NEURAL NETWORKS: AN EMPIRICAL NEUROSCIENCE APPROACH TOWARD UNDERSTANDING COGNITION

Review of Joaquín M. Fuster's *Cortex and Mind: Unifying Cognition.* ISBN 0-19-514752-9, Oxford University Press, 2003, 310 pages. Price: US $55.00, UK £33.50

Theories of brain organization focus on two distinct, but complementary principles: *modularity,* the existence of neuronal assemblies with intrinsic functional specialization and *network connectivity*, the integration of information from distributed brain regions resulting in organized behavior. While the modular model may be reasonable to describe fundamental features of primary cortices, it is insufficient to explain complex cognitive processes that cannot be localized to isolated brain regions. Rather, cognitive abilities emerge from contextual relations that are subserved by widespread cortical connections. As espoused by Joaquin Fuster, in his book *Memory and the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate* (1995) and extended in the recently published *Cortex and Mind: Unifying Cognition* (Fuster, 2003), we are now at the forefront of a "revolution in contemporary neuroscience" as the delineation of the neural substrate of cognition through the study of neural networks has begun to be addressed experimentally. Fuster describes his book, *Cortex and Mind,* as "a chronicle of that shift of paradigms and of the new rules of discovery that it entails." He contends that such an empirical shift from a reductionist, modular model to a holistic, network model offers promise of accomplishing our long-term goal of resolving the mind-brain question.

Both the modular and network model have a long and rich history. The concept of modularity likely has it origins at the end of the 18th century in the study of phrenology (Wyhe, 2004). Phrenology was a practice based on the theories of Vienesse physician, Franz Joseph Gall, which attempted to evaluate cognitive traits using skull surface features and was for a time considered a legitimate scientific discipline. In a letter from Gall to a censorship official in 1798, he outlined the tenets of his theories, two of which are: "The faculties and propensities of man have their seat in the brain"… which… "are essentially distinct and independent: they ought, consequently, to have their seat in parts of the brain distinct and independent of each other" (Wyhe, 2004). Although the practice of phrenology as based on the examination of the skull has been discredited, the concept of functional localization within the brain has continued to thrive and accumulate empirical evidence. Notably, in 1861, just as phrenology was falling into disrepute, Paul Broca localized expressive language to the inferior frontal gyrus by studying brain lesions in stroke patients and thus offered the first anatomical evidence of the localization of specific brain functions (Berker et al., 1986). Empirical support for the modular model of cognition still exists largely through three venues: neurophysiological approaches, which characterize processing characteristics of single neurons or local populations, neuroanatomical approaches, which describe cytoarchitectonic differences and other structural distinctions between brain regions, and neuropsychological approaches, which attempt to link focal brain lesions to specific behavioral or cognitive functions (Fuster, 2003).

One of the earliest opponents of Gall's localizationalist views was a contemporary of his, Pierre Flourens, who argued that some forms of knowledge were dispersed throughout the entire cortex (Sabbatini, 2003). His views were based on what might be considered the first empirical studies of brain function. By ablating discrete regions of cortex in experimental animals and observing subsequent behavior, Flourens was able to ascribe several aspects of function to distinct brain regions, but was unable to localize "higher" cognitive abilities and so presumed these to be distributed (Tizard, 1959). As described in *Cortex and Mind,* distributed views of cognition have had influential advocates in the 20th century, with support from individuals such as John Hughlings Jackson, Constantin Von Monakow, Karl Lashley, Donal Hebb and Freidrich Hayek (Fuster, 2003). Despite an empirical background, the formal investigation of neural networks had its origins in the philosophical and computational field with the development of connectionism. While connectionists have sought the elucidation of cognitive processes by studying artificial neural networks, simplified models of brain region interactions (Garson 2002), the experimental investigation of actual neural networks within the field of neuroscience has lagged and only relatively recently has it been considered a viable empirical research strategy. However, momentum has been gaining as anatomical studies that characterize extensive cortico-cortical and subcortical connectivity with axonal degeneration and tracer techniques, and physiological studies that assess widespread cortical activity using microelectrodes, EEG and functional imaging, have begun to accumulate evidence of the structure and function of genuine neural networks.

Two centuries after Gall's letter, in which he presents the tenets of phrenology and thus the basis of the modular model, Fuster outlines the conceptual basis of the network model in *Cortex and Mind,* with seven points: 1) cognitive information is represented in wide, overlapping, and interactive neuronal networks of the cerebral cortex; 2) such networks develop on a core of organized modules of elementary sensory and motor functions, to which they remain connected; 3) the cognitive code is a relational code, based on connectivity between discrete neuronal aggregates of the cortex (modules, assemblies, or network nodes); 4) the code's diversity and specificity derive from the myriad possibilities of combination of those neuronal aggregates between themselves; 5) any cortical neuron can be part of many networks, and thus of many percepts, memories, items of experience, or personal knowledge; 6) a network can serve several cognitive functions; and 7) cognitive functions consist of functional interactions within and between cortical networks (page xi). It is this incisive summary and its support throughout the text, which we feel represents the greatest strength of *Cortex and Mind.*

We are in absolute agreement with Fuster's thesis and believe that he effectively establishes the conceptual framework necessary for an empirical approach to investigating the network model. Importantly, we wish to emphasize a critical aspect of his tenets; they do not encourage the rejection of the vast pool of evidence that has accumulated in association with cortical modules. A noteworthy example of how the network and module approach can be integrated was presented by Mesulam over 20 years ago, in which he describes the unique contribution of four distinct brain regions in a network that underlies directed spatial attention (Mesulam, 1981). He further describes how lesions limited to one component of this network results in a partial neglect syndrome, while lesions involving all network components result in deficits that far exceed the impact of larger more constrained lesions. Thus, although modules in themselves may not completely represent cognitive processes, they do serve as building blocks, establishing the complexity and diversity of hierarchical networks.

Throughout *Cortex and Mind*, Fuster presents an insightful account of physiological and anatomical evidence that unite his tenets in support of a network model of cognition. He begins with a detailed description of the development and organization of cortical networks, and follows with a systematic discussion, highlighting evidence for the network basis of perception, memory, attention, language and intelligence. His approach represents a distinct departure from many other monographs that address the mind-brain question through cognitive psychology or theoretical approaches. We were particularly satisfied by the manner in which he presented the network basis for the integration of top-down and bottom-up processing. By its very nature, the resolution between goal-directed decisions regarding environmental stimuli (top-down modulation) and the intrinsic impact of stimuli based on their novelty and salience (bottom-up processing) involves the interaction of widely distributed brain regions. This concept was well organized and effectively merged with his view of the *perception-action cycle*. Its underlying role in perception, attention and memory was stressed by its repetition throughout the book, and in our opinion, represents the most cohesive example of his argument.

We feel that Fuster was largely successful in accomplishing the objectives he outlined in the preface of his book, and was mostly limited by the often scant evidence that currently exists of the role of cortical networks in cognitive processes. However, we would have advocated the inclusion of more references to the growing list of methods already in practice, or in a state of development, which attempt to more directly address the network model. For example, there have been successful efforts to use electrophysiological techniques in experimental animals to investigate regional connectivity by employing simultaneous recordings from more than one neuron (Funahashi and Inoue, 2000; Constantinidis et al., 2001; Constantinidis and Goldman-Rakic, 2002; Wilson et al., 1993). Although, we do concede that these techniques are often limited to assessments of local interactions and not functional connectivity between multiple, widely distributed regions.

Functional neuroimaging in humans, however, is ideally and uniquely suited to explore neural networks, since it simultaneously records correlates of neural activity throughout the entire functioning brain with high spatial resolution. Unfortunately, in *Cortex and Mind*, Fuster underemphasizes the contribution of neuroimaging, due to its limitations in spatial and temporal resolution and the incompletely characterized relationship between blood flow and neural activity. While we acknowledge the validity of these concerns, advances in the physics of neuroimaging, understanding of the physiological basis of the imaging signal and the analysis and interpretation of data are progressing rapidly. In Fuster's discussion on perception, he alludes to the summary of neuroimaging findings from two review articles as revealing only "the tip of an iceberg" and further comments that "the graphic display of activation maxima exaggerates localization" (pagg. 104-105). In this regard, we

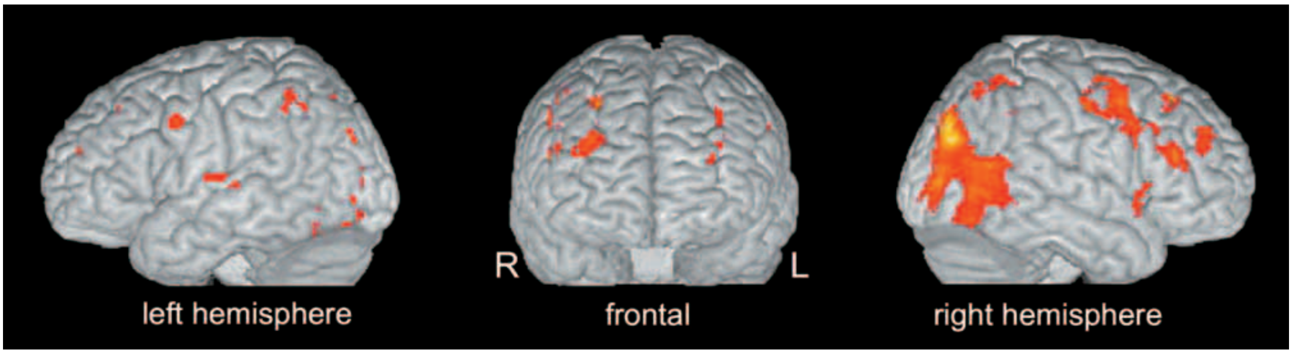

Fig. 1 – *Working memory maintenance network map. Three-dimensionally rendered brain, revealing the network of brain regions during the maintenance of a visual stimulus in mind. Univariate analysis revealed a much more restricted map (data not shown). Each of these regions, including the visual association cortex, prefrontal cortex, anterior cingulate, supplementary motor area, intraparietal sulcus, thalamus, caudate, and hippocampus, exhibited high functional connectivity with a region in the right visual association cortex (fusiform face area (FFA)) during the maintenance of a face representation across a short delay interval. We propose that it is the concerted activity and interaction between these regions within the framework of the distributed system that permits the retention of an internal representation of a stimulus and successful performance of the working memory task. Adapted from (Gazzaley et al., 2004).*

are in complete agreement that many studies utilizing functional neuroimaging have not fully explored the data that was collected. Moreover, functional neuroimaging, which has been accused by critics as being no more than modern-day phrenology (Utall, 2001), often does take a strong, localizationalist perspective.

Most functional imaging studies utilize univariate analyses, permitting only the independent assessment of activity within each brain region in isolation of all others. However, it is important to recognize that in response to this limitation of univariate data, there has been the steady development of multivariate approaches to analyze neuroimaging data in a manner that more directly addresses the network model of the cognition (McIntosh, 1998; Friston et al., 1993). Although originally applied to PET data (McIntosh et al., 1996; Nyberg et al., 1996; Kohler et al., 1998), multivariate analyses have now been successfully used to evaluate networks in functional MRI (fMRI) studies of learning (Buchel et al., 1999; Toni et al., 2002), working memory (Gazzaley et al., 2004), attention (Friston and Buchel, 2000; Rowe et al., 2002), long-term memory (Maguire et al., 2000) and language (Hampson et al., 2002). Multivariate analyses of imaging data allow the generation of functional and effective connectivity maps of brain regions that interact within the framework of a distributed system to underlie emergent cognitive processes. These methods do not trivialize the functional specialization of brain regions, but rather emphasize the role of brain regions in the context of other covarying, anatomically connected, active brain regions, as well as the specific cognitive process that is being engaged (see Figure 1). Comparisons between the univariate activation map and the multivariate network map of brain regions that interact with the visual association cortex during the maintenance of a visual stimulus, highlights the extensive, untapped information contained in fMRI datasets that are not being revealed by traditional univariate analysis (Gazzaley et al., 2004). The multivariate analytical approach has now begun to capitalize on the extensive network data obtained by simultaneously recording from the entire functioning human brain with high spatial resolution, a unique feature of functional imaging that distinguishes it from single-unit recordings in experimental animals, which is limited in this regard. In recent years, there has been a strong commitment to develop new multivariate methods to assess neural networks using fMRI (Rissman et al., 2004; Sun et al., 2004; Penny et al., 2004; Lin et al., 2003).

On a whole, we found *Cortex and Mind* to be nothing short of inspiring, encouraging a new generation of exploration into the neural basis of cognition. However, it is important to speculate on specifically how the future of cognitive neuroscience might be guided by Fuster's book. After all, it is insufficient empirical evidence that has created the imbalance between theoretical and practical approaches in understanding the network nature of cognition. The future of the neural network model now requires a focused and organized experimental approach by which to narrow this gap. How are networks to be characterized using currently available anatomical and physiological techniques? We take this opportunity to suggest some possible scenarios – and the theme is *integration*. First, as described in the discussion above, the continued development of effective multivariate techniques is essential, not just within the field of functional neuroimaging, but for all physiological methods. This will involve the integration of diverse research skills focused on the same project and will require the recruitment of students and post-doctoral fellows, as well as collaboration with colleagues, whose interests and expertise do not lie directly in neuroscience, but rather signal processing, physics and statistics. Second, an integration of the varied methodologies that can be used to explore networks will be necessary in order to generate richer and more informative datasets. Studies have already begun to capitalize on the respective strengths in spatial and temporal resolution of fMRI and ERP/MEG by coupling them in the same experiment (Bledowski et al., 2004; Schulz et al., 2004)). Furthermore, while fMRI, ERP, and MEG are promising techniques that can identify network regions associated with cognitive processes, they are correlational methods that cannot on their own establish the necessity of component regions. To establish the functional necessity of identified network nodes, it will be necessary to integrate these methods with those capable of inducing focal reversible lesions, such as cortical cooling in animals and transcranial magnetic stimulation in humans, as well as perform network studies on neurological/neurosurgical patients with chronic lesions. This will allow an assessment of the necessity of specific network regions for cognitive performance, as well as the plasticity of the network in response to their absence. Lastly, and perhaps most importantly, it is essential that we continue to integrate the reductionist – modular approach with the holistic – network approach. Parallel experiments should be planned from the onset to generate complementary information from these two approaches. Only by using data from both of these models to inform one another can we really hope to obtain a comprehensive understanding of neural mechanisms underlying cognitive processes.

Adam Gazzaley and Mark D'Esposito

References

Berker EA, Berker AH and Smith A. Translation of Broca's 1865 report. Localization of speech in the third left frontal convolution. *Archives of Neurology, 43:* 1065-1072, 1986.

Bledowski C, Prvulovic D, Goebel R, Zanella FE and Linden DE. Attentional systems in target and distractor processing: A combined ERP and fMRI study. *NeuroImage, 22:* 530-540, 2004.

Buchel C, Coull JT and Friston KJ. The predictive value of changes in effective connectivity for human learning. *Science, 283:* 1538-1541, 1999.

Constantinidis C, Franowicz MN and Goldman-Rakic PS. Coding specificity in cortical microcircuits: A multiple-electrode analysis of primate prefrontal cortex. *Journal of Neuroscience, 21:* 3646-3655, 2001.

Constantinidis C and Goldman-Rakic PS. Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *Journal of Neurophysiology, 88:* 3487-3497, 2002.

Friston KJ and Buchel C. Attentional modulation of effective connectivity from v2 to v5/mt in humans. *Proceedings of the National Academy of Science USA, 97:* 7591-7596, 2000.

Friston KJ, Frith CD, Liddle PF and Frackowiak RS. Functional connectivity: The principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow Metab, 13:* 5-14, 1993.

Funahashi S and Inoue M. Neuronal interactions related to working memory processes in the primate prefrontal cortex revealed by cross-correlation analysis. *Cerebral Cortex, 10:* 535-551, 2000.

Fuster JM. *Memory in the cerebral cortex*. Cambridge, MA: MIT Press, 1995.

Fuster JM. *Cortex and mind: Unifying cognition*. New York: Oxford University Press, 2003.

Garson J. Connectionism. In Zalta EN (Ed), *The Stanford Encyclopeida of Philosophy*, Winter 2002 Edition, http://plato.stanford.edu/archives/win2002/entries/connectionism.

Gazzaley A, Rissman J and D'Esposito M. Functional connectivity during working memory maintenance. *Cognitive Affectice Behavioral Neuroscience*: Submitted, 2004.

Hampson M, Peterson BS, Skudlarski P, Gatenby JC and Gore JC. Detection of functional connectivity using temporal correlations in mr images. *Human Brain Mapping, 15:* 247-262, 2002.

Kohler S, McIntosh AR, Moscovitch M and Winocur G. Functional interactions between the medial temporal lobes and posterior neocortex related to episodic memory retrieval. *Cerebral Cortex, 8:* 451-461, 1998.

Lin FH, McIntosh AR, Agnew JA, Eden GF, Zeffiro TA and Belliveau JW. Multivariate analysis of neuronal interactions in the generalized partial least squares framework: Simulations and empirical studies. *NeuroImage, 20:* 625-642, 2003.

Maguire EA, Mummery CJ and Buchel C. Patterns of hippocampal-cortical interaction dissociate temporal lobe memory subsystems. *Hippocampus, 10:* 475-482, 2000.

McIntosh AR. Understanding neural interactions in learning and memory using functional neuroimaging. *Annals of the New York Academy of Science, 855:* 556-571, 1998.

McIntosh AR, Bookstein FL, Haxby JV and Grady CL. Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage, 3:* 143-157, 1996.

Mesulam MM. A cortical network for directed attention and unilateral neglect. *Annals of Neurology, 10:* 309-325, 1981.

Nyberg L, McIntosh AR, Cabeza R, Nilsson LG, Houle S, Habib R and Tulving E. Network analysis of positron emission tomography regional cerebral blood flow data: Ensemble inhibition during episodic memory retrieval. *Journal of Neuroscience, 16:* 3753-3759, 1996.

Penny WD, Stephan KE, Mechelli A and Friston KJ. Comparing dynamic causal models. *NeuroImage, 22:* 1157-1172, 2004.

Rissman J, Gazzaley A and D'Esposito M. Measuring functional connectivity during dstinct stages of a cognitive task. *NeuroImage, 23:* 252-263, 2004.

Rowe J, Friston K, Frackowiak R and Passingham R. Attention to action: Specific modulation of corticocortical interactions in humans. *NeuroImage, 17:* 988, 2002.

Sabbatini RME. Phrenology, the history of brain localization. *Brain and Mind*: http://nib.unicamp.br/~sabbatin, 2003.

Schulz M, Chau W, Graham SJ, McIntosh AR, Ross B, Ishii R and Pantev C. An integrative meg-fmri study of the primary somatosensory cortex using cross-modal correspondence analysis. *NeuroImage, 22:* 120-133, 2004.

Sun FT, Miller LM, and D'Esposito M. Measuring inter-regional functional connectivity using coherence and partial coherence analyses of fmri data. *NeuroImage, 21:* 647-658, 2004.

Tizard B. Theories of brain localization from Flourens to Lashley. *Medical History, 3:* 132-145, 1959.

Toni I, Rowe J, Stephan KE and Passingham RE. Changes of cortico-striatal effective connectivity during visuomotor learning. *Cerebral Cortex, 12:* 1040-7, 2002.

Utall W. *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press, 2001.

Wilson FA, Scalaidhe SP and Goldman-Rakic PS. Dissociation of object and spatial processing domains in prefrontal cortex. *Science, 260:* 1955-1958, 1993.

Wyhe J. The history of phrenology on the web. *http://pages.britishlibrary.net/phrenology/*, 2004.

Adam Gazzaley, University of California, Berkeley 32 Barker Hall, Berkeley, CA 94720-1650
e-mail: adamgazz@comewander.com http://comewander.com/; http://bic.berkeley.edu/despolab/people/position.php?rid=4

Mark D'Esposito, Henry H. Wheeler Jr. Brain Imaging Center, Helen Wills, Neuroscience Institute and Department of Psychology, University of California, Berkeley, 3210 Tolman Hall, Berkeley,California 94720-1650, USA.
e-mail: despo@socrates.berkeley.edu http://bic.berkeley.edu/despolab/