INTRODUCTION

This chapter addresses some critical issues that arise when performing brain imaging experiments on patients with neurological insult and psychological impairment. In the 1990s, neuropsychology per se has been fundamentally augmented by the ability to measure the neurophysiological correlates of cognitive processing. This has led to a revision of some cognitive models and a shift in emphasis from cognitive science to cognitive neuroscience. Despite the abundant literature from imaging studies and neuropsychology, there are many issues of interpretation that remain unresolved. In this chapter, we consider some basic questions about how neuroimaging can be used to inform the neuropsychological characterization of patients and some of the logical limitations or restrictions on the inferences that can be made. In particular, we focus on the conditions that are necessary to draw tenable conclusions when a patient’s brain activates in an abnormal way relative to normal subjects. The chapter is divided into three sections. In this introductory section, we review the expectations of neuropsychology and neuroimaging. The second section provides some examples of how neuroimaging experiments have been used to inform normal and abnormal models of brain function, and the third section addresses some of the implicit assumptions and limitations that are encountered.

Neuropsychological Studies of Cognitively Impaired Patients

Neuropsychology is the study of patients with functional deficits in which the neuronal pathophysiology is known to a lesser or greater extent. Neuropsychological investigations have contributed to our understanding of normal brain function by informing models of cognition and functional anatomy. Typically, models of cognition are engendered or modified by neuropsychological studies when patients demonstrate a double dissociation in the impairment of selective functions. For instance, different types of dyslexia point to a double dissociation in reading processes: some patients retain the ability to read words with regular spelling-to-sound correspondence but fail to read words that do not follow spelling rules. In contrast, other patients suffer the reverse dissociation. This particular double dissociation has been used to infer the independence of two routes to reading (see Coltheart, 1981).

With respect to normal functional anatomy, the inferences that can be drawn from brain-damaged patients are based on the lesion deficit model. To be informative, the lesion deficit model requires a patient with a selective brain lesion and a selective cog-
nitive deficit. The function of the damaged brain region is simply equated with the lost cognitive skill. Some classic examples of the lesion deficit model, as applied to neuropsychological patients, were documented by the 19th-century neurologists. For instance, postmortem studies demonstrated that a patient who had been impaired at articulating language had damage encompassing the third frontal convolution (Broca, 1861) and a patient with a deficit in speech comprehension had damage to the left posterior temporal cortex (Wernicke, 1874). By deduction, Broca’s area was associated with speech production and Wernicke’s area was associated with speech comprehension. Wernicke developed the model further to predict that patients could have intact speech comprehension and production but a deficit in integrating these regions in order to repeat what was heard. This type of disconnection syndrome, referred to as conduction aphasia, was demonstrated by Lichtheim (1885) in a patient who had damage to the white matter tract that connects Broca’s area with Wernicke’s area (the arcuate fasciculus). By clinical descriptions and localization of lesions, Wernicke and Lichtheim were able to demonstrate that disorders of language arose either from damage to the “centers of memory images” or from disconnections between the so-called centers.

Limitations of the Lesion Deficit Model

The shortcomings of the lesion deficit model are becoming increasingly appreciated. For a number of reasons, it is very difficult to ascribe a function to a particular region that has been damaged. Perhaps the most obvious is that pathological (as opposed to experimental) lesions seldom conform to functionally homogeneous neuroanatomical systems. Furthermore, the neuropsychological profile is usually complicated, involving more than one functional deficit, and these deficits can be obscured by the compensatory measures adopted by the patient to overcome them. Any reasonable relationship between the functional deficit and the brain systems involved is therefore usually impossible to establish. Another problem with the lesion deficit model, which has a history dating from the 19th century (Goltz, 1881) is that the results of lesion studies are properly interpreted only by referring to the connections between cortical areas: damage to a selected area may impair nearby connections, and therefore the responsiveness of undamaged areas. Indeed, it is impossible to distinguish between the impact of a lesion due to the loss of neuronal infrastructure per se and the more pervasive dysfunction of distributed systems of which the lesioned area is a component. These considerations mean that all that can be concluded from a lesion deficit study is that the neuronal systems intrinsic to the lesioned area, or the connections passing through this area, were necessary for the cognitive function. One cannot say
that this region was either sufficient for, or uniquely identifiable with, the function in question.

**Expectations of Neuroimaging with Brain-Damaged Patients**

With the advent of functional neuroimaging, it is hoped that some of the incompleteness of the lesion deficit model can be remedied by studying brain activity in normal and brain-damaged subjects. Functional imaging offers several fundamental advantages over the lesion deficit model. The most obvious is that brain activity can be observed, noninvasively, in subjects who have normal psychological and physiological responses. The other major advantage is that, unlike the lesion deficit model, functional imaging is not limited to a particular region of the brain that has been damaged; rather, the system of distributed cortical areas that sustain sensory, motor, or cognitive tasks can be identified. This systems-level approach has several important implications. First, unlike the lesion deficit model, it does not assume that cognitive processes or operations are localized in discrete anatomical modules, but allows for functional specialization that is embodied in the interactions among two or more areas. In relation to patient studies, the systems-level approach enables one to identify where there is abnormal function in the absence of structural damage and where the responsiveness of an undamaged region is context dependent (responds normally) in some tasks and abnormally in others.

Most neuroimaging studies assume that different tasks will be associated with a different set of cortical areas, and experiments aim to identify the areas where there are changes in regional cerebral activity in response to changes in task or to pathology. Irrespective of whether a cognitive function is localized in one or more than one brain region, the existence of a distinct set of regions for one task relative to another is embodied in the concept of functional segregation. A different perspective on functional brain systems is functional integration. Whereas functional segregation refers to the specialization evident when different cognitive processes are associated with activity in different brain regions, functional integration refers to the integration of these regions where the interactions among regions may be profoundly task dependent (see chapter 3 in this volume). This distinction between studies of functional segregation and functional integration is crucial for imaging patients because some patients suffer from abnormal functional segregation (the function of a discrete cortical area is abnormal) and some patients suffer from abnormal functional integration (abnormal interactions among different brain regions). More specifically, as described by Wernicke and Lichtheim (see above), patients may behave abnormally following (1) damage to an area of gray matter with a particular specialization (e.g., Broca’s
area or Wernicke’s area), (2) damage to white matter that connects gray matter regions (e.g., the arcuate fasciculus), or (3) no detectable pathological damage but a failure to integrate activity during particular tasks. The latter two cases correspond to anatomical and functional disconnection syndromes, respectively, and in these cases a functional deficit will be revealed only by looking at how different regions interact.

Functional imaging can assess these functional interactions in terms of correlations between the responses of two regions. These correlations may be expressed during some tasks but not during others. Functional neuroimaging therefore provides a potential means to test directly the disconnection syndromes described by the 19th-century neurologists. Current advances in diffusion weighted magnetic resonance imaging may allow one to measure lesions in anatomical connections directly. This will, it is hoped, confer a more integrative perspective on the lesion deficit model.

Functional neuroimaging experiments aim to characterize the basic relationship between the cognitive processes elicited by a task and the neuronal responses that underpin them. In this chapter, the discussion of this relationship will be in terms of the task analysis, cognitive architectures, and neuronal architectures. A task consists of the decomposition of a task into these component processes, whose existence is inferred on the basis of neuropsychological, psychophysiological, and psychophysical studies. The cognitive architecture is any particular set of processes and the serial or parallel interactions among them. The neuronal architecture relates to how a particular cognitive architecture is implemented in the brain by neuronal dynamics in distributed cortical areas, subareas, and neuronal populations. It is the relationship between the neuronal and cognitive architecture that underlies the interpretation of brain imaging data in the context of neuropsychological impairment. This requires a comprehensive and valid task analysis that specifies the cognitive architecture and enables mapping from the cognitive to the neural domain. The problems of specifying a complete task analysis are addressed below in the section “Issues.”

With respect to neuroimaging studies of cognitively impaired patients, the most seemingly straightforward application is the investigation of how a neuropsychological deficit is characterized in terms of abnormal brain function. The idea is that brain activity observed when patients perform a task can be compared with that when normal subjects perform the same task. Differences between patients and normals can then be ascribed to the neuropsychological syndrome. In other words, it might be expected that the alteration of neuronal responses in the damaged brain will shed some light on the physiological underpinnings of a cognitive deficit. There are fundamental limitations when using neuroimaging in this way. We will discuss these limita-
tions in the section “Issues.” The most critical point relates to task performance. By definition, patients will have reduced performance relative to normals on tasks that reveal a cognitive impairment. In severe cases, the patients may not be able to perform a task at all. In this case it is meaningless to perform neuroimaging experiments that attempt to compare normal and abnormal brain activity because failure to activate could be due either to a loss of neuronal responsiveness or to a failure to perform the task. Neuroimaging studies of patients are therefore interpretable relative to normal subjects only if (1) the patients retain the ability to perform the task normally or (2) the patients activate normally (i.e., one variable is kept constant).

Of course, these restrictions severely limit the range of experiments that can be performed in an ideal fashion with patients. They also have a critical impact on the types of questions that can be addressed. Nevertheless, it may still be possible to equate cognitive symptoms to abnormal neuronal responses when patients perform significantly above chance but significantly below normal controls. For example, in the case of reduced accuracy, analysis of functional imaging data can model correct responses separately from incorrect responses. Patients can then be compared to normals on trials (or blocks of trials) where performance is matched. This is most effectively achieved by analyzing single events (trials) using an event-related design (in fMRI or ERP studies), but if performance is only mildly impaired, it is also possible to analyze blocks of trials (in PET or blocked fMRI designs) where the mean normal and patient performances are the same. Another approach to discounting performance differences is to enter normal and patient response measures (e.g., reaction times) into the analysis as confounds.

Once performance is matched, abnormal neuronal responses can be attributed to changes in either the cognitive architecture (cognitive reorganization) or changes in the neuronal architecture (neuronal reorganization). Cognitive reorganization occurs when a patient uses a different set of cognitive processes to perform the same task either because a new cognitive procedure has been learned or because of increased demands on normal cognitive processes, particularly attention. Neuronal reorganization is mediated by changes in the strength of preexisting connections. It does not indicate a rewiring of the neuronal system because in the mature brain, neuronal systems and extrinsic connections are fully established. Changes in the strength of pre-existing connections may result from learning-dependent plasticity or may simply be a direct consequence of brain damage that can disrupt neuronal responses at, or distant from, the lesion site.

Evidence for either cognitive or neuronal reorganization (in the context of normal task performance) has implications for both abnormal and normal models of processing. In either case, abnormal activations fall into three categories: (1) activation is
greater for normals than for patients (underactivity); (2) activation is greater for patients than normals (overactivity); and (3) the same regions activate but there are alterations in the effective connectivity between regions. Table 11.1 summarizes how abnormal activation in the context of normal performance can inform abnormal models of cognitive processing.

Table 11.1
Informing Models of Abnormal Cognitive Processing from Abnormal Brain Activation

<table>
<thead>
<tr>
<th>Underactivation</th>
<th>Overactivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Damage to local tissue</td>
<td></td>
</tr>
<tr>
<td>2. Damage to distant tissue (e.g., a diaschisis)</td>
<td></td>
</tr>
<tr>
<td>3. Cognitive reorganization</td>
<td></td>
</tr>
<tr>
<td>1. Learning-related plasticity</td>
<td></td>
</tr>
<tr>
<td>2. Disinhibition of duplicated neuronal system</td>
<td></td>
</tr>
<tr>
<td>3. Cognitive reorganization</td>
<td></td>
</tr>
</tbody>
</table>

Table 11.2
Informing Models of Normal Cognitive Processing from Abnormal Brain Activation

<table>
<thead>
<tr>
<th>Underactivation</th>
<th>Overactivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Redundant area (area not necessary for task performance)</td>
<td></td>
</tr>
<tr>
<td>2. Connections between damaged and underactive area</td>
<td></td>
</tr>
<tr>
<td>1. Duplicated neuronal system</td>
<td></td>
</tr>
<tr>
<td>2. Alternative cognitive strategies</td>
<td></td>
</tr>
</tbody>
</table>

greater for normals than for patients (underactivity); (2) activation is greater for patients than normals (overactivity); and (3) the same regions activate but there are alterations in the effective connectivity between regions. Table 11.1 summarizes how abnormal activation in the context of normal performance can inform abnormal models of cognitive processing.

Table 11.1 indicates that areas of underactivity in the patients could indicate damage to the tissue itself, damage to the inputs into the underresponsive region (a diaschisis), or changes in cognitive strategy (cognitive reorganization). In contrast, areas of overactivity in patients can reflect a disinhibition of a duplicate neuronal system that can implement the same task, learning-related plasticity, or cognitive reorganization. Table 11.2 summarizes how abnormal activation in the context of normal performance can inform normal models of cognitive processing.

Table 11.2 indicates that areas of underactivity (in the context of normal task performance) can be used to infer that normal activation was not necessary to perform the task (activation was redundant). If there is a known site of brain damage that is distant from a region of underactivity, we can also infer that the damaged region encompassed inputs to the underactive region (i.e., make implications about the normal connections between regions). Areas of overactivity in the patient (in the context of normal task performance) can be used to infer which regions of the brain are able to take over the function of damaged regions. The more difficult issue is to distinguish whether the functional reorganization is cognitive or neuronal in nature (see the section “Issues”). The ideas summarized in tables 11.1 and 11.2 are reprised in the next
section, which provides examples of the types of neuroimaging studies that can be performed with neurologically damaged patients.

FUNCTIONAL NEUROIMAGING OF NEUROPSYCHOLOGICALLY IMPAIRED PATIENTS

This section is divided into two parts. The first, “Informing Models of Abnormal Functional Anatomy,” provides examples of how functional imaging can inform models of abnormal processing; the second, “Informing Models of Normal Functional Anatomy,” provides examples of how functional imaging can inform models of normal processing.

Informing Models of Abnormal Functional Anatomy

The advantage of functional neuroimaging, as discussed above, is that it can identify distributed brain systems responding to a particular task. This means that, unlike the lesion deficit model, functional neuroimaging can detect (1) normal functionality at a site of brain damage, (2) abnormal functionality distant from the site of known brain damage, and (3) abnormal functionality in the absence of brain damage. Critically, some of these effects may be task specific, depending on which inputs are being used. In patients, for example, a nondamaged region can perform normally in some tasks and abnormally in other tasks, depending on whether input from a disconnected region is required.

Below we report functional imaging studies of (1) semantic dementia and (2) Broca's aphasia, where dysfunctional integration can be attributed directly to structural damage; (3) studies of schizophrenia where there is no obvious anatomical damage to account for the functional disintegration; and (4) studies of language recovery following aphasia, where functionality is preserved by activation of peridamage tissue, duplicated neuronal systems, or cognitive reorganization.

“Semantic dementia” is the term used to describe patients suffering from a progressive deterioration in semantic knowledge and name retrieval while other cognitive and language functions remain relatively intact (Warrington, 1975; Sasanuma and Monoi, 1975; Hodges et al., 1992). The anatomical correlates, according to the lesion deficit model, lie in bilateral temporal lobes; cortical atrophy commences in the anterior temporal poles and then spreads posteriorly as the disease progresses. Application of the lesion deficit model leads to the inference that the damaged anterior temporal lobes are the site of the impaired semantic and naming processes. Functional neuroimaging allows us to evaluate whether patients show response changes in the damaged anterior temporal lobe or whether damage to the anterior temporal lobe results in abnormal activation in other (intact) regions of the language system.
This approach was used by Mummery et al. (1999) with four semantic patients and six age-matched control subjects. Semantic similarity judgments were used to activate the semantic system, and perceptual (visual size) judgments were used as a control condition. In order to exclude the possibility that the abnormal responses detected in the patients were a direct consequence of impaired performance (see above), scans in which patients had impaired performance relative to the normals were removed from the analysis. The results revealed that both patients and normals activated the left inferior frontal, left temporoparietal cortex, left middle temporal cortex, anterior cingulate, and right cerebellum. Unlike the normals, the patients did not activate the left posterior inferior temporal cortex or the right temporoparietal junction. Neither of these regions evidenced any structural damage on structural MRI scans, but damage to the left posterior inferior temporal region in other patients has been associated with naming deficits (Foundas et al., 1998) and may account for the impairment all the patients had in naming.

Remarkably, the regions where there was structural damage (the anterior temporal cortices) were more active in the patients than in the normals. These results demonstrate that the site of reduced activation was distant from the site of structural damage and suggest a functional disintegration between the anterior and posterior temporal cortices. If this is the case, we might expect the posterior temporal region to perform normally in other tasks that do not require input from the anterior temporal region. This has not yet been investigated with the semantic dementia patients. However, an example of how an underresponsive region can respond normally in one task but abnormally in another is given below in the context of patients with damage to Broca’s area.

Broca’s aphasia was first described at the end of the 19th century (Broca, 1861) in the context of a patient who had impaired speech production but relatively intact speech comprehension following damage to the posterior inferior frontal cortex in the left hemisphere (see above). By scanning such patients during language tasks they are able to perform, the effect of damage to Broca’s area on other undamaged cortical regions can be assessed. For this purpose, we (Price, Warburton, et al., 1999) assessed regional activation when four patients with lesions to Broca’s area performed a simple visual feature detection task on words, relative to the same task on consonant letter strings. Like the normal controls, all four patients activated a region of the left middle temporal lobe associated with semantic processing (Vandenberghhe et al., 1996) but failed to activate Broca’s area or a region in the left posterior inferior temporal cortex that is activated when subjects name visual or tactile stimuli (Buechel et al., 1998; Price, 1998).

The lack of activation in Broca’s area confirmed that damage to this region had rendered it dysfunctional. More interesting were the abnormal responses in the left
posterior inferior temporal cortex. This region is seldom damaged by stroke because it is supplied by the posterior and middle cerebral arteries. Therefore its function has not, until recently, been associated with naming using the lesion deficit model (Raymer et al., 1997; Foundas et al., 1998). Certainly, the effect of damage to Broca’s area on the left posterior inferior temporal cortex could not have been inferred without functional neuroimaging. The question that remains concerns whether damage to Broca’s area renders the left posterior inferior temporal cortex permanently inactive or whether dysfunctionality depends on the involvement of Broca’s area. This was evaluated in two ways. First, we note that it was not simply the case that the left posterior inferior temporal cortex failed to activate; the abnormal responses were characterized by a deactivation. In other words, the responses were determined by the task with decreased activation when interaction with Broca’s area was required. The second approach was to scan one of the patients again, using the semantic similarity paradigm described for the patients with semantic dementia (see above). In this paradigm, which emphasizes temporoparietal interactions, the region that showed underactivity during the implicit reading paradigm activated normally (and the interaction between paradigm and pathology was significant). The inference from both accounts is that responses in the left posterior inferior temporal cortex were context dependent, with abnormality expressed only when input from Broca’s area was required. In summary, functional neuroimaging can be used to assess the effect of brain damage on brain activity at, and distant from, the site of damage. The context-sensitive abnormalities detected in the left posterior inferior temporal cortex speak to the integrative nature of neuronal architectures.

We turn now to studies of schizophrenia where functional deficits have not been attributed to obvious anatomical damage. Schizophrenia is a psychiatric condition in which patients suffer from the intermittent recurrence of one or more of at least three symptom types: psychomotor poverty (a lack of volition, e.g., to initiate motor movements or speech); reality disorder (hallucinations and delusions) in the context of normal cognition; and disorganized and incoherent speech and thought. Studies that have attempted to correlate the syndrome, or particular symptoms of the syndrome, with changes in cortical responsiveness (see above) have revealed inconsistent results. For instance, some studies have shown reduced responses in the frontal lobes (Weinberger et al., 1988; Berman et al., 1988, 1995), whereas others have shown normal prefrontal responses (Frith et al., 1995; Fletcher et al., 1996).

The inconsistency has been related in part to the finding that hypofrontality correlates only with the expression of psychomotor poverty (Liddle et al., 1992) and in part to the fact that the performance of patients is not always matched with that of normals (Frith et al., 1995). An alternative explanation is that some of the symptoms of schizophrenia do not necessarily reflect a regionally specific pathology; rather, there
is abnormal integration between different regions that may function normally when they are not required to interact. This would account for why, in schizophrenia, a region such as the frontal lobe might show normal activity in some contexts and abnormal activity in others: the responsiveness of the frontal region depends on the neuronal architecture and implicitly on the cognitive architecture engendered by the task.

Studies of abnormal functional integration can be assessed by measuring changes in the functional connectivity between regions. Essentially these measurements are based on temporal correlations between activity in distant cortical regions. In electrophysiological studies, which record spike trains of neural activity, the temporal scale is on the order of milliseconds. In functional neuroimaging, which measures hemodynamic changes, the temporal scale is on the order of seconds and a significant correlation simply implies that activity (pooled over the time scale) goes up and down together in distant regions. Such temporal correlations imply functional connectivity that could be of two critically different types. One type results from direct connections between the correlated regions (i.e., activity changes in one region cause activity changes in another region). The second type does not imply direct connections between correlated regions, but different regions may share connections from a region that is the source of correlated activity. The important point of this distinction is that functional connectivity does not necessarily imply direct connections between correlated regions.

Studies of functional connectivity in schizophrenia have shown that there are abnormal correlations between activity in the prefrontal and temporal regions during word generation tasks. More specifically, in normal subjects, activity in bilateral superior temporal cortices during word generation (relative to word repetition) is negatively correlated with activity in the prefrontal cortex, but in three groups of patients with schizophrenia, activity in the left superior temporal cortex was positively correlated with prefrontal activity (see Friston & Frith, 1995). These results, illustrating a complete reversal of the large-scale prefronto-temporal interactions in the schizophrenics, indicate consistent abnormalities in regionally specific functional connectivity. The reversed correlations can be regarded as a failure of prefrontal cortex to suppress activity in the temporal lobes (or vice versa). One behavioral interpretation that has been offered by Friston and Frith (1995) is that the prefrontal regions are necessary for intrinsically generated behavior and the bilateral temporal lobes are sensory perception regions that register the consequences of behavior (Frith et al., 1991). A failure to integrate these two regions may impair (1) intrinsically generated action, as in psychomotor poverty, and (2) perception, as in hallucinations and delusions, when self-induced sensory changes are attributed to an external cause. In other words, coherent interactions between prefrontal cortices and cortices devoted to perceptual
representations may be crucial for the integration of intrinsically generated behavior and perception. Further studies are being conducted to explore the prefronto-temporal disintegration in schizophrenia. As noted above, correlated activity could result from direct causal connections between the frontal and temporal regions or from shared influences from a third region. One hypothesis (Dolan et al., 1995) is that normal frontotemporal integration during the word generation paradigm is modulated by activity in the anterior cingulate (i.e., the anterior cingulate governs the negative correlations between frontal and temporal regions). This hypothesis could be tested with studies of effective connectivity (see Friston et al., 1997) and structural equation modeling. (See chapter 3 in this volume for further details on this sort of network analysis.)

Finally, we consider mechanisms of functional recovery. How patients might recover a lost function is one of the most crucial questions that needs to be addressed by imaging studies of brain-damaged patients. Structural indices of lesions (from conventional use of CT and MRI scanners) do not necessarily imply a complete loss of function, and it is sometimes surprising when a patient with a large lesion makes an unexpectedly good recovery. Functional neuroimaging, by contrast, can detect areas where a degree of functional responsiveness has been maintained even in areas that appear damaged in structural images. Typically these areas are around the region of insult (e.g., peri-infarct tissue) and sometimes within the lesion. Recovery of a lost function results either from the reactivation of tissue that was initially incapacitated (e.g., due to a reduction in edema) or from increases in the capacity of viable tissue until it can support a function that was originally executed by lost cells.

Functional imaging has an important role to play in evaluating the contribution of these mechanisms. However, to date it has probably been grossly underestimated. This is because functional imaging studies have usually been able to make inferences only by pooling data from different patients into one group and then comparing the patient group to a group of normal subjects (e.g., Weiller et al., 1995). Since peri-infarct activity inevitably varies from patient to patient, depending on the size and location of the lesion, it will not be detected in group-to-group comparisons. The demonstration of peri-infarct activity therefore relies on studies where each subject is analyzed individually.

In a study by Warburton et al. (1999), six patients with large left temporoparietal lesions who had lost, and then recovered, the ability to generate words were scanned six times during a word generation task and six times during rest. Data from each patient were analyzed independently and compared to a group of nine control subjects. In normal subjects, the word generation task (relative to rest) consistently activates a widely distributed system of language regions in the left hemisphere (in
particular the left prefrontal and posterior temporal cortices). By analyzing data from each subject individually it was possible to ascertain that half the normal subjects also activate the same set of regions in the right hemisphere. All six recovered aphasics also evidenced activation in the left prefrontal regions, all but one activated the damaged left temporal lobe, and half activated the right prefrontal and temporal cortices. The consistent activation in the damaged left temporal lobe in all but one patient demonstrated varying degrees of peri-infarct activity. This left temporal activity was not detected when the patients were pooled together for a group analysis. The conclusions of this study were that activations associated with cued word retrieval in the recovered aphasics were indistinguishable from those of the normal controls, except that in the presence of a lesion the activations were perilesional. Similar results have been obtained in a single case study of a patient who had recovered from auditory agnosia (Engelein et al., 1995). Furthermore, Heiss et al. (1997) have demonstrated, in a longitudinal study, that the recovery from aphasia is related to the reactivation of left hemispheric speech areas surrounding the area of infarction.

In contrast, other studies have suggested that recovery occurs following a laterality shift, with homologous regions in the contralateral cortex assuming the functions of the damaged region (e.g., Weiller et al., 1995; Buckner et al., 1996). Such a mechanism, when a different neuronal architecture sustains the same cognitive architecture, indicates that there may be a duplicated language system in the right hemisphere which can be brought into action, perhaps by disinhibition, following damage to the left hemisphere language system (see “Introduction”). However, there are a number of methodological issues that need to be resolved before such conclusions can be resolved, in particular, whether the patients (1) performed the task using the same cognitive architecture as normal subjects and (2) activated outside the normal range of controls. For example, the Warburton et al. study (1999) illustrated that almost half the normal subjects respond to language tasks by activating the right hemisphere. If these are the subjects who are most likely to recover following damage to the left hemisphere system, then patients who have recovered language abilities should ideally be contrasted to normals who also activate in the right hemisphere. Only when the patients (1) perform like normals and (2) activate more than normals with bilateral language function can inferences regarding disinhibition of a right hemisphere language system be made. To our knowledge, these criteria have not yet been met.

Informing Models of Normal Functional Anatomy

In the “Introduction,” we discussed how the lesion deficit model (with neuropsychological patients) is limited because although it is able to identify regions that are...
necessary to perform a task, it does not establish the premorbid sufficiency of the
damaged regions. For instance, a cognitive function can be impaired if the connec-
tions between two vital cortical areas are damaged (the connections are necessary
for performance but they are not sufficient). In contrast, functional neuroimaging in
normal subjects reveals distributed brain systems that can be considered sufficient to
perform a task, but it does not distinguish the relative contributions of the subcom-
ponents involved. Some activated regions may be superfluous to the task require-
ments (Price et al., 1996).

The joint and complementary use of neuroimaging and neuropsychology offers a
fundamental advantage over either technique in isolation. Neuroimaging in normal
subjects defines the sufficient set of regions (the neural architecture) for performing
one task relative to another. Neuropsychology establishes the necessity of component
brain areas in one of three ways. The first, most conventional, approach has been
described above and involves identifying the lesion site associated with a functional
deficit. By implication, this region was necessary for the specified function. The sec-
ond approach looks at the effect of a lesion on a region identified in a neuroimaging
paradigm. For example, a behavioral study of a patient with a right cerebellar infarct
(Fiez et al., 1992) was motivated by the observation that functional imaging studies
show activity in the right cerebellum during verbal fluency (Petersen et al., 1988).
On nonmotor tasks, the patient showed deficits in completing and learning a word
generation task but had normal or above normal behavior when performing stan-
dardized language tasks. In this instance, a neuroimaging study motivated a neuro-
psychological study and the neuropsychological study allowed inferences to be made
regarding a subcomponent of the language system. The third approach involves
inferences from patients who are not functionally impaired on a specified task but
nevertheless have damage to parts of the system defined by neuroimaging. Here the
damaged regions can be construed as not necessary. By designating each region in the
sufficient system as necessary or not necessary, the critical system can be identified.
However, the caveat is that some patients may be able to perform a task by activat-
ing peri-infarct tissue that appears to be damaged in routine structural imaging (see
above). Another possibility is that functionality is preserved due to neuronal re-
organization (e.g., involving the homologue region in the contralateral hemisphere;
Weiller et al., 1995; Buckner et al., 1996) or cognitive reorganization. To discount
these possibilities, functional imaging of the patient is a prerequisite.

Two attempts have been made to harness neuroimaging and neuropsychology in
order to investigate the role of the prefrontal cortex in linguistic processing. Neuro-
imaging studies of patients with frontal lobe lesions are able to determine whether (1)
patients retain the ability to perform some linguistic tasks without the left prefrontal
cortex; (2) there is peri-infarct activation in the left prefrontal cortex; or (3) there is compensatory activity in the right prefrontal cortex. In one study, Buckner et al. (1996) used functional neuroimaging and a stem completion task (generate words from word stems such as “TRO”) to demonstrate that a patient with left frontal lobe damage retained the ability to perform the task by activating the right inferior frontal cortex. In the other study, Price, Mummery, et al. (1999) used semantic similarity judgments and found that a patient with a large frontoparietal lesion can perform the task accurately by activating temporoparietal regions in the absence of either left or right frontal activity. By discounting peri-infarct activity in the left frontal cortex and cognitive or neuronal reorganization in the right frontal cortex, this result indicates that the left prefrontal activity, consistently seen in normals, is not necessary to perform the semantic similarity judgment task.

If the left prefrontal activation is not necessary for semantic similarity judgments, why is it activated normally? One possibility is that it relates to implicit memory processes, but to confirm this we would need to show that the patient was less able to remember the stimuli at a later date. Another question that is raised relates to whether other regions of the semantic system are not necessary for task performance. An indication that the left posterior inferior temporal cortex is not required comes from the observation (reported above) that the patients with semantic dementia failed to activate this area during the same task. Is it possible, then, that there is no critical semantic area but, rather, that other components of the system are able to compensate if one area is malfunctioning? This hypothesis needs to be tested explicitly, but it appears that some of the left extrasylvian temporal regions are critical, because damage to the ventral anterior temporal cortex results in semantic deficits (Hodges et al., 1992), and damage to the posterior inferior parietal cortex can result in speech comprehension deficits (Alexander et al., 1989). The set of regions where normals activate, and damage results in an inability to perform semantic tasks, are those which will constitute the necessary and sufficient neural system. Ideally, in order to delineate the complete necessary and sufficient brain system involved in semantic similarity judgments (or any other task), we need to image patients with lesions to each component of the system identified in normal subjects. In this way, functional imaging and neuropsychology can be combined to make inferences about functional anatomy that could not be done with either alone. Figure 11.1 illustrates the inferences that can be drawn from neuropsychology, neuroimaging on normal subjects, and neuroimaging on patients.

The other important ways that functional imaging studies of patients can inform models of normal functional anatomy have been summarized in table 11.2. One relates to the effect that regionally specific brain damage has on the responses in...
Figure 11.1
Examples of inferences that can be drawn from neuropsychology and the lesion deficit model, neuroimaging studies on normal subjects, and neuroimaging studies on patients. In the top row, the extent of the lesion is illustrated in black on models of the left and right hemispheres. In the second and third rows, respectively, the areas activated when normal subjects and the patient perform semantic similarity judgments relative to perceptual tasks are indicated in black. In the bottom row, the areas that are activated by each normal control but not by the patient are shown in black. Since the patient could perform the task within normal limits, we conclude that the left inferior frontal cortex is not necessary for semantic similarity judgments (see Price, Mummery, et al., 1999).
undamaged areas of the brain. For instance, the finding that the left posterior inferior temporal lobe failed to activate during semantic similarity judgments in the context of anterior temporal damage (Mummery et al., 1999), but not in the context of left frontoparietal damage (Price, Mummery, et al., 1999), indicates that during semantic judgments, inputs to the posterior temporal region come from the anterior temporal region. Similarly, the finding that in the context of left frontoparietal damage the left posterior inferior temporal cortex activated normally during semantic similarity judgments, but not during reading (Price, Warburton, et al., 1999), indicates that this region relies on inputs from the damaged frontal region during reading but not during semantic similarity judgments.

ISSUES

In this section, we address the issues encountered when designing and interpreting experiments that measure the brain function of patients with neuropsychological impairment. The key to interpreting neuroimaging results lies in experimental design and task analysis that allow one to disambiguate between a cognitive change, a neuronal change, or both.

The first consideration in designing experiments to study a brain-damaged patient is how the results are going to inform our understanding of the neuronal and cognitive architectures pertinent to that patient or the normal population. The critical observation in neuropsychological imaging is a differential pattern of activation in the patient relative to a normal group (a subject group-by-task interaction). As described above, this can be attributed to either neuronal or cognitive reorganization.

Distinguishing between Neuronal and Cognitive Reorganization

Neuronal reorganization (or plasticity) refers to the changes in a task-specific neuronal architecture that take place during learning or relearning in the normal or damaged brain. A key distinction can be made between plasticity that is “enduring” (the time course varies from hours to years) and plasticity that is “dynamic” (the time course ranges from several milliseconds to minutes). In the postdevelopmental period both are mediated by changes in the strength of preexisting connections (the efficacy of existing synapses is altered). Dynamic changes in connection strengths can be mediated by factors that are intrinsic or extrinsic to the neuronal processes themselves, such as the recent history of neuronal firing (e.g., facilitation, adaptation, and potentiation) or inputs from other neuronal populations that modulate connection strengths either by postsynaptic mechanisms or, more directly, by release of modulatory neurotransmitters such as noradrenaline. Enduring plastic changes consolidate
the dynamic changes into permanent changes. During development, new extrinsic connections (axons that traverse the white matter from area to area) can be formed through axonal sprouting and synaptic remodeling. However, in the mature brain, the gross schema of extrinsic connectivity is thought to be fairly fixed.

The point that neuronal reorganization results from changes in a preexisting system is an important factor for interpreting changes in neuronal activity following brain damage. Another important point is that dynamic plasticity can be expressed in many natural and experimental contexts, ranging from changes that underlie attentional modulation to profound changes in the organization of somatosensory fields shortly after deafferentation. For example, rapid neuronal reorganization can occur when a system is disinhibited following deafferentation of the inhibiting inputs (Buonomano & Merzenich, 1998). It follows that it would be incorrect to ascribe changes in attentional set or disinhibition to substantial remodeling of the anatomical connections. In summary, to demonstrate that neuronal reorganization of a dynamic or enduring nature has occurred, it is necessary to show that the cognitive architecture (including the attentional set and performance level) is the same in both patients and normal subjects. This may not be possible, particularly because the cognitive architecture is likely to change as the neuronal implementation changes. Similarly, if the cognitive architecture changes, the underlying neuronal architecture is likely to change.

Cognitive reorganization takes place when a patient uses a different set of cognitive processes to implement the same task, for instance, when a patient learns a particular strategy to recover the ability to perform a lost function. A more specific example is when dyslexic patients adopt a serial letter-by-letter reading strategy to compensate for a deficit in parallel letter processing. In this case cognitive, but not neuronal, reorganization has taken place. In order to demonstrate, with functional imaging, that patients are using a different cognitive architecture to perform the same task, reorganization at a neuronal level needs to be discounted. This would involve demonstrating that an equivalent activation pattern is elicited in normals when they are forced to use the same cognitive strategy as the patients. To our knowledge this experimental technique has not been explored, and it is likely to be extremely hard to implicate. We are therefore left with a dilemma. In order to demonstrate that differential activations reflect cognitive reorganization, it is necessary to show that there has been no change in the neuronal architecture, and in order to infer neuronal reorganization, it is necessary to exclude a change in the cognitive architecture. The distinction rests on the task analysis. However, a task analysis is seldom sufficiently detailed to ensure a constant cognitive architecture and thereby to demonstrate plasticity. Nevertheless, there is a continuum of task analysis depth—the more detailed the task analysis, the more valid the inferences about changes in neural implementation.
Task Analysis

There are several levels of task specification. At one extreme, there must be a one-to-one mapping between the cognitive architecture and its implementation; at this level, a change in neuronal dynamics implies a respective change in the functional correlates. At progressively coarser levels of analysis, where the elemental operations are less specified, many different functional operations could be employed to achieve the same task performance; a complicated relationship between the cognitive architecture and its neuronal implementation ensues. This may be seen as a one-to-many mapping (the same cognitive process can be implemented by different neuronal systems, e.g., in the right and left hemispheres) or a many-to-one mapping (different cognitive process are implemented by the same neuronal system). The point being made here is that although any task analysis may be valid at its own level of specification, it is useful in discriminating between cognitive and neuronal reorganization only if it is as detailed as possible. If the task analysis is very unspecified in terms of the elemental operations, then implied differences in neuronal implementation (plasticity) between the normals and the patient become specious. We mean this in the sense that although differential activations may emulate a plastic reorganization of the neuronal implementation, this may simply reflect the fact that different subprocesses are being called upon at a cognitive level which have not been addressed in the task analysis.

Unfortunately, in practice, a task analysis will never be as refined or as comprehensive as one would like because certain attributes of cognitive processing are not amenable to measurement. For example, subtle changes in attentional or cognitive set may bring about plastic changes in the neuronal architecture through neuromodulatory mechanisms. One of the more important factors of this sort is time, because performing a given task continuously means that the temporal context is changing, and this may evoke time-dependent plasticity in the neuronal implementation. Another example is incidental and implicit processing that may not be required for a task (see Price et al., 1996). A complete task analysis therefore should include not only the processing components but also all the contextual factors that may influence them.

To summarize, in order to make sense of neuropsychological studies with imaging, one has to have a sufficiently comprehensive task analysis to enable one to say that the cognitive architecture elicited by task performance in the patient and the control group are identical. Only when this is the case does a significant interaction imply that plastic changes in neuronal implementation have occurred. If the cognitive architectures are not demonstrably equivalent, then cognitive reorganization may be an explanation. The nature of this reorganization and the underlying neuronal architecture can be defined when normal subjects are coerced into adopting cognitive strategies
that emulate those used by the patient (rendering the cognitive architectures the same). Clearly, demonstrating interactions when the cognitive architectures are not identical is still useful in highlighting candidate brain regions. However, it is logically impossible to ascribe these interactions to plasticity or to change in cognitive strategy unless the constraints described above are applied.

Finally, the interpretation of abnormal activations shown by a particular type of patient will be informative only when the normal functional anatomy of the cognitive task, and all the compensatory measures that might be used to perform it, are understood. The usefulness of patient studies can therefore grow only in relation to our understanding of normal functional anatomy, which in turn depends on neuroimaging studies of normal subjects.

REFERENCES


