



PERGAMON

Neural Networks 13 (2000) 861–870

Neural  
Networks

www.elsevier.com/locate/neunet

2000 Special Issue

## Towards a network theory of cognition

A.R. McIntosh\*

Rotman Research Institute — Baycrest Centre, Department of Psychology, University of Toronto, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1

Received 13 June 2000; accepted 23 June 2000

### Abstract

For cognitive neuroscience to go forward a more explicit effort is needed to use neurophysiology to constrain how the brain produces human mental functions. This review begins with the suggestion that two fundamental features may be critical for this effort. The first is the connectivity of the brain, which occupies an intermediate position between complete redundant interconnections and independence. The term *semiconnected* is presented as a designation, which is an obvious derivation of the term semiconductors as used in engineering. The second is *transient response plasticity* where a given neuron or collection of neurons may show rapid changes in response characteristics depending on experience. Response plasticity is a ubiquitous property of the brain rather than a unique characteristic of “neurocognitive” regions. These two properties may be brought together when brain areas interact such that their *aggregate* function embodies cognition. Three examples are used to illustrate these general principles and to develop the idea that a particular region in isolation may not act as a reliable index for a particular cognitive function. Instead, the *neural context* in which an area is active may define the cognitive function. Neural context emphasizes that the particular spatiotemporal pattern of neural interactions may hold the key to bridge between brain and mind. © 2000 Elsevier Science Ltd. All rights reserved.

*Keywords:* Cognitive neuroscience; Connectivity; Plasticity

Throughout the history of cognitive neuroscience, there has been an ongoing debate as to whether mental operations are localized to discrete brain regions or represented by distributed networks of brain areas. Modern neuroimaging tools, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), provide an opportunity to assess these alternatives because activity can be assessed across the entire human brain. The activity maps produced from neuroimaging studies hint that cognitive function is a distributed property. In fact, contemporary theoretical writings seem to agree that cognitive functions arise from the action of distributed networks in the brain (Baars, 1993; Friston, 1997a; Heinze, Matzke, Dorfmueller, & Smid, 1997; John, Easton, & Isenhardt, 1997; Mesulam, 1990; Mountcastle, 1979), yet few have clearly worked out how to use this idea to study the link between brain and cognition.

One possible reason for the difficulty in building the link is that many cognitive neuroscience investigations do not take neurophysiology into account when interpreting data. Inconsistencies in the identification of “neurocognitive” systems across studies may reflect a fundamental character-

istic of the nervous system rather than an error in experimental design (Poeppel, 1996). This problem is not a new one. Russian physiologists, like Pavlov, early in the 20th century noted that few attempts at mapping brain structure to function attempted to account for the physiology of the system. Bethe (1931) suggested that quite a different picture of the functional organization of the brain would emerge if the physiology of the nervous system were considered instead of the behaviour that emerges from it. A similar critique was put forth by Lashley (1933) and has seen some revival in recent works (Friston et al., 1996; Gaffan, 1996; Magistretti, Pellerin, Rothman, & Shulman, 1999; McIntosh, 1998).

If we were to take account of the properties of the nervous system in the study of cognition, what features need to be considered? The purpose of this paper is to suggest that two dominant features, connectivity and plasticity of response, are key to the understanding of how cognition results from the operations of the brain. The paper is structured by first discussing the physiological expression of these features and how they may underlying cognitive functions when the actions of several brain regions are combined through their interactions. It then goes on to illustrate how these features serve to guide analysis and interpretation of neuroimaging studies with three examples: one concerns additivity

\* Tel.: +1-416-785-2500, ext. 3522; fax: +1-416-785-2862.

E-mail address: mcintosh@psych.utoronto.ca (A.R. McIntosh).

of cognitive and behavioural functions, one concerns cross-modal sensory learning and the final concerns the relation of awareness and learning to cortical network operations. Common through all these examples is the theme of dynamic network operations where the functional relevance of regional involvement depends on the patterns of interactions within large-scale networks.

### 1. Connectivity maximizes flexibility of information processing and representation

One prominent feature of the brain is connectivity. Neurons are linked to one another both locally and at a distance. Most other systems in the body show some capacity for cell to cell communication, but the nervous system appears to be specialized for rapid transfer of signals. Physiologically, this means that a single change to the system is conveyed to several parts of the brain simultaneously and that some of this will feed back onto the initial site. There are obvious extremes to just how “connected” a system can be. The nervous system occupies some intermediate position in the spectrum from complete interconnectivity to complete independence. Local cell networks are highly interconnected, but not completely so, and this means that adjacent cells can have common and unique connections. The term *semiconnected* is used to designate this particular property of local cell networks. The networks themselves can be thought of as semiconnectors, especially in as much as their function, as discussed below, is not only to mediate the signal between different cerebral regions but also to modulate the signal, in keeping with the specific properties of different semiconnectors. In information-theoretic terms, local cell networks (semiconnectors) act as noisy communication channels.

One consequence of the semiconnectivity of a local network is a certain degree of redundancy of responses (“degeneracy” in Edelman’s terms (1978) — see also Tononi, Sporns, and Edelman (1999). The term ‘semiconnectivity’ is preferred to ‘degeneracy,’ because in biology the latter term implies reversion to a less highly organized or simpler state of affairs.) The concept of semiconnectivity allows, at the level of local circuits, for adjacent neurons to have the same response properties (e.g. orientation columns in primary visual cortex) whereas neurons slightly removed may possess overlapping, but not identical, response characteristics. Broad tuning curves are characteristic for most sensory system cells and cells in motor cortices (Bakin & Weinberger, 1990; Desimone & Schein, 1987; Georgopoulos, Schwartz, & Kettner, 1986). The width of the curves, or degree of overlap, is modifiable; I will return to this feature in the next section. The broad tuning curves result from semiconnectivity, where cells share some similar and some unique connections. The noise fluctuations inherent to the biological systems also shape responses (Collins, Chow, & Imhoff, 1995; Douglass, Wilkens, Pantazidou, &

Moss, 1993), but that is not critical to the point developed here.

Connections between local ensembles are more sparse than the intra-ensemble connectivity. Estimates of the connections in the primate cortical visual system suggest that somewhere between 30–40% of all possible connections between cortical areas exist (Felleman & Van Essen, 1991). Recent simulation studies show that this sparseness is a computation advantage for the nervous system in that it allows for a high degree of flexibility in responses at the system level (Friston, 1997b; Tononi, Sporns, & Edelman, 1992). A system that has reciprocal and sparse connections, like the brain, is able to integrate a great deal more information than a system that is completely interconnected or one where regions are arranged hierarchically.

### 2. Transient response plasticity in the CNS is ubiquitous

Neural plasticity is an established phenomenon. Following central or peripheral damage there is profound reorganization of the nervous system (Hubel & Wiesel, 1965; Merzenich et al., 1983; Pons et al., 1991). Reorganization also can be observed after prolonged training (Karni et al., 1995). The plasticity considered here is more short-lived. Cells can show a rapid shift in response to afferent stimulation that is dependent on the context in which they fire. This *transient response plasticity* occurs over a much shorter time-scale compared to recovery from damage. Transient plasticity is typically considered to be a property of higher-order brain regions only, which thus specializes them for cognition (Mesulam, 1998). However, physiological investigation has consistently shown transient plasticity in the earliest parts of the nervous system (Morrell, 1961). If an auditory stimulus acquires some meaning, auditory cortex cells will respond more vigorously to the tone (Weinberger & Diamond, 1987). Even cells that do not respond to the tone prior to learning become more responsive to the tone after learning. Tuning curves that previously peaked at a different frequency, shift towards that of the conditioned tone. Transient plasticity of responses in relation to learning and memory have been observed in several parts of the brain, from single cells in isolate spinal cord preparations (Wolpaw & Lee, 1989) to primary sensory and motor structures (Donoghue & Sanes, 1994; Recanzone, Schreiner, & Merzenich, 1992). The changes can occur within a few stimulus presentations (Edeline, Pham, & Weinberger, 1993; Molchan, Sunderland, McIntosh, Herscovitch, & Schreurs, 1994). Transient plasticity may be a ubiquitous property of the central nervous system (Wolpaw, 1997).

Neurophysiological studies show that the brain is specially designed to modify its responses depending on prior experience at all levels of organization. Thus, one feature rudimentary to cognitive operations, namely plasticity, can be observed in many parts of the brain. This poses a problem for the idea that what distinguishes a cognitive

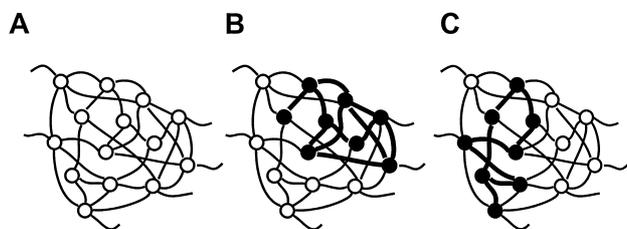


Fig. 1. Overlapping response properties of cell ensembles. Panel A shows the configuration with the anatomical connectivity between neuronal ensembles. Panel B shows the interactivity/activity pattern of the ensembles when processing one type of information and Panel C shows the interactivity/activity pattern of the ensembles processing a different type of information. The key is that some of the same ensembles may be involved in coding both sets of events, but the distinguishing feature is pattern of interactivity and not simply whether a particular ensemble is active. (Adapted from Sakurai, 1996).

from a non-cognitive brain area is plasticity of response (e.g. Mesulam, 1998). Taken further, it would also contradict the view that higher cognitive functions, such as memory, are mediated by specific systems in the brain that are dissociable from systems for sensory and motor functions. Cognitive function may instead be determined by how the properties of different regions are combined, or *aggregated*, through interregional interactions rather than by the involvement of any specific region.

### 3. Aggregate properties of neural populations

The idea of “neural aggregates” has some history and has been used to characterize brain theories that strike an intermediate position between strict localization of function and a holistic approach. Aggregate theories acknowledge that certain functions, usually sensory and motor, may be localizable, but higher-order function results from the combined, or aggregate, operations of several areas (Lashley, 1929). Recent examples of aggregate theories state that cognitive processes result from the integration of functionally specialized areas (Friston, 1997a; Tononi et al., 1992). There does not seem to be much disagreement in the cognitive neuroscience community about this statement. Most imaging researchers will make general statements about integrated functional systems (Fletcher et al., 1995; Haxby, Ungerleider, Horwitz, Maisog, & Grady, 1994; Kanwisher, McDermott, & Chun, 1997; Posner & Raichle, 1994; Shulman et al., 1997). But, for the most part, the results tend to then be discussed in a manner more consistent with localization by attempting to attribute a cognitive function to activated areas rather than the aggregate.

An aggregate function results from some combination of subordinate elements. A symphony provides a good illustration. When a collection of musical instruments is played in a coordinated manner, the aggregate function is the symphony that flows and has a recognizable structure and coherence. Each instrument of the orchestra plays a particular part of

the symphony, and elements of the symphony are detectable in the isolated instruments. However, the aggregate is much richer than a particular element. This same relation is likely to be in operation in the nervous system. Single neurons may show distinct patterns of activity that separately may not seem to represent anything cohesive. If the examination is moved to the level of small neuronal groups, responses reflecting fundamental properties such as learning, memory, and attention emerge, albeit with a limited repertoire. Activity patterns that seem to be consistent with learning and memory functions can be observed in the hippocampus (Wilson, Riches, & Brown, 1990; Wilson & McNaughton, 1994) or prefrontal cortex (Fuster, 1995; Goldman-Rakic, 1990; Watanabe, 1992), although the complete expression of this operation requires the contribution of other connected areas. Attentional modulation can be observed in visual cortices, but the modulatory effect likely comes from the influence of other brain regions (Buchel & Friston, 1997). Because neural ensembles can communicate, the combined action of a neural population of several thousands, or millions, results in a coherent pattern that “represents” features of the environment. Moreover, the population will have a richer repertoire since subtle changes in the spatiotemporal pattern would signal a different representation. Returning to the symphony allusion, if the instruments played the identical parts, but in a different order, the nature of the symphony changes. So too would be the case for the nervous system, which emphasizes the temporal dependency of neurocognitive function.

An aggregate property should not be confused with an emergent property. Emergent properties are those that are not evident in the subordinate constituents. For example, temperature is an emergent feature of molecule movement, but a given molecule alone does not have a temperature. For an aggregate, the constituents will show some features of the larger aggregate.

Finally, some symphonies will emphasize particular instruments in solos or as the primary part of a work. A violin, for instance, may play a lead in one orchestral piece, but in another it serves to set the “tone” of the work. In isolation, the sound of the instrument has not changed, but in the *context* of the entire orchestra, that same instrument may play several different roles depending on the accompaniment of other instruments. For the brain, the role of the region in cognition is determined by how it interacts with other regions—the *neural context*. The neural context is a direct result of the organizational features of the nervous system. Fig. 1 demonstrates this latter feature where cell ensembles that respond to one feature overlap with ensembles that code another.

There is a critical point to be made here with regard to neural context and aggregate functions. The allusion of symphony is only so useful because of the tendency to assume that there must be a conductor to guide the development of the symphony. *There is no conductor in the brain*. Unlike the orchestra, a given pattern of neural

interactions may be initiated from several different parts of a constituent system. A collection of neural ensembles that depict a memory trace, for example, may be activated through an external stimulus that evokes the memory trace or through internally initiated activity that similarly engages the trace. Such a property avoids the difficult situation of postulating a “neural conductor” or “executive” that controls the unfolding of cognition. The temporal and spatial unfolding of cognition results from the dynamic interactions among several areas of the brain and may be initiated from many different points (e.g. Hebb, 1949).

Electrophysiological work in motor and sensory cortices has provided some examples of aggregate operations. The aggregate functions increase the sensitivity and flexibility of cell populations over that of the individual neurons or cell groups (Georgopoulos, Taira, & Lukashin, 1993; Young, 1992). Individual cells have a limited, or non-specific, set of responses. When combined into cell populations a vast capability for distinguishing events ensues that is not obvious when each cell in the population is examined independently (Lindsey, Morris, Shannon, & Gerstein, 1997).

If perceptual and motor operations in the nervous system are best represented as aggregate features of neural population action, it is very likely cognitive phenomena also are aggregate features. Functional neuroimaging has shown that several areas are activated when the brain is performing a task that puts a great load on cognitive processes (Cabeza et al., 1997; Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991). There is a temptation to suggest that the activated areas form some sort of cognitive system apart from basic sensory or motor functions. A different perspective builds on the capabilities of groups of neurons. Population coding observed for sensory and motor functions can be extended to higher-order cognitive functions. However, instead of population actions at a regional level, the population activity from several regions would need to combine or interact. When neural populations interact with one another, their rudimentary functions combine and the aggregate is the cognitive processes. Cognitive operations are not localized in an area or network of regions, but are the consequence of dynamic network interactions that depend on the processing demands for a particular operation (Duncan, Humphreys, & Ward, 1997). Simply put, if a process involves the integration of visual and auditory information with a vocal output, regions that will be involved will be auditory and visual and motor regions that can control the vocal apparatus. This seems a bit mundane, but it pulls the focus away from the search for specialized speech regions to examining brain networks based on their anatomically-constrained interactions. The anatomy puts the limits on the repertoire a region may have (Gaffan, 1996; Mountcastle, 1979) — a sort of anatomical determinism. Within these limits, response plasticity of the cell and hence the ensemble means that the population has the potential to modify its response characteristics depending on afferent information. Since the ensemble response changes, so too will its efferent

influence and thus the plastic effects are carried to several other areas. The same collection of areas may be engaged across a variety of cognitive operations because of the semi-connectedness of regions and transient plasticity. This flexibility of function that defines the basic response properties of the central nervous system also defines human cognition.

### 3.1. Methodological note

The final part of this review presents empirical examples of neural interactions in cognition derived from functional imaging work. The illustrations rely on different types of multivariate analyses where the interregional covariances of activity are used to identify the dominant functional and/or effective connections (Aertsen, Gerstein, Habib, & Palm, 1989; Friston, Frith, & Fracowiak, 1993; Gerstein, Perkel, & Subramanian, 1978) during the performance of a cognitive operation. *Functional connectivity* is a statement that two regions show some non-zero correlation of activity, but does not specify how this correlation comes about. *Effective connectivity* is a statement about the direct effect one region has on another, accounting for mutual or intervening influences. There are several review and technical papers on methods to measure functional and effective connectivity in neuroimaging to which the interested reader is referred (Friston, 1994; Horwitz, Soncrant, & Haxby, 1992; McIntosh & Gonzalez-Lima, 1994). The important point is that the methods do not focus on activity differences but on the relation of activity between areas. Because the goal is to quantify and understand neural interactions, the methods of analysis must focus on the relations between neural elements. The exploration of functional and effective connectivity has been labeled *network analysis* (McIntosh & Gonzalez-Lima, 1994; McIntosh et al., 1994).

### 3.2. Semantic processing and response mode

Most neuroimaging studies have been designed using the hierarchical subtraction, or cognitive subtraction, procedure. An often cited example of this approach is the verb-generation experiments performed by Petersen, Fox, Posner, Mintun, and Raichle (1988). One task had subjects view a single noun and read it aloud while the verb-generate task had subjects generate a verb that matched the presented noun. Subtracting the reading task from the verb-generate task would isolate those regions more active when semantic processing of the word was required. The subtraction paradigm has been often criticized (Friston et al., 1996; Sergent, Zuck, Levesque, & MacDonald, 1992), but it has yielded some consistent results (Cabeza & Nyberg, 1997; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994).

We sought to examine the additivity of behavioural processes using a rather simple design (Jennings, McIntosh, Kapur, Tulving, & Houle, 1997). PET regional cerebral blood flow (rCBF) measures were taken of the same two cognitive processes, but under three different means of responding. Subjects carried out a semantic processing

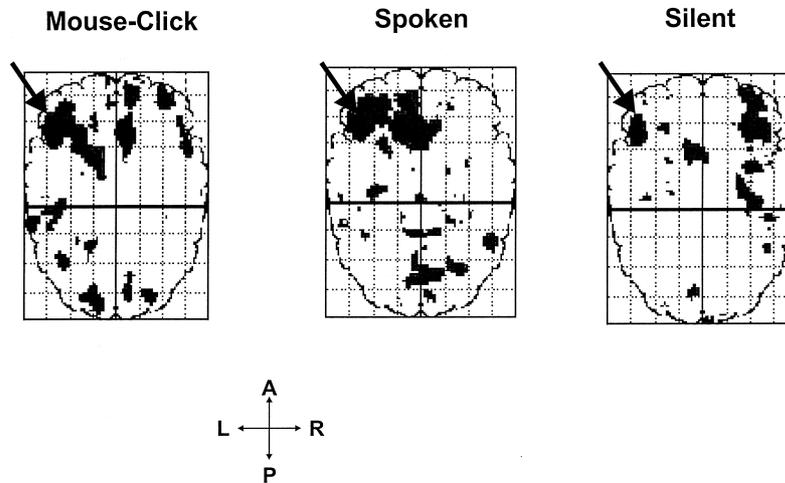


Fig. 2. Statistical parametric maps showing areas that were significantly more active during semantic processing than letter identification within each of the three response modes. If there was no effect of response mode on these differences, the maps should be identical for all three response modes. The maps are maximum intensity projections as would be viewed from top of a “glass” brain. The orientation (left, right, anterior, posterior) is indicated at the bottom of the figure.

task and a letter scanning control task. In all tasks, subjects were presented visually with concrete nouns. During the semantic processing task, subjects judged whether the noun was a living thing. During the letter scanning task

the subjects judged whether the word contained the letter ‘a’. In separate scans, subjects responded by pressing a mouse button (mouse-click), or by speaking (spoken), or kept the response to themselves (silent). In terms of basic experimental design, this was a two-by-three crossed design. One could question whether the difference between the semantic and letter task represents a single process, but a subtraction of the two tasks within each of the response modes would be expected to yield the same pattern of differences. There is no reason to expect the manner of response to affect the operations of the putative neurocognitive systems that support semantic processing and letter identification.

Pairwise comparisons of semantic and letter tasks within each response mode showed some common results, but there were striking differences. The statistical images from the pairwise subtractions (Fig. 2) show that activity in left inferior prefrontal cortex (indicated by arrows) was greater in the semantic task for all three response conditions. However, the magnitude of the left prefrontal effect was strongest for the mouse-click condition and weakest for the silent condition. Anterior cingulate was most strongly active in the mouse-click and spoken response tasks. Unique areas of activation also are obvious from the figure, where the semantic processing with a spoken response activate more of anterior cingulate and left prefrontal cortex, while the mouse-click and silent condition both identified right prefrontal cortex activation for semantic processing.

This outcome poses somewhat of a problem for the mapping of cognitive functions. It implies that a functional map would have to include the manner of responding, so that a “memory retrieval map” would become a “memory retrieval map when a spoken response is used”. When the organization of the brain outlined above is taken into account, the result in Fig. 2 is not unexpected. The connections between areas allow for feedback effects to modulate

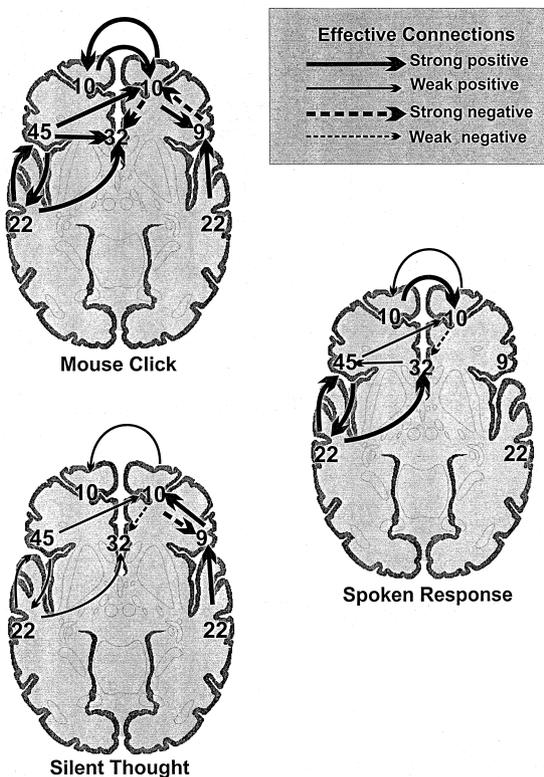


Fig. 3. Summary of effective connections for the semantic processing task within each of the three response modes. The dominant effects are shown (see Jennings et al., 1998 for complete functional models). Arrows represent the effect one region has on another with the size and sign of the effect indicated in the legend.

### Tone-elicited Effective Connections

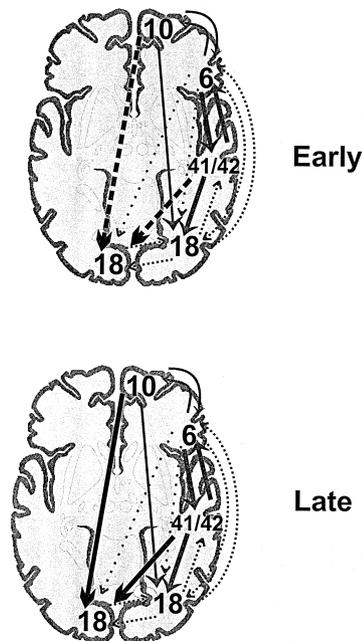


Fig. 4. Functional networks from early and late phases of training in a sensory associative learning task. The effective connections are in response to a tone that was associated with a visual stimulus. Only the effects from BA 10 and BA 41/42 on left BA 18 changed significantly with learning. (Adapted from McIntosh et al., 1998). Line thickness are interpreted according to the legend in Fig. 3.

activity patterns. Even if the same area, or collection of areas, makes a semantic decision, because these areas will have some unique links to regions more allied with response, feedback from the response-related areas may result in a unique activity pattern producing a unique map for semantic processing when making a spoken response. If this is true, it should be possible to identify common set of regions that interact similarly among themselves during semantic processing, but may engage in some unique interactions depending on response (Jennings, McIntosh, & Kapur, 1998).

By examining effective connectivity between several of the areas identified in Fig. 2, there did appear to be a “core” network for semantic processing that modulated its interactions in relation to response. The core network consisted of the influences between left Brodmann Area (BA) 22 and left BA 45; left BA 22 and BA 32; left BA 45 and right BA 10; right BA 10 and left BA 10; and right BA 10 and BA 32. As with the subtractions, the interactions within the core network were strongest for the mouse-click condition and weakest for the silent task (Fig. 3). Core regions also showed unique interactions for each response mode. Anterior cingulate (BA 32) showed the strongest interactions when an overt response was required. Right prefrontal interactions were suppressed during the spoken-response condition. Dorsal right prefrontal cortices showed similar activity patterns in the mouse-click and silent conditions, but the

interaction patterns for these regions differed between mouse-click and silent response tasks. Thus, the same area may show similar patterns of activity between conditions, but different network interactions.

### 3.3. Sensory associative learning

One example of the use of network analysis to test specific hypotheses comes from a study of sensory learning (McIntosh, Cabeza, & Lobaugh, 1998). The task had subjects learn an association between a tone and a visual stimulus. We measured brain activity in response to the tone by itself across acquisition of the association. The expectation was that as the tone acquired behavioural significance, presentation of the tone would elicit activity in visual areas. Since the activation of visual areas would occur without overt visual stimulation, the second hypothesis was that this activation would be mediated through effects from higher-order cortical areas, likely posterior association or prefrontal cortices. By using an analysis of functional connectivity, we identified four areas that could have influenced the activity of these visual areas. Then effective connections were quantified to determine which of these candidates exerted the strongest influence on these visual areas.

Two areas in particular seem to exert the dominant influence on the visual area as the association was learned. Superior temporal cortex (auditory association, BA 41/42) and prefrontal cortex near BA 10 (Fig. 4) both changed their effect on visual cortex from suppressive to facilitatory as the association was learned. This outcome emphasizes the need to explicitly quantify effective connections. The regions in the functional network were selected because they all showed a similar pattern of functional connections with visual cortex. As stated earlier, functional connectivity is a statement that two or more regions are functionally linked without any attempt to specify whether the link is from direct interactions or mediated by a third region. All regions included in the functional network have the anatomical potential for mediating the observed change in visual cortex activity, but when the effective connections were estimated, only two of the areas showed a significant change in effects.

The involvement of prefrontal cortex in this simple task may seem surprising since there is no strong cognitive component. From subject debriefing, there was no obvious cognitive mediation that subjects could verbalize, i.e. subjects were not aware of the relation between stimuli. Considering the anatomical connectivity, this result is not inconsistent with the configuration of the system. It is entirely possible for the same region to be engaged across several tasks because of different patterns of neural interactions.

### 3.4. Awareness and prefrontal cortex interactions

As a follow-up to the simple associative learning study, we sought to investigate further the neural interactions

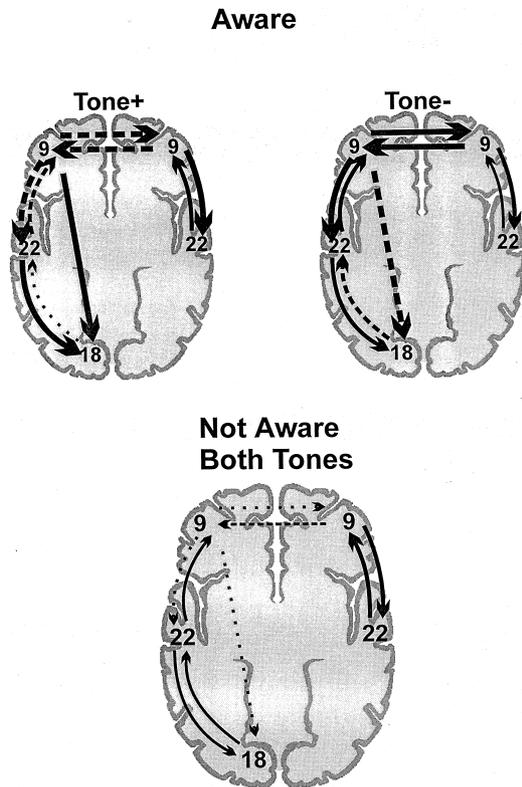


Fig. 5. Functional networks from late phases of training in a differential sensory conditioning task. Networks from two groups are shown. Aware subjects showed strong difference in effective connections involving left prefrontal area 9 and other regions that distinguished between the two tones. Conversely, the network for the Unaware subjects did not differ between tones and showed no strong left prefrontal involvement. Line thickness may be interpreted according to the legend in Fig. 3.

subserving cross-modal learning using differential conditioning (McIntosh, Rajah, & Lobaugh, 1999). In this study, two tones were used having differential relations to the visual stimuli. One tone was a strong predictor of the presentation of a visual stimulus (Tone +), and the other tone a weak predictor (Tone -). In this PET rCBF study, scans alternated between isolated presentations of the Tone + and Tone - as subject's learned. Much to our surprise, subjects in our sample divided perfectly in half into those who were aware of the stimulus associations and those who were unaware. The index of awareness came from debriefing questionnaires. Furthermore, only Aware subjects learned the differentiation between the tones, while Unaware subjects showed no behavioural evidence of learning.

In examining the underlying brain activity that supported learning and awareness, we observed that the strongest group difference in brain activity elicited by the tones was in left prefrontal cortex (LPFC) near BA 9. In aware subjects, LPFC activity showed progressively greater activity to Tone - than Tone +. Ventral and medial occipital cortices ( $X = 22, Y = -86, Z = -20$ ;  $X = -6, Y = -88, Z = 0$ ) and right thalamus ( $X = 16, Y = -22, Z = 16$ )

showed progressively greater activity to Tone + than to Tone -. In Unaware subjects, no consistent changes were seen in LPFC or in any of the other regions. At first, these results seem to confirm the prominent role of PFC in monitoring functions (Burgess & Shallice, 1996; Stuss & Benson, 1987), and especially its putative role in awareness (Knight, Grabowecky, & Scabini, 1995; Petrides, 1985). However, PFC activation also has been found in tasks where there was no overt awareness, such as in the previous sensory learning task, and in implicit novelty assessment (Berns, Cohen, & Mintun, 1997). It was thus possible that interactions of PFC with other brain regions, present in Aware but not in Unaware subjects, would better describe the neural system underlying awareness in this task.

When the interactions of LPFC were assessed between the two groups, we observed a remarkable difference in the strength and pattern of functional connections among several brain areas including right PFC, bilateral superior temporal cortices (auditory association), occipital cortex, and medial cerebellum. These areas were much more strongly correlated in Aware, than Unaware, subjects. To explore some of the network interactions within an anatomical reference, structural equation models were constructed for a subset of regions identified in the PLS analysis. The models were constructed for the last two pairs of tone scans since this corresponded to the greatest change in neural interactions and behaviour. With the small number of subjects relative to the number of peak voxels, we selected four regions to model, in addition to the LPFC voxel. Given the task involved auditory and visual stimuli, it seemed reasonable that structures associated with these modalities would be candidates for a model. The right PFC area also was included.

As may be expected from the correlations, there were significant changes in the effective connections for Aware subjects, including robust interactions involving LPFC (Fig. 5). During the last Tone + scan, feedback to occipital cortex was positive from temporal and prefrontal cortex, which may reflect implicit and explicit expectancy of the upcoming visual discrimination (McIntosh et al., 1998). In the Tone - scan, this feedback switched to negative, which could reflect the knowledge that there would be no visual event following presentation of the Tone -. The functional network for Unaware differed from Aware subjects, but there were no significant changes in effective connections across experiment for the Unaware group. There were non-zero interactions in the functional network, but the involvement of LPFC was weak. This confirms that LPFC was not interacting systematically across subjects in the Unaware group.

The integration among areas in this study underscores the importance of examining neurocognitive organization in terms of large-scale networks rather than discrete loci. The effective connections between areas in the functional networks suggested reliable changes in reciprocal influences of LPFC on posterior cortices. Such reciprocity in the

functional interactions emphasizes the cooperative nature of brain operations subserving different aspects of cognition. While awareness of stimulus associations may require the integrity of prefrontal cortex, the neurobiological process that underlies awareness arises from the interactions of prefrontal and posterior regions. More generally, emerging data emphasize that the same area may be equally engaged across different tasks or groups, but serve different cognitive functions that are determined the interactions with other parts of the brain — the *neural context* (D'Esposito, Ballard, Aguirre, & Zarahn, 1998; McIntosh, 1999; McIntosh, Grady, Haxby, Ungerleider, & Horwitz, 1996; Rajah, Lobaugh, Danjoux, & McIntosh, 1999).

#### 4. Neural context

The idea of a neural context is meant to underscore the importance of considering activity of the entire brain rather than individual regions. Activity of the area may be equivalent across several seemingly different cognitive tasks. What distinguishes tasks is the pattern of spatiotemporal activity and interactivity more than the participation of any particular region. Through its connectivity and basic response properties, large neural systems are engaged starting from the sensory systems, through the so-called association areas, then to the primary motor systems. As an organism moves through its environment, these basic input–output functions occur continuously. Where they differ depends on the specifics of the sensory events, whether these events have been encountered before, how the organism responds, and its consequences. It is easy to dismiss this account as too mechanistic or behaviourist in the tradition of psychologists like Watson or Skinner. It is not the intent to reduce cognition to a series of Stimulus–Response chains, but rather to point out that the nervous system is organized around this basic associative link. Sensory systems, association systems, and motor systems can all impact upon one another through the reciprocal connections. Because of this, the constituents for seemingly related functions can change dramatically, and it is also possible that similar constituents can be engaged across seemingly disparate functions. This reflects the dynamic and adaptive feature of the nervous system.

In some ways, this position would make the enterprise of neuroimaging rather unrewarding if there were no specificity to the patterns of neural activity. Clearly, the replicability of certain activity patterns suggest this is not true (Lepage, Habib, & Tulving, 1998; Nyberg, Cabeza, & Tulving, 1996; Tulving et al., 1994). There are some immediate constraints to what parts of the brain may participate in cognitive functions: the primary one, as stated before, is the anatomy. The anatomy acts as both the catalyst and the constraint on neurocognitive systems (Gaffan, 1996). It is rather unlikely that the nuclei of the brainstem mediate spinal reflexes and autonomic functions are

recruited for cognitive functions. This is not because the area do not possess the necessary response plasticity, but rather that they are not anatomically related to areas that are, at present, likely candidates for neurocognitive systems. It is no coincidence that the areas most typically related to cognitive functions are those that appear to have a diverse set of afferents that come from several sensory modalities, such as the frontal lobes and hippocampus. This does not mean that cognition cannot begin until these areas are engaged. It means that compared to regions allied to sensory or motor functions, frontal cortices and hippocampus may be involved in several more neurocognitive networks. It is at this point that the overlap between specialization and integration comes into play. Regional specialization is, in part, determined by the connectivity of the area. But the functional relevance of that area cannot be realized unless it operates in conjunction with other parts of the brain.

#### 5. Implications for cognition

Neurophysiology and cognitive psychology have developed independently across the centuries. There is no necessity for the two disciplines to influence each other. Neurophysiology has evolved partly from study of the basic properties of neurons without much link to overt behaviour beyond simple reflexes. Conversely, cognitive psychology has developed through careful experimental investigation of overt behaviour and how the manipulations of putative cognitive processes change the measured behaviour. It is perfectly feasible for cognitive psychology to continue its theoretical development without any link to neurophysiology and for neurophysiology to continue with no attempt to understand the biology of cognition (for an interesting debate around this issue see Kosslyn & Intriligator, 1992 and Caramazza, 1992). When the two are brought together, there will likely to be some important implications for theories of cognition that will arise from its neuroscientific study. The examples presented here only begin to reveal this potential. If systems that are engaged in cognitive functions interact directly with systems allied with response coordination, this underscores the non-hierarchical arrangement of neurocognitive systems. If we find that similar brain regions can serve different cognitive functions, this would imply linkages between the cognitive functions. For example, prefrontal cortex involvement may span from simple perceptual and associative learning functions to higher-order cognitive functions, which may suggest a behavioural interdependency that is not easily appreciated from behavioural study. While it is convenient to consider attention, memory, language, and perception to be the domain of independent neural systems, the physiology of the brain would suggest strong, if not complete, overlap of these operations. Admittedly, not all psychological studies treat these processes as independent (e.g. Duncan et al., 1997; McClelland, 1979; Miller, Galanter, & Pribram,

1960). An explicit attempt to incorporate the physiology into psychological theories can only result in a richer understanding of human cognitive function.

## Acknowledgements

The contributions of several of my colleagues to the development of the ideas expressed here are gratefully acknowledged: Drs B. Horwitz, JM Jennings, E Tulving, L Nyberg, R Cabeza, and NJ Lobaugh. AR McIntosh is supported by the Natural Sciences and Engineering Council of Canada and the Medical Research Council of Canada (grant MT-13623).

## References

- Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of effective connectivity. *Journal of Neurophysiology*, *61*, 900–917.
- Baars, B. J. (1993). How does a serial, integrated and very limited stream of consciousness emerge from a nervous system that is mostly unconscious, distributed, parallel and of enormous capacity? *Ciba Foundation Symposium*, *174*, 282–290.
- Bakin, J. S., & Weinberger, N. M. (1990). Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. *Brain Research*, *536*, 271–286.
- Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science*, *276* (5316), 1272–1275.
- Bethe, A. (1931). *Plasticity and localization theory*, *Handbuch der normalen und pathologischen Physiologie*, Vol. 15 (pp. 1175–1220).
- Buchel, C., & Friston, K. (1997). Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modeling and fMRI. *Cerebral Cortex*, *7* (8), 768–778.
- Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34* (4), 263–272.
- Cabeza, R., Kapur, S., Craik, F. I. M., McIntosh, A. R., Houle, S., & Tulving, E. (1997). Functional neuroanatomy of recall and recognition: a PET study of episodic memory. *Journal of Cognitive Neuroscience*, *9* (2), 254–265.
- Cabeza, R., & Nyberg, L. (1997). Imaging cognition: an empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience*, *9* (1), 1–26.
- Caramazza, A. (1992). Is cognitive neuropsychology plausible? *Journal of Cognitive Neuroscience*, *4* (1), 80–95.
- Collins, J. J., Chow, C. C., & Imhoff, T. T. (1995). Stochastic resonance without tuning. *Nature*, *376* (6537), 236–238.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *Journal of Neuroscience*, *11*, 2383–2402.
- Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *Journal of Neurophysiology*, *57*, 835–868.
- D'Esposito, M., Ballard, D., Aguirre, G. K., & Zarahn, E. (1998). Human prefrontal cortex is not specific for working memory: a functional MRI study. *NeuroImage*, *8* (3), 274–282.
- Donoghue, J. P., & Sanes, J. N. (1994). Motor areas of cerebral cortex. *Journal of Clinical Neurophysiology*, *11* (4), 382–396.
- Douglass, J. K., Wilkens, L., Pantazidou, E., & Moss, F. (1993). Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature*, *365*, 337–340.
- Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, *7* (2), 255–261.
- Edeline, J. M., Pham, P., & Weinberger, N. M. (1993). Rapid development of learning-induced receptive field plasticity in the auditory cortex. *Behavioral Neuroscience*, *107* (4), 539–551.
- Edelman, G. M. (1978). Group selection and phasic re-entrant signalling: a theory of higher brain function. In V. Mountcastle & G. M. Edelman, *The mindful brain* (pp. 55–100). Cambridge, MA: MIT Press.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1* (1), 1–47.
- Fletcher, P. C., Frith, C. D., Grasby, P. M., Shallice, T., Frackowiak, R. S. J., & Dolan, R. J. (1995). Brain systems for encoding and retrieval of auditory-verbal memory: an in vivo study in humans. *Brain*, *118*, 401–416.
- Friston, K. (1994). Functional and effective connectivity: A synthesis. *Human Brain Mapping*, *2* (1/2), 56–78.
- Friston, K., Price, C., Fletcher, P., Moore, C., Fracowiak, R., & Dolan, R. (1996). The trouble with cognitive subtraction. *NeuroImage*, *4*, 97–104.
- Friston, K. J. (1997aa). Imaging cognitive anatomy. *Trends in Cognitive Sciences*, *1* (1), 21–27.
- Friston, K. J. (1997bb). Transients, metastability, and neuronal dynamics. *NeuroImage*, *5* (2), 164–171.
- Friston, K. J., Frith, C., & Fracowiak, R. (1993). Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping*, *1*, 69–79.
- Fuster, J. M. (1995). Memory and planning: two temporal perspectives of frontal lobe function. In H. H. Jasper, S. Riggio & P. S. Goldman-Rakic, *Epilepsy and the functional anatomy of the frontal lobe* (pp. 9–20). New York: Raven Press.
- Gaffan, D. (1996). Associative and perceptual learning and the concept of memory systems. *Cognitive Brain Research*, *5* (1-2), 69–80.
- Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233* (4771), 1416–1419.
- Georgopoulos, A. P., Taira, M., & Lukashin, A. (1993). Cognitive neurophysiology of the motor cortex. *Science*, *260*, 47–52.
- Gerstein, G. L., Perkel, D. H., & Subramanian, K. N. (1978). Identification of functionally related neural assemblies. *Brain Research*, *140*, 43–62.
- Goldman-Rakic, P. S. (1990). Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. In H. B. M. Uylings, C. G. Van Eden, J. P. C. De Bruin, M. A. Corner & M. G. P. Feenstra, *Progress in brain research* (pp. 325–336), Vol. 85. Amsterdam: Elsevier Science.
- Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., & Grady, C. L. (1994). Neural systems for encoding and retrieving new long-term memories: a PET-rCBF study. *Investigations in Ophthalmic Vision Science*, *35*, 1813.
- Hebb, D. O. (1949). *The organization of behavior*, New York: Wiley.
- Heinze, H. J., Matzke, M., Dorfmueller, G., & Smid, H. G. (1997). Flexibility in the structure of human information processing. *Advances in Neurology*, *73*, 359–375.
- Horwitz, B., Soncrant, T. T., & Haxby, J. V. (1992). Covariance analysis of functional interactions in the brain using metabolic and blood flow data. In F. Gonzalez-Lima, T. Finkenstaedt & H. Scheich, *Advances in metabolic mapping techniques for brain imaging of behavioral and learning functions* (pp. 189–217). Dordrecht: Kluwer Academic.
- Hubel, D. H., & Wiesel, T. N. (1965). Binocular interaction in striate cortex of kittens reared with artificial squint. *Journal of Neurophysiology*, *28* (6), 1041–1059.
- Jennings, J. M., McIntosh, A. R., & Kapur, S. (1998). Mapping neural interactivity onto regional activity: an analysis of semantic processing and response mode interactions. *NeuroImage*, *7*, 244–254.
- Jennings, J. M., McIntosh, A. R., Kapur, S., Tulving, E., & Houle, S. (1997). Cognitive subtractions may not add up: the interaction between semantic processing and response mode. *NeuroImage*, *5*, 229–239.

- John, E. R., Easton, P., & Isenhardt, R. (1997). Consciousness and cognition may be mediated by multiple independent coherent ensembles. *Conscious Cognition*, 6 (1), 3–39 (discussion 40-1, 50-5, 65-6).
- Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17 (11), 4302–4311.
- Karni, A., Meyer, G., Jezard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, 377 (6545), 155–158.
- Knight, R. T., Grabowecy, M. F., & Scabini, D. (1995). Role of human prefrontal cortex in attention control. *Advances in Neurology*, 66, 21–34.
- Kosslyn, S. M., & Intriligator, J. M. (1992). Is cognitive neuropsychology plausible? The perils of sitting on a one-legged stool. *Journal of Cognitive Neuroscience*, 4, 96–106.
- Lashley, K. S. (1929). *Brain mechanisms and intelligence*, New York: Hafner Publishing.
- Lashley, K. S. (1933). Intergrative functions of the cerebral cortex. *Physiological Reviews*, 13 (1), 1–42.
- Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET activations of memory encoding and retrieval: the HIPER model. *Hippocampus*, 8 (4), 313–322.
- Lindsey, B. G., Morris, K. F., Shannon, R., & Gerstein, G. L. (1997). Repeated patterns of distributed synchrony in neuronal assemblies. *Journal of Neurophysiology*, 78, 1714–1719.
- Magistretti, P. J., Pellerin, L., Rothman, D. L., & Shulman, R. G. (1999). Energy on demand. *Science*, 283 (5401), 496–497.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.
- McIntosh, A. R. (1998). Understanding neural interactions in learning and memory using functional neuroimaging. *Annals of the New York Academy of Sciences*, 855, 556–571.
- McIntosh, A. R. (1999). Mapping cognition to the brain through neural interactions. *Memory*, 7 (5/6), 523–548.
- McIntosh, A. R., Cabeza, R. E., & Lobaugh, N. J. (1998). Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus. *Journal of Neurophysiology*, 80, 2790–2796.
- McIntosh, A. R., & Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, 2 (1/2), 2–22.
- McIntosh, A. R., Grady, C. L., Haxby, J. V., Ungerleider, L. G., & Horwitz, B. (1996). Changes in limbic and prefrontal functional interactions in a working memory task for faces. *Cerebral Cortex*, 6, 571–584.
- McIntosh, A. R., Grady, C. L., Ungerleider, L. G., Haxby, J. V., Rapoport, S. I., & Horwitz, B. (1994). Network analysis of cortical visual pathways mapped with PET. *Journal of Neuroscience*, 14, 655–666.
- McIntosh, A. R., Rajah, M. N., & Lobaugh, N. J. (1999). Interactions of prefrontal cortex related to awareness in sensory learning. *Science*, 284 (5419), 1531–1533.
- Merzenich, M. M., Kaas, J. H., Wall, J. T., Sur, M., Nelson, R. J., & Felleman, D. J. (1983). Progression of change following median nerve section in the cortical representation of the hand in areas 3b and 1 in adult owl and squirrel monkeys. *Neuroscience*, 10, 639–665.
- Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, 28 (5), 597–613.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain*, 121 ((Pt 6)), 1013–1052.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*, New York: Holt, Rinehart, and Winston.
- Molchan, S. E., Sunderland, T., McIntosh, A. R., Herscovitch, P., & Schreurs, B. G. (1994). A functional anatomical study of associative learning in humans. *Proceedings of the National Academy of Science USA*, 91, 8122–8126.
- Morrell, F. (1961). Electrophysiological contributions to the neural basis of learning. *Physiological Reviews*, 41, 443–494.
- Mountcastle, V. B. (1979). An organizing principle for cerebral function: The unit module and the distributed system. In V. B. Mountcastle & G. M. Edelman, [*The mindful brain*] (pp. 7–50). Boston: MIT Press.
- Nyberg, L., Cabeza, R., & Tulving, E. (1996). PET studies of encoding and retrieval: the HERA model. *Psychonomic Bulletin and Review*, 3, 135–148.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331, 585–589.
- Petrides, M. (1985). Deficits in associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 23, 601–614.
- Poeppl, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language*, 55 (3), 317–351.
- Pons, T., Garraghty, P. E., & Ommaya, A. K., et al. (1991). Massive cortical reorganization after sensory deafferentation in adult macaques. *Science*, 252, 1857–1860.
- Posner, M. I., & Raichle, M. E. (1994). *Images of mind*, New York: Scientific American Library.
- Rajah, M. N., Lobaugh, N. J., Danjoux, N., & McIntosh, A. R. (1999). Brain-behavior interactions during acquisition and reversal of sensory associations in humans. *Society for Neuroscience Abstracts*, 25 (2), 2068.
- Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1992). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience*, 13 (1), 87–103.
- Sakurai, Y. (1996). Population coding by cell assemblies — what it really is in the brain. *Neuroscience Research*, 26 (1), 1–16.
- Sergent, J., Zuck, E., Levesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex*, 2, 68–80.
- Shulman, G., Corbetta, M., Buckner, R., Raichle, M., Fiez, J., Miezin, F., & Petersen, S. (1997). Top-down modulation of early sensory cortex. *Cerebral Cortex*, 7 (3), 193–206.
- Stuss, D. T., & Benson, D. F. (1987). The frontal lobes and control of cognition and memory. In E. Perecman, *The frontal lobes revisited* (pp. 141–158). New York: The IRBN Press.
- Tononi, G., Sporns, O., & Edelman, G. (1992). A measure of brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Science USA*, 91, 5033–5037.
- Tononi, G., Sporns, O., & Edelman, G. M. (1999). Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Science USA*, 96 (6), 3257–3262.
- Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proceedings of the National Academy of Science USA*, 91 (6), 2016–2020.
- Watanabe, M. (1992). Frontal units of the monkey coding the associative significance of visual and auditory stimuli. *Experimental Brain Research*, 89, 233–247.
- Weinberger, N. M., & Diamond, D. M. (1987). Physiological plasticity in auditory cortex: rapid induction by learning. *Progress in Neurobiology*, 29 (1), 1–55.
- Wilson, F. A., Riches, I. P., & Brown, M. W. (1990). Hippocampus and medial temporal cortex: neuronal activity related to behavioural responses during the performance of memory tasks by primates. *Behavioral Brain Research*, 40 (1), 7–28.
- Wilson, M., & McNaughton, B. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, 256, 676–679.
- Wolpaw, J., & Lee, C. (1989). Memory traces in primate spinal cord produced by operant conditioning of H-reflex. *Journal of Neuroscience*, 61 (3), 563–572.
- Wolpaw, J. R. (1997). The complex structure of a simple memory. *Trends in Neuroscience*, 20 (12), 588–594.
- Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, 256, 1327–1331.