

Opinion

Integrated Intelligence from Distributed Brain Activity

John Duncan,^{1,2,*} Moataz Assem,¹ and Sneha Shashidhara¹

How does organized cognition arise from distributed brain activity? Recent analyses of fluid intelligence suggest a core process of cognitive focus and integration, organizing the components of a cognitive operation into the required computational structure. A cortical ‘multiple-demand’ (MD) system is closely linked to fluid intelligence, and recent imaging data define nine specific MD patches distributed across frontal, parietal, and occipitotemporal cortex. Wide cortical distribution, relative functional specialization, and strong connectivity suggest a basis for cognitive integration, matching electrophysiological evidence for binding of cognitive operations to their contents. Though still only in broad outline, these data suggest how distributed brain activity can build complex, organized cognition.

Organizing Distributed Brain Activity

Organized cognition of any kind arises from widely distributed brain activity. An immediate question is how such activity is integrated, allowing just the right cognitive contents to be combined in just the right way for current purposes. Though much is certainly unknown, an outline view of the relevant psychological and physiological mechanisms is beginning to appear. In this opinion article, we describe recent progress towards a whole-brain understanding of cognitive integration.

We begin with recent work on the cognitive mechanisms of **fluid intelligence** (see [Glossary](#)). Theoretical accounts of fluid intelligence focus on processes of cognitive control [1–3] and cognitive integration [4] and, based on recent findings, we suggest a synthesis of these two approaches. Results from brain imaging [5–7] and lesion [8] studies relate fluid intelligence to a well-known control network in the brain, which previously we have called the **multiple-demand (MD) system** [9,10]. We describe recent studies on the detailed anatomy and physiology of MD activity and how they begin to illuminate the physiological underpinning of cognitive control and integration. In broad outline, these findings suggest how distributed brain activity builds organized cognition. We conclude with some of the many questions that this scheme raises for future work.

Fluid Intelligence and Attentional Integration

A fundamental psychometric discovery is positive manifold: to some extent, all tests of different cognitive abilities tend to have positive correlations [11,12], even those that on the surface are dissimilar. In his foundational work, Spearman [11] proposed that some general or *g* factor contributes to success in any task. If this model is fit to correlational data, novel problem-solving tests turn out to be excellent measures of *g*, reflecting the fact that, in a diverse task battery, it will be these tests that have the largest average correlations with a wide range of others. Well known examples are matrix problems ([Figure 1A](#)) [13,14], series completions [15], etc. The ability measured in such tests has been called ‘fluid intelligence’. Later, we consider several possible contributors to positive manifold, but meanwhile, the broad ability of fluid intelligence tests to predict success in many kinds of activity, from laboratory tests to life achievements, suggests cognitive mechanisms of widespread importance in mental life.

Highlights

Fluid intelligence tests predict success in many activities, suggesting cognitive mechanisms of broad importance.

We propose a core process of attentional integration. Complex problems must be segmented into simpler parts. Attention to each part integrates cognitive fragments into a computational structure.

Fluid intelligence is linked to the brain’s multiple-demand (MD) system, defined by common activity across different cognitive demands. Across the brain, MD patches shows anatomical and physiological properties adapted to attentional integration.

Neurophysiology of putative MD regions shows adaptive coding of task-relevant information. Suiting attentional integration, many neurons show conjunctive coding (e.g., binding cognitive operations to their target objects).

In broad outline, these results suggest how distributed brain activity builds organized cognition.

¹MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge CB2 7EF, UK

²Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, UK

*Correspondence:
john.duncan@mrc-cbu.cam.ac.uk
(J. Duncan).

We begin with models of fluid intelligence based on cognitive or executive control. Following early work linking fluid intelligence to frontal lobe functions [1,16,17], multiple aspects of control have been implicated in fluid intelligence models: maintenance in working memory in the face of distraction [2], avoiding lapses and mind-wandering [3,18], inhibiting unwanted mental content [2], and so on. Though this broad focus is appealing, and the link to frontal lobe functions is clear, conceptions of control often seem underdeveloped. Undoubtedly, simple concepts such as maintenance in working memory [19], attentional biasing [20–22], inhibition [23,24], and the like bear on important aspects of cognition. On their own, however, they fall far short of addressing what would be needed to generate even a simple everyday activity, such as planning and carrying out a trip to the grocery store.

A different perspective on cognitive control comes from classical work in artificial intelligence, going back to the early problem-solving programs of Newell *et al.* [25]. As this work made clear, most complex problems cannot be solved in a single processing step. Instead, they must be divided into simpler parts, with these simpler parts solved in turn to produce final success on the whole problem. A familiar example organized as a goal–subgoal hierarchy is shown in Figure 1B. It is not possible simply to set the goal of traveling to Japan and then move immediately to the question of how to move one’s hand; the statement of the goal brings too few constraints to lead immediately to motor commands. But beginning with the goal of travel to Japan, the problem solver can set progressive subgoals of flying, buying a ticket, and logging onto the internet, and now, sitting at a laptop, it becomes possible to plan a movement [26]. This reasoning suggests that a core aspect of ‘control’ must be cognitive segmentation: based on knowledge of the problem domain and moves available to the problem solver, a complex whole must be organized into separate, more easily soluble parts.

Early theoretical work [27] emphasized the importance of cognitive segmentation in fluid intelligence problems. In a recent study [28], we modified traditional matrix problems so that little was left beyond this segmentation requirement. As shown in Figure 1C, entries in each cell of the matrix had three parts (in this case, short line straight or curved, longer line center or right, arrow pointing left or right). To minimize the working memory demands of traditional problems (Figure 1A), the participant simply drew their answer into a response box, allowing each part of the answer to be drawn as soon as it was determined. The problem appears trivial once attention is focused on a single part of the figures (e.g., focusing just on the arrow) but, nevertheless, participants with low scores on a standard fluid intelligence test struggled to solve these simplified problems (Figure 1E, blue). To confirm that segmentation was the only significant requirement, we used problems that were presegmented, with each part of the figure presented in its own matrix (Figure 1D). Though component problems were exactly the same as those that would be produced by attention to a single part in the original figures, now all participants performed well (Figure 1E, red).

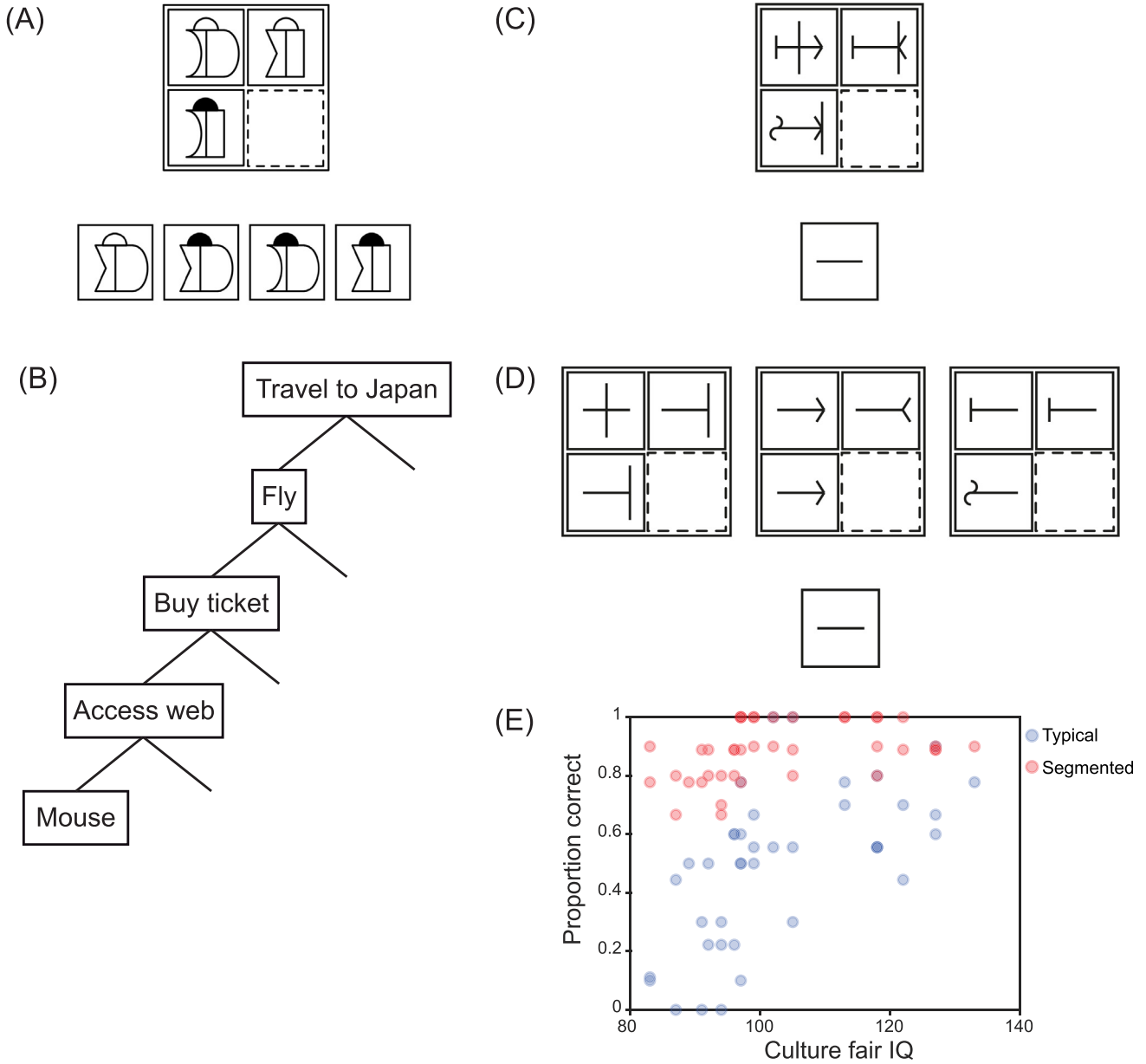
Sometimes, accounts of fluid intelligence based on cognitive control have been contrasted with an account based on cognitive integration [4,29]. According to integration accounts, the key process in fluid intelligence is binding together the different components of a cognitive process or representation. In this light it is instructive to consider what is needed to ‘attend’ to one part of the problem in Figure 1C (e.g., to determine that the arrow should point left). Evidently, information is needed on arrow direction in each cell of the matrix. This direction information must be correctly bound to positions within each figure and in the matrix as a whole. The layout of left and right arrows must be related to an overall conception of the problem to be solved and what it means for the solution to ‘look right’. Doubtless, problem solving is guided by internal reward signals bound to each successful step. To create attention, multiple cognitive fragments

Glossary

Fluid intelligence: the ability measured in psychometric tests of novel problem-solving, including matrices, series completions, etc. Though many tests use simple shapes and figures, fluid intelligence tests can also involve verbal or numerical materials.

Multiple-demand (MD) system: a distributed set of cortical regions showing widespread increase of activation associated with many different cognitive demands.

Nonlinear mixed selectivity: a neural response pattern in which firing rate is driven by a conjunction of task variables (e.g., a particular object presented only at a particular point in a memory list).



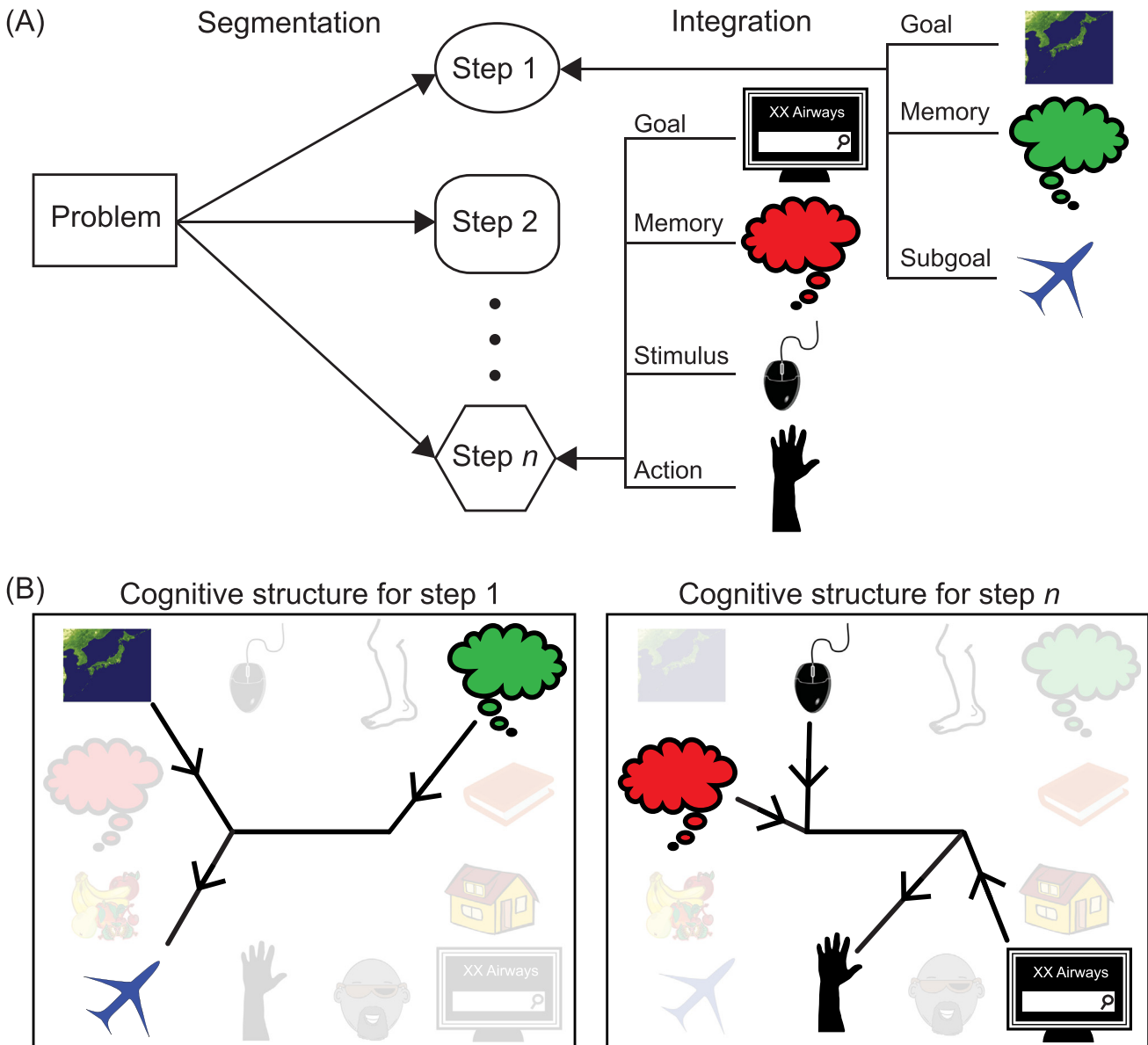
Trends In Cognitive Sciences

Figure 1. Fluid Intelligence and Cognitive Segmentation. (A) Matrix problem. The task is to choose which of the response alternatives (bottom) would correctly complete the matrix (top). (B) Goal-subgoal hierarchy. (C) Modified matrix problem in typical format. The task is to decide what figure would fill the empty cell of the matrix (broken outline), with the answer to be drawn in the response box below. (D) Same problem segmented into separate parts. (E) Proportion of problems correctly solved as a function of fluid intelligence. Blue, typical format; red, segmented format. Adapted, with permission, from [28].

must be integrated into precisely the correct combinations and relationships. As these computational structures are built, the task as a whole is segmented into useful parts, each representing a step closer to the overall goal.

The need for attentional integration – creating the steps of complex behavior, each consisting of components assembled into just the right computational structure – is obvious in fluid intelligence

problems. We would argue, however, that attentional integration lies at the heart of all organized cognition (Figure 2), helping to explain why fluid intelligence tests predict success in such a wide variety of behavior. Segmentation into parts is minimized in very simple tasks, but even here, just the right cognitive elements, organized in just the right way, must be assembled for behavioral control. Rather than conflicting, we suggest that accounts of fluid intelligence based on control and integration reflect two views of the same process. Focused attention, resistance to distraction, and integration are all important aspects of brain activity that defines and assembles the contents of



Trends In Cognitive Sciences

Figure 2. Integration and Segmentation in Complex Cognition. (A) Attention to one part of a complex task integrates multiple component fragments (right). As a series of steps is created, the problem is progressively segmented into simpler subproblems (left). This example is the 'travel to Japan' problem from Figure 1B. (B) To create each step, fragments must be selected from many potential candidates and assembled into precisely the correct computational structure.

an elementary cognitive operation. Our approach thus places popular accounts of fluid intelligence within a broader view of how ‘cognitive control’ in general should be conceived.

Integrating Distributed Brain Activity

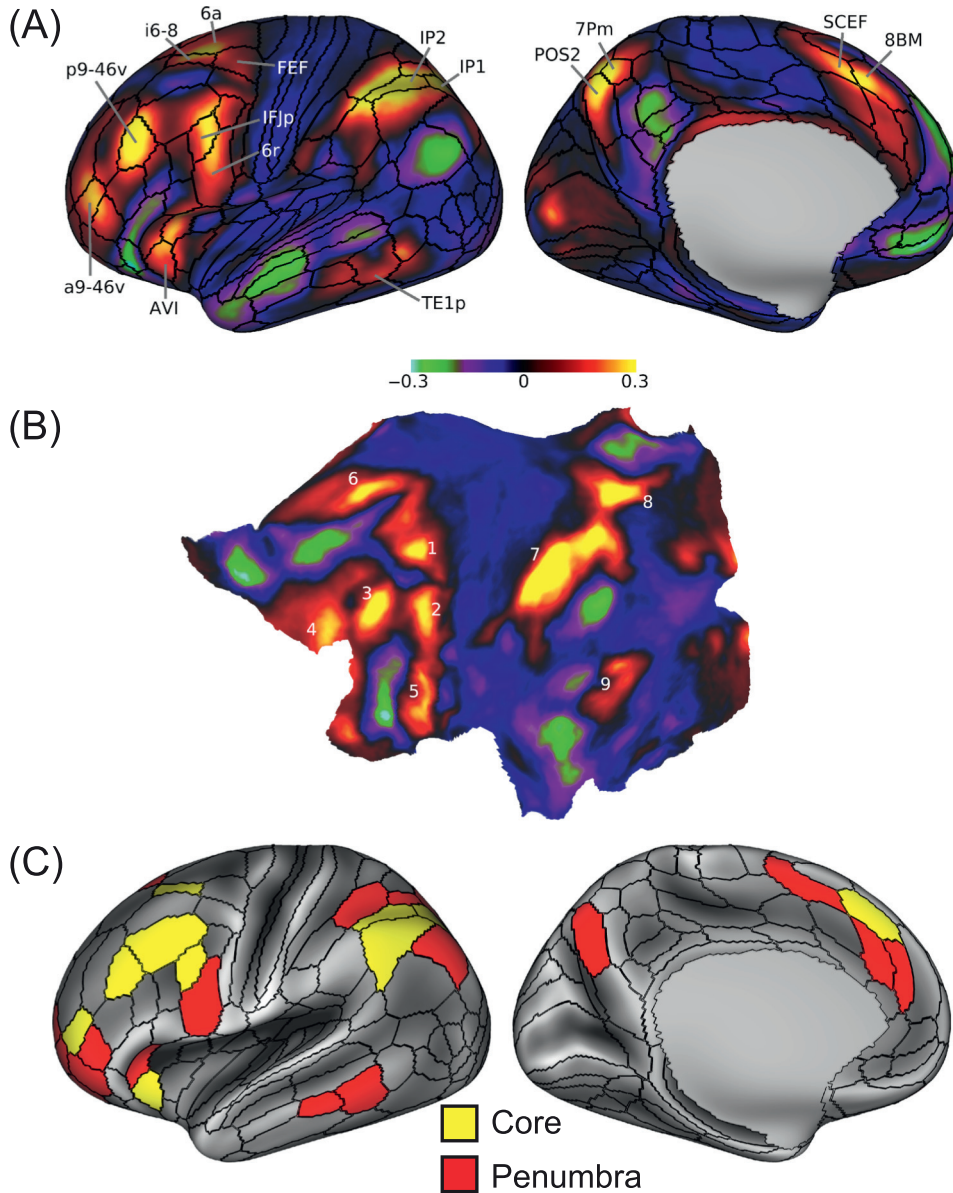
Of course, the material to be integrated in creating a cognitive structure will be represented in widely distributed brain activity. Even attention to a visual object requires integration of activity across multiple cortical and subcortical brain areas, representing that object’s different properties and action affordances [30,31]. For a complex cognitive structure such as a single problem-solving step, relevant activity may be distributed across much or perhaps most of the brain. This calls for an integration mechanism with widespread, flexible access to whatever neural activity is needed in current cognition.

Along with regions of brain activity linked to specific cognitive domains, brain imaging studies often show activity in a widely distributed MD system: a set of brain regions in which activity increases with almost any kind of cognitive load ([9,10,32,33]; for an exception see [34,35]). MD activity is generally seen in several regions on the lateral frontal surface, in and around the anterior insula, in dorsomedial frontal cortex, including presupplementary motor area and dorsal anterior cingulate, within and to either side of the intraparietal sulcus, and often also in a region at the occipitotemporal border. From the early days of brain imaging, MD activity has been linked to cognitive control [36,37] and integration [37,38]. A core role in fluid intelligence [9,39,40] is indicated by strong MD activity during work on fluid intelligence problems [5–7] and losses of fluid intelligence associated with MD damage [8,41–43]. In this section we link new findings on MD anatomy and physiology to the computational requirements of attentional integration.

Though MD-like activity has been known for many years, its precise anatomy has remained uncertain. To address this limitation, we turned recently to the methods and data of the Human Connectome Project (HCP) [44]. In the HCP approach, surface-based processing using multimodal MRI features improves brain coregistration, much sharpening delineation of cortical areas. Multimodal data are further used to parcellate the cerebral cortex of each participant into 180 distinct regions per hemisphere [45]. In a sample of 449 HCP participants, we used the conjunction of three cognitive contrasts – high versus low working memory load, relational reasoning versus perceptual matching, and arithmetic versus story comprehension – to examine MD regions and their properties.

The results clarify the anatomy of known MD regions, as well as suggesting new regions. Across the cortex we see nine tightly defined patches of MD activity, distributed across frontal, parietal, and temporo-occipital cortex (Figure 3A,B). The HCP parcellation further divides MD patches into 27 individual regions, which we separate into a ten-region core with strongest activity and a surrounding penumbra (Figure 3C). Five MD patches lie on the lateral frontal surface, the most dorsal lying just anterior to the frontal eye field (Figure 3A, i6-8/6a), the most ventral incorporating anterior insula and adjacent regions of the frontal operculum (Figure 3A, AVI), and between these, a chain of three extending from the inferior frontal junction posteriorly into the rostralateral frontal cortex (Figure 3A, IFJp/6r; p9-46v; a9-46v). An additional frontal patch lies on the dorsomedial frontal surface (Figure 3A, SCEF/8BM). In the lateral parietal lobe, the primary MD patch is centered in the depths of the intraparietal sulcus (Figure 3A, IP1/IP2), with an additional patch in dorsomedial parietal cortex (Figure 3A, POS2/7Pm). Confirming previous indications, and despite the fact that one of our three contrasts used auditory rather than visual materials, we find additional MD activity in a region at the junction of anterior occipital and posterior temporal cortex (Figure 3A, TE1p).

With this fine-scale definition of MD regions, we can ask also about functional differentiation. Despite very frequent coactivation of MD regions, there have also been many suggestions of



Trends in Cognitive Sciences

Figure 3. Anatomy and Physiology of the Multiple-Demand (MD) System. (A) Patches of cortical MD activity defined in data of 449 participants from the Human Connectome Project (HCP), using a conjunction of fMRI contrasts for working memory, reasoning, and arithmetic. Left hemisphere data are shown; largely similar patches are also seen on the right. Regional parcellation (black outlines) and selected anatomical labels are taken from [45]. (B) The same data shown on a flat map of the left hemisphere. Numbering shows nine MD patches distributed across lateral frontal (1–5), dorsomedial frontal (6), lateral (7) and medial (8) parietal, and temporo-occipital (9) cortex. (C) Individual MD regions using the HCP parcellation. Data are averaged from left and right hemispheres and, for illustration, projected onto the left. The extended MD system (27 regions) is divided into core (ten regions, yellow), with activity above the mean of all 27 regions in at least 2/3 contrasts, and penumbra (remaining 17 regions, red). Adapted, with permission, from [44].

functional differentiations (e.g., [46–48]), though with no clear consensus emerging. Across the full set of 27 MD regions, Figure 4A shows profiles of activity for our three task contrasts. The results illustrate both sides of coactivation–differentiation picture. On the one hand is a

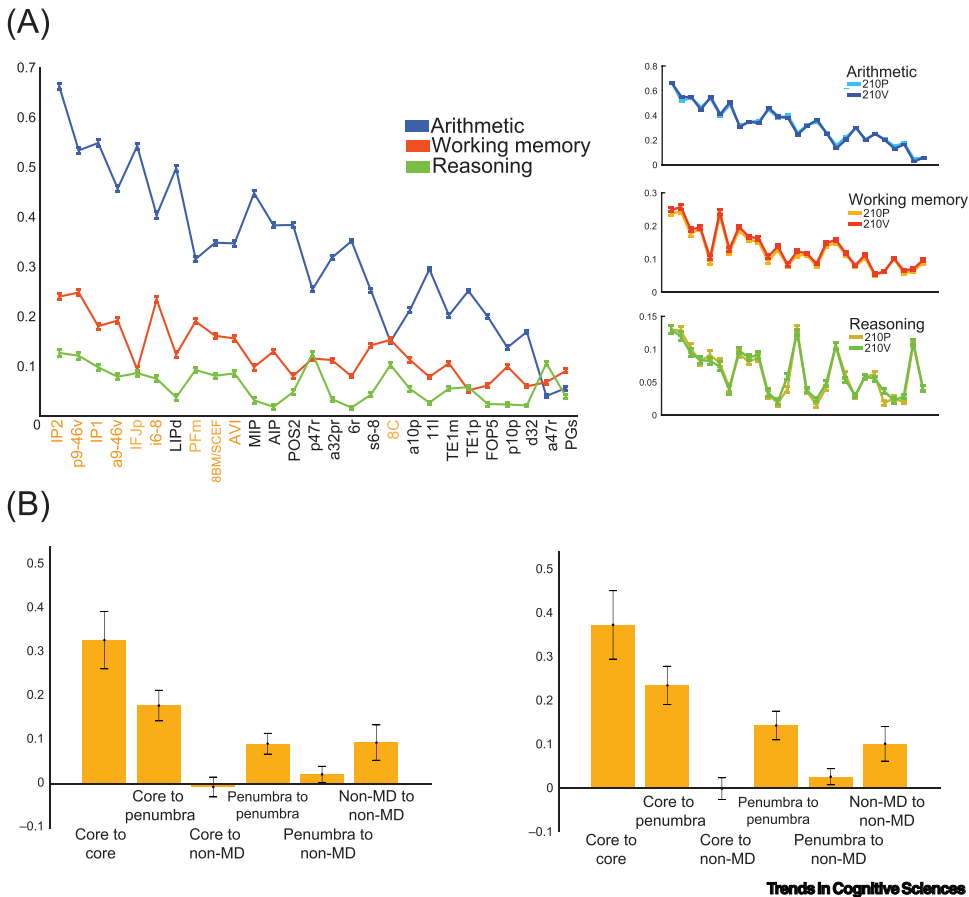


Figure 4. Functional Profiles and Connectivity of Multiple-Demand (MD) Regions. (A) Profiles of activation across the extended MD system for each task contrast. To show reliability, right panels show overlaid plots for two independent groups of 210 participants each. (B) Resting state connectivity (correlation of time series), calculated for every pair of cortical regions and then averaged for connections of each type. Left, left hemisphere; right, right hemisphere. Adapted, with permission, from [44].

strong pattern of coactivation, with strongest activity for arithmetic and weakest for relational reasoning, across almost all MD regions (Figure 4A, left). On the other is differentiation, with exact activity profiles differing between contrasts, and with this large group of participants, even small differences are highly reliable (Figure 4A, right). Though the MD system is commonly recruited as a whole, the exact pattern of this recruitment differs from task to task (see also [49]).

HCP resting state data also allowed us to examine patterns of functional connectivity. Commonly, resting state studies define a ‘frontoparietal control network’ (FPN), substantially overlapping with the MD system that is defined by activation data [50]. Based on HCP resting state data, Figure 4B shows average connectivity for all possible types of connection between core, penumbra, and non-MD regions. Strikingly, core–core connections were strongest, followed by connections of core to penumbra. Comparison with a canonical FPN, previously extracted by Ji *et al.* [51] from the same HCP data, showed that all regions in our MD core lie within the FPN and, within that network, core–core connections are especially strong. Penumbra regions, by contrast, are spread across several networks.

Recent analyses of brain connectivity have emphasized its small world structure [52], with strong connectivity within local modules and long-range connector hubs linking between modules. Graph theory and other approaches suggest that MD regions are among the brain's widely connected hubs [53,54], with dynamic, task-dependent connections to other brain networks [55,56]. Several previous accounts link fluid intelligence to this widespread MD connectivity [57], including a recent 'network neuroscience' proposal [58]. In line with such ideas, our results suggest that the MD system is well placed for attentional integration (Figure 5A). Because local connections dominate in the brain, a system needing access to many modules needs widely distributed parts. Locally, we suggest, different kinds of information are preferentially fed into different MD regions, accounting for partial functional differentiations. Similarly, local connections allow MD outputs to influence processing in many modules. Creating integrated cognitive structures, however, requires that all types of information can be creatively selected and combined, in precisely the roles and relationships required by a particular cognitive demand. Strong connectivity between MD regions, in particular between regions of the core, suggests a medium for information exchange and integration and an explanation for the strong element of coactivation seen across many different task demands.

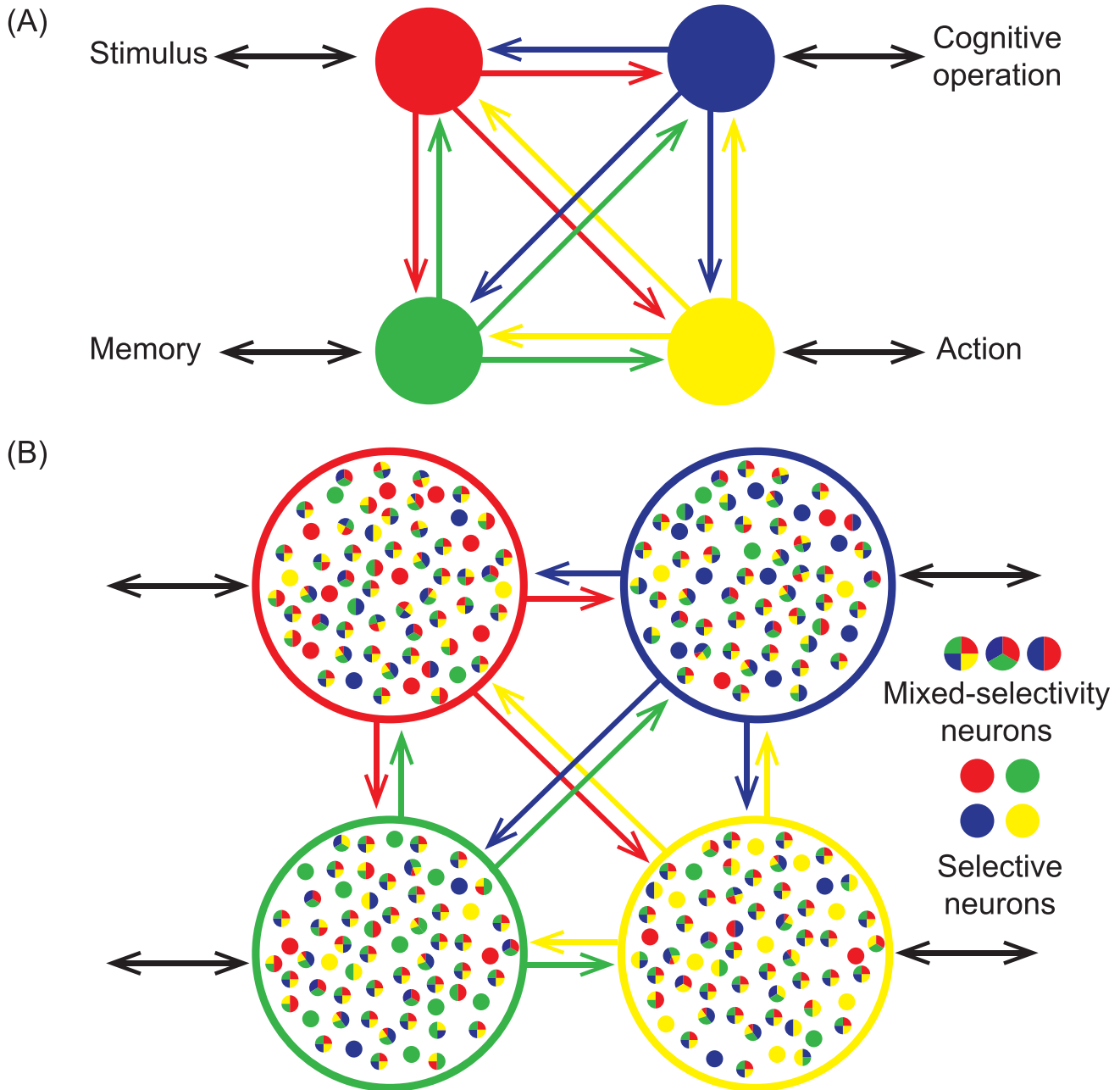
A final aspect of the HCP findings is worth emphasis. Though our primary analyses focused on cerebral cortex, MD activity was also seen in associated subcortical structures, including parts of the caudate and thalamus, and in specific regions of the cerebellum. Though details are unknown, it is likely that attentional integration is achieved through interacting cortical and subcortical activity.

Task Representation: Contents, Roles, and Relations

If MD regions are to create cognitive episodes, the contents of these episodes must be represented in the firing of MD neurons. In human brain imaging, multivoxel pattern analysis confirms widespread MD encoding of different aspects of a current task, including discrimination of relevant stimuli, rules, responses, and more [59–61]. In contrast to more dedicated regions, such as visual or auditory cortex, MD representation of relevant task contents is often weaker [62] but also broader [63]. Matching a role in selective attention, MD representations favor what is relevant to a current task [64–66]. Matching univariate findings, MD representations may become stronger as task difficulty is increased [61], likely related to the experience of more careful attention.

Corresponding findings come from electrophysiology in the behaving monkey. While exact homologies between human and monkey are uncertain, imaging data suggest a monkey MD network somewhat resembling that of the human brain [67,68], including regions of lateral frontal, dorsomedial frontal, and inferior parietal cortex. In behaving animals, neurons in these regions show properties strongly reminiscent of human findings, with firing rates of many neurons discriminating the important events of a current task and adjusting their properties to favor currently relevant over irrelevant information [69–74]. Again, these data suggest a distributed network with widespread access to task-relevant information and bidirectional communication between network nodes [75–78].

Most important for current purposes, the properties of MD neurons suggest a central role in integration. As we have discussed, formation of a cognitive episode requires combining relevant contents into exactly the required roles and relations. The rules of a task may determine how stimuli and responses should be paired, the required order of several responses, how alternatives are linked to available rewards, etc. Similarly, in a goal–subgoal tree like that of Figure 1B, subgoals must be bound to goals such that, if the subgoal fails, an alternative route to the higher



Trends In Cognitive Sciences

Figure 5. The Multiple-Demand (MD) System and Cognitive Integration. (A) Depending on its pattern of connectivity, each MD region (colored circles) has direct access to different information and brain operations, here illustrated with just a single link (black bidirectional arrows) for each region. Strong connectivity between MD regions allows assembly of these fragments into the required computational structure. The scheme suggests a basis for partial functional differentiations within a broad context of coactivation. (B) With their varying external connectivity, MD regions may show quantitative differences in neural coding for different task features. Against this background, however, communication between MD regions provides a strong basis for mixed selectivity.

goal can be sought. Linking contents to roles is the classic computational problem of variable binding [79] and, for neural representation, suggests a requirement for conjunctive coding of contents and roles. In recent years, conjunctive coding in frontal and other regions has come

to prominence under the heading of '**nonlinear mixed selectivity**': neural activity driven by nonlinear combinations of multiple task features [80,81].

The literature contains many remarkable examples of mixed selectivity. If an animal must remember a sequence of objects, for example, many frontal neurons may encode object identity; a neuron's object preference at one serial position, however, can be independent of its preference at another [82,83]. If a cue indicates that the animal should search for a particular target stimulus, activity in the delay between cue and choice display may encode the upcoming target, but this encoding can be independent of activity when the target finally appears [84]. In a study of frontal activity during problem-solving [85], animals solved an on-screen maze, planning movements of a cursor along open maze paths from a central start position to a peripheral goal. Once the path had been planned but before movement began, individual frontal cells could be selective for the direction of just the first, just the second, or just the third movement in the plan.

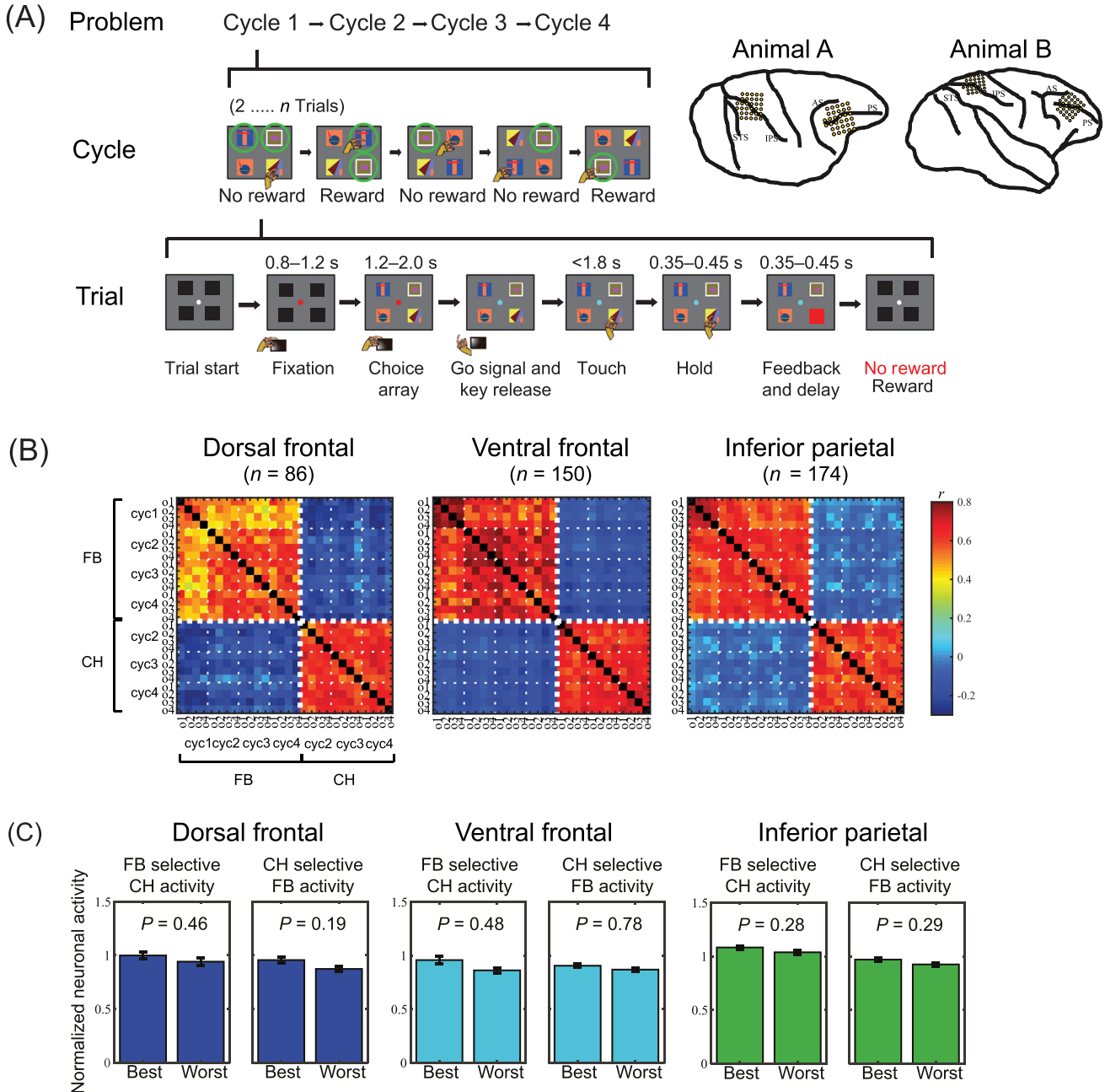
In line with the requirements of cognitive integration, information exchange between MD regions is well placed to create mixed selectivity (Figure 5B). Exchange between MD regions allows conjunctive coding, binding the components of a cognitive operation into the correct computational structure.

An example from our own recent work [86] illustrates binding of cognitive operations to their target objects. In this task (Figure 6A), displays of four objects were shown on a touchscreen and, on each trial, the monkey selected a single object by touching it. For each new problem, one or two objects were rewarded targets; the monkey had to find these targets by trial and error and then, once targets were found, could reselect them on subsequent trials. We analyzed data from two trial periods: the choice phase (CH), during which monkeys were shown the object display and awaited a go signal indicating that they could make their choice, and the feedback phase (FB), when, following the touch, a cue indicated whether the selected object was target or nontarget.

Recordings were made in two MD-like regions, lateral frontal and inferior parietal cortex (Figure 6A, upper right; recordings in superior parietal cortex not considered here). For each region, we used two kinds of analysis to compare activity at FB and CH. First, we considered patterns of activity across the whole recorded cell population (Figure 6B). Within one phase, FB or CH, population activity patterns were strongly correlated for different objects and cycles. Between phases, however, correlations were close to zero. For each neuron, in other words, firing rate in one task phase was approximately independent of firing rate in the other. Our second analysis showed that this same independence extended to object selectivity. At both CH and FB phases, there were frontal and parietal neurons that encoded the identity of the selected object. However, a neuron's object preference during one task phase, CH or FB, was unpredictable of its preference at the other phase (Figure 6C). As MD regions direct the cognitive operations of successive task stages (Figure 5B), orthogonal representations may minimize confusion, allowing the correct cognitive operations to be executed at each stage [87]. Meanwhile, mixed selectivity for conjunctions of object and phase binds object information to these different operations.

Positive Manifold

Earlier we noted positive manifold, the finding of ubiquitous positive correlations between different cognitive tests, and Spearman's original proposal that some general or *g* factor contributes to success in any cognitive activity [11,12]. One simple interpretation is that *g* reflects the attentional integration functions of the MD system and, in agreement with Spearman's hypothesis, we suggest that MD functions contribute very broadly to effective cognition. At the same time, it



Trends In Cognitive Sciences

Figure 6. Neurophysiology of Binding Object to Role. (A) Object selection task, two-target version. On each trial, the monkey touches a single object in a visual display. For each new problem, in the first set of trials (cycle 1), the monkey selects one object after another, learning which two objects (targets) are associated with reward. Targets are indicated here by green circles (not present on actual display). In subsequent trials (cycles 2–4), the animal can reselect the same targets for further rewards. After four cycles, targets are redefined for the next problem. In another task version, problems have only a single target. Recording areas in each animal are shown at upper right. (B) Correlation of population firing patterns for feedback (FB) and choice (CH) periods, separated by target object and cycle. Data from correct trials only. For CH, data are shown only for cycles 2–4, as correct choices were not known in cycle 1. Frontal lobe data are separated into dorsal and ventral regions (separated at fundus of principal sulcus). Small negative correlations between FB and CH patterns are an artefact of data normalization [74]. (C) Independence of object preference at CH and FB. For each recording area, left panel shows data for neurons identified as object-selective during FB. For each neuron, ‘best’ (highest firing rate) and ‘worst’ (lowest) were identified based on FB data. Plots show activity for these ‘best’ and ‘worst’ objects while selecting them at CH (normalized for each neuron, averaged across neurons). Across the population, ‘best’ and ‘worst’ objects defined at FB did not give significantly different responses at CH. Right panels show reverse analysis, defining ‘best’ and ‘worst’ at CH and plotting data from FB. Adapted, with permission, from [86].

seems likely that the full explanation for positive manifold is more nuanced, with multiple contributory factors [39,58].

For any task, creating the appropriate control structure is not simply a matter of MD function. Undoubtedly, MD activity combines with activity in multiple, more specialized systems involved in operations of this particular task, and it seems likely that the quality of the resulting computational structure will depend on all collaborators and their interaction. In many studies, fluid intelligence has been found to correlate with different aspects of brain structure, broadly distributed across many brain regions [88,89]. In a recent study, for example, intelligence correlated with the size and complexity of dendritic trees in tissue taken from the temporal lobe of patients undergoing epilepsy surgery [90]. Plausibly, many neural properties will be correlated across regions of the cortex, and large dendritic trees in the temporal lobe may be predictive of large dendritic trees in MD regions and elsewhere. In this case, a broadly distributed property of neural function may facilitate the general process of creating cognitive operations, whatever their particular content. Intelligence may also correlate with whole-brain functional properties, such as stability of functional networks [91,92]. Very likely, there is much overlap between the core cognitive mechanisms required in fluid intelligence tests, the functions of the MD system, and sources of individual differences leading to positive manifold, but these three are not likely to be identical.

The role of long-term knowledge also bears on positive manifold, and a distinction that is often drawn between fluid and crystallized intelligence. While fluid intelligence concerns current problem-solving, crystallized intelligence reflects the accumulated body of a lifetime's learning [93,94]. The two are generally correlated, as expected if more useful knowledge is acquired during better constructed, more focused learning episodes, and if a lifetime of learning leads to a large body of knowledge, with parts of this knowledge applicable to many new problems, then 'crystallized intelligence' will also contribute broadly to success in new activities and thus to positive manifold. A related perspective is provided by the idea of mutualism in development: that growth in one ability or domain may have positive influences on growth in others [95,96]. Often, it seems, this must be true; learning to represent relationships as graphs, for example, must surely encourage effective mental representation and hence learning in many future contexts. If positive manifold reflects the broad ability to construct good cognitive structures, it is likely influenced by many aspects of lifetime experience.

Concluding Remarks and Future Directions

Many issues are raised by the integration account. Here we discuss two: the interface of short-term cognitive activity and long-term knowledge, and the nature of attentional capacity limitations.

As implied by our discussion of positive manifold, a core question is interface between on-line cognition and long-term knowledge. As in classical symbolic artificial intelligence (e.g., [97]), a complex problem is divided into simple parts on the basis of long-term knowledge of the structure of the world and relations within it. It is knowledge that tells us how travel to Japan can be divided into component steps, how a useful move can be made in proving a mathematical theorem, or where we should look in seeking a solution to a spatial puzzle. In the brain, knowledge that might shape current cognition is distributed across multiple brain systems. Semantic memory, for example, may be based around a proposed hub in the temporal pole [98], while episodic memory, spatial knowledge, and social knowledge are linked to distinct components of the default mode network [99]. To understand MD activity in constructing solutions to cognitive problems, we need to know how multiple aspects of knowledge feed into this process. Again, this is reminiscent of the widespread connectivity of MD regions (Figure 5) and our finding that multiple networks have representatives in the MD penumbra.

Outstanding Questions

How do different MD regions interact? Across the extended MD system, we know little of the dynamics of information representation and exchange during task performance. The very different connectivities of MD regions imply separate functional contributions to cognitive integration, but in fMRI, the dominant picture is one of corecruitment. This picture may reflect the low temporal resolution of fMRI, rendering the method blind to high-speed information development and exchange. Elucidating how task-relevant information arises and is distributed across the network calls for simultaneous electrophysiological recordings in separate MD regions, either in experimental animals or patients implanted for intracranial recordings.

How does MD activity bind together coherent processing across multiple brain regions? Again, this calls for electrophysiological studies, addressing questions that include directional information flow at different stages of a cognitive operation, and the role of precise timing relations (e.g., oscillatory synchrony) across brain regions.

How are brief segments of cognition combined into complex goal-subgoal structures? For example, we know little of how sustained goal maintenance directs brain activity in pursuit of a series of subgoals. Especially for complex behavior, a critical question is interaction between immediate cognitive activity and long-term knowledge of goals, subgoals, and their relationships.

What is the role of prominent MD foci seen outside cerebral cortex, especially in caudate and cerebellum? Almost nothing is known of cortical-subcortical and cerebro-cerebellar interaction as cognitive operations are carried out. High-field imaging may bring the spatial resolution needed for studies of small subcortical structures.

In classical artificial intelligence, problem solutions were often built up in an unlimited working memory, keeping track of a progressively more complex structure of goals and subgoals. For biological cognition this is not plausible; for goals such as travel to Japan or solving a scientific problem, only a small fraction can be represented in active neural firing at any one time, with the rest of the structure in long-term memory, ready for retrieval when required. At the same time, the current active focus of attention must remain bound to the long-term structure, so that, for example, a failure to progress to a goal by one route can trigger a search for an alternative. The issue is reminiscent of recent biological accounts of working memory, combining active neural firing with storage through short-term synaptic change [100,101]. It is presently unknown how the focus of attention in active cognition can be situated within a complex, long-term representation of the larger-scale problem.

A further open issue concerns the well-known capacity limitations of 'attention', reflected in difficulty carrying out several tasks at once [102,103]. Shared demands on MD activity could provide an obvious basis for such limits and, indeed, various authors have linked capacity limitations to the functions of frontal and parietal cortex [16,22,104,105]. Such proposals find support in neurophysiological studies, showing that, in frontal and parietal cortex, there is interference between representations of different visual stimuli [106], working memory items [107], or task components [86,108]. Further work is needed, however, to understand the physiological basis of this interference. In the visual system, capacity limits in representing multiple stimuli are thought to arise through a process of competition or divisive normalization [109–111]. In such models, each stimulus attempts to drive the activity of a neuron to a particular value, appropriate to representing the properties of this stimulus; with multiple stimuli in the field, opposing forces bring activity to a compromise value, reducing the fidelity of representation for any one. Similar patterns can be seen in the visual responses of prefrontal neurons [112,113], raising the possibility that divisive normalization is a general principle in MD cortex. Recurrent neural networks have become popular as models of working memory and cognitive control (e.g., [114]), and in a recent model, divisive normalization is the basis for limited working memory capacity [115]. Further experimental work is needed to test whether divisive normalization models may be extended to the broader attentional limits of MD activity and cognitive control.

Of course, our account of cognitive integration leaves much unknown. That said, like an early map of the globe, it provides an outline sketch of how distributed brain activity can assemble complex cognition. This sketch, we suggest, provides the skeleton we need to guide future, more detailed physiological study (see Outstanding Questions).

Acknowledgments

This work was funded by the Medical Research Council (UK, programme SUAG/045.G101400), the Wellcome Trust (grant 101092/Z/13/Z), a Cambridge Trust-Yousef Jameel scholarship to M.A., and a Gates Cambridge Trust scholarship to S.S.

References

1. Duncan, J. *et al.* (1996) Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognit. Psychol.* 30, 257–303
2. Kane, M.J. and Engle, R.W. (2002) The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon. Bull. Rev.* 9, 637–671
3. Unsworth, N. and Robison, M.K. (2017) A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychon. Bull. Rev.* 24, 1282–1311
4. Wilhelm, O. *et al.* (2013) What is working memory capacity, and how can we measure it? *Front. Psychol.* 4, 433
5. Bishop, S.J. *et al.* (2008) COMT val158met genotype affects recruitment of neural mechanisms supporting fluid intelligence. *Cereb. Cortex* 18, 2132–2140
6. Duncan, J. *et al.* (2000) A neural basis for general intelligence. *Science* 289, 457–460
7. Prabhakaran, V. *et al.* (1997) Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognit. Psychol.* 33, 43–63
8. Woolgar, A. *et al.* (2010) Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14899–14902
9. Duncan, J. (2010) The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179
10. Duncan, J. (2013) The structure of cognition: attentional episodes in mind and brain. *Neuron* 80, 35–50

11. Spearman, C. (1904) General intelligence, objectively determined and measured. *Am. J. Psychol.* 15, 201–293
12. Spearman, C. (1927) *The Abilities of Man*, Macmillan
13. Raven, J.C. et al. (1988) *Manual for Raven's Progressive Matrices and Vocabulary Scales*, H.K. Lewis
14. Institute for Personality and Ability Testing (1973) *Measuring Intelligence with the Culture Fair Tests*, The Institute for Personality and Ability Testing
15. Thurstone, L.L. and Thurstone, T.G. (1949) *Examiner Manual for the SRA Primary Mental Abilities Test*, Science Research Associates
16. Duncan, J. (1995) Attention, intelligence and the frontal lobes. In *The Cognitive Neurosciences* (Gazzaniga, M.S., ed.), pp. 721–733, MIT Press
17. Duncan, J. et al. (1995) Fluid intelligence after frontal lobe lesions. *Neuropsychologia* 33, 261–268
18. Kane, M.J. et al. (2007) For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychol. Sci.* 18, 614–621
19. Baddeley, A.D. (1986) *Working Memory*, Oxford University Press
20. Cohen, J.D. and Servan-Schreiber, D. (1992) Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol. Rev.* 99, 45–77
21. Desimone, R. and Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222
22. Norman, D. and Shallice, T. (1980) *Attention to Action: Willed and Automatic Control of Behavior (Report No. 8006)*, University of California, Center for Human Information Processing
23. Anderson, M.C. et al. (2004) Neural systems underlying the suppression of unwanted memories. *Science* 303, 232–235
24. Miyake, A. et al. (2000) The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognit. Psychol.* 41, 49–100
25. Newell, A. et al. (1958) Elements of a theory of human problem solving. *Psychol. Rev.* 65, 151–166
26. Sacerdoti, E.D. (1974) Planning in a hierarchy of abstraction spaces. *Artif. Intell.* 5, 115–135
27. Carpenter, P.A. et al. (1990) What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychol. Rev.* 97, 404–431
28. Duncan, J. et al. (2017) Complexity and compositionality in fluid intelligence. *Proc. Natl. Acad. Sci. U. S. A.* 114, 5295–5299
29. Chuderski, A. (2019) Even a single trivial binding of information is critical for fluid intelligence. *Intelligence* 77, 101396
30. Duncan, J. et al. (1997) Competitive brain activity in visual attention. *Curr. Opin. Neurobiol.* 7, 255–261
31. Phaf, R.H. et al. (1990) SLAM: a connectionist model for attention in visual selection tasks. *Cognit. Psychol.* 22, 273–341
32. Duncan, J. and Owen, A.M. (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483
33. Fedorenko, E. et al. (2013) Broad domain generality in focal regions of frontal and parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 110, 16616–16621
34. Han, S.W. and Marois, R. (2013) Dissociation between process-based and data-based limitations for conscious perception in the human brain. *Neuroimage* 64, 399–406
35. Wen, T. et al. (2018) Response of the multiple-demand network during simple stimulus discriminations. *Neuroimage* 177, 79–87
36. Duncan, J. (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.* 2, 820–829
37. Cole, M.W. and Schneider, W. (2007) The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage* 37, 343–360
38. Naghavi, H.R. and Nyberg, L. (2005) Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration? *Conscious. Cogn.* 14, 390–425
39. Kovacs, K. and Conway, A.R.A. (2016) Process overlap theory: a unified account of the general factor of intelligence. *Psych. Inquiry* 27, 151–177
40. Euler, M.J. (2018) Intelligence and uncertainty: implications of cognitive predictive processing for the neuroscience of cognitive ability. *Neurosci. Biobehav. Rev.* 94, 93–112
41. Woolgar, A. et al. (2018) Fluid intelligence is supported by the multiple-demand system not the language system. *Nat. Hum. Behav.* 2, 200–204
42. Barbey, A.K. et al. (2012) An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain* 135, 1154–1164
43. Gläscher, J. et al. (2010) Distributed neural system for general intelligence revealed by lesion mapping. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4705–4709
44. Assem, M. et al. (2020) A domain-general cognitive core defined in multimodally parcellated human cortex. *Cereb. Cortex* 30, 4361–4380
45. Glasser, M.F. et al. (2016) A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178
46. Dosenbach, N.U. et al. (2006) A core system for the implementation of task sets. *Neuron* 50, 799–812
47. Yeo, B.T. et al. (2015) Functional specialization and flexibility in human association cortex. *Cereb. Cortex* 25, 3654–3672
48. Hampshire, A. et al. (2012) Fractionating human intelligence. *Neuron* 76, 1225–1237
49. Shashidhara, S. et al. (2019) Progressive recruitment of the frontoparietal multiple-demand system with increased task complexity, time pressure, and reward. *J. Cogn. Neurosci.* 31, 1617–1630
50. Yeo, B.T. et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165
51. Ji, J.L. et al. (2019) Mapping the human brain's cortical-subcortical functional network organization. *Neuroimage* 185, 35–57
52. Bullmore, E. and Sporns, O. (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198
53. Cole, M.W. et al. (2010) Identifying the brain's most globally connected regions. *Neuroimage* 49, 3132–3148
54. Power, J.D. et al. (2013) Evidence for hubs in human functional brain networks. *Neuron* 79, 798–813
55. Cole, M.W. et al. (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355
56. Higo, T. et al. (2011) Distributed and causal influence of frontal operculum in task control. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4230–4235
57. Cole, M.W. et al. (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* 32, 8988–8999
58. Barbey, A.K. (2018) Network neuroscience theory of human intelligence. *Trends Cogn. Sci.* 22, 8–20
59. Li, S. et al. (2007) Flexible coding for categorical decisions in the human brain. *J. Neurosci.* 27, 12321–12330
60. Haynes, J.D. et al. (2007) Reading hidden intentions in the human brain. *Curr. Biol.* 17, 323–328
61. Woolgar, A. et al. (2011) Adaptive coding of task-relevant information in human frontoparietal cortex. *J. Neurosci.* 31, 14592–14599
62. Bhandari, A. et al. (2018) Just above chance: is it harder to decode information from prefrontal cortex hemodynamic activity patterns? *J. Cogn. Neurosci.* 30, 1473–1498
63. Woolgar, A. et al. (2016) Coding of visual, auditory, rule, and response information in the brain: 10 years of multivoxel pattern analysis. *J. Cogn. Neurosci.* 28, 1433–1454
64. Woolgar, A. et al. (2015) Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *Neuroimage* 109, 429–437
65. Jackson, J. et al. (2017) Feature-selective attention in frontoparietal cortex: multivoxel codes adjust to prioritize task-relevant information. *J. Cogn. Neurosci.* 29, 310–321
66. Erez, Y. and Duncan, J. (2015) Discrimination of visual categories based on behavioral relevance in widespread regions of frontoparietal cortex. *J. Neurosci.* 35, 12383–12393
67. Ford, K.A. et al. (2009) BOLD fMRI activation for anti-saccades in nonhuman primates. *Neuroimage* 45, 470–476
68. Mitchell, D.J. et al. (2016) A putative multiple-demand system in the macaque brain. *J. Neurosci.* 36, 8574–8585

69. Freedman, D.J. *et al.* (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316
70. Freedman, D.J. and Assad, J.A. (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443, 85–88
71. Everling, S. *et al.* (2002) Filtering of neural signals by focused attention in the monkey prefrontal cortex. *Nat. Neurosci.* 5, 671–676
72. Sakagami, M. and Niki, H. (1994) Encoding of behavioral significance of visual stimuli by primate prefrontal neurons: relation to relevant task conditions. *Exp. Brain Res.* 97, 423–436
73. Hunt, L.T. *et al.* (2018) Triple dissociation of attention and decision computations across prefrontal cortex. *Nat. Neurosci.* 21, 1471–1481
74. Kuwabara, M. *et al.* (2014) Cognitive control functions of anterior cingulate cortex in macaque monkeys performing a Wisconsin Card Sorting Test analog. *J. Neurosci.* 34, 7531–7547
75. Buschman, T.J. and Miller, E.K. (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862
76. Salazar, R.F. *et al.* (2012) Content-specific fronto-parietal synchronization during visual working memory. *Science* 338, 1097–1100
77. Oemisch, M. *et al.* (2015) Interareal spike-train correlations of anterior cingulate and dorsal prefrontal cortex during attention shifts. *J. Neurosci.* 35, 13076–13089
78. Voloh, B. *et al.* (2015) Theta-gamma coordination between anterior cingulate and prefrontal cortex indexes correct attention shifts. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8457–8462
79. Smolensky, P. (1990) Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* 46, 159–216
80. Rigotti, M. *et al.* (2010) Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* 4, 24
81. Rigotti, M. *et al.* (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590
82. Warden, M.R. and Miller, E.K. (2010) Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15801–15810
83. Naya, Y. *et al.* (2017) Contributions of primate prefrontal cortex and medial temporal lobe to temporal-order memory. *Proc. Natl. Acad. Sci. U. S. A.* 114, 13555–13560
84. Stokes, M.G. *et al.* (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78, 364–375
85. Mushiaki, H. *et al.* (2006) Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* 50, 631–641
86. Kadohisa, M. *et al.* (2020) Focused representation of successive task episodes in frontal and parietal cortex. *Cereb. Cortex* 30, 1779–1796
87. Sigala, N. *et al.* (2008) Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 105, 11969–11974
88. Deary, I.J. *et al.* (2010) The neuroscience of human intelligence differences. *Nat. Rev. Neurosci.* 11, 201–211
89. Basten, U. *et al.* (2015) Where smart brains are different: a quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence* 51, 10–27
90. Goriounova, N.A. *et al.* (2018) Large and fast human pyramidal neurons associate with intelligence. *eLife* 7, e41714
91. Hilger, K. *et al.* (2020) Temporal stability of functional brain modules associated with human intelligence. *Hum. Brain Mapp* 41, 362–372
92. Schultz, D.H. and Cole, M.W. (2016) Higher intelligence is associated with less task-related brain network reconfiguration. *J. Neurosci.* 36, 8551–8561
93. Cattell, R.B. (1971) *Abilities: Their Structure, Growth and Action*, Houghton-Mifflin
94. Hebb, D.O. (1942) The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. *Proc. Am. Philos. Soc.* 85, 275–292
95. van der Maas, H.L. *et al.* (2006) A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychol. Rev.* 113, 842–861
96. Kievit, R.A. *et al.* (2017) Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychol. Sci.* 28, 1419–1431
97. Newell, A. (1990) *Unified Theories of Cognition*, Harvard University Press
98. Lambon Ralph, M.A. *et al.* (2017) The neural and computational bases of semantic cognition. *Nat. Rev. Neurosci.* 18, 42–55
99. Andrews-Hanna, J.R. *et al.* (2010) Functional-anatomic fractionation of the brain's default network. *Neuron* 65, 550–562
100. Stokes, M.G. (2015) 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405
101. Lundqvist, M. *et al.* (2016) Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164
102. Broadbent, D.E. (1958) *Perception and Communication*, Pergamon
103. Kahneman, D. (1973) *Attention and Effort*, Prentice-Hall
104. Dehaene, S. *et al.* (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8520–8525
105. Marois, R. and Ivanoff, J. (2005) Capacity limits of information processing in the brain. *Trends Cogn. Sci.* 9, 296–305
106. Kadohisa, M. *et al.* (2013) Dynamic construction of a coherent attentional state in a prefrontal cell population. *Neuron* 80, 235–246
107. Buschman, T.J. *et al.* (2011) Neural substrates of cognitive capacity limitations. *Proc. Natl. Acad. Sci. U. S. A.* 108, 11252–11255
108. Watanabe, K. and Funahashi, S. (2014) Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nat. Neurosci.* 17, 601–611
109. Chelazzi, L. *et al.* (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* 80, 2918–2940
110. Reynolds, J.H. *et al.* (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* 19, 1736–1743
111. Reynolds, J.H. and Heeger, D.J. (2009) The normalization model of attention. *Neuron* 61, 168–185
112. Lara, A.H. and Wallis, J.D. (2014) Executive control processes underlying multi-item working memory. *Nat. Neurosci.* 17, 876–883
113. Matsushima, A. and Tanaka, M. (2014) Differential neuronal representation of spatial attention dependent on relative target locations during multiple object tracking. *J. Neurosci.* 34, 9963–9969
114. Chaisangmongkon, W. *et al.* (2017) Computing by robust transience: how the fronto-parietal network performs sequential, category-based decisions. *Neuron* 93, 1504–1517
115. Bouchacourt, F. and Buschman, T.J. (2019) A flexible model of working memory. *Neuron* 103, 147–160