

Original Article

Contexts and Catalysts

A Resolution of the Localization and Integration of Function in the Brain

Anthony Randal McIntosh

Rotman Research Institute of Baycrest Centre, University of Toronto; 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada; E-mail: mcintosh@psych.utoronto.ca

Abstract

There has been a historical tension between theories of brain function emphasizing regional specialization and those focusing on integration across regions. This tension continues despite the pervasive use of functional neuroimaging, which enables testing of these theories in the human brain. There are instances of agreement, where regions thought to be critical for a given behavior (e.g., Broca's area and language production) do become more active when a person engages in that behavior. However, a number of disconcerting results have also been found. These include activation in areas not thought to be important for the behavior, and lack of activation in regions thought to be critical for particular behaviors based on studies of the damaged brain. A recently proposed Neural Context hypothesis of brain function provides a mechanism that can reconcile these apparently

disparate findings. The hypothesis states that the functional relevance of a brain area depends on the status of other connected areas—i.e., the context within which the region is operating. A region can participate in several behaviors through variations in its interactions with other areas. It is possible that certain critical nodes serve as Behavioural Catalysts, enabling the transition between behavioral states, without differential alterations in the measured activity. By virtue of its anatomical connections, an area could facilitate a shift in functional connectivity between one set of regions to another. An imaging study on the changing interregional interactions involving the hippocampus in learning and awareness serves as an example of neural context. In this case, the hippocampus may serve to catalyze the transition to awareness.

Index Entries: Neuroimaging; awareness; cognition; neural networks; learning.

Introduction

Neuroscience investigations into the link between mind and brain have received an

enormous boost with the advent of functional neuroimaging. As more scientists used neuroimaging, which was greatly facilitated by

the introduction of functional MRI, a great many studies have purportedly identified regionally specific changes across a vast number of cognitive operations. As these data were amassed, however, there was substantial overlap in the areas attributed to one function versus another, and the reliability of regional activation appeared less than ideal (Poeppel, 1996; McGonigle et al., 2000). Indeed, a skeptical eye would suggest that on any given day, any part of the cerebrum is likely to be active for a given task. Unfortunately, published meta-analyses of imaging data seem to corroborate this skeptical view (Cabeza and Nyberg, 2000).

Perhaps part of the problem with this mapping exercise is the prevalent model of how to make the link between brain and mind. Most studies have tried to identify which area is most active for the task of interest compared to other control tasks, but perhaps activity of an isolated region is not the critical factor. Instead, it is the status of the rest of the brain at that point in time that is the most direct link between the neural firing patterns and mental function. This is what I have called neural context (McIntosh, 1999; McIntosh, 2001). Stated differently, the contribution of an area (or any neural element) to a mental function depends on the activity of other anatomically related regions.

Figure 1 gives a simplistic illustration of neural context. The left figure shows a collection of neural ensembles that are interacting during a mental operation, and the right figure shows a different pattern of interactions during a different mental operation. In both cases, the same four sets of ensembles are engaged (shaded gray), but because they are interacting with different areas, they participate in two different operations. In other words, the neural context between the two networks change and thus the same four neural elements can contribute to different tasks.

Two primary features of the brain are central to neural context: anatomical connectivity

and response plasticity (McIntosh, 2000). Several authors have noted that brain connectivity shows dense local connections and sparse distal connections (Felleman and Van Essen, 1991; Scannell et al., 1999; Sporns et al., 2000; Stephan et al., 2000). When considered from a regional perspective, the system seems to waver between complete connectivity and complete independence. The term "semi-connected" or "degeneracy" has been used to characterize this connectivity structure (Edelman, 1978; Tononi et al., 1999; McIntosh, 2000; Friston and Price, 2003). The main point is that this pattern of anatomical connectivity enables many parallel routes of information flow between areas such that any two parts of the brain may share similar connections, but also have unique inputs and outputs. From the perspective of information theory, this allows the system to encode a great deal more information than it might through some other configuration (Tononi et al., 1994; Tononi et al., 1996; Sporns et al., 2000). A second feature is response plasticity, where neurons can show a rapid and transient shift in response to afferent stimulation that is dependent on the conditions under which they fire. Response plasticity in relation to learning and memory has been observed in several parts of the brain, from single cells in isolate spinal cord preparations (Wolpaw and Lee, 1989) to primary sensory and motor structures (Donoghue et al., 1990; Recanzone et al., 1992), and may be a ubiquitous property of the central nervous system (Wolpaw, 1997).

When these two factors, connectivity and plasticity, are considered together, it raises the potential that the relevance of the response to a given connection will vary depending on the current operations within a region. This response, in turn, can be shaped by the demands of the present situation. It is here that the internal, or neural, context and the external context meet. The external environment and the requirements placed on the organism will influence the internal state. Several authors

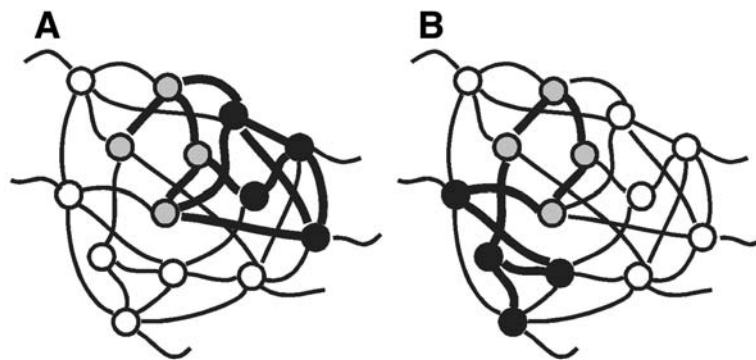


Fig. 1. Graphical representation of Neural Context. **(A)** Interactivity pattern (thick black lines) of the ensembles when processing one type of information. **(B)** Interactivity pattern of the ensembles processing a different type of information. Some of the same ensembles, colored gray, are involved in coding both sets of events, but the context under which they are interacting differs, i.e., the other areas that are engaged.

have emphasized this interplay between the external world and the brain's attempts to represent it (Picton and Stuss, 1994; Tononi et al., 1996; Friston and Price, 2001). Moreover, contextual dependencies have been demonstrated at the level of individual neurons and cell populations (Kozlov and Shabaev, 2000; Worgotter and Eysel, 2000). Considering across levels of organization, the combination of semi-connected anatomical structure and response plasticity provides great flexibility in the contributions different brain regions can make to the representation of the environment and the response to it.

If we now consider neuroimaging data as the expression of neural context, then the skeptical picture turns much more positive. Regarding the whole activity patterns measured in neuroimaging as the embodiment of cognition may be the most veridical translation of brain to mental function. Emerging work is demonstrating that regional activation alone may be ambiguous as an index of cognitive operation (D'Esposito et al., 1998). Other work has demonstrated that the overall pattern of activity (increase and decrease) may actually be a better index of specific mental function than the activation of a specialized area (Haxby et al.,

2001), which is congruent with similar observations of population coding in electrophysiological studies (Georgopoulos et al., 1986; Young and Yamane, 1992; Pouget et al., 2000).

A better appreciation of neural context can be obtained by explicit measures of interregional interactions, or "functional" or "effective" connectivity (Friston, 1994; Horwitz et al., 1999). Our imaging work across the past several years has provided hints of contextual dependencies. We have shown that the interactions of part of the prefrontal cortex (PFC) and the medial temporal lobe (MTL) can vary systematically depending on memory load (McIntosh et al., 1996) and the success of memory retrieval (McIntosh et al., 1997). In fact, it is quite possible for regions to show changes in their interactions without a concomitant change in activity (McIntosh et al., 1994; Kohler et al., 1998; Stephan et al., 2003). This is not surprising if one considers that the absolute amount of information processed within a region may not change between tasks, but the afferent source of the information would show task dependencies. For example, the collection of four core neural ensembles in Fig. 1 may receive identical amounts of afferent stimulation in both functional networks

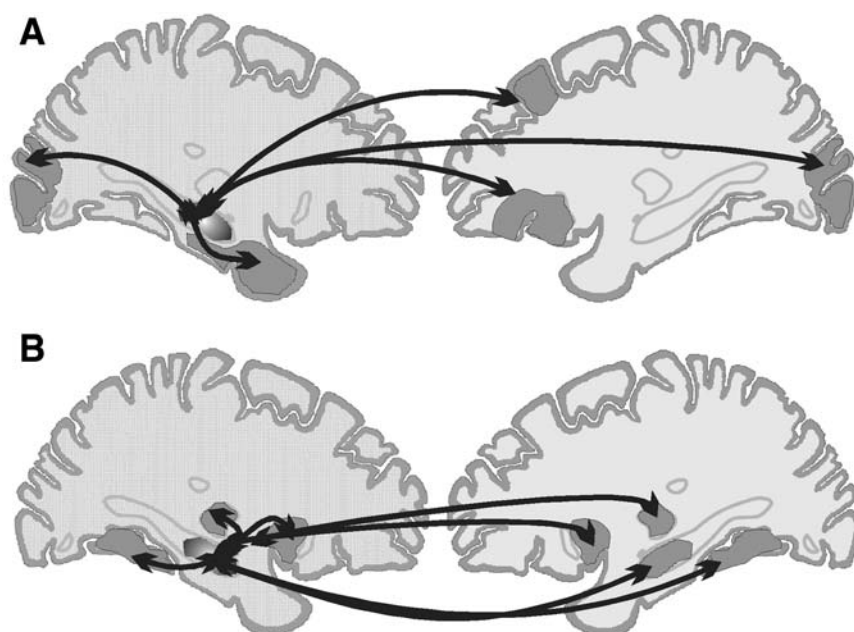


Fig. 2. Summary of the dominant functional connections of the left medial temporal lobe (MTL) in learning with **(A)** and without **(B)** awareness. Panel A shows dominant interactions, indicated by bidirectional arrows, between MTL and occipital, temporal and prefrontal cortices when learning proceeds with awareness. Panel B shows dominant interactions, between MTL with thalamus, basal ganglia, and contralateral MTL when learning proceeds without awareness. The key feature is the common involvement of the MTL, but because the regions to which it is functionally connected differ, the MTL can be related to learning in both situations.

and show similar levels of activity, yet depict quite different cognitive operations because the source of afferent input is different in the two networks.

A salient demonstration of neural context comes from a recent paper examining functional connectivity of the medial temporal lobe (MTL) in relation to learning and awareness (McIntosh et al., 2003). In a sensory learning paradigm, subjects were classified as Aware or Unaware based on whether they noted that one of two tones predicted a visual event. Only Aware subjects acquired and reversed a differential response to the tones, but both groups showed learned facilitation.

When we related brain activity (indexed by blood flow measured with positron emission tomography) to behavior in each group, we

observed that MTL activity was related to facilitation in both groups. This was curious given the suggestion the MTL is critical for learning with awareness, but not when learning proceeds without awareness (Clark and Squire, 1998; Clark and Squire, 2000). Considering the idea of neural context, it was possible that this common regional involvement in the two groups was an expression of contextual dependency. We then examined the functional connectivity of the MTL and observed completely different interactions of the MTL between groups (Fig. 2). In the Aware group, dominant MTL interactions were observed for prefrontal, occipital and temporal cortices, while in Unaware subjects, MTL interactions were more spatially restricted to inferotemporal, thalamus and basal ganglia. The MTL was related

to learning in both groups, but because the regions interacting with the MTL varied between groups, only in one case was learning accompanied by awareness. In other words, differences in neural context serve as a possible explanation for the involvement of the MTL in both groups.

Catalysts and Critical Functions

One problem for the notion of a neural context is the observation that the expression of certain behaviors or cognitive states appears to rely critically on specific brain areas. The overt expression of declarative memory, for example, depends on the integrity of the MTL (Squire, 1987). Such dependencies have led many researchers to speculate that such regions are parts of neurocognitive systems whose constituents subserve that function, but this implies a rather static view of brain function.

An alternative perspective, and one consistent with the idea of neural context, is to consider neural dynamics as a critical feature to understanding functional dependencies such as the MTL and memory. Most often studied at the cellular level, such dynamics are thought to be vital in enabling neurons to code for rapid temporal shifts in the environment and to make rapid adjustments of effectors at time scales much smaller than any single cell can achieve (Milton and Mackey, 2000). It is quite likely that these dynamics are at play across many levels of organization in the brain, with a similar general outcome: that it allows for rapid integration of information and responding. These same dynamics are also likely to underlie contextual effects between interacting populations which will manifest as changes in similarly large-scale behaviors, such as attentional states, perceptions, memory types, and quite likely consciousness (Bressler and Kelso, 2001). We have speculated (McIntosh et al., 2003), that shifting in behavioral states may require the integrity of certain key regions, which when damaged would result in a deficit

in that state. Such regions may not necessarily participate in the processing within that particular state, but rather enable the transition between states—it is a behavioral catalyst.

The likely feature of such catalysts is the anatomical relation of that region to those that are processing the primary information in the state in question. In the awareness study, the MTL was engaged in learning with or without awareness, interacting with regions that seemed to be related to learning in either attentive state. The MTL is anatomically connected with regions that were part of both patterns, providing the potential for the MTL to catalyze the transition between two different networks and thereby the movement from learning without awareness to learning with awareness. The critical point at which the MTL is needed is when learning moves to the conscious state. Prior to this, the MTL can be engaged, by virtue of its anatomical links, but not be critical for behavioral expression.

Considering regions that are critical for the expression of a function as potential catalysts emphasizes the dynamic nature of brain function. The temporal expansion of any behavior or cognitive function can be viewed as a series of transitions that require specific regions to be intact. In some cases, this dependence may reflect a network node that transmits information between regions (e.g., the lateral geniculate nucleus in the visual system). In other cases, areas enable the change in dominant interactions from one set of regions to another. These are the catalysts.

Neural context and catalysts are based on historical notions of dynamic systems and distributed functions (Kelso, 1995; Haken, 1996), but also are a new means of conceptualizing brain organization. The conflict between localization and distributed functions no longer exists in this view of brain function: functions are localized in the sense that their expression requires the integrity of certain regions or that a particular type of information is available to an area (e.g.,

visual, motor), but the actual expression of any function results from the actions of several regions. Such an integrated view of brain operations will likely bear more fruit as the notions of context and catalysts are tested and refined.

References

- Bressler, S. L. and Kelso, J. A. S. (2001) Cortical coordination dynamics and cognition. *Trends Cognit Sci* 5, 26–36.
- Cabeza, R. and Nyberg, L. (2000) Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J. Cognit. Neurosci.* 12, 1–47.
- Clark, C. M. and Squire, L. R. (2000) Awareness and the conditioned eyeblink response. In: *Eyeblink Classical Conditioning Volume I: Applications in Humans* (Woodruff-Pak D. S., Steinmetz J. E., eds.), Kluwer Academic Publishers, Norwell, MA, pp 229–253.
- Clark, R. E. and Squire, L. R. (1998) Classical conditioning and brain systems: the role of awareness. *Science* 280, 77–81.
- D'Esposito, M., Ballard, D., Aguirre, G. K., and Zarahn, E. (1998) Human prefrontal cortex is not specific for working memory: A functional MRI study. *Neuroimage* 8, 274–282.
- Donoghue, J. P., Suner, S., Sanes, and J. N. (1990) Dynamic organization of primary motor cortex output to target muscles in adult rats. II. Rapid reorganization following motor nerve lesions. *Exp. Brain Res* 79, 492–503.
- Edelman, G. M. (1978) Group selection and phasic re-entrant signalling: A theory of higher brain function. In: *The Mindful Brain* (Mountcastle V., Edelman G. M., eds.), MIT Press, Cambridge, MA, pp. 55–100.
- Felleman, D. J. and Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Friston, K. (1994) Functional and effective connectivity: A synthesis. *Hum. Brain. Mapp.* 2, 56–78.
- Friston, K. J. and Price, C. J. (2001) Dynamic representations and generative models of brain function. *Brain Research Bulletin* 54, 275–285.
- Friston, K. J. and Price, C. J. (2003) Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* 7, 151–152.
- Georgopoulos, A. P., Schwartz A. B., and Kettner R. E. (1986) Neuronal population coding of movement direction. *Science* 233, 1416–1419.
- Haken, H. (1996) *Principles of Brain Functioning: A synergetic approach to brain activity, behavior and cognition*. Springer, Berlin.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Horwitz, B., Tagamet, M. A., and McIntosh, A. R. (1999) Neural modeling, functional brain imaging, and cognition. *Trends Cognit. Sci.* 3, 91–98.
- Kelso, J. A. S. (1995) *Dynamic Patterns: The self-organization of brain and behavior*. MIT Press, Cambridge, MA.
- Kohler, S., McIntosh, A. R., Moscovitch, M., and Winocur, G. (1998) Functional interactions between the medial temporal lobes and posterior neocortex related to episodic memory retrieval. *Cereb. Cortex* 8, 451–461.
- Kozlov, A. P. and Shabaev, V. V. (2000) Analysis of the dynamics of interneuronal functional connections during conditioned reflex activity. *Neurosci. Behav. Physiol.* 30, 625–634.
- McGonigle, D. J., Howseman, A. M., Athwal, B. S., Friston, K. J., Frackowiak, R. S., and Holmes, A. P. (2000) Variability in fMRI: an examination of intersession differences. *Neuroimage* 11, 708–734.
- McIntosh, A. R. (1999) Mapping Cognition to the Brain Through Neural Interactions. *Memory* 7, 523–548.
- McIntosh, A. R. (2000) From location to integration: How neural interactions form the basis for human cognition. In: *Memory, Consciousness, and the Brain: The Tallinn Conference* (Tulving E., ed). Psychology Press, Philadelphia, PA.
- McIntosh, A. R. (2001) Towards a network theory of cognition. *Neural Netw.* 13, 861–876.
- McIntosh, A. R., Rajah, M. N., Lobaugh, N. J. (2003) Functional connectivity of the medial temporal lobe relates to learning and awareness. *J. Neurosci.* 23, 6520–6528.
- McIntosh, A. R., Nyberg, L., Bookstein, F. L., and Tulving, E. (1997) Differential functional connectivity of prefrontal and medial temporal cortices during episodic memory retrieval. *Hum. Brain Mapp.* 5, 323–327.
- McIntosh, A. R., Grady, C. L., Haxby, J. V., Ungerleider, L. G., and Horwitz, B. (1996) Changes in limbic and prefrontal functional interactions in a working memory task for faces. *Cereb. Cortex* 6, 571–584.

- McIntosh, A. R., Grady, C. L., Ungerleider, L. G., Haxby, J. V., Rapoport, S. I., and Horwitz, B. (1994) Network analysis of cortical visual pathways mapped with PET. *J. Neurosci.* 14, 655–666.
- Milton, J. G. and Mackey, M. C. (2000) Neural ensemble coding and statistical periodicity: speculations on the operation of the mind's eye. *J. Physiol. Paris* 94, 489–503.
- Picton, T. W. and Stuss, D. T. (1994) Neurobiology of conscious experience. *Curr. Opin. Neurobiol.* 4, 256–265.
- Poeppel, D. (1996) A critical review of PET studies of phonological processing. *Brain Lang.* 55, 317–351.
- Pouget, A., Dayan, P., and Zemel, R. (2000) Information processing with population codes. *Nat. Rev. Neurosci.* 1, 125–132.
- Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1992) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.
- Scannell, J. W., Burns, G. A., Hilgetag, C. C., O'Neil, M. A., and Young, M. P. (1999) The connectional organization of the cortico-thalamic system of the cat. *Cereb Cortex* 9, 277–299.
- Sporns, O., Tononi, G., and Edelman, G. M. (2000) Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Netw.* 13, 909–922.
- Squire, L. R. (1987) *Memory and Brain*. Oxford University Press, New York, NY.
- Stephan, K. E., Hilgetag, C. C., Burns, G. A., O'Neill, M. A., Young, M. P., and Kotter, R. (2000) Computational analysis of functional connectivity between areas of primate cerebral cortex. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 355, 111–126.
- Stephan, K. E., Marshall, J. C., Friston, K. J., Rowe, J. B., Ritzl, A., Zilles, K., and Fink, G. R. (2003) Lateralized cognitive processes and lateralized task control in the human brain. *Science* 301, 384–386.
- Tononi, G., Sporns, O., and Edelman, G. M. (1994) A measure of brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* 91, 5033–5037.
- Tononi, G., Sporns, O., and Edelman, G. M. (1996) A complexity measure for selective matching of signals by the brain. *Proc. Natl. Acad. Sci. USA* 93, 3422–3427.
- Tononi, G., Sporns, O., and Edelman, G. M. (1999) Measures of degeneracy and redundancy in biological networks. *Proc. Natl. Acad. Sci. USA* 96, 3257–3262.
- Wolpaw, J. and Lee, C. (1989) Memory traces in primate spinal cord produced by operant conditioning of H-reflex. *J. Neurosci.* 61, 563–572.
- Wolpaw, J. R. (1997) The complex structure of a simple memory. *Trends Neurosci.* 20, 588–594.
- Worgotter, F. and Eysel, U. T. (2000) Context, state and the receptive fields of striatal cortex cells. *Trends Neurosci.* 23, 497–503.
- Young, M. P. and Yamane, S. (1992) Sparse population coding of faces in the inferotemporal cortex. *Science* 256, 1327–1331.