal cortex, between frontal regions and other cortical areas, as well as between frontal regions and subcortical/noncortical areas. Some of these aspects of connectivity have been addressed more than others, but it remains obscure as to how they enable the function and computations performed by frontal cortex (for an overview, see Shenhav et al., Chapter 12).

Although there has been extensive research and knowledge gleaned regarding the organization and functioning of cells in sensory and motor regions, such heuristics may be unlikely to provide a suitable framework to understand the cellular organization and functioning of prefrontal cortex. For example, although the types of information being represented by cells in sensory cortex have been well delineated (e.g., contrast between light and dark in primary visual cortex, motion in area MT), prefrontal cortex is fundamentally different as prefrontal cells appear to have a multiplex coding of information, integrating multiple dimensions. Moreover, the same cell can flexibly change its coding scheme depending on task demands as well as code for abstract categories that are not constrained by physical characteristics (e.g., visual features). Hence, searching for “the element” of information that is encoded by prefrontal cells may be a futile or frustrating endeavor, yet this may be the exact reason why this region of cortex is so adept at modifying and modulating the functioning of other portions of the brain.

At the cellular level, at least part of (pre)frontal function, as is also true for other regions of the brain, is governed by its intrinsic connectivity patterns. It may be the connectivity pattern of the cells themselves that are important. For example, the ability to maintain representations over time, which are critical for working memory and maintaining goals, may depend on the ability of prefrontal cells to sustain activity via recurrent activity (see Koechlin and Wang, Chapter 10). On a different level, the ability of prefrontal cells to code abstract information may depend on the pattern of inputs from sensory, more posterior aspects of cortex, and subcortical regions. Once again, like other regions of cortex, the function of prefrontal cells is likely to vary by the input they each receive, the way that input is coded/weighted, and the context in which such information occurs (see Murray et al., Chapter 8). However, unlike portions of the visual system, we currently lack a detailed understanding of the “wiring

diagram” of frontal cortex. Hence a better understanding of the motifs of input to prefrontal cortex will be desirable, as well as a greater understanding of the degree to which such motifs may vary depending on the source of inputs—other prefrontal regions, more distant regions of cortex, or subcortical/noncortical regions—and in what contexts.

One notable aspect of (pre)frontal cells is that the same cells appear to be able to code different types of information flexibly at different times, often referred to as mixed selectivity. This characteristic may also distinguish them, in part, from other brain regions that act as convergence zones. For example, the hippocampus has some cells that code an animal’s location in space (place cells), and others that code the time at which events occurs (time cells). These populations of cells appear to be somewhat distinct as cells in the CA3 subfield of the hippocampus appear to encode information about space only, whereas those in CA2 encode information about time only (Eichenbaum 2017). (It should be noted that, nonetheless, at least some mixing of sensitivity to both of these dimensions does occur in the CA1 subfield).

The question that then arises is what schema or mechanism can allow the same cell to represent different information at different times. One topic highly relevant to this question is a consideration of how oscillatory firing may enable aspects of (pre)frontal function. At the cellular level, it is possible that firing of frontal neurons may be able to be disentangled into different frequency bands (e.g., via a Fourier transform), enabling separate simultaneous channels of communication to distinct target regions. For example, it has been suggested that high theta oscillations and gamma oscillations may play an important role in prefrontal functions including cognitive control (Cavanagh and Frank 2014), insight (Bieth et al. 2024), and problem solving (Bieth et al. 2024; Lin et al. 2023). This issue is a topic ripe for future investigation (for further information, see Weiner et al., Chapter 4).

With regards to connectivity patterns between prefrontal regions and the information flow between them, less is known. A number of models argue that connectivity is somewhat hierarchical in nature, with information computed at one level passed onto the next which overcomes the computational limitations at that prior level (see Koechlin and Wang, Chapter 10) or assumes that particular regions of prefrontal cortex (i.e., dlPFC) sit at the top of the hierarchy based on asymmetries of input and output connections (see Badre, Chapter 7). Aspects of such models are appealing as they mimic characteristics of the organization of posterior brain regions in which regions further along a hierarchy of a circuit aid in building up a representation (e.g., spots of light are detected in early visual areas, which are linked together to allow the detection of edges, which are linked together to detect lines of different orientations). In the case of prefrontal hierarchies, it is generally the case that the representations become progressively more abstract along a caudal-rostral lateral prefrontal axis; hence presumably elements of the earlier representations are lost as the abstract representation develops; in this sense, the hierarchies for visual and

prefrontal representation are parallel. However, exactly how the connectivity pattern of frontal cortex enables its specific functions is not well understood.

With regards to cortical-noncortical interactions, there were three main systems discussed at the Forum on which research has focused: the thalamus, the basal ganglia, and the amygdala. Understanding these interactions in more detail is likely to aid in understanding the influence of frontal regions on function. An interesting issue is whether the different regions of prefrontal cortex interact with these subcortical structures in an essentially similar manner. Although not much discussed at the Forum, another region of importance in this regard is the hippocampus. We consider each in turn.

While it has been known for quite some time that the role of the thalamus is to filter, sort and relay information reaching the cortex, more recent research has suggested that it may play a role in aspects of higher-level cognition through its connectivity with prefrontal cortex (Hwang et al. 2020). And through reciprocal loops of course, the cortex may in turn influence what information it “receives” from the thalamus.

Connectivity between the basal ganglia and the cortex has been argued to act a “gate” that is kept closed when the need to maintain information in working memory is high, but then opened when information in working memory needs to be updated (Hazy et al. 2007) with control over the gate learned through reinforcement learning driven by dopaminergic inputs. These basal ganglia inputs to cortex appear to have a specific topology, which could enable the gating of different types of information in parallel but interacting striatal-frontal loops (Rusu and Pennartz 2020) and has implications for diseases such as Parkinson disease (Wapstra et al. 2022) as well as for normal development (Parr et al. 2022).

Another aspect of prefrontal connectivity concerns connectivity with regions involved in emotional processing, most notably the amygdala. Prefrontal control over the amygdala can occur through several cortico-amygdala connections, which is of obvious importance for mental health and psychiatric disorders. For example, a dlPFC–vmPFC–amygdala circuit is thought to be involved in the ability to reframe information or experiences of emotional significance to enable more adaptive behavior (Denny et al. 2023). These connections are, of course, bidirectional; there are times when salience of information needed for survival detected by the amygdala necessitates more automatic action than a more thoughtful and planned response that would involve prefrontal control. Exactly which parameters govern such “interrupts” remain unknown.

An aspect of connectivity that was not discussed much at the Forum involves connectivity between frontal regions and the hippocampus. This is clearly of importance for the types of functions that frontal cortex supports, both in terms of higher-order learning and adaptive behavior. Once again, it is important to consider both directions of connectivity; that is, from the prefrontal cortex to the hippocampus as well as from the hippocampus to prefrontal cortex. With regards to the former, there is evidence that connectivity from frontal regions

to the hippocampus are involved in aiding to highlight specific features of an event at encoding (Kim 2011), in strategic retrieval of memories (Blumenfeld and Ranganath 2019), and even control including the inhibition of memory retrieval itself (dlPFC) (Anderson et al. 2016). Moreover, connectivity between inferior medial prefrontal regions and the hippocampus may allow for a contextual rubric that integrates information across diverse episodes that are then used for inference and higher-order thinking (Morton and Preston 2021). In the opposite direction, connectivity from the hippocampus provides a rich source of information that can be used for prefrontal function. For example, work suggests that connectivity from the hippocampus to prefrontal cortex aids in mental simulation of future events (Campbell et al. 2018), and with integration with information from vmPFC allows for the affective valuation of future possibilities (Benoit et al. 2019).

Finally, different portions of frontal cortex are linked to different sets of cortical areas. In the human brain, this is most notable with regards to different (pre)frontal regions belonging to distinct intrinsic connectivity networks (see Gratton et al., Chapter 11, and Duncan and Friedman, Chapter 9). These different networks have been linked to somewhat different functions (e.g., cognitive control by the frontoparietal network; salience detection and evaluation by the ventral attention/cingulo-opercular networks; and “internal cognition” by the default mode network). However, these broad functions appear rather vague and in need of further definition. It is also unclear how precisely they interact to mediate executive function. Nonetheless, multiple networks of this sort allows different regions of frontal cortex to participate in distinct networks that code different aspects of information, thus providing parallel representations of information used in service of higher-order thought (DeRosa et al. 2024). One issue for further investigation is how “hubs” (see Rowe, Chapter 16) within the prefrontal cortex may participate in several independent neural networks.

Consistent with the patterns observed for prefrontal connectivity, it has been suggested that such connectivity is important for top-down inhibitory control, which is often considered a core aspect of prefrontal function (Friedman and Miyake 2017). It has been suggested that prefrontal regions maintain a task set and use that information to modulate processing in distant brain regions that are relevant for the current goal. Maintenance of such task sets is of particular importance in “inhibitory” tasks in which the maintained task set must override more prepotent or automatic responses (Munakata et al. 2011). Empirical evidence suggests a specific role of different portions of the lateral prefrontal cortex of the right hemisphere in exerting such “inhibitory” control. For example, it was observed within the same individuals that functional connectivity between the dlPFC and the hippocampus is associated with an individual’s ability to inhibit memory retrieval, functional connectivity between this region and the amygdala is associated with the ability to suppress emotional responsivity and functional connectivity between this region and the inferior frontal

gyrus and subthalamic nucleus predicts the ability to inhibit motor responses (Depue et al. 2016).

Social Processing, Higher-Order Cognitive Skills, and Development

On display at the Forum were diverse and complicated aspects of cognitive and emotional processing supported by (pre)frontal mechanisms. Many other topics, however, did not receive much attention but make the frontal lobes an intriguing and important brain region to study and understand.

One such issue is how frontal neural mechanisms allow for a more developed understanding of the self, one's relationship to others (including Theory of Mind), and the ability to use emotions for a higher purpose than self-survival. These abilities include self-evaluation and extend as empathy, moral reasoning, and judgment. Evidence exists to suggest that these abilities rely at least in part on medial prefrontal regions that form part of the default mode network (Andrews-Hanna et al. 2010), but also most likely require an exquisite coordination of information from across the brain to be integrated in frontal regions. As such, a better understanding of the underlying motifs of frontal organization and function, as well as the essential computations performed by frontal regions, are likely to be necessary to understand how the frontal lobe contribute to those skills and abilities that make us uniquely human.

While some of these higher-order abilities can be examined in animal models (e.g., associative inference), many aspects of human cognition will not benefit from cross-species comparisons (Levy 2023). In general, these abilities derive from the role that the prefrontal cortex plays in temporal processing and integration (Fuster 2001). In humans, the more extended capacity to consider and integrate information over longer time spans allows for abilities such as higher-order planning, analogical reasoning, and likely even historical perspective.

While certain aspects of prefrontal function can be explained by reinforcement learning algorithms (see Shenhav et al., Chapter 12), in which learning is based on whether an action leads to a reward or not, some do not rely on having a prior experience. These aspects of human frontal function allow not only for simulations based on past experience but for the novel, innovative, and creative aspects of thought. They cannot be guided by past experience of reward (or lack thereof), but only by one's imagination, insight, or conceptual vision.

Finally, another issue of great importance is the development of prefrontal function and the factors that influence its development (Rowe et al., Chapter 16). Prefrontal cortex undergoes protracted development through the late teens into the early 20s and supports many of the abilities generally considered to make one an "adult." At a cellular level, these processes includes cell division,

migration axonal connections, synaptogenesis, pruning, and myelination. At a larger scale, these processes include changes in network coherence and connectivity. Understanding how various insults, such as stress, impacts these processes is critical for evaluating functional deficits and developing strategies for early intervention. For example, stress during early development can cause structural and connectivity changes. While touched on briefly by Rowe et al. (Chapter 16), this topic needs further discussion. Developmental issues were not emphasized, in part, because they are currently the focus of much work internationally, including the Adolescent Brain Cognitive Development (ABCD) Study in the United States (<https://abcdstudy.org>) and the IMAGEN project in Europe (Schumann et al. 2010). Nonetheless, such work is likely to be a rich source of information to address many of the issues raised above, from how cortical and subcortical interaction lead to increase cognitive control, to the understanding of how we as individuals can infer other's likely internal thoughts.

Summarizing Statement

The Forum provided excellent and much needed discussion and analysis of where we stand today in our scientific understanding of this critically important brain region for humans as well as other animals. Moreover, the concluding chapters of each section describing the group discussions provide a roadmap for the conceptual and practical issues that will need to be addressed to further enhance our understanding of the frontal lobes. A general conclusion across many of the group discussions at the Forum was the realization that, while the frontal lobes are a very active area of research, much of this work occurs within “silos,” and there is not as much communication amongst researchers as might be desirable. We may be divided by species, prefrontal region, specific function or technique. Animal researchers may focus on a particular species and not be integrating knowledge with findings found in other species; individuals may focus on their particular frontal region of interest and not consider how that region interfaces with other regions of the frontal cortex or indeed the rest of the brain; researchers may focus on a particular function (e.g., language) of a region of prefrontal cortex (e.g., left ventrolateral prefrontal cortex), and neglect whether that region might be involved in other aspects of cognitive or emotional processing. Finally, it is evident that the methods we employ for analyzing the frontal cortex vary enormously in terms of their spatial and temporal resolution, but that more effective cross-disciplinary integration of the data they acquire is required, for example to understand how neural networks measured using fMRI in humans relate to anatomical and electrophysiological findings at the cellular level. The need to engage with, and resolve, the complexities of the techniques themselves also lead to methodological “silos,” that are obstacles to broader discussion.

A take home message for us as convenors of the Forum, therefore, was that more cross-talk would go a long way to speeding up our understanding of the frontal lobes, and we were so very grateful that the Forum provided an opportunity to take a meaningful step in this direction.

