

Mapping Cognition to the Brain Through Neural Interactions

Anthony Randal McIntosh

University of Toronto, Canada

Brain imaging methods, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), provide a unique opportunity to study the neurobiology of human memory. As these methods can measure most of the brain, it is possible to examine the operations of large-scale neural systems and their relation to cognition. Two neuroimaging studies, one concerning working memory and the other episodic memory retrieval, serve as examples of application of two analytic methods that are optimised for the quantification of neural systems, structural equation modelling, and partial least squares. Structural equation modelling was used to explore shifting prefrontal and limbic interactions from the right to the left hemisphere in a delayed match-to-sample task for faces. A feature of the functional network for short delays was strong right hemisphere interactions between hippocampus, inferior prefrontal, and anterior cingulate cortices. At longer delays, these same three areas were strongly linked, but in the left hemisphere, which was interpreted as reflecting change in task strategy from perceptual to elaborate encoding with increasing delay. The primary manipulation in the memory retrieval study was different levels of retrieval success. The partial least squares method was used to determine whether the image-wide pattern of covariances of Brodmann areas 10 and 45/47 in right prefrontal cortex (RPFC) and the left hippocampus (LGH) could be mapped on to retrieval levels. Area 10 and LGH showed an opposite pattern of functional connectivity with a large expanse of bilateral limbic cortices that was equivalent for all levels of retrieval as well as the baseline task. However, only during high retrieval was area 45/47 included in this pattern. The results suggest that activity in portions of the RPFC can reflect either memory retrieval mode or retrieval success depending on other brain regions to which it is functionally linked, and imply that regional activity must be evaluated within the neural context in which it occurs. The general hypothesis that learning and memory are emergent properties of large-scale neural network interactions is

Requests for reprints should be sent to A.R. McIntosh, PhD, Rotman Research Institute of Baycrest Centre, 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada. Email: rmcintosh@rotman-baycrest.on.ca

The contributions of several of my colleagues to the development of the ideas expressed here are gratefully acknowledged: Drs B. Horwitz, C.L. Grady, E. Tulving, L. Nyberg, R. Cabeza, and N.J. Lobaugh. A.R. McIntosh is supported by the Natural Sciences and Engineering Council of Canada (grant OGP017034) and the Medical Research Council of Canada (grant MT-13623).

discussed, emphasising that a region can play a different role across many functions and that role is governed by its interactions with anatomically related regions.

INTRODUCTION

The past decade has seen tremendous advances in human cognitive neuroscience. Undoubtedly, part of the reason for this has been the explosion in the use of functional neuroimaging to examine the neural basis of cognition. Methods that image metabolic processes, such as positron emission tomography (PET) and most recently functional magnetic resonance imaging (fMRI), have provided wide access to the neural substrates of cognitive operations. The preferred method for determining how the brain represents cognition has been to use psychological theories of cognition to design studies for mapping the brain mechanisms. As the psychological models are by and large quite successful at explaining behaviour and guiding research, it seems reasonable that the brain areas supporting cognition should look something like these models. With the advent of the computer age, the prevailing metaphor is to view cognitive processes as a series of hierarchically organised computational modules (e.g. Baddeley, 1992; Fodor, 1983). The processes investigated are usually conceived as separable, and cognition is usually thought of as the additive result of these processes (although there are some notable exceptions, e.g. McClelland, 1979; Miller, Galanter, & Pribram, 1960).

When using these models to study the brain, it is important to keep in mind that they were developed without regard to the biology of the nervous system. Modularity in the nervous system is observed mainly in the sensory and motor systems, and even there, after a few synapses beyond the receptor organs, modularity becomes fuzzier as the cross-talk between systems increases. Within sensory and motor systems there is a great deal of redundancy, such that brain regions close to one another show overlapping functional characteristics. The partial redundancy in the nervous system allows for a great deal of flexibility in the responses of relatively small collections of cells to the internal and external world.

Another important characteristic of the nervous system is the connectivity between cells. Whereas most other systems in the body show some capacity for cell to cell communication, the nervous system is specialised for rapid transfer of chemoelectric signals. Although it is not true that every part of the brain is connected to every other part of the brain, there is a great deal of parallel and reciprocal connectivity. Physiologically, this means a single perturbation to the system is conveyed to several parts of the brain simultaneously and some of this will feed back onto the initial perturbation site. Such feedback can modify the response of the system to subsequent influence and lead to nonlinear responses despite linear changes in input.

When the question of mapping cognition to the brain is viewed in this way, the complexity of the problem seems obvious. Is it the case that cognitive processes, such as attention and memory, are localised to discrete portions of the brain, or does the brain represent cognition through the action of the whole? Is there some reasonable compromise between these views?

The purpose of this paper is to suggest that cognitive and behavioural phenomena, here learning and memory, are the direct result of the interactions among anatomically connected brain areas: neural interactions. The notion that cognition results from the operations of neural networks has a long history (Finger, 1994; Lashley, 1929), and has been put forth in recent theoretical work (e.g. Mesulam, 1998). Central to the idea of large-scale networks for learning and memory is that learning and memory are seen as a ubiquitous property of nervous tissue (Gonzalez-Lima, 1989), while the phenomenological diversity of learning and memory are emergent features of the interactions among brain regions (e.g. Edelman, 1978). A development out of the network approach is the idea of a “neural context”. Neural context simply means that the same activity change in an isolated region may be observed across several behavioural and cognitive operations, but the activity and interactivity of related regions may be quite different. It is the relation of the activated region to other areas that determines the cognitive operation. The empirical demonstrations presented here come from studies concerning neural interactions as revealed from metabolic imaging methods, but the theoretical implications extend across all temporal and spatial scales of investigation.

MEASURING NEURAL INTERACTIONS

Neural interactions refers, in a general sense, to influences that different elements in the nervous system have on each other via synaptic communication (the term “elements” refers to any constituent of the nervous system, either a single neuron or collections thereof). The typical approach to the understanding of neural interactions has been to see if the activity varies systematically with some manipulated parameter. However, activity changes in one neural element usually result from a change in the influence of other connected elements, so focusing on activity in one area will miss the change in afferent influence. Furthermore, it is logically possible for the influences on an element to change without an appreciable change in measured activity. The simplest example would be where an afferent influence switches from one source to another, without a change in the strength of the influence. Monitoring regional activity alone would miss this critical shift (for experimental evidence see: Lindsey, Morris, Shannon, & Gerstein, 1997; McIntosh et al., 1994).

The alternative approach to quantifying neural interactions is to measure the relation of activity between elements. One way to do this is by measuring the covariance of activity. The foundation for covariance analysis in neuroimaging

was laid by Horwitz in a number of papers that looked at regional interrelations in a pairwise manner (Horwitz, 1989; Horwitz, Duara, & Rapoport, 1984, 1986; Horwitz et al., 1991; Horwitz, Soncrant, & Haxby, 1992b). Since then, covariance analyses have been extended to the exploration of interacting neural systems (McIntosh & Gonzalez-Lima, 1994), and to the identification of spatial and temporal clustering with various multivariate techniques (Friston, 1994; McIntosh, Bookstein, Haxby, & Grady, 1996a; McLaughlin et al., 1992).

The measurement of neural interactions in neuroimaging has proceeded under two general approaches. The first emphasises pairwise interactions, often in terms of correlations or covariances. The second incorporates additional information, such as anatomical connections, and considers interactions of several neural elements simultaneously to explicitly quantify the effect one element has on another. These two approaches are known as *functional* and *effective connectivity* respectively. Both terms were introduced in the context of electrophysiological recordings from multiple cells (Aertsen, Gerstein, Habib, & Palm, 1989; Gerstein, Perkel, & Subramanian, 1978). More recently, they have been used in reference to neuroimaging data (Friston, 1994). Two methods that will be discussed here typify the use of covariance tools in the analysis of neuroimaging data: structural equation modelling, which provides a measure of effective connectivity, and partial least squares (PLS), which can be used to assess functional connectivity (for technical details see McIntosh et al., 1996a; McIntosh & Gonzalez-Lima, 1992, 1994; McIntosh et al., 1994). Both methods have been used in other scientific disciplines and are discussed here as they pertain to neuroimaging. The present focus is on how to use the methods to ask specific questions about the functional organisation of the nervous system. An example using structural equation modelling is described first.

CHANGING NETWORKS IN WORKING MEMORY

The neural basis of working memory has been researched extensively in functional neuroimaging and in monkey electrophysiological and lesion studies (e.g. Goldman-Rakic, 1990; Jonides et al., 1993; Kirkby, Van Horn, Ostrem, Weinberger, & Berman, 1996; McCarthy et al., 1994; Petrides, 1994). (For the present experiment, working memory is defined as the process of maintaining an active representation of visual information for ongoing information processing [Baddeley & Hitch, 1974].) Consistent across most of these neurophysiological studies is a strong involvement of prefrontal cortex. A feature that has been difficult to appreciate is how the areas connected with prefrontal cortex influence working memory operations. From the perspective of interacting neural systems, working memory as a neurobiological process should engage several regions depending on task requirements. Although it may be the case that for the operation of working memory to occur, prefrontal cortex must be involved, the actual process is the interactions of prefrontal cortex with other

brain regions. So rather than the function of prefrontal cortex, working memory may be best appreciated as one of the emergent properties of the interactions of prefrontal cortex with other brain areas. This possibility was examined using structural equation modelling in the study presented next (McIntosh et al., 1996b).

Structural Equation Modelling

Structural equation modelling, or path analysis, is a multivariate analytic tool that is used to test hypotheses about the causal influences among measured or latent variables. One of the main purposes is to determine whether a hypothesised set of causal relationships is consistent with the observed data. The covariances among the variables are used to provide weights to proposed causal relationships in a manner similar to a multiple linear regression, which then indexes how well the proposed causal structure represents the observed covariance. It is used extensively in psychology and other social sciences (Bollen, 1989). For example, structural equation modelling has been used to distinguish between inherited and environmentally determined influences on certain personality traits (Loehlin, 1987) and whether performance on several memory tasks is best accounted for by the influence of a unitary or bidimensional memory system (Nyberg, 1994). As applied to neuroimaging data, structural equation modelling combines interregional covariances and neuroanatomy. This is an important feature of the application of structural equation modelling to neuroimaging. The causal structure is determined from the anatomy, rather than hypothesised, and the major goal is to evaluate the experimentally induced changes in the effective connections between regions. The basic steps for structural equation modelling are illustrated in Fig. 1 and described next:

- *Select regions or nodes of the network:* The selection of areas for the model is driven by a combination of univariate analysis of changes in mean rCBF, multivariate analyses (e.g. PLS), and theoretical guidance. If there is a particular theoretical model under consideration, such as a putative network for visuospatial attention (Mesulam, 1981), the areas forming that model should be included.

- *Obtain the anatomical model:* The anatomic connectivity between selected brain regions is derived from the neuroanatomy literature. This is not a trivial step and it is at this stage that the theoretical persuasion of the investigator must guide the decision as to which connections to include in a model. Any system of equations where there are unknowns to be solved benefits from constraints to possible solutions. Using the anatomy of the system helps to constrain solutions. However, if all major and minor paths were included most models would contain reciprocal loops at nearly every level with some interconnections

Steps in Structural Equation Modeling

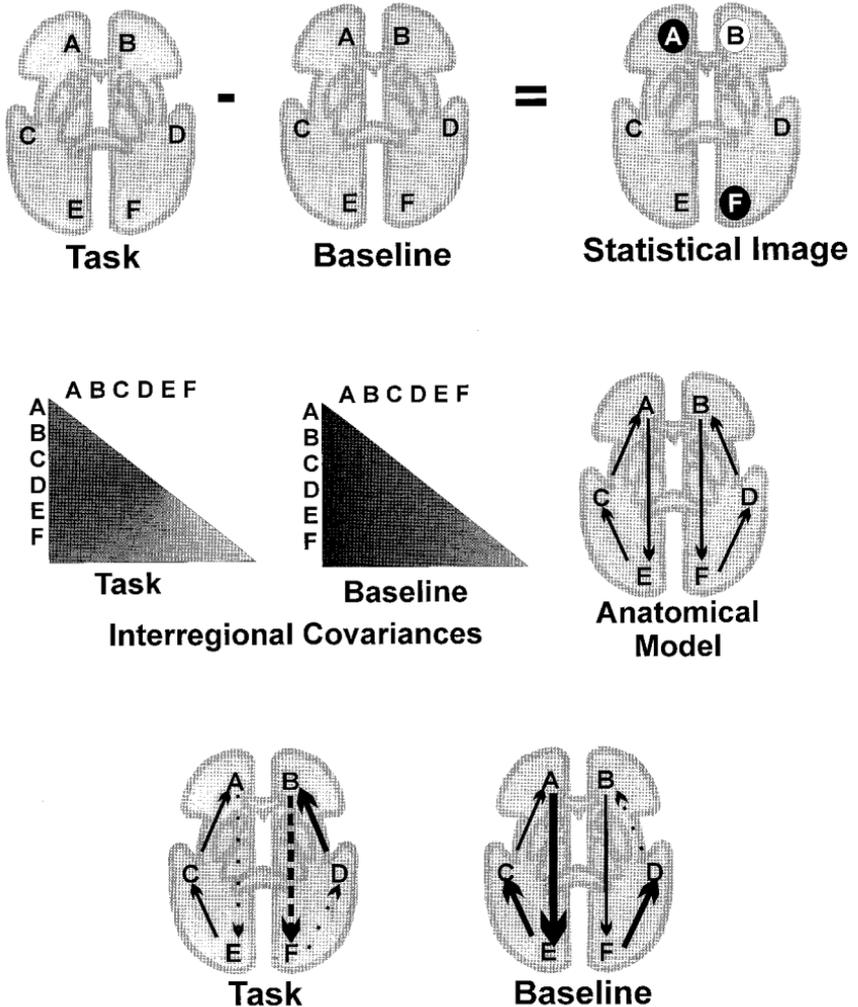


FIG. 1. Graphic representation of essential steps in structural equation modelling for imaging data. In the top row, regions for the model are selected based partly on the statistical differences between task and baseline (regions A and F showing deactivation and B showing activation) and partly on theoretical considerations, such as anatomical completeness or a theoretical model (regions C, D, and E). In the middle row, the activity correlations between areas are computed within-task and the anatomical model is constructed based on published neuroanatomical work. Finally, in the bottom row, the correlations are used to derive path coefficients for each anatomical connection within-task yielding two functional models. Positive weights are solid and negative dashed with the thickness of the line indicating the size of the weight. The “deactivations” identified from the subtraction image in region A correspond to a reduced involvement of region A, while the deactivation of F is because of strong negative feedback from region B. The activation in region B corresponds to its increased (suppressive) influence on F plus the stronger afferent influence from D.

between levels, both feedforward and feedback. When all possible anatomical connections are included, it is likely that an underdetermined system of equations would result, where there are either the same number of known and unknown elements or more unknown elements. In either case, unique solutions are not obtainable. In most cases, some compromise between anatomical accuracy and interpretability may be needed. There have been several published accounts where the compromises in model building have been made explicit (for further discussion see McIntosh & Gonzalez-Lima, 1992, 1994). For instance, in the example to be described next, feedforward connections were modelled first, then the path coefficients fixed at those estimates and feedback effects computed. The stability of the solution was guaranteed by ensuring the estimates were the same regardless of the order that the models were constructed. Interhemispheric effects were then estimated based on the constraints from the feedforward and feedback effects. This is obviously a compromise of what the reality of the interactions may be, and these compromises need to be made explicit for a complete understanding of the final model. Such specificity is often lacking in neurocognitive models that are loosely based on activation patterns. Any modelling effort, whether based on simulations, data fitting, or intuition, is necessarily a simplification and represents an approximation of reality. It is the degree of simplification that determines the utility of the model.

- *Calculate of interregional covariances:* For most published accounts of structural equation modelling in neuroimaging, the interregional covariances were computed within a condition and across participants. For fMRI and data from electrical or magnetic recordings (ERP or MEG), covariances can be computed for an individual subject across tasks or across trials of the same task, as many more within-subject measures can be made (for an application to fMRI see Buchel & Friston, 1997).

- *Calculate the path coefficients and comparison of functional models:* Path coefficients represent the proportion the activity in one area determined by the activity of other areas that project to it. When the coefficients are based on functional activity measured across participants, they reflect what could be thought of as an average functional influence within a given task and index the reliability and sign of the influence. The final step includes the comparison of path coefficients across tasks to determine if the interactions within the functional network differ.

The data set was obtained from a rCBF PET study of working memory using a delayed match-to-sample task for faces with a parametric manipulation of the delay interval (Haxby et al., 1995). The experiment consisted of a match-to-sample task with no delay (perceptual matching), and five scans with increasing delay (1, 6, 11, 16, and 21 s). The data from delay conditions were averaged: the 1 and 6 s delay conditions were averaged into a short delay condition, the 11 and 16 s delay into an intermediate delay condition, and the 21 s delay was the long

delay condition. The grouping of conditions was based on the patterns of regional mean differences (Haxby et al., 1995).

For the working memory models, the regions chosen were those showing either a significant increase or decrease in rCBF with increasing delay as determined with the linear trend analysis and a task PLS (McIntosh et al., 1996b). The areas included ventral occipital and temporal regions (Brodmann areas [BA] 18, 37, and 21), inferior prefrontal cortices (BA 47; middle and medial prefrontal regions are part of the larger model presented in the original publication), anterior and posterior cingulate (BA 24 and 23), and the hippocampal gyrus (GH). Generally speaking, the posterior occipital and temporal areas showed decreasing rCBF with increased delay and the prefrontal and cingulate regions showed increasing activity with increasing delay. Measures of activity were obtained from both hemispheres, and interregional covariances were computed within-task and across participants.

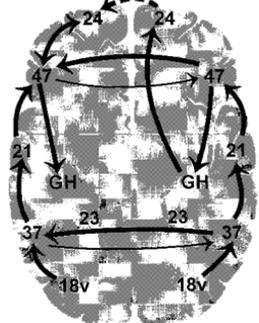
Results

Figure 2 shows a summary of the dominant interactions for the four conditions: perceptual matching, short, intermediate, and long delay. In the perceptual matching network (top of Fig. 2), strong interactions were present bilaterally in the ventral cortical visual stream (Ungerleider & Mishkin, 1982), extending from occipital BA 18v, to occipitotemporal BA 37, to temporal BA 21 and then into ventral prefrontal BA 47. Prefrontal influences from BA 47 on GH were strong bilaterally. Cross-hemispheric effects were strongest between BA 47 and 37 with a slight asymmetry favouring the right hemisphere. The ventral stream and interhemispheric part of the network is extraordinarily similar to that of a previous path analysis from a similar perceptual matching task (McIntosh et al., 1994), which indicates that the functional networks are replicable across experiments and subjects.

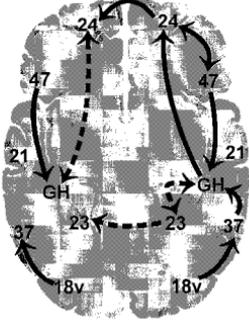
The main trends of interest across the delay conditions were a change in the interactions along the ventral stream in the right hemisphere, more corticolimbic interactions in the short and intermediate delays, and more frontal involvement in the long delay. Of particular interest was the observation that several key areas, especially prefrontal cortex and anterior cingulate, showed strong interactions across all tasks, but the nature of these interactions changed as a function of delay.

The general feature of the short delay network, compared to matching, was an increase in the interactions right hemisphere involving BA 37, inferior frontal BA 47 and GH, and reduced influences in the left hemisphere (Fig. 2, bottom left). The pathway from BA 37 to BA 21 was weaker in both hemispheres and the influence of BA 37 appeared to be rerouted into the GH on the right. Within the right hemisphere, there was a strong corticolimbic trinodal loop set up with GH, BA 47, and BA 24. This loop changed at the longer delay conditions.

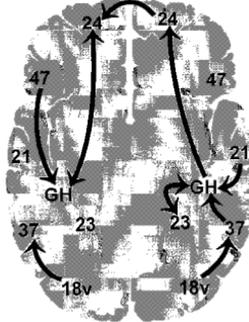
Perceptual Matching



Short Delay



Intermediate Delay



Long Delay

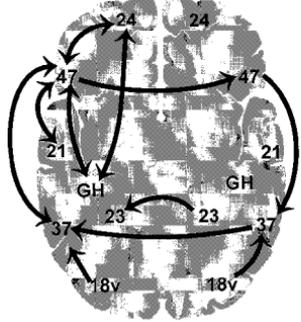


FIG. 2. Summary diagram representing some of the dominant changes in functional interactions for the working memory tasks. Interactions are depicted on a horizontal schematic. To maintain clarity, the locations of these areas are not completely accurate and the reader is referred to the original publication for the exact location of the regions and their effects within these functional networks (McIntosh et al., 1996b). Positive path coefficients are shown as solid arrows, whereas negative ones are shown as segmented arrows. Two-headed arrows represent reciprocal and symmetric functional interactions.

The intermediate delay network (Fig. 2, bottom centre) could be characterised as showing less right inferior frontal involvement, compared to the short delay network, but still maintaining the right corticolimbic interactions with a positive loop between posterior cingulate and GH. The intermediate delay model also showed some changes in left hippocampocingulate interactions relative to the short delay model.

The long delay condition was characterised by a shift of what had been primarily right hemisphere interactions to stronger left hemisphere interactions. Right corticolimbic interactions were essentially zero, while left corticolimbic interactions increased, and a trinodal loop emerged similar to that observed in the short delay for the right hemisphere. Prefrontal effects from BA 47 seemed to play a stronger role in the long delay model compared to the other functional

models (Fig. 2, bottom right). Frontal effects on occipitotemporal and temporal areas were especially strong in the long delay model. This feedback, along with the changes in interhemispheric interactions, resulted in an interesting interhemispheric circuit involving prefrontal BA 47 and posterior cortical BA 37.

Discussion

The interpretation of the change in the functional networks across delay was informed by the subject debriefing reports. The participants indicated that they could maintain a mental image of the sample face at short delays, but that it tended to degrade at the longer delays. From this, it was felt that the strategy at shorter delays was to rely on maintenance of the iconic representation of the face, whereas at longer delays an elaborative rehearsal and encoding strategy was employed. The interactions of BA 37, GH, and BA 47 during shorter delays, therefore, may partly reflect the attempt to hold the image of the face, whereas the left hemisphere interactions at the longer delay may represent elaborative rehearsal and encoding (McIntosh et al., 1996b).

The task-related network changes were not simply that certain areas interacted in one task and not in another, but in the nature of the interactions. It is important to note that the changes in neural interactions were much more striking than changes in regional mean activity. Anterior cingulate, prefrontal cortices, and GH showed strong interactions across all tasks, but the activity of these regions was not consistently higher than the control condition (Haxby et al., 1995). Even in cases where an area is not "active" relative to baseline, it may still show strong interactions. Activation analysis detects differences only, whereas covariance analyses reveal interesting changes in brain interactions not detectable through activation analysis.

Working memory studies have implicated three core regions: the anterior cingulate, dorsolateral prefrontal cortex, and the hippocampus (Goldman-Rakic, 1990). These regions were part of the functional networks for working memory discussed here, but they were also involved in a perceptual matching task with no obvious working memory component. An argument could be made that the interactions during perceptual matching also reflect some working memory component related to task execution, but this would require further experimental validation. Structural equation modelling results bring the suggestion that regions may play a part in more than one functional network, and that it is the interactions with other brain regions that determine what operations are being served at that time. Frontohippocampal interactions were present across all functional models. What changed between tasks was the nature of these interactions (indicated by the sign of the path coefficients). The hippocampus has been an enigma for brain imaging studies, but when examined in the context of covariance relationships, part of the reason for lack of consistent changes in

activity may be that the interactions of this area are always strong and what changes with memory demand is the nature of these interactions. If this is true for the rest of the nervous system, much is to be gained by evaluating neural activity *and* interactivity related to cognition.

NEURAL CONTEXT AND RETRIEVAL MODE

One concept that arises from the network approach is that of a *neural context*. Most brain regions receive inputs from many areas and then send projections to several others. At any instance, the interactions through anatomical connections may shift from one afferent/efferent source to another, resulting in a change in cognition or behaviour. Across several different tasks, a brain area may show the same activity pattern but serve different functions because of the relation of that activity with other brain regions (for empirical examples see Chafee & Goldman-Rakic, 1998; D'Esposito, Ballard, Aguirre, & Zarahn, 1998; Zhang, Riehle, Requin, & Kornblum, 1997). The important factor is not that a particular event occurred at a particular site, but rather under what *neural context* did that event occur—in other words what was the rest of the brain doing? Neural context is closely related to the idea of “functional pluripotentialism”, put forth by Filimonov (cited in Luria, 1962, pp. 24–25), which states that no formation in the central nervous system is responsible solely for a single function, and under certain conditions a given formation may be involved in other functional systems and may participate in performance of other tasks.

The notion of neural context can be illustrated by considering the activity of right prefrontal cortex (RPFC) in episodic memory retrieval (McIntosh, Nyberg, Bookstein, & Tulving, 1997). One of the most reliable results in neuroimaging of human memory is that of increased RPFC activity during episodic memory retrieval (Cabeza & Nyberg, 1997; Tulving et al., 1994). Experiments following from this observation suggest that the activation of RPFC may reflect “retrieval mode” or the act of searching memory without regard to the success of this search (Nyberg et al., 1995). In isolation, the interpretation seems reasonable. However, if cognitive states like retrieval mode are subserved by a large-scale neural network that includes RPFC, then the common pattern of RPFC activation across the retrieval tasks should reflect part of the operation of a general network for retrieval mode. Another possibility, which ties into the idea of neural context, is that RPFC activity may be similar across retrieval tasks, but result from the interactions with different brain regions depending on the act of retrieving and the success of the retrieval. Is the activation of RPFC reflecting a general retrieval mode network, or are there several networks, differentially engaged, whose interactions result in a similar pattern of RPFC activity?

One way to examine whether the same region has a consistent pattern of interactions across retrieval tasks is to explore change in the correlation of that region, or functional connectivity, with other parts of the brain. If increased

RPFC activity during retrieval represents retrieval mode, then there should be a pattern of functional connectivity for the RPFC that is similar across memory retrieval tasks. The question was addressed using PLS.

Partial Least Squares Analysis

PLS is a multivariate tool that can be used to describe the relation between a set of exogenous measures, like experimental design or behavioural measures, and a set of functional brain images (McIntosh et al., 1996a; McIntosh et al., 1998a; Schreurs et al., 1997). What results from a PLS analysis are sets of images that may be interpreted as nodes of neural systems representing some experimental effect or relating to some behaviour measure. The same analytic tool can be used to explore whether a part of the brain, represented by an image voxel, shows any task-related changes in its relation to the rest of the brain.

A highly idealised graphical description of the PLS procedure used to analyse changes in regional correlation is presented in Fig. 3. It can be regarded as an extension of the “seed voxel” correlation analysis proposed by Horwitz et al. (1992a). In panel A of Fig. 3, activity from a particular seed voxel (middle right of the “image”) is correlated with the activity from the rest of the image in three tasks. This produces a correlational map for each task depicting areas that are correlated with the seed voxel.

In interpreting the correlation maps, a researcher would usually identify where the maps differ across three tasks. For example, in Fig. 3A the seed voxel shows an opposite pattern of correlations with posterior areas in tasks 2 and 3, and correlates with its contralateral homologue in task 1 only. Common correlations across the three tasks are in contralateral prefrontal regions, and, of course, of the seed voxel with itself. The extraction of commonalities and differences in the correlation maps is what the PLS analysis does. The PLS analysis of the correlation maps, or seed PLS, operates on the three correlation maps together and, through singular value decomposition (SVD), provides sets of mutually orthogonal latent variable (LV) pairs. One element of the pair contains numerical weights for each task creating a profile that depicts either a common correlation pattern or a task-related difference (i.e. a contrast)¹. The other element of the LV pair identifies the parts of the image that show the profile across tasks and can be displayed in image space. Because it is derived from SVD, it is called a *singular image*. Within the singular image are numerical weights for each voxel, and their variation across the image shows which areas are maximally expressed on the particular LV. The weights for both the singular

¹ Another way to think about this is to consider an experiment with only two tasks and thus only two correlation maps. The sum of these two maps (or average) will represent areas of common correlation and a subtraction of the two maps would identify differences in correlations (McIntosh et al., 1996a). This is essentially what occurs in PLS and whether a singular image depicts common correlations or differences would be identified in the task profiles.

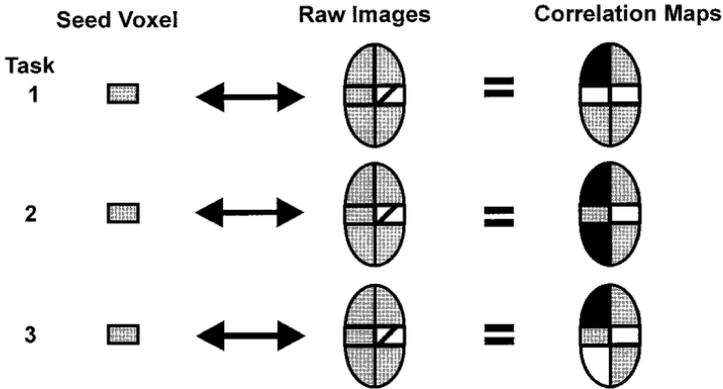
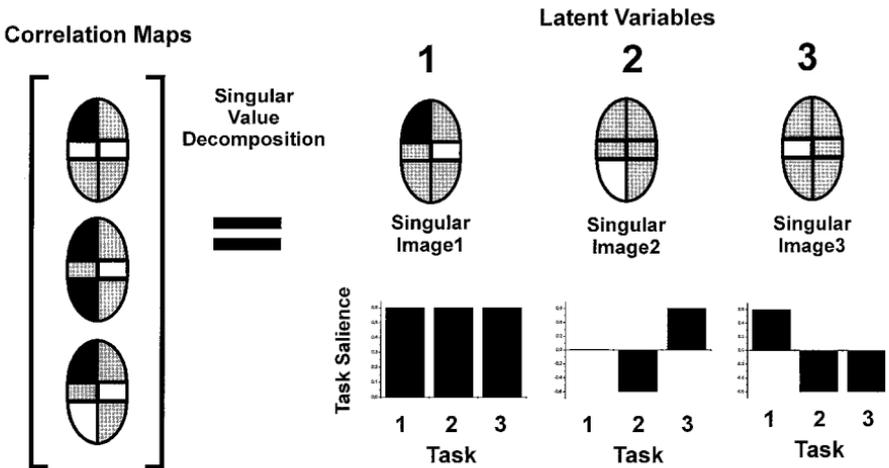
A**B**

FIG. 3. Graphic representation of the steps involved in a partial least squares analysis of seed-voxel correlations. Panel A: a seed voxel is selected from the middle of the right side of the image (indicated by the black hash mark) within each of three tasks and correlated with the rest of the six-voxel image resulting in one correlation map per task. Black represents a strong negative correlation and white a strong positive. In panel B the three correlation maps are stacked into one large matrix and decomposed with singular value decomposition resulting in three latent variable (LV) pairs. Each pair consists of the singular image, which is an image representation of the weights (saliences) for voxels, and the weights or saliencies across tasks or the task profile (bottom of each LV). As with the correlations, black represents a strong negative saliency and white a strong positive. LV1 is a common pattern of seed voxel correlations, LV2 distinguishes the correlation maps from tasks 2 versus 3, and LV3 contrasts the map for task 1 versus tasks 2 and 3.

image and the task profile are called *saliences*. In reference to Fig. 3B, the singular image from the first LV shows positive saliencies at the location of the seed voxel (middle right of the “image”) and negative saliencies at the contralateral prefrontal voxel. In the bar graph below the singular image, the saliencies across tasks are equal, so the first LV is the common correlation of the seed voxel with itself and the contralateral frontal regions. The singular image in LV2 has a negative salience at a posterior location, and profile across the tasks contrasts tasks 2 and 3, so the LV depicts a posterior difference in seed-voxel correlation patterns. Finally, LV3 shows the salience at the contralateral homologue for the seed voxel and the task profile contrasts task 1 with 2 and 3. Although the results of the seed PLS analysis in Fig. 3 could have easily been distilled from the examination of the within-task correlation maps, it is much more difficult when images contain several thousand more voxels are analysed, as with PET or fMRI data (and real data are seldom as clean). The seed PLS analysis can also be extended to include more than one seed voxel, as in the example described next. The interpretation of the singular image and task profiles are the same as for the single seed voxel analysis, but a given singular image may show different task profiles for each seed voxel (e.g. task commonalities for one seed voxel, task differences for another).

The question of neural context and retrieval mode was examined with data from a PET rCBF study of episodic memory retrieval (Nyberg et al., 1996b; Nyberg et al., 1995). Three retrieval tasks were used with differing levels of retrieval success. Before scanning, participants were presented with two lists of words, some spoken by a male and some by a female. For one list, participants were asked to identify the gender of the speaker (shallow processing) and for the other list they were asked to decide whether the word represented a living thing (deep processing or semantic encoding). The retrieval conditions consisted of yes/no recognition (indicated by a button-press) for visually presented word lists that were either: (1) all unstudied (New); (2) from the shallow processing list (Shallow); or (3) from the deep processing list (Deep). Retrieval success would be highest in condition 3, but in all tasks participants would be in retrieval mode. A baseline task was also used, where a subject read a single word and pressed a button. Each condition was scanned twice.

Results and Discussion

For the seed voxel PLS analysis, representative voxels from three areas were selected: the two RPFC areas (BA 45/47 and BA 10, atlas [Talairach & Tournoux, 1988] XYZ coordinates: 32, 22, 0; 28, 44, 4,) and the left hippocampal gyrus (LGH: -24, -36, -8). The RPFC regions were activated to a similar degree in all retrieval tasks compared to baseline (Nyberg et al., 1995). The LGH location was activated only in the Deep condition (Nyberg et al., 1996b). The RPFC regions, assuming their activation reflects part of a retrieval

mode network, should show common correlation patterns across the three retrieval conditions. The LGH, which may be more related to retrieval success, should show a pattern of correlations unique to the Deep condition. These hypotheses are presented graphically in Fig. 4 in terms of expected task profiles from a seed PLS. A profile representing retrieval mode (Fig. 4, top) would contrast the read baseline task against the three retrieval conditions, whereas a profile favouring retrieval success (Fig. 4, bottom) would contrast the New task against the Deep task with Shallow at some intermediate level.

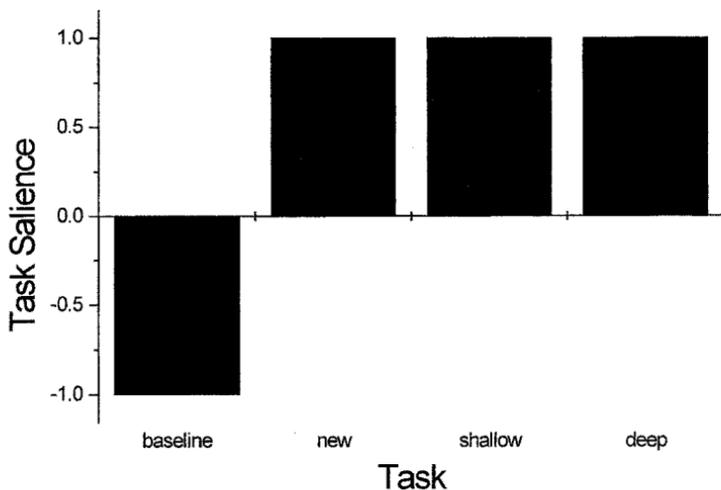
None of the task profiles from the seed PLS was consistent with a pure retrieval mode interpretation. The singular image from the first LV is displayed in Fig. 5. Large areas of positive saliences extend across bilateral inferior temporal lobes, hippocampal and parahippocampal gyri, and retrosplenial cortex. Negative saliences are noted in left and right middle and dorsolateral prefrontal, medial occipital cortices, and midbrain. The task profiles for each seed voxel in the bar graph at the bottom of Fig. 5 indicate that the first LV was mainly a common trend in correlations for LGH and BA 10. LGH was positively correlated with areas that were positively salient in the singular image, while BA 10 was positively correlated with negatively salient areas. There was, however, an interesting addition in the Deep condition. Along with LGH and BA 10, the Deep condition showed a strong positive salience for the seed voxel at BA 45/47.

One may interpret this LV to suggest that there is a common pattern of covariances between BA 10, LGH, and the other limbic and frontal areas identified on the singular image. In the Deep condition, involving high memory retrieval following semantic encoding, another right prefrontal region, BA 45/47, shows strong covariances with limbic areas. That is to say, these three seed voxels are bound by a common pattern of functional interrelations with themselves and other brain areas only when there is successful episodic retrieval.

This interpretation of the LV may seem incongruent with the results from the activation findings, which deal with mean differences, not covariances of voxels. Recall that RPF areas used here were relatively activated in all retrieval tasks, while the LGH activity was highest in the Deep condition. Patterns of covariation across other LVs (not presented graphically) suggested that the common task-related activation of RPF may have arisen from different functional interrelations, as none of the LVs distinguished the retrieval tasks from baseline (e.g. Fig. 4, top). For example, LV 2 was salient for the right prefrontal regions for the Shallow and Deep conditions. The third LV was salient for BA 45/47 in the reading baseline only. In the fourth LV, the salience for BA 45/47 was strong in the New and Shallow tasks.

The results from the PLS analysis suggest that similar patterns of activation in RPF came about through different patterns of interregional interactions. In LV 1, prefrontal BA 10 and the LGH showed common correlations across all conditions, but only when there was high memory retrieval following semantic encoding was BA 45/47 incorporated into the pattern. In light of the activation

Retrieval Mode



Retrieval Success

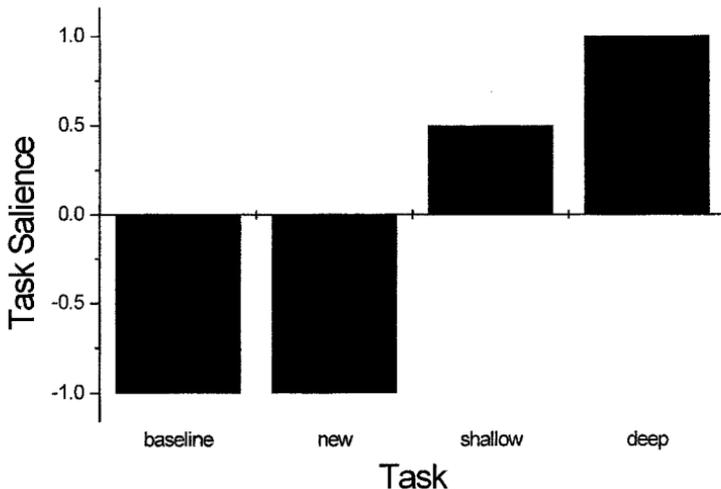


FIG. 4. Possible task profiles from a seed-voxel PLS that reflect two competing hypotheses for memory retrieval. In the top of the figure, the task profile contrasts the three retrieval tasks with the reading baseline task. This profile would be consistent with a pattern of functional connectivity for retrieval mode, or the act of retrieving independent of retrieval success. The bottom figure shows a task profile that contrasts the baseline and new task, where retrieval success is lowest, with Deep, and to a lesser degree, Shallow, where retrieval success is greatest. This is consistent with a pattern of functional connectivity representing retrieval success.

Singular Image

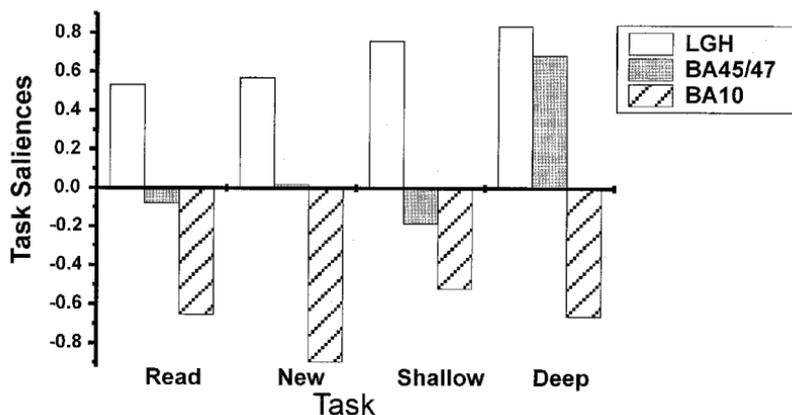
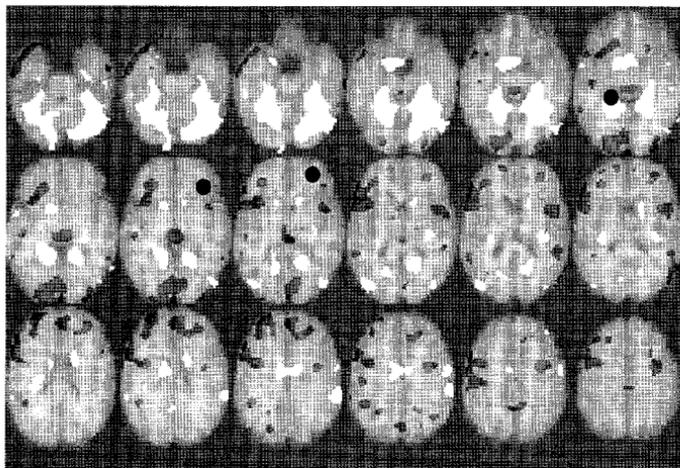


FIG. 5. First latent variable from the PLS analysis of seed voxel correlations for left hippocampus (LGH), right BA 45/47, and right BA 10. The bottom of the figure shows the task profile for each seed voxel by task, which is an index of how strongly the voxel covaries by task with the singular image at the top. This profile can be interpreted as a common pattern of correlations for LGH versus BA10 for all tasks, but Deep versus all other tasks for BA 45/47. The singular image depicts peak voxel that as a collection covaried with the task profile. Positive saliences (voxels positively correlated with LGH, i.e. greater values for higher LGH values) are drawn in white, negative saliences (voxels positively correlated with BA10 or greater values for higher BA 10 values) in black, both thresholded at a salience greater than 0.5 (absolute value). The image is displayed on horizontal sections from a structural MRI that conforms to the atlas space of Talairach and Tournoux (1988). Slices start at -28mm from the AC-PC line at the top left slice and move in increments of 4mm to $+40\text{mm}$ at the bottom right, and left is left and the top is anterior in the image. Black circles within the singular image mark the location of the three seed voxels and the surrounding voxels that were removed to reduce the spatial autocorrelation.

analysis, this may imply that the greater activation of the three seed voxels in the Deep condition arises through their common functional connections. For the Shallow and New tasks, however, the common activation of the two right prefrontal regions came about through other patterns of connectivity, perhaps due to increased suppressive influences on other regions (Nyberg et al., 1996a) and stronger interactions with regions identified on subordinate LVs.

The exploration of neural interactions in this study is a compelling example of the idea of neural context. Activity patterns in RPFC were common for the three retrieval conditions compared to baseline, but there was no evidence for a general retrieval mode network involving RPFC. One implication is that the equivalent activation level arose from different neural interactions across the three retrieval tasks. What distinguished the three retrieval conditions was the neural context in which the RPFC activation occurred. This result raises an interesting possibility that *neural interactions representing retrieval mode vary depending on the success of memory retrieval*. This speculation could be formally tested using methods that afford better temporal resolution, like ERP/MEG or fMRI, with the expectation that the regions that are functionally connected with RPFC would depend on successful memory retrieval (e.g. Duzel et al., 1999).

Neural context allows regions to be part of more than one functional network. This is contrary to notions that cognitive processes like attention and memory are subserved by specific neural systems (c.f. Mesulam, 1990; Posner & Petersen, 1990). Most parts of the brain possess the rudimentary properties necessary for cognition. For instance, modification of responses related to learning and memory have been observed in several parts of the brain from single cells in isolate spinal cord preparations (Wolpaw & Lee, 1989; Wolpaw, 1997), to primary sensory and motor structures (Donoghue & Sanes, 1994; Gonzalez-Lima, 1992; Molchan et al., 1994; Recanzone, Schreiner, & Merzenich, 1992; Weinberger & Diamond, 1987). When several brain regions interact at a larger scale, these rudimentary features will combine to produce a particular cognitive function (Bressler, 1995). Whether or not a region is part of a neurocognitive system depends on the specifics of the processing demands (what is the person doing?) and the interactions with other regions (what is the rest of the brain doing?). Just as an instrument in an orchestra may switch from a lead to a support role in different pieces of music, some regions may play a more prominent role in certain cognitive functions, and then play a supporting role in others.

METHODOLOGICAL CONSIDERATIONS AND CAVEATS

The techniques presented in this paper represent two of several possible approaches to the estimation of functional and effective connectivity. Partial least squares and structural equation modelling may be employed separately or

together to address specific questions about neural interactions. Indeed the combination of the two is a powerful means to go from activity changes, to functional connectivity, and then finally to the derivation of effective connectivity (e.g. McIntosh, Cabeza, & Lobaugh, 1998b). However, each approach does have its limitations. The examination of functional connectivity provides a means to address whether a region or collection of regions is interacting, but gives no information on the direction of these interactions. As such, inferences regarding the modulation of one area by another are not warranted. Estimation of effective connectivity provides some basis for inferences regarding the direction of influences, but in order to do this, assumptions on how these influences can occur are required. For example, in structural equation modelling the anatomical model constrains how regions may interact. Moreover, because it is a modelling technique certain simplifications are required as stated earlier. Both methods as presented here are restricted to exploration of linear relationships between brain areas. Although it is reasonable to assume that the scales measured in PET the predominant relationship is linear, the explicit incorporation of nonlinearities will be required for a full appreciation of neuronal components of cognition. Indeed there is some suggestion that in the time scales relevant for cognition (50–200ms) there are significant nonlinear components (Friston, 1997). Nonlinear extensions to structural equation modelling are rather straightforward (Kenny & Judd, 1984) and there has been some progress towards this in fMRI (Buchel & Friston, 1997).

There is some concern about which source of variance, across tasks or across subjects, is best for the estimation of neural interactions (see Friston, 1995; Strother, Kanno, & Rottenberg, 1995). The issue of which source of variability is “correct” is not unique to neuroscience (Mandler, 1959), and there is no necessity for a logical connection between covariances computed across tasks within-subjects and those computed across subjects within-task. However, there is also no justification for preferring one source of variability to another, particularly in cases where both can be examined as in fMRI or ERP studies. Within-subject analysis assesses the direct relation between regions, whereas across-subjects analysis provides an indication of the stability that relation. These are complementary, not contradictory, pieces of information.

For illustration, say we take 10 people of varying heights and weight, and ask them to pull on a potentiometer by flexing their arm (an arm curl). If you measured activity of the muscles in the arm for each subject, say through blood flow, and correlated them, you would probably find a strong correlation with the biceps and brachialis muscles. Although each person would differ in the amount of blood flow to the muscles, from the correlation based on this variance, you would conclude that the muscles on the ventral surface of the arm have something to do with flexion. If instead you measured muscle activity in a single subject with a progressive increase in the resistance to arm flexion, you would

find a correlation between muscle activity in the ventral part of the arm. Replicating the measurement by running different subjects would lead you to the same conclusion that you had reached by using the between-subjects covariance. The point here is that computing covariances between or within subjects can lead to complementary conclusions so long as there are adequate experimental controls, and the statistical analysis ensures the answers are reliable.

CONCLUSIONS

The paper began by presenting the possibility that with neuroimaging we can investigate the operations of neural networks supporting cognition. Creating activity maps can give hints about the constituents of these distributed systems, but an appreciation for their interactions can only come from investigation of something like interregional covariances. Two covariance-based methods, structural equation modelling and partial least squares, were presented as ways to quantify neural interactions. Structural equation modelling was used to demonstrate the dependence of the functional network interactions on memory load. Partial least squares was used to show how right prefrontal cortical activity could be associated with the act of searching memory (retrieval mode) or successful retrieval depending on its relation to other brain regions, or neural context. Common in both examples is the possibility that learning and memory, very broadly defined, may emerge from neural interactions rather than being the responsibility of particular brain areas.

The idea that learning and memory are ubiquitous properties of the brain is counter to the parcellated view of nervous system functional organisation. However, physiological data have consistently shown that almost all parts of the nervous system show the capacity for learning and memory. There may, therefore, be several “types” of learning and memory depending on which regions are interacting. This is not to say that the brain contains several memory systems that are independent of perceptual operations, as has been implied by some theories in cognitive psychology (Schacter & Tulving, 1994). Instead, higher-order memory will involve interactions among different brain regions depending on the stimuli, the process, and what must be done with the information. This idea is similar to psychological theories that view memory as the dynamic combination of several processes rather than a discrete system (Jacoby, 1991). The interactions between brain regions lead to formation of memory. Memory *per se* is an emergent behaviour of the central nervous system—it is something the brain does.

At first pass, this may seem in direct opposition to data demonstrating severe and permanent deficits in learning and memory following damage to specific parts of the brain. However, it should be clear that the idea of emergent properties from neural interactions is neutral with respect to the lesion findings. Lesion findings give clues about functional organisation, but are silent as to how

the function is carried out. This point is underscored by the following excerpt (John, 1961, p.480):

The fact an animal can learn or retain the [conditioned] response after a lesion does not of itself warrant the conclusion that the structure normally played no role in the response. The engram, we suspect, is wily enough to elude the subcortical shot as easily as the cortical knife. Memory seems more likely to be set processes, which define a state, than a “bit” in a place. That places participate in process is apparent, but we may more legitimately expect the lesion of a region to alter the process than to abolish the state. That lesions can alter process we know, witness the effect of septal lesions on slow hippocampal waves. The problem that confronts us is to unravel how the condition response is produced by the process.

Substituting the words “conditioned response” with something like “memory trace” makes this point quite relevant to human cognitive neuroscience.

Some valuable insights could be gained from the exploration of network interactions in the damaged nervous system to see how the networks reorganise themselves in an attempt to compensate. When interpreted in light of interactions in the normal brain, such research could provide the important link between lesion data and normative experiments by explaining the lesion effect. For example, is the lesioned area a convergence site for functional interactions (Damasio, 1989)? Some strides towards this end have been made in examining ageing and Alzheimer’s disease effects from a network perspective (Horwitz et al., 1995).

The idea of a neural context has important implications for how neuroimaging data are evaluated. If one is willing to accept that brain regions communicate with one another in the course of cognitive operations, then what one brain area does must be determined by what other areas connected to it are doing. When an area is more active in a cognitive task relative to a control task, that change must arise from neural interactions. Undoubtedly, there are consistencies in these activations as the realisation of the Hemispheric Encoding/Retrieval Asymmetry (HERA) model (Tulving et al., 1994), and the more recent Hippocampal Encoding/Retrieval model (HIPER, Lepage, Habib, & Tulving, 1998) attest. The next reasonable step would be to determine whether such consistent activations arise from the interactions of common neural systems.

Neuroimaging studies often present data in terms of “activated” regions that are considered independent of deactivated regions. However, relative deactivations are just as important as activations in setting up a neural context. An activation that occurs in one case may mean a different thing in the company of different deactivations. An illustrative example is selective attention. In PET studies of visual selective attention, rCBF in auditory and somatosensory regions decreases while visual regions are activated (Haxby et al., 1994). Consider another study where the same visual activation occurs without concomitant

deactivation of auditory cortex. This pattern may signal quite a different cognitive process, perhaps that both visual and auditory domains are being attended to. If activations alone were examined, there would be no difference between studies. Deactivations are problematic for those who strictly adhere to the cognitive subtraction paradigm, but from a physiological perspective, increases and decreases in neural activity are equally important for nervous system operation. Deactivations can result from an active suppressive influence or from reduced interactions in the task of interest (as with regions F and A in Fig. 1, respectively; it is worth noting that these interpretation ambiguities are also there for activations). A recent paper from our lab demonstrates how it is possible to evaluate the source of deactivations with structural equation modelling (Nyberg et al., 1996a), leading to network models of cognitive operations that are more congruent with neurophysiology (Abeles, 1982; Douglas et al., 1995; Somers, Nelson, & Sur, 1995; Tsodyks, Skaggs, Sejnowski, & McNaughton, 1997; Turova, 1997; van Vreeswijk & Sompolinsky, 1996).

The complexity of data analysis and interpretation is great when covariance-based approaches are used. However, most questions in neuroimaging are phrased such that the answer is best provided by some sort of covariance analysis. Data interpretations are often made from the perspective of interacting networks, even if the network structure is never specified (Fink et al., 1996). To appreciate neural interactions does not require a drastic change in experimental questions, rather it requires a change in how the answers are provided. For example, top-down modulation in studies of selective attention is best explored through a covariance-based approach. Activation of higher-order areas (prefrontal) may arise from increased input or output, and without further analysis these possibilities are not distinguishable (e.g. Shulman et al., 1997). Covariance analyses like structural equation modelling can sort out these possibilities. Neuroimaging is in the peculiar position of having the answers to experimental questions constrained by the analytic methods. With the intellectual and financial investment needed from functional imaging studies, it is a shame that the fate of an entire experiment rests on a t-test. The subtraction method in PET, its carryover to fMRI, and the related statistical tools, have become the default for imaging experiments, but neuroimaging has matured beyond this. The choice of the analytic tool should be dictated by the particular experimental question. If the question is about neurocognitive networks, then the answer should come from something like a covariance-based analyses. If questions are phrased at a regional level (e.g. does an area contribute to an operation?), then activation analyses will provide an appropriate answer. Removed from the constraints of convention, the theoretical issues addressable by neuroimaging expand greatly.

REFERENCES

- Abeles, M. (1982). Role of the cortical neuron: Integrator or coincidence detector? *Israel Journal of Medical Science*, 18(1), 83–92.
- 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.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559.
- Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–90). San Diego, CA: Academic Press.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bressler, S.L. (1995). Large-scale cortical networks and cognition. *Brain Research Reviews*, 20, 288–904.
- 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.
- Cabeza, R., & Nyberg, L. (1997). Imaging cognition: An empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience*, 9(1), 1–26.
- Chafee, M.V., & Goldman-Rakic, P.S. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *Journal of Neurophysiology*, 79(6), 2919–2940.
- Damasio, A.R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, 1, 123–132.
- 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.
- Douglas, R., Koch, C., Mahowald, M., Martin, K., & Suarez, H. (1995). Recurrent excitation in neocortical circuits. *Science*, 269, 981–985.
- Duzel, E., Cabeza, R., Picton, T.W., Yonelinas, A.P., Heinze, H.-J., Scheich, H., & Tulving, E. (1999). Task-related and item related processes in episodic and semantic retrieval: A combined PET and ERP study. *Proceedings of the National Academy of Science USA*, 96, 1794–1799.
- Edelman, G.M. (1978). Group selection and phasic re-entrant signalling: A theory of higher brain function. In V. Mountcastle & G.M. Edelman (Eds.), *The mindful brain* (pp. 55–100). Cambridge, MA: MIT Press.
- Finger, S. (1994). *Origins of neuroscience: A history of explorations into brain function*. New York: Oxford University Press.
- Fink, G., Markowitsch, H., Reinkemeier, M., Bruckbauer, T., Kessler, J., & Heiss, W. (1996). Cerebral representation of One’s Own Past: Neural Networks Involved in Autobiographical Memory. *Journal of neuroscience*, 16(13), 4275–4282.
- Fodor, J.A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Friston, K. (1994). Functional and effective connectivity: A synthesis. *Human Brain Mapping*, 2(1&2), 56–78.
- Friston, K.J. (1995). Statistical parametric mapping: Ontology and current issues. *Journal of Cerebral Blood Flow and Metabolism*, 15, 361–370.
- Friston, K.J. (1997). Another neural code? *Neuroimage*, 5(3), 213–220.
- 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 (Eds.), *Progress in brain research* (Vol. 85, pp. 325–336). Amsterdam: Elsevier Science Publishers.

- Gonzalez-Lima, F. (1989). Functional brain circuitry related to arousal and learning in rats. In M.A. Arbib & J.P. Ewert (Eds.), *Visuomotor coordination: Amphibians, comparisons, models and robots* (pp. 729–765). New York: Plenum Press.
- Gonzalez-Lima, F. (1992). Brain imaging of auditory learning functions in rats: Studies with fluorodeoxyglucose autoradiography and cytochrome oxidase histochemistry. In F. Gonzalez-Lima, T. Finkenstaedt, & H. Scheich (Eds.), *Advances in metabolic mapping techniques for brain imaging of behavioral and learning functions* (Vol. 68, pp. 39–109). Dordrecht: Kluwer Academic Publishers.
- Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., & Grady, C.L. (1994). The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience*, *14*, 6336–6353.
- Haxby, J.V., Ungerleider, L.G., Horwitz, B., Rapoport, S.I., & Grady, C.L. (1995). Hemispheric differences in neural systems for face working memory: A PET-rCBF Study. *Human Brain Mapping*, *3*, 68–82.
- Horwitz, B. (1989). Functional neural systems analyzed by use of interregional correlations of glucose metabolism. In J.-P. Ewert & M.A. Arbib (Eds.), *Visuomotor coordination* (pp. 873–892). New York: Plenum Press.
- Horwitz, B., Duara, R., & Rapoport, S.I. (1984). Intercorrelations of glucose metabolic rates between brain regions: Application to healthy males in a state of reduced sensory input. *Journal of Cerebral Blood Flow and Metabolism*, *4*, 484–499.
- Horwitz, B., Duara, R., & Rapoport, S.I. (1986). Age differences in intercorrelations between regional cerebral metabolic rates for glucose. *Annals of Neurology*, *19*, 60–67.
- Horwitz, B., Grady, C., Haxby, J., Schapiro, M., Carson, R., Herscovitch, P., Ungerleider, L., Mishkin, M., & Rapoport, S.I. (1991). Object and spatial visual processing: Intercorrelations of regional cerebral blood flow among posterior brain regions. *Journal of Cerebral Blood Flow and Metabolism*, *11* (Suppl. 2), S380.
- Horwitz, B., Grady, C.L., Haxby, J.V., Ungerleider, L.G., Schapiro, M.B., Mishkin, M., & Rapoport, S.I. (1992a). Functional associations among human posterior extrastriate brain regions during object and spatial vision. *Journal of Cognitive Neuroscience*, *4*, 311–322.
- Horwitz, B., McIntosh, A.R., Haxby, J.V., Furey, M., Salerno, J.A., Schapiro, M.B., Rapoport, S.I., & Grady, C.L. (1995). Network analysis of PET-mapped visual pathways in Alzheimer type dementia. *Neuroreport*, *6*, 2287–2292.
- Horwitz, B., Soncrant, T.T., & Haxby, J.V. (1992b). Covariance analysis of functional interactions in the brain using metabolic and blood flow data. In F. Gonzalez-Lima, T. Finkenstaedt, & H. Scheich (Eds.), *Advances in metabolic mapping techniques for brain imaging of behavioral and learning functions* (pp. 189–217). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- John, E. (1961). High nervous functions: Brain functions and learning. *Annual Review of Physiology*, *23*, 451–484.
- Jonides, J., Smith, E.E., Koeppel, R.A., Awh, E., Minoshima, S., & Mintun, M.A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, *363*, 623–625.
- Kenny, D.A., & Judd, C.M. (1984). Estimating nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201–210.
- Kirkby, B., Van Horn, J., Ostrem, J., Weinberger, D., & Berman, K. (1996). Cognitive activation during PET: A case study of monozygotic twins discordant for closed head injury. *Neuropsychologia*, *34*(7), 689–697.
- Lashley, K.S. (1929). *Brain mechanisms and intelligence*. New York: Hafner Publishing Co., Inc.
- 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.
- Loehlin, J.C. (1987). Latent variable models. An introduction to factor, path, and structural analysis. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Luria, A.R. (1962). *Higher cortical functions in man*. New York: Basic Books.
- Mandler, G. (1959). Stimulus variables and subject variables: A caution. *Psychological Review*, 55, 145–149.
- McCarthy, G.M., Blamire, A.M., Puce, A., Nobre, A.C., Bloch, G., Hyder, F., Goldman-Rakic, P., & Shulman, R.G. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. *Proceedings of the National Academy of Science USA*, 91, 8690–8694.
- 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., Nyberg, L., Bookstein, F., & Tulving, E. (1997). Differential functional connectivity of prefrontal and medial temporal cortices during episodic memory retrieval. *Human Brain Mapping*, 5(4), 323–327.
- McIntosh, A.R., Bookstein, F.L., Haxby, J.V., & Grady, C.L. (1996a). Spatial pattern analysis of functional brain images using Partial Least Squares. *Neuroimage*, 3, 143–157.
- McIntosh, A.R., Cabeza, R.E., & Lobaugh, N.J. (1998b). Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus. *Journal of Neurophysiology*, 80, 2790–2796.
- McIntosh, A.R., Cabeza, R., Lobaugh, N.J., Bookstein, F.L., & Houle, S. (1998a). Convergence of neural systems processing stimulus associations and coordinating motor responses. *Cerebral Cortex*, 8, 648–659.
- McIntosh, A.R., & Gonzalez-Lima, F. (1992). The application of structural modeling to metabolic mapping of functional neural systems. In F. Gonzalez-Lima, T. Finkenstadt, & H. Scheich (Eds.), *Advances in metabolic mapping techniques for brain imaging of behavioral and learning functions* (pp. 219–255). Dordrecht, Germany: Kluwer Academic Publishers.
- 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. (1996b). 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.
- McLaughlin, T., Steinberg, B., Christensen, B., Law, I., Parving, A., & Friberg, L. (1992). Potential language and attentional networks revealed through factor analysis of rCBF data measured with SPECT. *Journal of Cerebral Blood Flow and Metabolism*, 12, 535–545.
- Mesulam, M.M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, 10, 309–325.
- 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), 1013–52.
- Miller, G.A., Galanter, E., & Pribram, K.H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart, & 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.
- Nyberg, L. (1994). A structural equation modeling approach to the multiple memory systems question. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2), 485–491.

- Nyberg, L., McIntosh, A.R., Cabeza, R., Nilsson, L.-G., Houle, S., Habib, R., & Tulving, E. (1996a). Network analysis of positron emission tomography regional cerebral blood flow data: Ensemble inhibition during episodic memory retrieval. *Journal of Neuroscience*, *16*, 3753–3759.
- Nyberg, L., McIntosh, A.R., Houle, S., Nilsson, L.-G., & Tulving, E. (1996b). Activation of medial temporal structures during episodic memory retrieval. *Nature*, *380*, 715–717.
- Nyberg, L., Tulving, E., Habib, R., Nilsson, L.-G., Kapur, S., Houle, S., Cabeza, R., & McIntosh, A.R. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport*, *7*, 249–252.
- Petrides, M. (1994). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 59–82). Amsterdam: Elsevier Science B.V.
- Posner, M.I., & Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.
- 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.
- Schacter, D.L., & Tulving, E. (1994). What are the memory systems of 1994? In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). Cambridge, MA: MIT Press.
- Schreurs, B.G., McIntosh, A.R., Bahron, M., Herscovitch, P., Sunderland, T., & Molchan, S.E. (1997). Lateralization and behavioral correlation of changes in regional cerebral blood flow with classical conditioning of the human eyeblink response. *Journal of Neurophysiology*, *77*, 2153–2163.
- 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.
- Somers, D., Nelson, S., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, *15*, 5448–5465.
- Strother, S.C., Kanno, I., & Rottenberg, D.A. (1995). Principal components analysis, variance partitioning, and “functional connectivity”. *Journal of Cerebral Blood Flow and Metabolism*, *15*, 353–360.
- Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain [Mark Rayport, Trans.]. New York: Thieme Medical Publishers, Inc.
- Tsodyks, M., Skaggs, W., Sejnowski, T., & McNaughton, B. (1997). Paradoxical effects of external modulation of inhibitory interneurons. *Journal of Neuroscience*, *17*(11), 4382–4388.
- 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.
- Turova, T. (1997). Stochastic dynamics of a neural network with inhibitory and excitatory connections. *Biosystems*, *40*(1–2), 197–202.
- Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, *274*, 1724–1726.
- Weinberger, N.M., & Diamond, D.M. (1987). Physiological plasticity in auditory cortex: Rapid induction by learning. *Progress in Neurobiology*, *29*(1), 1–55.
- 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.
- Zhang, J., Riehle, A., Requin, J., & Kornblum, S. (1997). Dynamics of single neuron activity in monkey primary motor cortex related to sensorimotor transformation. *Journal of Neuroscience*, *17*(6), 2227–2246.