

INTRODUCTION TO FUNCTIONAL NEUROIMAGING

Luis Hernandez

Tor Wager

John Jonides

The University of Michigan

Address correspondence to:

Dr. John Jonides

Department of Psychology

University of Michigan

525 E. University Ave.

Ann Arbor, MI 48109-1109

Phone: 734-764-0192

Fax: 734-994-7157

E-mail: jjonides@umich.edu

There has been explosive interest in the use of brain imaging to study cognitive and affective processes in recent years. Examine Figure 1, for example, to see the dramatic rise in numbers of publications from 1992 to 1999 in which the term, “functional Magnetic Resonance Imaging” (fMRI) appears in the title. Because of the surge of empirical work that now relies on a combination of behavioral and neuroimaging data, it is critical for students of the mind also to be students of the brain because data about each informs the other. Our goal in this chapter is to provide an introduction to the growing field of neuroimaging research for those not expert in it. The chapter provides general coverage of the various steps involved in conducting a neuroimaging experiment, from task design to interpretation of the results. We begin by detailing several reasons one might want to use neuroimaging data to understand cognitive and other processes. Having provided this motivation, we then trace out several techniques that are used in the design and execution of imaging experiments. Finally, in the last section of the chapter, we provide a detailed overview of positron emission tomography (PET) and fMRI: a review of the physics underlying each technique and the analysis tools that can be used to work with the resulting data. In these three sections, we hope to illustrate to the reader the why, the what, and the how of functional neuroimaging.

[Insert Figure 1 about here.]

I. The Why: Uses of data from functional neuroimaging

a. Brain mapping.

Perhaps the most obvious rationale for conducting functional neuroimaging experiments is to correlate structure with function. While some psychologists in the last century argued that the brain operated by the principle of mass action (Lashley, 1950), we now know that there is substantial localization of many functions in the neural tissue of the brain. Knowing this, many investigators have

sought to map out the primitive processes that are engaged when various brain structures are active. In a certain gross sense, modern neuroimaging is similar to the 18th century practice of phrenology, whose practitioners read a patient's personality traits from bumps on their skulls. To be sure, both modern functional imaging and phrenology are attempts to map out the localization of function in the brain. But the similarity ends there. Modern neuroimaging measures processes within the brain that are replicable and have been extensively cross-validated with other neuroscientific methodologies. Phrenology, of course, turned out to be wrong. However, it is instructive to compare the assumptions of phrenology with those of modern neuroimaging. Phrenologists believed that a lump at a certain place on the head corresponded with a particular personality trait; the larger the lump, the more of that trait. So, for example, a larger bump might indicate more agreeableness, or a better memory. In neuroimaging, by contrast, it is assumed that complex psychological processes are best described in terms of combinations of constituent elementary operations. The elementary processes may not be localized in single locations in the brain. Rather, they are often the result of networks of neurons (often spatially distributed) acting together. And unlike phrenology, most modern researchers do not assume that skill at one mental operation is a function of the sheer size of the underlying neural tissue involved. The assumptions of neuroimaging lead naturally to a search for the brain activations that accompany elementary psychological processes. Mapping these elementary processes onto regions and functional networks in the brain is a major goal of modern brain imaging research.

We should note that once certain regions of the brain have been identified with certain psychological processes, this provides an opportunity to go beyond simple assignment of structure to function. It provides an opportunity to examine circuits of activation that might be involved in a complex psychological task using statistical techniques such as inter-regional correlations, factor analysis, and structural equation modeling, which we shall review below. The added value of these techniques is that they permit us to go beyond the functions of any single region or small set of regions involved in an elementary cognitive operation. These tools can be used to help analyze what combinations of elementary processes are involved in a psychological task. Thus, we can go from the elementary to the complex by examining patterns of activation and knowing the functions of the structures that are activated in a pattern.

Overall, the sort of behavioral neurology that is provided by studies of functional neuroimaging is quite helpful on several fronts. A detailed mapping of the functions of various brain structures will give us solid evidence about the primitive psychological processes of the brain. It will also provide detailed information for neurosurgical planning. And, it will allow us to predict what functions will be lost on the occasion of brain injury, whether focal or diffuse. Thus, if there were no other reason to conduct studies that use functional neuroimaging, mapping the brain would be sufficient reason. However, there *are* other reasons as well.

b. Dissociating psychological processes

One of the great benefits of having data on the patterns of activation due to one psychological task versus another is that it permits one to examine whether the two tasks doubly dissociate (Smith and Jonides, 1995). The logic is this: Suppose there is some brain region A that mediates some cognitive process a. Suppose, similarly, that there is some other brain region B that mediates some other cognitive process b. Now imagine that we can devise two psychological tasks, 1 and 2, such that Task 1 requires cognitive process a but not b, and Task 2 requires cognitive process b but not a. If we have subjects perform these two tasks while we image the activations in their brains, we should find activation of region A during performance of Task 1, but not performance of Task 2, and vice versa for region B. This pattern of evidence would permit one to argue that there are two separable psychological processes involved in the two tasks, as there are two brain regions that are activated. This logic applies, by the way, whether regions A and B are single sites in the brain or networks of sites, thus generalizing the method to a wide variety of circumstances.

Now consider a similar but more complex case. Suppose that Task 1 requires several psychological processes as does Task 2. By the assumptions outlined above, we should find activations in various regions of the brain when subjects engage in Task 1 and Task 2. If Task 1 activates some group of sites that are wholly different from those activated while subjects engage in Task 2, we would have

evidence of differing processes in the two tasks. However, the two tasks may activate some quite different sites as well as some sites in common. In this case, we get leverage in accounting for the processes involved in the two tasks by noting the sites whose activation they share and the sites whose activations are unique to each task. If we knew the functions of each site from other research, we would then have a more complete understanding of the processes involved in these tasks, both those that they share in common and those that differ between the tasks.

The use of imaging data to evaluate double dissociations has become quite widespread. These data go beyond previous demonstrations of double dissociations that have involved behavioral data on either normal or brain-injured subjects. In the case of behavioral data on normal subjects, double dissociations can be established by finding two experimental variables, one of which affects performance on Task 1 but not Task 2, and another of which affects performance on Task 2 but not Task 1. This pattern permits one to argue that the two tasks differ in their engagement of some set of psychological processes, although it is not very specific about the particular processes that are engaged. In the case of behavioral data on brain-injured subjects, one seeks two patients, one of whom can perform Task 1 but not Task 2, and the other of whom can perform Task 2 but not Task 1. This pattern again allows one to argue that the tasks differ in the underlying processes they recruit, but there are weaknesses to this approach: Often damage in patients is not tightly localized; sometimes patients develop compensatory mechanisms for their deficits; and studies of this sort require one to make conclusions about normal performance from patients who have selective deficits, perhaps compromising the generality of the conclusions one can reach. Because double dissociations in neuroimaging have a different set of weaknesses, they complement neuropsychological dissociations, making neuroimaging another important point of leverage in distinguishing psychological processes.

[Insert Figure 2 about here.]

To see how successful this double-dissociation technique can be, consider an example. For some time, it has been suspected that working memory may consist of at least two sub-systems, one

concerned with spatial information and the other concerned with verbal information. This was originally proposed by Baddeley and his colleagues (e.g., 1986, 1992), and the proposal has received support from behavioral studies of normal and brain injured adults (see Jonides et al., 1996 for a review). A critical finding that helps seal the case for two sub-systems of working memory comes from a pair of experiments that compared the brain regions activated by parallel spatial and verbal working memory tasks (see Jonides et al., 1993; Awh et al., 1996; Smith et al., 1996 for details). A schematic that illustrates the two tasks is shown in Figure 2. In the spatial case, subjects had to encode three locations marked by dots on a screen and to retain these in memory for 3 seconds. Following the retention interval, a single location was marked and subjects had to indicate whether this location matched one of the three in memory. The verbal task was similar in that subjects had to encode 4 letters and to retain these in memory for 3 seconds, following which a single letter was presented and subjects had to decide whether it matched one of the three in memory.

[Insert Figure 3 about here.]

PET scans of subjects engaged in these two tasks revealed a striking dissociation in the circuitry that underlies them, as shown in Figure 3. The figure includes left and right lateral as well as a superior view of the brain, superimposed on which are the activations for each of the memory tasks, shown in color scales. The spatial task recruited mechanisms of neocortex predominantly of the right hemisphere while the verbal task recruited mechanisms predominantly of the left hemisphere. The details of which regions were activated and what these activations might signal for the processes in each task are reported elsewhere (Smith et al., 1996). For the present, it is sufficient to note that this pattern of results provides quite nice support for the claim that working memory is composed of separable systems for different sorts of information, a claim that relies on the sort of double dissociation shown in the figure.

c. Convergence of neuroimaging and behavioral data in normal adults

One of the great strides forward that will be possible with the advent of neuroimaging arises because of the opportunity for convergence between behavioral data and neuroimaging data drawn from normal experimental participants. The leverage that is gained from this convergence is large. If we have data from behavioral studies that suggest a dissociation between two different psychological processes, we have the opportunity to study whether these processes are represented in separable neural tissue. If so, this greatly strengthens the case for separable processing systems.

[Insert Figure 4 about here.]

Consider the following example from work in our laboratory (Badre et al., 2000). This work has been concerned with identifying executive processes and their neural implementation. One executive process that has been a focus of study is task-management, the ability to manage multiple tasks simultaneously. We have constructed a situation that requires task management of two sorts, illustrated in Figure 4. Subjects see a series of computer displays as illustrated in figure that have two panels, one on the left and one on the right. Each display contains a single arrow that points up or down. Subjects begin each series of trials with two counters set at “20” each, and each time an arrow appears on one side or the other, they are to change that count up or down, depending on whether the arrow points up or down. At the end of a run of trials, subjects are queried about each of the counter values to be sure they have kept the counts accurately. Notice that in this task, there are two counters that must be managed. On successive trials, subjects might have to access the same counter or they may have to switch counters from one to the other. Notice also that there are two types of counting operations that are required, incrementing or decrementing. Again, on successive trials, subjects may be using the same operation or they may have to switch from one operation to the other. Behavioral data about the time it takes subjects to complete each trial (measured by subjects’ depressing a response button when they are ready to accept the next stimulus display) show a clear effect: There is a cost in switching between counters and there is a cost in switching between operations. Importantly, these two costs are independent of one another, as shown in Figure 5. The time cost of each type of switch does not influence the time cost of the other. This result leads to the

implication that there may be two mechanisms involved in the two types of switch (two dissociable executive processes); if so, we may be able to find neural evidence of the two mechanisms.

[Insert Figure 5 about here.]

In fact, a follow-up experiment that studied subjects performing this task in an fMRI environment found just this evidence. Some data from this experiment are shown in Figure 6. The figure shows that there is a region of lateral frontal cortex that is activated by the switch in counters, but not the switch in operations; similarly, there is another region of medial frontal cortex, more anterior than the first, that is activated by a switch in operations but not a switch in counters. This sort of double dissociation follows the behavioral data nicely in suggesting two at least partially independent mechanisms for the two executive processes. So, here is a case in which the behavioral data about a task led nicely to an imaging experiment, whose data converged with the behavior in normal adults.

[Insert Figure 6 about here.]

d. Convergence of neuroimaging and behavioral data in patients

It is possible to extend this hunt for convergence beyond the study of normal adults as well. An excellent example comes from the study of memory processes. It is by now well-documented that there are two memory systems subserving long-term memory in adults. The distinction between the two types is often called a distinction between explicit and implicit memory. Take the concept of a bicycle, for example. You may be able to remember where you parked your bicycle this morning or yesterday, or where you bought that bike. These would be examples of explicit memory because you are explicitly retrieving a piece of information that you have previously stored. By contrast, most adults can ride a bicycle with little trouble, a skill that any young child will tell you is quite difficult. The skill to ride a bicycle reflects that adults have stored some information that translates into motor movements that make balancing, pedaling, turning, and so forth possible. This is a kind of implicit memory because while riding

a bike, there is no sense of explicitly retrieving information from memory; rather, information is retrieved in the course of executing the required behavior.

As it happens, the distinction between explicit and implicit memory is well-supported by studies of patients with brain lesions that reveal a double dissociation between these two types of memory. Some patients with medial temporal lobe lesions, including extensively studied ones such as H.M. (Milner et al., 1968), have an inability to acquire new information and retrieve that information explicitly, but they have intact implicit memory for motor skills and other procedural knowledge. By contrast, Gabrieli et al. (1995) have reported the result of a patient with damage to the right occipital lobe, M.S., who has an intact and functioning explicit memory system, but impaired implicit memory. Taken together, pairs of patients such as this suggest the existence of two memory systems that dissociate in their functions and in the neural tissue that subserves them. This claim leads naturally to the prediction that testing normal adults on explicit and implicit memory tasks ought to result in different patterns of brain activation as the signatures of these two memory systems. By now, various reports have surfaced that support this contention (see, e.g., Schacter and Buckner, 1998). In general, explicit memory tasks compared to a control condition cause increased activation of medial temporal lobe structures; in general, implicit memory tasks cause decreased activation in association cortex of posterior regions of the brain. Why these two particular patterns of increase and decrease of activation should occur in response to explicit and implicit tasks respectively is a question beyond our scope here; but the result nicely illustrates how imaging evidence and evidence from patient populations can be used in tandem to converge on a view of cognitive processing.

e. Convergence of neuroimaging data in humans with behavioral data in animals

Invasive or recording studies on animals other than humans have raised important hypotheses about the layout of various cognitive systems residing in sensory, motor, or association cortex. Neuroimaging studies with humans now permit tests of these hypotheses. One care that must be taken to achieve this convergence between animal and human studies has to do with homology. It is often difficult

to determine just what structure in a monkey brain, for example, is homologous to a structure in a human brain. Sometimes this homology can be approached cytoarchitectonically, by examining the morphology of cells in brain regions of the two species in question. Sometimes, there are functional data from other studies that give good leads about which areas in the brains of two species are performing related functions. Regardless of the approach one takes to the problem of homology, one must take this approach with care to be sure that a case can be made for a structural or functional similarity.

Sometimes the homology is reasonably straightforward, however, as it appears to be for a leading case that has exploited the opportunity to relate data from monkeys and humans concerning visual function. Since the pioneering work of Ungerleider and Mishkin (1982), it has become increasingly clear that early visual processing proceeds along two streams. A ventral stream of information flows from primary visual cortex to temporal cortex, with increasingly complex computations performed on the information in the service of revealing the form, color, and identity of objects in the environment. A dorsal stream also flows from primary visual cortex to structures of the parietal lobe, responsible for processing information about the spatial location and movement of objects. The data from which this view of the visual system derive come from studies of lesioned monkeys performing tasks of object recognition or spatial localization, and they also come from single-cell recording studies of the functions of temporal and parietal systems. Both sorts of studies have provided quite strong support for the duality of the visual processing stream.

Much more recently, evidence from human neuroimaging studies has provided nice convergence with the data from monkeys. Perhaps the seminal such study was one by Haxby et al. (1994). They had human volunteers perform a matching-to-sample task under two conditions. In one, subjects compared a sample face to two alternatives and picked which alternative matched the sample. In the other, subjects compared the position of a dot in a frame to the positions two other dots in frames to see which of the two ways was identical to the first. The first task required the processing of information about shape and form, while the second required the processing of information about spatial position. As predicted by the data from monkeys, the two tasks resulted in activation of separable regions in cortex: The task involving form caused activation of occipital and temporal cortex while the task involving location caused activation

of occipital and parietal cortex. Here, then, is an illustration of how data from cognitive studies with animals can be used to motivate the examination of cortical function in humans using neuroimaging techniques.

Let us summarize. We have devoted significant space at the opening of this chapter to a detailed examination of why one would want to conduct research using neuroimaging techniques, especially PET and fMRI. The motivations for these techniques are several in number, as we have elaborated. Overall, there is good reason to believe that neuroimaging methods will become centerpieces in the array of tools available to cognitive psychology (and to other fields in psychology as well). So, it is well worth the effort for the student of cognition to learn about the techniques available and how they can be applied to the study of psychological tasks. We turn now to these issues.

II. The What: Neuroimaging techniques and task design

a. Neuroimaging techniques

Imaging methods for human studies include a number of alternatives: fMRI, PET, single positron emission computerized tomography (SPECT), event-related potentials (ERP), electroencephalography (EEG), magnetoencephalography (MEG), and near-infrared spectroscopy. A number of other brain imaging techniques are available for use in animals using radiolabeling, histological, or optical imaging techniques.

Although all these techniques are in frequent use and provide important insights into brain function, we shall focus on the two techniques most commonly used in current human research concerned with localization of function: PET and fMRI. The main advantages of these techniques are that they can be used on humans, they offer a balance between spatial resolution and temporal resolution, and they can be used to create images of the whole brain. This last feature offers a great potential for synergy with animal research. Single-cell recording in animals, for example, offers spatial resolution down to a single neuron and millisecond temporal resolution. Its main weakness is that testing usually occurs within single,

isolated brain regions, while other regions important to performance of some task may be missed. Neuroimaging using PET and fMRI is well suited to providing exploratory analyses of brain processes, allowing new hypotheses about specific brain areas to be developed and tested in animal models. In addition, neuroimaging with PET and fMRI offers a broad view of how remote brain regions interact in particular psychological functions, complementing the detailed analysis of individual cell behavior that is possible using animal models.

1. What PET and fMRI can measure

The number of techniques for imaging brain processes with PET and fMRI is growing. Although a thorough discussion of all of these is far beyond the scope of this chapter, it is important to realize what sorts of processes can be imaged using these techniques. Some of the alternatives are described here briefly; our subsequent discussions of task design will focus on measures of regional brain activation because these are the ones used most often to study human cognition and affect. Table 1 shows a summary of the various methods available using PET and fMRI as measurement tools. Following is a brief description of each method.

[Insert Table 1 about here.]

Structural scans. fMRI can provide detailed anatomical scans of gray and white matter with resolution well below 1 mm^3 . This can be useful if one expects structural differences between two populations of individuals, such as schizophrenics versus normal controls (Andreasen et al., 1994), or changes in gross brain structure with practice or some other variable. An example is a recent study that reported larger posterior hippocampi in London taxi drivers who had extensive training in spatial navigation (Maguire et al., 2000). Another structural scanning technique is diffusion tensor imaging, described below. This technique allows one to identify white matter tracts in the human brain, which is useful for studying structures such as the corpus callosum and changes in these structures as a function of

some variable, such as age.

Regional brain activation. Perhaps the most frequent use of both PET and fMRI, and the one that is the focus of this chapter, is the study of changes in some property of metabolism or the vasculature that accompany changes in neural activity. With PET, one may separately measure glucose metabolism, oxygen consumption, and regional cerebral blood flow (rCBF). Each of these techniques allows one to make inferences about the localization of neural activity based on the assumption that neural activity is accompanied by a change in metabolism, in oxygen consumption, or in blood flow. Functional MRI using the Blood Oxygen Level Dependent method (BOLD) measures the concentration of deoxygenated hemoglobin in the blood across regions of the brain. The rationale is that a) more deoxygenated blood in an area causes a decrease in BOLD signal, and b) neural activity is accompanied by increased blood flow, which dilutes the concentration of deoxygenated hemoglobin and produces a relative increase in signal (Hoge et al., 1999). In that fMRI takes advantage of changes in blood flow with changed neural activation, and PET measurements of rCBF does the same, there should be good correspondence between these two measures for the same tasks, and this is generally the case (Joliot et al., 1999; Kinahan & Noll, 1999; Ramsey et al., 1996). One difference appears to be that fMRI activations are usually located several millimeters dorsal to those of PET, consistent with the idea that fMRI is sensitive to deoxygenated hemoglobin in the capillaries and draining venules surrounding synapses.

Anatomical connectivity. A new methodology being developed to map the white matter tracts that connect regions of the brain is diffusion tensor imaging. Several current methods use standard MRI scanners configured to be sensitive to the diffusion of water to estimate water diffusion tensors in each area of the brain (Peled, Gudbjartsson, Westin, Kikinis, & Jolesz, 1998). Think of a tensor as a measure of motion in the x, y, and z dimensions (a vector is a special kind of tensor). Researchers are interested in the shapes of the tensors in different brain locations. Water diffuses equally easily in all directions in the ventricles and other fluid spaces, producing a spherical tensor. At the edges of the brain and in other areas, water may be restricted from diffusing in one direction, producing a planar tensor. Near a white

matter tract, however, water diffuses most easily along the tract, producing a diffusion tensor that is large along the axis of the tract and small in the other dimensions. These 'linear' tensors mark the existence and direction of a white matter tract in the brain.

Factors that affect the shape of a tensor are the density of axon fibers in the tract, the degree of myelination, the fiber diameter, and the similarity in the directions of the fiber projections.

In the published literature, diffusion tensor images are usually labeled with different colors for the x, y, and z components of motion; a solid block of one color indicates fiber tracts running along either the x, y, or z-axis of the image. While most studies of diffusion tensor imaging have so far focused on the methodology itself, there are many potential applications to the study of brain function, including combined studies of structure and brain activation to help define functional networks.

Receptor Binding. The affinity of particular chemicals for specific types of neurotransmitter receptors offers researchers a leverage point for investigating the functional neurochemistry of the human brain. Radioactive labels are attached to carefully chosen compounds, which are then injected into the arteries of a subject by either a bolus (a single injection) or continuous infusion of the substance until the brain concentrations reach a steady state. This method can be used to image the density of a specific type of receptor throughout the brain. It can also be used to image the amount of binding to a particular type of receptor that accompanies performance of a task, as it was used in one study of dopamine binding during video game playing (Koepp, 1998).

The most common radioligands and transmitter systems studied are dopamine (particularly D2 receptors) using [^{11}C]raclopride or [^{123}I]iodobenzamide, muscarinic cholinergic receptors using [^{11}C]scopolamine, and benzodiazepines using [^{11}C]flumazenil. In addition, radioactive compounds that bind to serotonin, opioid, and a number of other receptors have been developed. Because the dynamics of radioligands are complex, a special class of mathematical models, called kinetic models, have been developed to model them. Kinetic modeling can allow the researcher to estimate how much of the radiolabeled compound is in the vasculature as opposed to in the brain, how much is freely circulating in brain tissue, how much is bound to the specific receptor type under investigation, and how much is bound

at nonspecific sites in the brain. Estimation of all these parameters requires a detailed knowledge of the properties of the specific substances used and how they act in the brain over time.

Gene Expression. Very recently, new methods in both PET and MRI have allowed researchers to investigate local gene expression within the living brain. PET can be used to image the distribution of an enzyme in the brain by radiolabeling one of its substrate compounds. When the labeled substrate is converted into the enzyme, the label becomes “trapped” in tissue and emits a persistent signal that can be detected by the PET camera. One recent study used this method to label a substrate of an adenoviral enzyme that directs the expression of a particular gene in mice, thereby indirectly indexing gene expression (Gambhir et al., 1999).

MR spectroscopy provides a different way to image enzymes and biochemicals related to gene expression. The arrangement of atoms in their constituent molecules gives rise to very small inhomogeneities in the scanner’s magnetic field. These magnetic variations alter the spectrum of energy that the atoms will absorb, giving rise to a characteristic frequency ‘signature’ for various types of atoms. One research group took advantage of the fact that fluorine is an easy element to identify by its spectrum; they used spectroscopy to quantify the amount of containing-containing compounds related to expression of a particular gene (Stegman et al., 1999). A combination of creativity and specific knowledge of the relevant physics and biochemistry can lead to imaging solutions for a very large number of experimental questions.

Having provided a brief summary of these various techniques, we shall now concentrate on PET and fMRI as they are used to measure changes in blood flow and oxygenation.

2. Limitations of PET and fMRI

Spatial limitations. There are limitations restricting what both PET and fMRI can measure. Neither technique is good for imaging small subcortical structures or fine-grained analysis of cortical activations. The spatial resolution of PET, on the order of 1-1.5 cc³, precludes experiments testing for

neural activity in focused areas of the brain (e.g., mapping receptive fields of cells in visual cortex).

fMRI has a much greater spatial resolution, as low as 1 mm^3 but often on the order of 3 mm^3 for functional studies. The impact of this limitation in spatial resolution is that activation in some structures may be missed entirely or mislocated, although recent fMRI studies have found activity in structures as small as the nucleus accumbens (Breiter et al., 1997).

Artifacts. Artifactual activations (i.e., patterns that appear to be activation arising from non-neural sources) may arise from a number of sources, some unexpected. One study, for example, found a prominent PET activation related to anticipation of a painful electric shock in the temporal pole (Reiman, Fusselman, Fox, & Raichle, 1989). However, it was discovered some time later that this temporal activation was actually located in the *jaw* – the subjects were clenching their teeth in anticipation of the shock!

Functional MRI signals are especially susceptible to artifacts near air and fluid sinuses and at the edges of the brain. Testing of hypotheses related to activity in brain regions near these sinuses, particularly orbitofrontal cortex and inferior temporal cortex among neocortical regions, is problematic using fMRI. Functional MRI also contains more sources of signal variation due to noise than does PET, including a substantial slow drift of the signal in time and higher frequency changes in the signal due to physiological processes accompanying heart rate and respiration (the high frequency noise is especially troublesome for imaging the brainstem). The low-frequency noise component can obscure results related to a psychological process of interest and it can produce false positive results, so it is usually removed statistically prior to analysis. The low frequency source of noise also makes it difficult to test hypotheses of slow changes during a session (e.g., effects of practice during scanning) although careful design still allows such issues to be tested (Frith & Friston, 1997).

Temporal resolution and trial structure. Another important limitation of scanning with PET and fMRI is the temporal resolution of data acquisition. The details of this are discussed in subsequent sections, but it is important to note here that PET and fMRI measure very different things, over different

time scales. Because PET computes the amount of radioactivity emitted from a brain region, at least 30 seconds of scanning must pass before a sufficient sample of radioactive counts is collected. This limits the temporal resolution to blocks of time of at least 30 seconds, well longer than the temporal resolution of most cognitive processes. Functional MRI has its own temporal limitation due largely to the latency and duration of the hemodynamic response to a neural event. Typically, changes in blood flow do not reach their peak until several seconds after a neural event, so the locking of neural events to the vascular response is not very tight.

Duty Cycle. A final limitation for both PET and fMRI has to do with what is often called the duty cycle of a task. In order for a neural event to create a measurable hemodynamic response, the neural event must occupy a substantial proportion of the time taken in any measurement period. For example, if only a small number of nerve cells fires for some process or if the duration of firing is small with respect to the temporal resolution of the measurement technique, then the signal-to-noise ratio for that event is low and it may be difficult to detect. Although processes that elicit very brief neural activity, such as brief flashes of light, can be detected using neuroimaging, as a general rule experiments need to be designed so that the process of interest occupies a substantial proportion of the measurement window of time.

3. Summary of advantages of PET and fMRI

Above, we have commented on the limitations of PET and fMRI, but we also need to point out their advantages when used as tools to measure the vascular response to neural events. Each has some unique features that makes it apt for certain types of experiment. Table 2 summarizes these advantages.

[Insert Table 2 about here.]

An inspection of the table shows that PET and fMRI each have different characteristics that make it

particularly suited to certain types of imaging questions.

b. A road map of a neuroimaging experiment

Before starting a neuroimaging experiment, several important decisions must be made. First, a hypothesis must be chosen, much as we described several hypotheses in the introduction to this chapter and how these led to imaging experiments. Second, appropriate methods must be selected; these choices will be constrained by the nature of the task chosen, the available imaging technology and its limitations, and the types of inference one wishes to draw from the study. Third, an experiment must be conducted, analyzed, and interpreted. Here is an overview of some of the highlights in each of these steps, with details to follow.

The design of a task limits the ultimate interpretability of the data. Tasks must be chosen that yield theoretical insight into the neural and psychological processes under investigation, and they must avoid the influence of nuisance variables. Nuisance variables may be neural processes unrelated to the question of interest—either prescribed by the task or unrelated to it—they may be technological artifacts such as slow drift in the signal from an MRI scanner, or they may be artifacts due to heart rate, respiration, eye movements, or other physiological processes. To the extent that nuisance variables influence the brain activations in a task, they will mitigate the uniqueness of an interpretation that one may place on the data. That is, one would like to claim that neuroimaging activations are related to psychological process X, not that activations are related to process X, or process Y, or some physiological artifact such as irrelevant eye movements during a task. Constructing adequate tasks can be quite challenging, and it may not be possible in some situations.

Once a task is designed and data are collected, analysis of those data is composed of two important substages: preprocessing of the images and statistical analysis of the resulting activations. Preprocessing consists of several steps. Before statistical tests are performed, the various images in a set of data must be aligned to correct for head motion that may have occurred from one image acquisition to the next. Following alignment, images are often normalized to a standard template brain so that results

from several subjects can be combined into averages and plotted in standard coordinates for comparison with other studies. Many researchers also smooth images, averaging activity levels among neighboring voxels so that smooth regions of activation result. Although smoothing decreases the spatial resolution of the images, it helps to estimate and control for statistical noise.

Following these preprocessing stages, statistical tests are performed on the data. Most analyses are essentially variants of the general linear model (GLM). Studies are often analyzed using t-tests that compare one or more experimental conditions of interest with a control condition. Slightly more complicated designs may use analysis of variance (ANOVA) with one or more factors. An increasingly popular technique uses multiple regression to model the processes of interest as well as to take into account the influence of nuisance covariates. In this kind of analysis, a covariate is constructed that contrasts periods when activity of interest is supposed to occur (i.e., the experimental condition in a blocked study or a neural response-evoking event in an event-related study) with control periods. Nuisance covariates are constructed that model activity related to processes of no interest, such as heartbeat or processes included in the task that are of no theoretical relevance. The effects of the nuisance covariates are statistically removed from the data during the analysis, decreasing the statistical error and increasing the power of the analysis. With fMRI, signals at low spatial frequencies—essentially slow, random drift due to variations in the magnetic field—may produce artifacts in the data, and these are usually modeled as nuisance covariates as well or filtered out before beginning the analysis.

With this brief summary, we are ready to launch into a more thorough treatment of experimental design. We do this by reviewing the various designs that have become popular in experiments using PET and fMRI measurement techniques. Following our description of these designs, we review techniques that can be used with these designs to contrast different experimental conditions.

c. Types of experimental designs

1. Blocked designs

Because long intervals of time (30 seconds or more) are required to collect sufficient data in a PET experiment to yield a good image, the standard experimental design used in PET activations studies is the blocked design. A blocked design is one in which different conditions in the experiment are presented as separate blocks of trials, with each block representing one scan during an experiment. Thus, the activations of interest in a PET experiment are ones that accumulate over the entire recording interval of a scan. If one is interested in observing the neural effect of some briefly occurring psychological process (e.g., the activation due to a briefly flashed light stimulus), in a PET experiment one would have to iterate the event repeatedly during a block of trials so that activations due to it accumulated over the recording interval of a scan. One could then compare the activations in this scan to an appropriate baseline control scan in which the event did not occur. Given the temporal limitation of this technique, PET is not well suited to examining the fine time course of brain activity that may change within seconds or fractions of a second.

The blocked structure of PET designs is a major factor in the interpretability of results. Activations related to slowly changing factors such as task-set or general motivation are captured in the imaging study. This is an advantage if one wishes to image such effects. However, if one wishes to image the neural responses to individual stimuli, PET is not a technique suited to that goal. Even if such slowly changing processes are of interest, one must take care to elevate their duty cycle within a scan so that their neural signatures form a significant portion of the entire scan's processes.

Some researchers have made good use of differences in duty cycle as a way to circumvent some of the limitations of blocked designs (Garavan, Ross, Li, & Stein, 2000). These studies have used trial-blocks with different percentages of certain trial types to capture a process of interest. For example, one might conduct a blocked study of a particular process of interest, but parametrically vary the number of trials within the block that recruit that process. Rather than comparing blocks of the active task with rest, one might compare blocks in which the task of interest was performed on 80% of the trials with blocks in which the task of interest was performed on 20% of the trials.

Many studies using fMRI also have made good use of blocked designs. One advantage to a blocked design is that it offers more statistical power to detect a change—some estimates are that it

offers four times the power of a single-trial design. As with PET, the ability to examine brain activations due to single trials is lost. Because the time to acquire a stable image is substantially less with fMRI than with PET, fMRI does allow one to conduct experiments in which activations due to single trials can be collected in a stable way. A sample of the MRI signal in the whole brain can be obtained in 2-3 seconds on average, depending on the way data are acquired and depending on the required spatial resolution of the voxels that are imaged. For studies that do not sample the whole brain, acquisition can be much more rapid – as low as 100 ms for single-slice fMRI. In fact, the limiting factor in the temporal resolution of fMRI is not the speed of data acquisition, but the speed of the underlying hemodynamic response to a neural event, which begins within a second after neural activity occurs and peaks 5-8 seconds after that neural activity has peaked.

2. Individual-trial, event-related fMRI

To take advantage of the rapid data-acquisition capabilities of fMRI, an ‘event-related’ fMRI technique has been developed to create images of the neural activity related to specific stimuli or cognitive events within a trial. The technique involves spacing stimuli far enough apart in time so that the hemodynamic response to a stimulus or cognitive event is permitted to return to baseline before the onset of the next stimulus or event. Most researchers consider 14-16 seconds enough time for this to occur, although some data have revealed that the hemodynamic response can persist as long as 30-60 seconds (Turner & Friston, unpublished data). Using this technique, signals from individual trials of the same task can be averaged together, and the time course of the hemodynamic response within a trial can be determined. This technique permits the randomization of trials from different conditions, which is essential for certain tasks. It also allows researchers to analyze only selected types of trials in a mixed trial block, enabling the study of error monitoring (to name one example) and a number of other processes that only occur on some trials.

Selective averaging provides one way around the temporal limitations imposed by the hemodynamic response function. By averaging across trials of the same type and by comparing these

averages across different conditions, the time course of fMRI signals differing by as little as 100 ms can be distinguished. An example comes from the work of Aguirre, Singh, and D'Esposito (1999), who studied activation in the fusiform gyrus in response to upright and inverted faces. When they compared trials from the two conditions, they found that the BOLD response was shifted 100 ms later for inverted faces, paralleling increased reaction times to recognize inverted faces.

Another creative example of the added hypothesis-testing power of event-related fMRI comes from studies of episodic memory. Buckner et al. (Buckner et al., 1998) studied people encoding lists of words, and they subsequently tested the subjects to see which words they remembered correctly. Functional MRI scanning during the learning of each word allowed the researchers to compare activity during the encoding of words that were successfully retrieved with the encoding of words that were later forgotten, revealing important differences.

3. Rapid event-related fMRI

More recent developments in event-related fMRI designs have made experimental trials more similar to those found in standard behavioral experiments (Zarahn, Aguirre, & D'Esposito, 1997). The main problem with the event-related design discussed above is that trials are very slow in pacing (with 12-16 seconds required between successive trials). It is possible to speed this up substantially by making use of knowledge of the precise shape of the hemodynamic response function to a pulse of neural event. Various investigators have measured the nature of this response, and good models of it now are used routinely. Using prior knowledge of the typical hemodynamic response function, or measuring it individually for each subject, one can now perform experiments in which successive stimuli or cognitive events can be presented with as little as 750 ms intervening (Burock, Buckner, Woldorff, Rosen, & Dale, 1998; Dale & Buckner, 1997). Closely packed trials of a number of experimental conditions can then be presented in random order in a scanning interval. One then creates a model function that includes the timing of critical stimuli or cognitive events convolved with a model of the known or hypothesized hemodynamic response function. This convolved predictor function can then be used as a regressor in a

multiple regression analysis, and the fit of the actual data to the expected pattern of BOLD signal can be measured. Of course, one would have several regressors to fit to the data, each one designed to predict the effect of one type of cognitive event (i.e., one condition). In this way, one can compare different regressors to examine which one or ones fit the data best, thereby accounting for the pattern of obtained activations. We describe this technique in more detail below.

This method has led to several important advances. One is the ability to closely space trials in time, resulting in a pacing that is more in line with the large body of literature in experimental psychology. Another is that the effects of fatigue, boredom, and systematic patterns of thought unrelated to the task during long inter-trial intervals are avoided with such designs. In addition, the capability to obtain images of more trials per unit time compared to individual event-related designs counters the loss of power that occurs when using a single-trial design as opposed to a blocked design, making designs with closely packed trials more efficient than those with long inter-trial intervals. Of course, if trials are spaced too closely, the ability to tell which part of the signal came from which type of trial is decreased, so there is a trade-off between the number of trials one can include and how much resolving power is lost. Some researchers have estimated that a 4 s inter-trial interval is optimal for detecting task-related activations (Postle & D'Esposito, 1999), although much more research needs to be done on the specifications and limitations of this new technique.

An important element of these rapid event-related designs is that the inter-trial interval must be varied from trial to trial. The ability to separate signal coming from different trial-types when the hemodynamic responses to each trial overlap in time depends on jittering the time between trials and on either randomly intermixing trials of different experimental conditions or carefully counterbalancing their order. To get an intuition about how rapid designs allow one to discriminate the effects of different conditions, consider this: With a randomized and jittered design, sometimes several trials of a single type will occur in a row, and because the hemodynamic response to closely spaced events sums in a roughly additive fashion (although there are minor nonlinearities; e.g., Boynton, Engel, Glover, & Heeger, 1996; Dale & Buckner, 1997), the expected response to that trial-type will build to a high peak. Introducing longer delays between some trials and shorter ones between others allows peaks and valleys in activation

to develop that are specific to particular experimental conditions. A regression model will be more sensitive to a design with such peaks and valleys than it will to a design that has a uniform spacing of trials because one with peaks and valleys will create a unique signature for that type of trial. The effect of jittering is essentially to lower the effective temporal frequency of the design, so it is particularly appropriate for rapidly presented trials. Without jittering the inter-trial interval, the neural events would occur too rapidly to be sampled effectively.

One problem with the hemodynamic response-convolution technique used in rapid event-related designs is that it is based on a predicted shape of the hemodynamic response. So, if one mis-specifies this response function, one will lose significant power in this experimental technique. This problem is especially acute when comparing different subject populations (e.g., older versus younger adults, or patients and normal controls) because their hemodynamic response functions may differ from one another (D'Esposito, Zarahn, Aguirre, & Rypma, 1999). One approach that has been used to avoid this problem is the measurement of hemodynamic responses in each individual subject, often by presenting brief flashes of light and measuring the BOLD response over the seconds following the stimulation in the primary visual cortex or by measuring the hemodynamic response in motor cortex to simple finger movements (Aguirre, Zarahn, & D'Esposito, 1998). Of course, the best use of this technique is when the region of interest in an experiment corresponds to the region in which the hemodynamic response is measured. If it does not, one must assume that the measured hemodynamic response in one region of the brain is equivalent to that in another region.

d. Techniques for contrasting experimental conditions

For a psychologist, the main value of neuroimaging data is that they provide new tools for understanding psychological processes. For example, finding that premotor cortex is activated during the identification of tool-like objects opens up a new set of hypotheses about the nature of object recognition (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995). Or, finding that visual cortex is activated in blind individuals when they perform tactile tasks suggests a set of hypotheses about the extent of plasticity in

the sensory nervous system (Sadato et al., 1996).

Of course, the value of neuroimaging data to psychological inference depends on an accurate assessment of which brain regions are activated in any task. The problem with making inferences about cognitive processes from neuroimaging data is that nearly any task, performed alone, produces changes in most of the brain. To associate changes in brain activation with a particular cognitive process requires that we isolate changes related to that process from changes related to other processes. In short, it requires that we have contrasting experimental conditions that isolate the processes that interest us. You can understand how these contrasts can be designed without understanding details of data acquisition and analysis, topics that we treat in the final section of this chapter. However, it is useful to know one fact about data acquisition: Data in neuroimaging experiments are in the form of a matrix of signal intensity values in each region of the brain. The brain is divided up into voxels, typically 60,000 - 100,000 small volumes of brain tissue, whose size and number vary from study to study depending on the acquisition methods that were used to gather the data. These voxels are the elementary units of data; we assume that the signal in a voxel represents the neural activation in that region of the brain (more on the biophysics of that assumption below). It is the behavior of these voxels that is the focus of an imaging experiment. There are four techniques that are most frequently used to study the behavior of brain voxels: subtraction, parametric variation, factorial design, and correlational studies.

1. Subtraction

The first method that was devised for making inferences about psychological processes from neuroimaging data involves statistically comparing activations derived from an experimental condition with activations from a control condition that is putatively identical except that it does not recruit the process of interest. This is the 'subtraction method,' the logic of which dates back to Donders (1868). The technique was first used by Posner et al. (Petersen, Fox, Posner, Mintun, & Raichle, 1988; Posner, Petersen, Fox, & Raichle, 1988) in a study of reading processes. The logic of subtraction is this: If one tests two experimental conditions that differ by only one process, then a subtraction of the activations of

one condition from those of the other should reveal the brain regions associated with the target process. This subtraction is accomplished one voxel at a time. Together, the results of the voxel-wise subtractions yield a three dimensional matrix of the difference in activation between the two conditions throughout the scanned regions of the brain. T-tests can be performed for each voxel to discover which of the subtractions is reliable (of course, one needs to correct for the fact that multiple comparisons are being conducted—more about this below). The resulting parametric map of the t-values for each voxel shows the reliability of the difference between the two conditions throughout the brain, and images of t-maps or comparable statistics (z or F maps) are what generally appear in published reports of neuroimaging studies.

As an example of the implementation of subtraction logic, consider an experiment from our laboratory (Reuter-Lorenz et al., 1999) that was similar to the task shown in Figure 2. In the experiment of Reuter-Lorenz et al. (1999), subjects had to encode the locations of three target-dots on a screen and store these in memory for three seconds, following which a single probe-dot appeared and subjects had to decide whether the probe-dot was in the same spatial position as one of the previous three target-dots. In order to isolate processes of spatial storage, we constructed a control condition that was identical to this experimental condition, but with one difference: In the experimental condition, the retention interval was 3 sec, whereas in the control condition it was 200 ms. We reasoned that a subtraction of the activations from the control condition from those of the experimental condition would then reveal the brain regions responsible for the extra storage required in the experimental condition.

In the case of our experiment, the logic of the subtraction method is fairly safe. Note, however, that in general, subtraction logic rests on a critical assumption, what has been called the assumption of “pure insertion” (Sternberg, 1969). According to this assumption, changing one process does not change the way other processes are performed. Thus, by this assumption, the process of interest may be ‘purely’ inserted into the sequence of operations without altering any other processes. Although violations of subtraction logic have been demonstrated experimentally (Zarahn et al., 1997), the logic is still widely used because it greatly simplifies the inference-making process. If this assumption is violated, then a difference in the observed neuroimaging signal between an experimental and a control condition may be

due to one of these other altered processes rather than the process of interest.

To appreciate the difficulty of implementing subtraction logic in an experimental setting, consider a hypothetical study. In the experimental condition, subjects must press a button every time they see a red stimulus; in the control condition, they passively view the same stimulus sequence as in the experimental condition. The experimenter might assume that activity related to visual processing is the same in both conditions, and the two tasks differ only in that the first requires the execution of a response. Thus, when activations from the control condition are subtracted from those of the experimental condition, the experimenter may attribute the significantly activated areas to response-execution processes.

There are a number of flaws with this conclusion, the nature of which provide some insight into the assumptions of subtraction logic. First, several processes vary at once, since the experimental condition includes an overt manual response as well as a cognitive decision to execute a response. We cannot know whether activated areas are related to the decision, to response preparation, to response execution, or to an interaction between two or more of those processes. In addition, the experiment may violate the assumption of pure insertion. When we assume pure insertion in this case, we assume that adding decision and response processes will not change the nature of the perceptual processing of the stimuli. However, making a stimulus relevant, and causing attention to be directed to the stimulus, alters perceptual processing in very early areas of visual cortex (Hopfinger, Buonocore, & Mangun, 2000). Our naive experimenter may assume that activations in the occipital lobe revealed by the subtraction are related to the response process, when in fact those areas may be involved in processing the color of the stimuli, modulated by changes in attentional focus.

This example illustrates the difficulty in selecting experimental and control conditions appropriately. It also illustrates another point about subtraction logic. Several researchers have argued that pure insertion can be tested within the experiment, and that violations of pure insertion will appear as significant *decreases* in signal when the control task is subtracted from the experimental task (Petersen, van Mier, Fiez, & Raichle, 1998). While this may be true in many cases, there are two problems with assuming that pure insertion is only violated in cases where 'deactivations' occur. First, it is difficult to

tell whether decreases in signal are due to a violation of pure insertion (i.e., the control task includes a process that the experimental task does not) or due to an actual inhibition of a certain brain area related to the process of interest. Second, our hypothetical example illustrates a case in which pure insertion may be violated, but the violation would produce no decreases in activity, just an increase unrelated to the process under investigation. Clearly, inferences about cognitive processes that rely on subtraction of activation in two conditions must be interpreted with caution.

2. Parametric variation

Several approaches have been used to improve upon subtraction logic and to strengthen the credibility of inferences drawn from differences between conditions. One of these is parametric variation over several levels of a particular process of interest. Examples of experimental parameters to vary incrementally might be the number of words to remember in a memory experiment, the percentage of a certain type of trial, or the time on task.

An example of this is the studies of working memory by Jonides et al. (1997) using the n-back task in a PET context. In the n-back task, a subject sees a string of letters appearing one at a time, and he must match each letter to the one that appeared “n” items back in the series. In separate conditions, the values of n varied from 1 to 2 to 3. Another ‘0-back’ control condition required subjects to indicate a match each time a fixed letter (e.g., “G”) appeared. Common to all tasks are encoding and response processes, and what differs is the working memory load and the requirement to update information stored in working memory. Jonides et al. (1997) found that several regions varied in their activation systematically with variations in working memory load, as compared to other regions that showed no systematic variation. In a succeeding experiment by Cohen et al. (1997) using fMRI, a finer dissociation was documented among the regions showing variation with working memory load. Some regions, such as posterior parietal cortex, showed monotonic increases in activity with increases in load, whereas dorsolateral prefrontal cortex showed a step-function increase in activation from 1-back to 2-back, with no other differences in activation. Thus, the parametric technique permitted a fine discrimination of areas

involved in working memory from other brain regions, and it permitted an examination of the details of activation differences even among the regions involved in working memory.

Another example of parametric variation is a study of the Tower of London task by Dagher et al. (Dagher, Owen, Boecker, & Brooks, 1999). The Tower of London task requires subjects to make a sequence of moves to transfer a stack of colored balls from one post to another in the correct order. The task requires subjects to plan out a number of moves which subjects had to devise and store in memory in advance of completing the task. The experimenters varied the number of moves incrementally from 1 to 6. Their results showed linear increases in activity in dorsolateral prefrontal cortex across all 6 conditions, suggesting that this area subserved the planning operations critical for Tower of London performance.

[Insert Figure 7 about here.]

The power of parametric variation lies in two features. First, the reliance on pure insertion is weakened. Rather than assuming that insertion of a process does not change other processes, the logic assumes that altering the load on one process does not change other component processes in the task. Second, the results are more highly constrained because, unlike a subtraction study, parametric variation permits one to make multiple comparisons among multiple levels of a variable (e.g., 1- vs. 2- vs. 3-back in the n-back task). This feature permits researchers to search for a pattern of change in activation across all the levels of the variable, such as a monotonic increase. Such a pattern renders more unlikely false positives and spurious activations due to improperly controlled variables.

3. Factorial designs

Factorial designs are an extension of subtraction logic. Whereas the foundation of subtraction studies is the t-test, factorial studies rely on factorial analysis of variance. Consider a simple factorial design from our studies of task switching presented in the introduction to this chapter. Our studies

contained two types of switching, each varied independently: switching which of two mental counters was to be updated, and switching which of two operations (add or subtract) should be applied. This design is a simple 2 X 2 factorial, with two levels of counter switching and two levels of operation switching. The neuroimaging data from this experiment can be analyzed with a factorial ANOVA on a voxel-by-voxel basis and subsequently corrected for multiple comparisons. By testing for main effects, we would then be asking the question: Is each voxel sensitive to switching counters, switching operations, or both? By testing for the interaction, we would be asking whether activity in the voxel was affected by both kinds of switch in a non-additive fashion. For example, a voxel might be activated only when both counter-switch and operation-switch are required, signaling that this brain area might be involved in the coordination of two kinds of executive processes.

Factorial designs allow one to investigate the effects of several variables on brain activations. They also permit a more detailed characterization of the range of processes that activate a particular brain region – e.g., counter-switch only, operation-switch only, either, or both. Factorial designs also permit one to discover double dissociations of functions within a single experiment. To restate, a double dissociation occurs when one variable affects one brain region but not another, and a second variable affects a second region but not the first. A factorial design is required in order to infer that a manipulation (e.g. counter switching) affected dorsolateral prefrontal cortex, but a second manipulation (e.g. operation switching) did not.

Factors whose measurements and statistical comparisons are made within subjects, as are those described above, are within-subjects factors. When differences are examined between older and younger subjects, between normal individuals and members of a patient population, or between other groups, the subject-group becomes a between-subjects factor because different levels of the factor are represented by different subjects. Because there are many reasons that two groups of subjects might differ in brain activation, researchers typically compare between subjects differences in activation related to a specific task. This comparison involves first subtracting a control task from a task of interest within subjects and then comparing the difference images between subjects. As an example, consider the fact that older and younger subjects differ in the amount of atrophy present in their brains, with older subjects typically

showing some 15% more atrophy than the young (Raz, unpublished manuscript). To mitigate this difference in comparing activations between old and young, Reuter-Lorenz et al. (2000) tested older and younger subjects in a working memory task and a control condition. They then subtracted the activations of the control from the memory task and compared older and younger subjects on their differences in these subtracted activations. This technique allows one to statistically remove any effects of the differences in brain atrophy between the groups.

4. Correlational studies

Correlational designs are often considered a weaker type of design from the perspective of making inferences because a correlation between two variables does not carry any information about the causal relationship between them. However, correlations have been used effectively in neuroimaging studies in several ways.

The most straightforward way is to examine the correlation of regional activation with behavioral performance variables. For example, Casey et al. (1997) found correlations between anterior cingulate activation and errors in a go/no-go task in children, suggesting that the anterior cingulate plays a role in response selection or inhibition. As another example, Lane et al. (1998) found that a region of anterior cingulate correlated with self-ratings of emotional awareness in women.

Another important way that researchers use correlations is by examining the inter-regional correlations among brain areas. A high correlation between two voxels is taken to be a measure of functional or effective connectivity – the tendency for two brain areas to be coactive (Frith & Friston, 1997). One recent trend is to examine the effects of different tasks on functional connectivity. For example, a study by Coull, Buchel, Friston, & Frith (1999) found that connectivity patterns were different between an attention-demanding task and rest, suggesting that attention changes the functional connectivity of the brain.

While functional connectivity is often taken to mean the degree to which one brain area activates another, caution must be taken in the interpretation of such data, as with all correlational data. The data do

not indicate which of two “functionally connected” areas sends output to the other, or if both are influenced by a third area as the cause of the correlations between the two.

Although functional neuroimaging data are often analyzed in terms of separate regions that are differentially active among conditions, most psychological processes that we may want to study do not map 1:1 onto unique brain regions. They are often subserved by processing in distributed networks of interconnected areas some of which overlap and some of which do not. The mapping of regional correlations in conjunction with principal components analysis, described below, can be used to identify separate distributed networks that are related to different processes. Intuitively, voxels whose signal is correlated are grouped together to define a functional network in the brain, or “spatial mode,” which then becomes the unit of analysis for task-related effects. For example, Frith and Friston (Frith & Friston, 1997) describe a study in which they identified three distributed networks of brain areas that tended to be co-activated. One area was related to task performance, a second was related to the effects of practice within a session, and a third was related to magnetic artifact during the initial scans.

Another technique for examining networks of connectivity using correlational data is structural equation modeling. To use this technique, one creates an a priori model of expected patterns of connectivity and determines how well the data fit the pre-specified theoretical model. This approach can be very useful for testing hypotheses about networks of activations that may be involved in a task. Marschuetz et al. (1999), for example, investigated the ability of a tripartite model of working memory to account for imaging data in working memory tasks, and they were able to compare this model to several others to determine which provided the best fit.

One generalization of factorial design that is particularly useful was employed by Kanwisher, McDermott, and Chun (1997) to study face recognition. They identified an area on the fusiform gyrus that responded to pictures of faces and drawings of faces, but not to houses, scrambled faces, partial faces, facial features, animal faces, and other control stimuli. By presenting a large number of control stimuli of various types, Kanwisher et al. were able to infer that the brain area they studied was specific to the perception of faces. In general terms, they presented a number of different kinds of stimuli (each one a sort of factor, but without clearly defined levels) in an attempt to define both which stimuli elicit a

response from a region and which stimuli do *not* elicit a response. In the case of face recognition, it was very important to use a wide variety of control stimuli, as it could be argued that ‘face specific’ activations are really related to the color, general shape, or spatial frequency of the stimuli. This technique is particularly powerful for ruling out alternative explanations based on variables of no interest (e.g., spatial frequency of visual stimuli) that are confounded with processes of interest (e.g., face perception).

We have now provided broad coverage of the motivation for using neuroimaging data and the various techniques that can be used, with PET and fMRI as the imaging tools. Having covered this ground, we are now prepared to examine the details of these two imaging modalities. We do so in the next section.

III. The How: Data acquisition and analysis

a. The physics of PET and fMRI

Currently, functional neuroimaging techniques are based on the assumption that neuronal activity will cause changes in regional blood flow and metabolism that can be detected by the imaging technique of choice. If one discovers a regional change in blood flow or metabolism, then it is inferred that the neural activity in that region has changed. These changes in blood flow and metabolism are usually elusive to detect, and they require sophisticated statistical analyses to distinguish a real signal from the surrounding statistical noise. Very often, the statistical analysis of functional imaging data requires corrections for different effects that are specific to the acquisition technique, so it is quite important that the investigator understand the physics and the details of the experiment. There exist an array of methods for functional neuroimaging, each constituting an area of research in itself. Unfortunately, an in-depth review of all the available techniques would be too exhaustive for the scope of this chapter, so we shall focus on PET and fMRI. Our aim is to provide the reader with the background necessary to understand the data acquisition process, and the relationship between the acquisition and the analysis of functional data.

1. A brief summary of the physics of PET

Positron Emission Tomography is based on the detection of positrons emitted by a radioactive tracer that is injected into the subject. The scanner does not directly detect the positrons themselves; rather, energy is detected that is released by their annihilation. Some man-made isotopes decay by emitting positrons (subatomic particles having the same mass but opposite charge as an electron -- they are "anti-matter electrons"). Some isotopes that emit positrons include ^{75}Br , ^{18}F , ^{11}C , ^{15}O , ^{13}N , ^{68}Ga , ^{82}Rb , and they are usually made by bombardment of the atoms with accelerated particles. The decay rate of such isotopes is quite fast, and their half-lives are on the order of a few hours or less. Oxygen-15, for example, is the isotope used most frequently in studies of blood-flow using PET, and its half-life is approximately 2 minutes. This makes PET scans quite expensive in that it is necessary to have a cyclotron nearby in order to obtain a fresh supply of isotopes for the tracer.

When an emitted positron encounters an electron (from either the same isotope or from a neighboring atom), they collide. The result of this collision is that the positron and the electron are annihilated and two photons get ejected in opposite directions from one another. The laws of conservation of energy and mass dictate that the energy of the emitted photons be equal to the added masses of the electron and the positron. The law of conservation of momentum predicts that the momenta of the emitted photons be equal, but in exactly opposite directions. These two facts are very important because they imply that each of the emitted photons can be detected at around 511 KeV (the equivalent mass of an electron), and they may be detected *simultaneously* and *in pairs* by two detectors situated opposite one another, due to the opposite flight of the photons after annihilation.

[Insert Figure 8 about here.]

Thus, in order to establish the location of an annihilation event as well as to make sure that the detected photons indeed came from an annihilation event, one needs a set of detector pairs placed around

the source (i.e., - the subject's head). Additionally, each pair of detectors must be wired to a "coincidence detector" circuit, as illustrated in Figure 8. The coincidence detector counts only the photons that are detected pair-wise within a few nanoseconds of each other, and it dismisses other photons as background radiation. Ideally, the only photons that are detected are those that emerge from the annihilation of positrons in the tissue that is directly between the detectors in the pair. Unfortunately, photons from other locations can also be counted if they arrive simultaneously at the pair of detectors by sheer chance. In order to avoid detecting events that happen outside of the column of tissue between a given pair of detectors, there are usually small tubes (called "septa" or "collimators") placed around the detectors to shield them from radiation from the sides, while letting in the radiation that is coming in from the front. Depending on the design, most PET scanners are made up of an array of detectors that are arranged in a circle around the patient's head, or in two separate flat arrays that are rotated around the patient's head by a gantry.

2. Using PET for neuroimaging

When we inject a tracer into a subject, the tracer will distribute itself through the brain and accumulate in some locations more than others, depending on the tissue and the nature of the tracer. Let's use a two-dimensional function $\mathbf{D}(\mathbf{r})$ to describe the density of the tracer in a given slice of brain, where \mathbf{r} is a vector that indicates a location in space. The coincidence detectors simply count the number of coincidences (and therefore the number of emitted positrons) detected by a pair of detectors during the scan time. Thus, the number of positrons that are counted by each pair of detectors around the subject is proportional to the amount of tracer in a column of tissue running between the two detectors, as shown in Figure 9. In essence, the raw data from a PET scanner are a set of projections of the function $\mathbf{D}(\mathbf{r})$ onto the detector array at different angles, and the objective is to reconstruct the function $\mathbf{D}(\mathbf{r})$ from the projections.

[Insert Figure 9 about here.]

An intuitive way to think about the image formation process is to start out with a blank image, where all the pixels have a value of zero. Next, we take the individual intensities (number of counts) in one of the projections along a given angle, and we add these values to the image pixels along a line perpendicular to the projection, as illustrated in Figure 10a. We then move on to the next projection angle and repeat the procedure, adding the counts from the detectors along the new projection, and so on. The result is that different areas of the image will accumulate different numbers of counts from the projections, depending on the original distribution of the tracer in the plane, as shown in Figure 10b and Figure 10c. This distribution of tracer density is what constitutes an image. Now, since the number of projections is not infinite, and neither is the number of pixels in the image, some severe artifacts will occur in the image, and they must be compensated for by applying different filters to the data. This method is referred to as filtered backprojection.

[Insert Figure 10 about here.]

In practice, this procedure is usually implemented by using a Fourier transform. More rigorously, the projections can be described by

$$P(\mathbf{q}) = \sum_r D(r) \cdot \Delta r$$

where $\mathbf{P}(\mathbf{q})$ is the projection, or sum of the counts through the columns at the angle \mathbf{q} . Each $\mathbf{D}\mathbf{r}$ constitutes a portion of the object along the projection, as we saw back in Figure 9. It can be shown that the function $\mathbf{D}(\mathbf{r})$ can be reconstructed from all the projections, $\mathbf{P}(\mathbf{q})$, by computing the inverse Fourier transform of the data. Thus, the two-dimensional function describing the density of the tracers in a "slice" of tissue being imaged is given by

$$D(r) = FT^{-1} \{ r \cdot P(\mathbf{q}) \}$$

The reader should be aware that there are a number of other methods to reconstruct PET images, as well as corrections for scattering and other nuisances, that are beyond the scope of this book. We refer the user to the texts by Macovski (1983), and Sandler et al. (1995), and Bendriem and Townsend (1998) for greater details.

Thus, PET allows you to determine a map of the density of a radioactive tracer by reconstructing an image from the projections of the different angles. The tracers are usually physiologically relevant molecules that are labeled radioactively. One can label tracers that flow through the tissue, like water, or specific ligands that will bind to specific sites. This is where the strength of PET resides: it allows the researcher to measure a number of parameters with spatial specificity depending on the choice of tracer. There are three classes of technique in which PET is used, as summarized above. One is tracking regional cerebral blood flow, a second is tracking regional metabolism, and the third is tracking the binding of neurotransmitters to their receptors.

In most blood flow studies, radioactive water (H_2O-15) is injected intravenously, permitting measurement of blood flow by monitoring the passage of the labeled water through the tissue and measuring the uptake rate of the water into the tissue. Metabolism is measured using 18-fluorodeoxyglucose (FDG), a deoxyglucose molecule labeled with a radioactive Fluorine-18 atom. Just like glucose, it is taken up by tissue for energy production; by monitoring its uptake rate, one can identify regions of activity. For studies of receptor binding, radioactive labels have been developed for several hundred compounds related to specific neurochemical systems in the brain. The major neurotransmitter systems are most commonly studied, and this is accomplished by attaching radioactive labels such as ^{11}C or ^{13}C (carbon) or ^{123}I (iodine) to a receptor agonist or antagonist. Great care must be exercised in selecting and imaging radiolabeled compounds, as the observed level of signal depends on the concentration of the radiolabeled substance in the blood, the blood flow and volume, the binding affinity

of the substance to receptors, the presence of other endogenous chemicals that compete with the labeled substance, the rate of dissociation of the substance from receptors, and the rate at which the substance is broken down by endogenous chemicals.

3. A brief summary of the physics of MRI

MRI evolved from Nuclear Magnetic Resonance (NMR), a technique employed by chemists and physicists since the 1940's to study quantum mechanics, and to identify or characterize the structure of molecules (Bloch 1946; Purcell, Torrey, and Pound 1946; Hahn 1950). The raw signals in both NMR and MRI are produced the same way. As we will soon explain in more detail, a sample is placed in a strong magnetic field and radiated with a radiofrequency (RF) electromagnetic field pulse. The nuclei absorb the energy only at a particular frequency, which is dependent on their electromagnetic environment, and then return it at the same frequency. The energy is in turn detected by the same antenna that produced the RF field. In NMR experiments, the types of nuclei present in a molecule can then be identified and quantified by analyzing the frequency characteristics of the returned signal. It was not until the 1970's that it was realized that one could obtain spatial information about the nuclei emitting the radiation by manipulating the magnetic fields around the sample (Mansfield 1978; Lauterbur 1973).

Let us now examine more closely the production of a signal in an NMR experiment, and then proceed to how one can obtain spatial information from that signal, in order to obtain an image. As most people know, the human body is made of mostly water. The brain is no exception; so let us consider the hydrogen atoms that are present in a water molecule. They are made up of a single proton and a single electron. Every proton has its own magnetic dipole moment represented by a vector. The magnetic moment is the amount of "magnetization" of an object, and it determines how strongly it interacts with magnetic or electric fields (a bar magnet is a dipole, and a very strong one would have a very large dipole moment).

When they are placed in a magnetic field, such as that of an MR scanner, a portion of the protons, or "spins," will align with or against the magnetic field. A couple of things should be kept in mind about this alignment: first, the larger the magnetic field, the greater the proportion of spins that are aligned, making the alignment easier to detect. Second, whether the spins are aligned with or against the field is determined by their spin quantum number, which can have values of $+1/2$ and $-1/2$. Being aligned with the magnetic field takes less energy than being aligned against it, so a greater number of the spins will be aligned in the direction of the field.

[Insert Figure 11 about here.]

The interaction between the main magnetic field (usually labeled \mathbf{B}_0) with the proton dipole produces a set of forces that result in the *precession* of the dipole. Precession of a vector is a movement that takes place such that the origin of the vector stays fixed, while the tip spins describing a circle around a vertical axis, as shown in Figure 11. The vectors representing the magnetic moment of the $+1/2$ spins will precess about the magnetic field, and the $-1/2$ spins will precess about the opposite direction of the magnetic field. The rate of precession, w_0 (i.e. - the angular velocity of the spins' precession) is proportional to the magnetic field \mathbf{B}_0 , as described by

$$w_0 = g \cdot B_0$$

where g is a constant called the gyromagnetic ratio. This nicely linear relationship between the precession frequency and the magnetic field is a key factor that will enable us to obtain spatial information about the sample by simple manipulations of the magnetic field. The gyromagnetic ratio is specific for the nucleus in question (Hydrogen nuclei's g is 42.58 Mhz/Tesla.), which can allow us to obtain NMR signals from specific nuclei without interference from other nuclei. The molecular environment around the nuclei (the number of electrons present, the proximity of other nuclei, ...etc) can change the \mathbf{B}_0 field around the nuclei and thus alter their precession rate, as predicted by the equation

above (that is how one can make inferences about the molecular structure of a molecule that contains protons).

Let's now consider the *net* magnetization vector of a population of spins. Together, the spins' magnetization vectors add up to a single magnetization vector that is aligned with the magnetic field (see Figure 12). Since the x and y components of the magnetic moments are randomly oriented at any given time, they cancel each other when all the vectors in a large population are added together. Thus, all that remains is the component that is parallel to the magnetic field along the z axis (remember that a greater number of the spins align with the field than against it).

[Insert Figure 12 about here.]

Now that we have a picture of the behavior of the magnetic moments of water protons in a large magnetic field (\mathbf{B}_0), let us consider what happens when a second magnetic field (\mathbf{B}_1) is applied in a direction perpendicular to the main magnetic field. This \mathbf{B}_1 field is generated by the RF coil in MR experiments, and it rotates at a particular frequency. If the \mathbf{B}_1 field rotates at the precession frequency of the spins, it 'looks' to them like a stationary magnetic field, since they are both rotating at the same rate. In fact, we can look at the whole system from a rotating frame of reference, to simplify things (consider how things look when you ride in a carousel. The other children don't seem to be moving, but their parents, and anything outside the carousel does). In this rotating frame of reference, we now have a magnetization vector, \mathbf{M} , which is aligned with the main magnetic field, \mathbf{B}_0 , and a second stationary magnetic field \mathbf{B}_1 . According to classical physics, the \mathbf{B}_1 field will exert a torque on the magnetization vector such that it is rotated onto the x - y plane at an angular velocity determined by the magnitude of \mathbf{B}_1 . This is illustrated in Figure 13.

[Insert Figure 13 about here.]

In our rotating frame of reference, after we turn the \mathbf{B}_1 field off, the magnetization vector is stationary on the x - y plane, but relative to the real world, the magnetization vector is rotating about the z -axis on the x - y plane at an angular velocity ω_0 . A property of classical electromagnetism is that changes in a magnetic field will induce electrical currents in a wire coil. The antenna used for transmission of the RF pulse is such a coil, and when the magnetization vector rotates through it, it induces a current in it. This current induced in the coil is the NMR signal that we observe. The current oscillates at the frequency of the angular rotation of the magnetization vector (this is the same frequency that we used to transmit the RF pulse, also called the "resonance" frequency)

When the transmitter is turned off after the application of a pulse, the magnetization vector will "relax" back to its equilibrium position. This relaxation happens through several mechanisms: 'Spin-lattice' relaxation occurs as the spins give away their energy, and they return to their original quantum state. This translates into the longitudinal (i.e. - along the z -axis) component returning to its equilibrium value at a rate T_1 . Spin-spin relaxation happens along the transverse (i.e., on the x - y plane) component of the magnetization vector. It is due to the ensemble of spins falling out of phase with each other and thus adding destructively to the net magnetization vector as illustrated in Figure 14. These two mechanisms are often referred to as T_1 and T_2 relaxation, respectively.

[Insert Figure 14 about here.]

Another kind of relaxation is caused by inhomogeneities in the magnetic field at the microscopic level. If there are variations in the magnetic field, there will also be variations in the individual protons' precession frequency, which causes the ensemble to lose phase coherence faster than expected due to simple T_2 . This change is referred to as T_2^* (pronounced 'T2 star'). The relaxation rate constants T_1 , T_2 , and T_2^* are dependent on a number of properties of the nuclei themselves and of their environment at the molecular level. This is quite useful for several things. The relaxation constants can be used to identify the nuclei in an NMR spectroscopy experiment, or they can provide a mechanism for image contrast

between different tissues, such as white and gray matter, or lesions, when performing an imaging experiment. For example, T_1 weighted images are images that are acquired with parameters such that the image contrast between tissues is mostly determined by their T_1 relaxation rate. An example of the same slice of tissue imaged with T_1 and T_2 weighting can be seen in Figure 15. As you can see, the images look strikingly different. Changing the contrast mechanism can be very useful in differentiating brain structures or lesions, since some structures will be apparent in some kind of images but not in others. For example, multiple sclerosis lesions are virtually invisible in T_1 weighted images, but appear very bright in T_2 weighted ones.

[Insert Figure 15 about here.]

4. From the NMR signal to neuroimaging

Now that we have an idea of how a signal is produced, let us take a look at how we can extract spatial information from it so that we can form an image. We mentioned above that the precession frequency of the spins (and thus their resonance frequency) was proportional to the strength of the magnetic field. Now, consider what happens when we apply another magnetic field in the direction of \mathbf{B}_0 , except that this field varies linearly with location along the x -axis (This is referred to as a magnetic field gradient in the x direction). What we have now is a magnetic field whose intensity changes in direct proportion to the location in space along the x -axis. Since the magnetic field strength varies with the position in space, the resonance frequencies of the spins also vary with their position in space (recall the equation $\omega_0 = \gamma \mathbf{B}_0$).

Thus, if we tip the spins onto the x - y plane with a \mathbf{B}_1 pulse and then turn on a magnetic field gradient, the signal that we get back from the sample will not simply oscillate at the resonant frequency, as we described earlier. It will be a more complex signal made up of the sum of the signals generated by the tissue at different locations along the x -axis, and thus oscillating at different frequencies. This

technique is usually called 'frequency encoding'

The contribution of each frequency component of the signal is proportional to the magnitude of the magnetization vector at the corresponding location. So, if we can separate the different frequency components of the signal, we'll get the distribution of magnetization across the x -axis in space. Luckily, there is a mathematical technique designed to do exactly that: the Fourier Transform separates a function into its frequency components.

In reality, things are a bit more complex. Since the spins at different locations along the x -axis are precessing at different rates in the presence of the gradient, their magnetization vectors get out of phase with each other, causing the net x - y magnetization to decay quickly. The spins can be refocused in two different ways. One could apply a gradient of equal but opposite intensity to un-do the phase gain caused by the first gradient, such that the spins will regain their phase coherence and form what's called a "gradient-echo". Alternatively, one could also apply another RF pulse to rotate the magnetization 180 degrees, then re-apply the original x gradient, such that the spins regain their phase coherence, as shown in Figure 16.

[Insert Figure 16 about here.]

We have seen how we can obtain spatial information along a single dimension, but, in order to form an image, we need to extract the distribution of densities along at least two dimensions. We need to come up with a method to encode the spatial information along the both the x - and y -axes. The way to that is to perform frequency encoding along the x -axis as before and, at the same time, apply a brief gradient along the y -direction, such that the precession of the spins gets a little bit ahead depending on where they are along the y -axis. Recall that applying a gradient causes the spins to precess faster or slower depending on their location. Thus, a short gradient pulse causes them to change their precession rate briefly, resulting in a phase gain that depends on the location of the spins along the y -axis and the duration of the gradient

pulse.

The sequence is repeated a number of times, so that we can get a whole distribution of phase gains along the y -axis. The end result is a two dimensional data set that contains the x -direction distribution of densities encoded in frequency along the x -axis and the y -direction distribution of densities encoded in phase along the y -axis. The two-dimensional Fourier Transform of this raw data image is an image of the magnetization vector across the imaging plane.

It is very useful to have an image of the magnetization vector across tissue, because there are a number of tissue specific properties that affect the magnetization and thus provide with a contrast between different kinds of tissues. These include the water content, and the T_1 , T_2 , and T_2^* relaxation rates. As mentioned, the pulse sequence parameters can be manipulated to emphasize the contrast due to any of those properties individually.

There are many different ways to form an image using MR, and we have discussed only one of them in order to give the reader an idea of the principles underlying the formation of an MR image. For further information on the many techniques available, and a more rigorous treatment of the subject of MR imaging, we refer the reader to excellent texts such as Nishimura's (1996), or Elster's (1994).

5. Functional MRI using the BOLD effect

Let us now explore how we can take advantage of MR imaging to obtain functional images by taking advantage of the BOLD (Blood Oxygenation Level Dependent) effect. Functional studies can be made because the intensity of the detected water signal depends on many parameters, as we mentioned earlier. Among them, it is dependent on the water density and T_2^* relaxation rate of the tissue. The hemoglobin in blood can take two different conformations, depending on whether it is oxygenated or not. In its deoxygenated state, the iron atoms are more exposed to the surrounding water, creating small distortions in the \mathbf{B}_0 field. The magnetic susceptibility of a substance is its ability to distort a magnetic field, and it affects the relaxation constant T_2^* . Thus, the magnetic susceptibility of hemoglobin is lower when it is in its deoxygenated state, and this change in susceptibility translates into a shortening of the T_2^*

of the deoxygenated blood (Ogawa, Lee, Kay, and Tank, 1995).

When brain tissue becomes active, it requires more oxygen than when it is at rest. In order to accommodate this need, the amount of oxygenated blood to the tissue increases via a blood flow increase. During periods of activation, the increase in blood flow brings in more oxygenated blood, decreasing the concentration of deoxyhemoglobin. Thus, the increase in blood flow and blood volume contribute to an increase in signal, and the increase in magnetic susceptibility increases the amplitude of the water signal. The net result is an increase in signal following activation. It is important to realize that the degree to which the blood flow and the deoxyhemoglobin content are coupled can vary, and modeling the exact properties of the BOLD response is currently a topic of intense research (Vazquez, and Noll, 2000; Buxton, Wong, and Frank, 1997; Frahm, Merboldt, Hanicke, Kleinschmidt, and Boecker, 1994).

An alternative technique is to measure changes in blood flow alone using "Arterial Spin Labeling" (ASL) techniques (Williams, Detre, Leigh, and Koretsky, 1990; Detre, Leigh, Williams, and Koretsky, 1992; Kim, 1995). ASL is an MRI-based technique that mimics PET blood flow techniques by tracking the passage of a tracer through the tissue. In the case of PET, the tracer is a radioactive substance that is injected intravenously. In the case of ASL, the tracer is simply the water in the arterial blood, which is magnetically labeled by an RF pulse. The label consists of tipping the arterial water's magnetization vector all the way to the negative z-axis by a B_1 pulse that is applied somewhere upstream from the tissue of interest. As those "inverted spins" flow through the tissue, they can be detected by changes in the signal intensity of the image. There are a number of limitations that render the technique impractical for many applications, but overcoming these limitations is a growing area of research (Kim 1995, Gonzalez, Alsop, and Detre, 2000; Wong, Buxton, and Frank, 1997) and the technique will likely soon become a powerful tool for functional studies.

6. Diffusion tensor imaging

Diffusion tensor imaging can be used to explore questions about connectivity among brain regions by identifying the orientation of white matter tracts. The technique produces images that are

dependent on the direction of diffusion of the tissue water, which yields information about the orientation of the tissue. For example, in the case of a pot of water, the water molecules are equally likely to diffuse in all directions, except near the walls, which restrict the movement of the water molecules. If we were to put a bunch of lasagna noodles into the pot of water such that they lay flat on top of each other, the water would be more likely to move horizontally than vertically because the water molecules would be more likely to bump against the lasagna noodles when they try to move vertically than when they try moving horizontally. Similarly, the geometry of the white matter tracts in the brain running parallel between two different structures restrict the diffusion of water molecules along all directions *perpendicular* to the direction of the tracts. As we will soon discuss, the diffusion of the spins can affect the MR signal, and we can take advantage of this phenomenon to obtain images that are sensitive to the microscopic geometry of the tissue even though MR images do not afford the resolution to see the actual microscopic structures.

When a proton moves in the presence of a magnetic field gradient, its precession frequency varies depending on its location along that gradient, as we saw earlier. That means that it acquires a phase difference in its rotation relative to the rest of the ensemble of stationary spins (here, "acquiring phase" means that the magnetization vector for that spin gets ahead of the rest). If all spins move together in the same direction and at the same speed, then they all acquire the same amount of phase coherently, and the magnitude of the net magnetization vector would not be altered. However, in the diffusion process, movement happens random and incoherently among the spins in the ensemble, so the net effect is a signal loss, since some of the moving spins will acquire phase, and some will lose phase, depending on which direction they move.

The degree of attenuation seen in the signal is related to the freedom of movement of the water molecules in the direction of the applied gradient. This has found a number of applications for clinical imaging, such as providing information about the brain tissue cells' membrane integrity. It can also give information about the orientation of the tissue through acquiring diffusion tensor images. Diffusion

tensor images are produced by applying diffusion gradients during the imaging process in different combinations of the x, y, and z gradients. The result is a set of images that are weighted according to the restriction of water movement along the direction of the applied gradient combination. (Le Bihan 1995, Moseley et al., 1990)

Consider again the shape of white matter tracts. If two areas are functionally connected, then one can expect that there are a large number of tracts running between the two areas. If in a region of tissue there is a large number of tracts running parallel in a particular direction, then the diffusion of water is less restricted along that direction, since the water molecules are more likely to collide against fibers when they move in directions other than those of the fibers. Thus, by obtaining images whose intensity is proportional to the diffusion coefficient of water in a particular direction in space, one can obtain information about how the tissue is structurally laid out, giving information about what regions of the brain are structurally interconnected.

b. The biophysics of PET and fMRI

We have described several ways that changes in blood flow and oxygenation can be detected by neuroimaging scanners. Critical to the undertaking is the assumption that these changes reliably result in a signal that can be detected by a scanner. However, before this signal change can be interpreted as neural activation, there is another critical assumption that must be justified—the assumption that changes in blood flow and oxygenation reflect changes in neural activity.

Roy and Sherrington (1890) were the first to hypothesize a connection between blood flow and neural activity. Since then, the mechanism behind the relationship between blood flow and neural activity has been investigated at length. For example, Shulman and Rothman (1998) have proposed that increased glucose uptake is controlled by astrocytes, whose end-feet contact the endothelial cells lining the walls of blood vessels. Glutamate, the primary excitatory neurotransmitter in the brain, is released by some 60-

90% of the brain's neurons. When glutamate is released into synapses, it is taken up by astrocytes and transformed into glutamine. When glutamate activates the uptake transporters in an astrocyte, it may signal the astrocyte to increase glucose uptake from the blood vessels. Vasodilation resulting in increased blood flow and increased oxygen consumption may be coupled to neural activity through similar mechanisms. If it is only glutamate release that triggers the vascular, oxygen uptake, and glucose uptake effects, then "activation" is excitatory. However, release of GABA or other inhibitory neurotransmitters could trigger these responses as well, and further research is needed before firm conclusions are reached about exactly what changes we may infer are taking place when we measure changes in blood flow or BOLD signal.

Also, importantly, the relationship between neural activity and glucose uptake indicates that the neuroimaging signal reflects activity in the neuropil, at the synapses where neurotransmitters are released, and not in the brain regions containing cell bodies. A neuroimaging signal may therefore be related to increased *input* into an area, which may or may not lead to increased output from that area to other local or remote brain regions. If a task activates dorsolateral prefrontal cortex (DLPFC), for example, it means that DLPFC is receiving substantial input from other areas. That input could be excitatory or inhibitory.

While there is still uncertainty about the exact mechanism by which a neuroimaging signal is produced, sufficient evidence has been collected that we may proceed forward with reasonable confidence. In the end, all the available indices of neural activation – rCBF, oxygen uptake, or glucose utilization – may be suitable for most studies of psychological function.

c. Statistical analysis of neuroimaging data

Armed with a general understanding of the physical information contained in PET and fMRI images, we are now in a position to extract information about brain function from our imaging studies. In the case of PET, there will be images acquired under different experimental conditions, whose signal

intensity will be dependent on the amount of tracer present in each voxel. In the case of fMRI, the signal intensity will be dependent on the BOLD response (based on oxygenation and blood flow). The task at hand for analysis of the signal is primarily to identifying those voxels whose activity matches a predicted model, be it a model due to subtraction logic, parametric variation, factorial manipulation, or correlation. Because fMRI has become the dominant modality for the collection of data about regional cerebral blood flow, we will focus our discussion below on the analysis of data from an fMRI experiment. However, most of the principles are applicable to PET as well.

1. The General Linear Model.

Consider an *ideal* experiment in which a subject's brain is inactive when there is no task to perform, and it is activated by an experimental task of interest. Each time the task is performed, a set of physiological events takes place in a functional region resulting in a BOLD response, as described earlier. By the General Linear Model, we think of the brain as a linear, time-invariant system, whose input is the task, and whose output is the BOLD response. Thus, for our ideal experiment, the input function can be considered as a train of spikes corresponding to the psychological events of the task (e.g., encoding stimuli, making decisions, executing responses), where the events happen at times $t_1, t_2, t_3 \dots$. Each of these events may cause neural tissue somewhere in the brain to become activated, which in turn causes hemodynamic changes. The change in signal intensity in any voxel of the brain is described by the convolution of a function describing the train of events with a function that describes the hemodynamic response to a single stimulus. Figure 17 shows the response to a single stimulus, as well as the response to a train of stimuli occurring at random times.

[Insert Figure 17 about here.]

By design, psychological events from different conditions of an experiment may be intermingled, as in an event-related design, or they may be grouped into epochs, as in a blocked design.

Under the assumptions of the General Linear Model, different tasks constitute different input functions that give rise to their own BOLD responses. So, if one constructed an experiment with different conditions (i.e., different tasks) intermingled in an event-related design, one could construct different input functions to model the output function. In order to create a more realistic model for the observed signal, one must also include other input functions for such variables as drift in the scanner signal, effects due to motion of the subject, effects of respiration and heart rate, and other nuisance variables. Thus, the observed signal from a voxel can be thought of as a sum of weighted functions corresponding to the different predicted effects. Some of these effects are of interest, some are not, and the weights of those functions are a set of scalar parameters that determine the amplitude of those functions. For example, if our model is made up of effects that are represented by functions $x_1(t)$, $x_2(t)$, $x_3(t)$..., and these are weighted by the coefficients β_1 , β_2 , β_3 ..., then the *predicted* signal is given by

$$y(t) = \mathbf{b}_1 \cdot x_1(t) + \mathbf{b}_2 \cdot x_2(t) + \mathbf{b}_3 \cdot x_3(t) + \dots + \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \dots$$

Where the \mathbf{e} terms represent the error. All of this can be expressed in matrix form as:

$$Y = \mathbf{b}X + E$$

Where \mathbf{Y} is the observed signal in a given voxel expressed as a vector whose elements are the individual time observations, \mathbf{X} is a matrix whose columns contain the individual functions that make up the model, \mathbf{b} is a vector containing the weights of the individual component functions, and \mathbf{e} is the residual noise in the measurement. The matrix \mathbf{X} is often referred to as the "design matrix" and displayed as an image whose intensities correspond to the values of the elements, as shown in Figure 18. Our task in the analysis is to obtain an estimate for \mathbf{b} , and to identify the voxels that fit the estimated model.

[Insert Figure 18 about here.]

It can be shown that \mathbf{b} can be estimated by:

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

Note that $\hat{\mathbf{b}}$ is used to represent the estimated value of \mathbf{b} . Now, what remains is to test each individual voxel to see which ones fit the model described by \mathbf{Y} and $\hat{\mathbf{b}}$. Assuming normally distributed noise, this can be accomplished by computing the correlation between the model and the data for every voxel. Depending on the statistical approach taken, one can obtain a T or an F-score for the correlation.

One could ask many questions using the same model; in fact, one can test for the presence of any given linear combination of the covariates contained in the design matrix's columns by multiplying the parameter estimates by a *contrast* vector \mathbf{c} . This vector contains additional weights to be multiplied by the parameter estimates in the vector \mathbf{b} . In the design matrix shown in Figure 18, for example, we could test for which voxels the activity in the first condition is greater than in the second condition while disregarding all the other effects as nuisance covariates, by using a contrast [1 -1 0 0 0 0]. Conversely, we could test whether the activity in the second condition is greater than in the first by using a contrast [-1 1 0 0 0 0].

2. Statistical Inference from the General Linear Model

As usual with all statistical procedures, we must calculate the significance of the correlation between the model and the data in each voxel (whether it is a correlation coefficient, T, F, or Z-score). This is usually done by calculating the probability of obtaining that value by sheer chance, given the probability distribution of the statistic. In neuroimaging, however, we must beware of the fact that there are hundreds of thousands of voxels in an image, so a number of them are bound to be correlated to the design matrix by chance. If the voxels' signals were independent of each other, we could compute a

Bonferroni correction of the significance level, but that is not usually the case because the voxels tend to be correlated with their neighbors. Instead, we have to come up with a method of examining the statistical image (made up of T values or such), and calculating the likelihood of having a cluster of voxels above a given threshold. Random Fields Theory (RFT) does exactly that. Based on the assumption of Gaussian-distributed background noise in an image, we can measure the spatial characteristics of the distribution in three dimensions, and from those measurements we can make predictions about the number of clusters that can be expected to appear significant in the statistical image just by chance. The Euler Characteristic of a solid geometric figure is a measure of how many of its elements are connected together and how many holes exist within it. As it turns out, the expected Euler Characteristic of a thresholded statistical image is a good approximation of the likelihood that a cluster of voxels above a certain threshold will occur by chance in a random image. The calculation of the expected Euler Characteristic is based on a calculation of the "smoothness" of the image and is beyond our scope. The smoothness of an image is a measurement of how many independent measurements exist within the image. These independent measurements are referred to as "resels" (short for resolution elements). This sort of technique has great applications in the analysis of noisy imaging data, when the objective is to identify significant clusters (not just functional neuroimaging, but astronomy as well). For greater detail on the calculation of the smoothness of the image and the Euler Characteristic, we refer the reader to Worsley (1992, 1995, and 1996), Peterssen (1999), and Friston (1999).

3. Assumptions

The main assumption underlying the General Linear Model approach is that the BOLD response to a set of neuronal processes is a time-invariant, linear combination of those processes. A time-invariant system is one whose response to a given input is always the same, regardless of the previous events. Linearity means that if two separate inputs are applied to the system, its response will equal to the sum of the individual responses to those inputs. It is becoming increasingly clear that the BOLD response is neither linear nor time-invariant (Vazquez and Noll, 1998; Boynton, Engel, Glover, and Heeger, 1996;

Buxton, Frank, 1998) but these violations are not severe within reasonable boundary conditions.

Because of the necessity to evaluate significance for the computed statistics, the General Linear Model is also heavily dependent on the theory of Gaussian Random Fields, whose main underlying assumption is that the residual variance in the images after applying a model is distributed normally, and that each voxel's signal is independent of the signal in other voxels. Unfortunately, for both PET and fMRI, the signal intensity in one voxel is always contaminated by the signal of other voxels. In the case of PET, the correlation is due to scattering of the positrons and to smoothing of images during data preprocessing. In the case of fMRI, it is largely a function of limitations in resolution and of any smoothing that is done during preprocessing. Thus, one must take care to be sure that these correlations are not serious contaminants of the data, and the data must be spatially smoothed as described below.

4. Pitfalls

A major concern in the analysis of fMRI data is that the BOLD effect is a vascular one, not an electrochemical one. It is a response to the underlying neuronal activity, and it distorts that neuronal activity to the extent that it does not mimic it directly. The limitations here are ones of time and space. In time, the hemodynamic response lags behind the neuronal response by as much as several seconds, and it is stretched out longer than the neuronal response as well. In space, the blood flow changes that are measured may or may not be in the immediate neighborhood of the underlying neuronal response that caused them because the vasculature is not tuned precisely to the spatial location of the neural tissue to which it is responding. Thus, to have a good idea of when and where a neural response occurred, one needs to have a good idea of the nature of the hemodynamic response in that part of the brain; this is currently a matter of extensive study.

Using the General Linear Model also poses a number of limitations on the analysis of neuroimaging data. The General Linear Model is used to ask whether the data fit a set of predictions.

Thus, one must have a set of predictions. If these are wrong, then one might repeat an analysis with a different model, looking for a better fit. There are other approaches to data analysis, such as Principal Components Analysis and Independent Components Analysis. These extract the underlying functions from the response without an a priori guess (we'll discuss these briefly below). At the same time, though, these approaches yield no information about which component corresponds to which process.

When building a design matrix for the General Linear Model, one must be very careful to include all the effects present in the data, including confounds. At the same time, one must also be careful not to include too many effects in a single experiment. If one “under-parameterizes” the analysis, the confounds’ variance can overwhelm the signal, making the effects of interest insignificant. If one “over-parameterizes,” one expands the search space for the \mathbf{b} coefficients, making it more likely to get erroneous results.

As we have discussed earlier, the BOLD response is not always linear and at the present time, experiments must be designed such that the BOLD responses will be in the near-linear range. Otherwise, the regressors in the model will not fit the data well enough to yield accurate results. To approximate linearity, one must ensure that the intervals between trials are within one to fifteen seconds long. In addition, longer stimulus durations tend to produce more linear responses (Vazquez, 2000).

5. Preprocessing Requirements

There are several conditions about the fMRI images that must be met in order to carry out a successful data analysis. Most analyses are based on the assumption that all the voxels in any given image of the series of images taken over time were acquired at the same time. They also assume that each data point in the time series from a given voxel was collected from that voxel only. Another assumption is that the residual variance (i.e. - variance remaining after removing all the effects of interest) will have a Gaussian distribution. Additionally, when carrying out analyses across different subjects, we assume that

any given voxel will correspond to the same brain structure in all the subjects in the study. Without any pre-processing, none of these assumptions hold entirely true, and they will introduce errors in the results, so there are several steps to carry out before diving into an analysis so that the data will meet (or at least approximate) the assumptions.

Slice Timing. Since most image-acquisition sequences acquire brain images slice by slice, there can be a difference of one to three seconds between the acquisition of the first slice and acquisition of the last slice. The problem is that an analysis assumes that all the voxels in an image acquired at a given time point of the time series are acquired at the same time. In reality, the data from different slices are shifted in time relative to each other. Thus, we need to calculate the signal intensity of all the slices at the same moment in the acquisition period. This is done by interpolating the signal intensity at the chosen time point from the same voxel in previous and subsequent acquisitions. A number of interpolation techniques exist, from bilinear to sinc interpolations, with varying degrees of accuracy and speed. Event-related experiments require a more precise control over the onset time of the stimulus than blocked-design experiments, so the interpolation is often not necessary in blocked designs, where the epochs can last many seconds (>30sec). Because of the long length of epochs, there will not be much sensitivity lost by not having had the slices all collected at the same time.

Realignment. A major problem in most time-series experiments is movement of the subject's head during acquisition of the time series. When this happens, the image voxels' signal intensity gets "contaminated" by the signal from its neighbors. Thus, one must rotate and translate each individual image to compensate for the subject's movements.

The coordinates of a point in 3D space (x,y,z) can be expressed as a vector. It can be shown that the coordinates of a given point in space after any given translation, rotation, or combination of both, can be calculated by multiplying a matrix by the original vector. Such a matrix is called an *affine transformation* matrix. Thus, in order to undo the rotation and translation of the head, we must calculate

the elements in this affine transformation matrix, and apply it to all the voxels in the image. Usually, this is done by a least squares approximation that will minimize the difference between the image to be corrected and the first image in the time series.

Smoothing. Random Field Theory requires that each voxel is independent of the other voxels, and that the images have normally distributed noise. This is not the case in most experiments, since the signal is often correlated among different voxels, especially in fMRI experiments. In order to make the noise in the images meet the assumption, the images are convolved with a Gaussian kernel, which gives the noise a more Gaussian distribution. This smoothing of images also effectively produces a weighted average of the signal across neighboring voxels, which gives the smoothed images a blurry appearance. A side effect of smoothing is a reduction of the amount of high frequency spatial noise present in the data. This can work to our advantage by increasing the overall signal-to-noise ratio of the individual images in the time series, making the tests more sensitive at the expense of spatial resolution.

Normalization. In order to make quantitative comparisons across subjects, the corresponding brain structures must have the same spatial coordinates. Of course, this is usually not the case, since most people's brains are different. We can however, stretch and compress the images (not the actual brains, of course!) in different directions so that the brain structures are in approximately the same locations. Usually we normalize all the brain images so that they will match a standard brain (e.g., the Talairach brain or the Montreal Neurological Institute brain templates).

The normalization process includes an initial realignment of a set of images so that they approximate the template in orientation. Additionally, the images are transformed by multiplying them by a series of cosine basis functions, whose coefficients must be estimated by a least squares error-minimization approach. For those familiar with Fourier or Taylor series, it is as if we search for some function that will give the right transformation of the image. Since we do not know what the function is, we search for coefficients in the lower order terms that would make up the unknown function. For more

information on the techniques to estimate the parameters, we refer the reader to Frackowiak (1997).

6. Random effects

One approach to analyzing multi-subject data is to normalize all images from all subjects and concatenate them into the design matrix, while including additional regressors for each subject. The result is a massive analysis including all trials from all subjects, which is quite expensive from a computational point of view. This is referred to as a fixed-effects analysis. Such an analysis would answer the question: "If we repeat this experiment many times on the same subjects, what is the likelihood that we'll get the same significant voxels?"

If, on the other hand, we were interested in making a statistical inference about the *population* from which those subjects were taken, we would need to first analyze each subject separately, then look for commonalities across the statistical maps obtained in this first level of analysis. It has been shown that one can make statistical inferences across subjects by simple statistical tests performed on the statistical parameter maps (Friston et al, 1999; Holmes and Friston, 1998). The tests (usually t- tests) can be carried out on the maps of \mathbf{b} estimates calculated in the general linear model to search for those voxels that give the same magnitude of response to the stimulus. Those tests can also be carried out across the T-statistic maps obtained from the analysis of individual subjects, in order to search for voxels with the same level of significance.

Thus, we can perform a multi-subject analysis by two stages: first, the estimation of parameters at the individual-subject level, and then another test of the individuals' statistical maps across subjects to see which voxels show the same level of activation across subjects. When we do this, we are assuming that 1) the images have been spatially normalized such that the tests are conducted on corresponding structures from subject to subject, 2) the global intensity of the images has been scaled to a common level, and 3) all brains have similar BOLD responses to the same activity. These assumptions are not

always met perfectly and the introduce some variance into across-subjects analyses.

7. Other approaches: Principal Components and Independent Components

Analysis.

An analysis based on the General Linear Model lets you identify the voxels whose brain activity matches your model, but it doesn't reveal any additional information about the activations. Additionally, in the presence of unknown or non-linear confounds, analyses based on the General Linear Model are not effective in removing the variance due to those confounds. A few methods based on the temporal signal have been designed in order to identify the major task-related patterns of activation in the brain without any a priori knowledge of the stimulation paradigm. Principal Components and Independent Components Analyses are among these.

Principal components analysis calculates spatial patterns that account for the greatest amount of variability in a time series of images. This is done by obtaining the eigenvectors of a matrix containing the covariance among all voxels of the time series images. The eigenvectors, \mathbf{x} , of a matrix, \mathbf{A} , are those that satisfy the condition

$$\mathbf{Ax} = \lambda \mathbf{x}$$

where λ is a scalar value called the eigenvalue. The usefulness of the eigenvectors of a matrix is that they provide a set of basis functions for the original matrix.

Usually, the eigenvectors are calculated through single value decomposition (SVD) of a covariance matrix. The result of SVD of a matrix is a set of three matrices whose columns are orthogonal vectors, \mathbf{U} , \mathbf{S} , and \mathbf{V} . In the framework of neuroimaging, \mathbf{U} is interpreted as the temporal patterns present in the time series, \mathbf{V} as the spatial patterns of covarying voxels, and \mathbf{S} is a diagonal matrix whose elements

are a measure of how much variance is accounted for by a particular spatial pattern. The columns of the matrix V can be shown to be the eigenvectors of the original data matrix. For details on eigenvectors, eigenvalues and SVD, we refer the reader to a linear algebra text, such as Strang's (1988).

Independent Components Analysis is akin to Principal Components Analysis in that the Independent Components algorithm also produces a set of components of the signal. However, in Independent Components Analysis, there is a constraint that the components be statistically independent, and not necessarily orthogonal. Orthogonality, which characterizes Principal Components Analysis, implies that the voxel-values are uncorrelated between all pairs of components, whereas statistical independence, which characterizes Independent Components Analysis, implies that the joint probability of all the components is the same as the product of the individual probabilities, and that higher-order correlations between the components are also zero. Thus, Independent Components Analysis involves a more stringent criterion (McKewon 1998a). The algorithm to extract the independent components is an iterative procedure based on information theory and is beyond the scope of this chapter. Let it suffice to say that the algorithm searches for a solution that will maximize the entropy (or minimize the mutual information) between the components. For more details, we refer the reader to McKewon 1998a, McKewon 1988b, Bell and Sejnowski, 1995; Petersen et al., 1999.

d. Summary

We have completed our tour of the Why, the What, and the How of neuroimaging. There are many reasons one might delve into neuroimaging, both for an understanding of brain mechanisms, and for an understanding of psychological mechanisms. Having recognized this, cognitive neuroscientists have developed a number of techniques that allow one to implement neuroimaging techniques in experimental contexts of interest to psychology. Understanding the physics of how these techniques work is important to understanding what they offer and what constrains them. Equally important is understanding how experiments are designed to maximize their inference-making power, and what analysis methods are

available. Having surveyed these issues, we have proffered a tour of the highlights. The interested student of cognitive neuroscience will benefit from deeper analyses of all the topics we have surveyed, which are available from several sources. Other excellent introductory papers include those by Frith and Friston (1997), Aguirre and D'Esposito (1999), and Buckner and Braver (1999). Some useful texts include Elster (1994) for a clear explanation of MRI principles, Strang's (1988) linear algebra text, and Frackowiak et al.'s Human Brain Function (1997). There is also a vast amount of information on functional imaging on the internet, at sites such as the CBU's (<http://www.mrc-cbu.cam.ac.uk/Imaging/>) or the FIL's (<http://www.fil.ion.ucl.ac.uk/spm/>) or many others.

References

- Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for brain fMRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), Functional MRI. New York: Springer.
- Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. Neuroreport, *10*, 189-94.
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. Neuroimage, *8*, 360-9.
- Andreasen, N. C., Arndt, S., Swayze, V., 2nd, Cizadlo, T., Flaum, M., O'Leary, D., Ehrhardt, J. C., & Yuh, W. T. (1994). Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. Science, *266*, 294-8.
- Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppel, R.A., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from PET. Psychological Science, *7*, 25-31.
- Baddeley, A.D. (1986). Working memory. Oxford: Oxford University Press.
- Baddeley, A.D. (1992). Working memory. Science, *225*, 556-559.
- Badre, D. T., Jonides, J., Hernandez, L., Noll, D. C., Smith, E. E., & Chenevert, T. L. (2000). Behavioral and neuroimaging evidence of dissociable switching mechanisms in executive functioning. Poster presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco.

- Bell, A.J., Sejnowski, T.J. (1995) An Information-Maximization Approach to Blind Separation and Blind Deconvolution. Neural Computation, 7, 1129-1159.
- Bendriem B, Townsend D: The Theory and Practice of 3D PET, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Bloch, F (1946). Nuclear Induction. Physics Review, 70, 460-473.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. Journal of Neuroscience, 16, 4207-21.
- Breiter, H. C., Gollub, R. L., Weisskoff, R. M., Kennedy, D. N., Makris, N., Berke, J. D., Goodman, J. M., Kantor, H. L., Gastfriend, D. R., Riorden, J. P., Mathew, R. T., Rosen, B. R., & Hyman, S. E. (1997). Acute effects of cocaine on human brain activity and emotion. Neuron, 19, 591-611.
- Buckner, R. L., & Braver, T. S. (1999). Event-related functional MRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), Functional MRI. New York: Springer.
- Buckner, R. L., Koutstaal, W., Schacter, D. L., Dale, A. M., Rotte, M., & Rosen, B. R. (1998). Functional-anatomic study of episodic retrieval. II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. Neuroimage, 7, 163-75.
- Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. Neuroreport, 9, 3735-9.
- Buxton R. B., Frank L. R. (1997). A model for coupling between cerebral blood flow and oxygen metabolism

during neural stimulation. Journal of Cerebral Blood Flow and Metabolism, *17*, 64-72.

Buxton R. B., Wong E. C., Frank L. R. (1998). Dynamics of Blood Flow and Oxygenation Changes during Brain Activation: The Balloon Model. Magnetic Resonance in Medicine *39*, 855–864.

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., Castellanos, F. X., Haxby, J. V., Noll, D. C., Cohen, J. D., Forman, S. D., Dahl, R. E., & Rapoport, J. L. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. Journal of Cognitive Neuroscience, *9*, 835-847.

Cohen, J.D., Perlstein, W.M., Braver, T.S., Nystrom, L.E., Noll, D.C., Jonides, J., and Smith, E.E. (1997). Temporal dynamics of brain activation during a working memory task. Nature, *386*, 604-608.

Coull, J. T., Buchel, C., Friston, K. J., & Frith, C. D. (1999). Noradrenergically mediated plasticity in a human attentional neuronal network. Neuroimage, *10*, 705-15.

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (1999). Mapping the network for planning: a correlational PET activation study with the Tower of London task. Brain, *122*, 1973-87.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. Human Brain Mapping, *5*, 329-340.

D'Esposito, M., Zarahn, E., Aguirre, G. K., & Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. Neuroimage, *10*, 6-14.

Detre J. A., Leigh J. S., Williams D. S., Koretsky A. P. (1992). Perfusion Imaging. Magnetic Resonance In Medicine, *23*, 37-45.

- Edelstein W. A., Hutchinson J. M. S., Johnson G., Redpath T. (1980). Spin Warp NMR Imaging and Applications to Human Whole-body Imaging. Physics in Medicine and Biology, *25*, 751-756.
- Elster A.D. (1994). Questions and Answers in MRI, Mosby Yearbook, Inc. St. Louis, MO.
- Frahm J., Merboldt K. D., Hanicke W., Kleinschmidt A., Boecker H. (1994). Brain or Vein - Oxygenation or Flow? On Signal Physiology in Functional MRI of Human Brain Activation. NMR in Biomedicine, *7*, 45-53.
- Friston K. J., Holmes A. P., Price C. J., Buchel C., Worsley K. J. (1999). Multisubject fMRI Studies and Conjunction Analyses. Neuroimage *10*, 385-396.
- Frith, C. D., & Friston, K. (1997). Studying the brain with neuroimaging. In M. D. Rugg (Ed.), Cognitive Neuroscience, 169-195. Cambridge, MA: MIT Press.
- Gabrieli, J., Fleischman, D., Keane, M., Reminger, S., and Morell, F. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. Psychological Science, *6*, 76-82.
- Gambhir, S. S., Barrio, J. R., Phelps, M. E., Iyer, M., Namavari, M., Satyamurthy, N., Wu, L., Green, L. A., Bauer, E., MacLaren, D. C., Nguyen, K., Berk, A. J., Cherry, S. R., & Herschman, H. R. (1999). Imaging adenoviral-directed reporter gene expression in living animals with positron emission tomography. Proceedings of the National Academy of Sciences of the United States of America, *96*, 2333-8.
- Garavan, H., Ross, T. J., Li, S., & Stein, E. A. (2000). A parametric manipulation of central executive functioning. Cereb Cortex, *10*, 585-92.

- Gonzalez J. B., Alsop D. C., Detre J. A. (2000). Cerebral Perfusion and Arterial Transit Time Changes During Task Activation Determined with Continuous Arterial Spin Labeling. Magnetic Resonance in Medicine, 43, 739.
- Hahn E. L. (1950). Spin Echoes, Physics Review, 80, 580-594.
- Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., and 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.
- Hoge, R. D., Atkinson, J., Gill, B., Crelier, G. R., Marrett, S., & Pike, G. B. (1999). Investigation of BOLD signal dependence on cerebral blood flow and oxygen consumption: the deoxyhemoglobin dilution model. Magnetic Resonance in Medicine, 42, 849-63.
- Holmes A. P., Friston K. J. (1998). Generalisability, Random Effects and Population Inference. Neuroimage 7, S754.
- Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. Nature Neuroscience, 284-91.
- Joliot, M., Papathanassiou, D., Mellet, E., Quinton, O., Mazoyer, N., Courtheoux, P., & Mazoyer, B. (1999). FMRI and PET of self-paced finger movement: comparison of intersubject stereotaxic averaged data. Neuroimage, 10, 430-47.
- Jonides, J., Reuter-Lorenz, P., Smith, E.E., Awh, E., Barnes, L., Drain, M., Glass, J., Lauber, E., Patalano, A., Schumacher, E.H (1996). Verbal and spatial working memory, In D. Medin (Ed.), The Psychology of

Learning and Motivation, 43-88.

Jonides, J., Schumacher, E.H., Smith, E.E., Lauber, E.J., Awh, E., Minoshima, S., and Koeppe, R.A. Verbal-working-memory load affects regional brain activation as measured by PET (1997). Journal of Cognitive Neuroscience, 9, 462-475.

Jonides, J., Smith, E.E., Koeppe, R.A., Awh, E. Minoshima, S., & Mintun, M.A. (1993) Spatial working memory in humans as revealed by PET. Nature, 363, 623-625.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. Journal of Neuroscience, 17, 4302-4311.

Kim S. G., Quantification of Relative Cerebral Blood Flow Change by Flow Sensitive Alternating Inversion Recovery (FAIR) Technique (1995) Application to Functional Mapping. Magnetic Resonance in Medicine, 34 293-301.

Kinahan, P.E., & Noll, D.C. (1999). A direct comparison between whole-brain PET and BOLD fMRI measurements of single-subject activation response. Neuroimage, 9, 430-8.

Koepp, M. J. (1998). Evidence for striatal dopamine release during a video game. Nature, 393, 266-268.

Kumar A., Welti D., Ernst R.R. (1975). NMR Fourier Zeugmatography, Journal of Magnetic Resonance, 18, 69-83.

Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A., & Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness. Evidence of an interaction between emotion and attention in the anterior cingulate cortex. Journal of Cognitive Neuroscience, 10, 525-35.

Lashley, K.S. (1950). In search of the engram. In Symposium of the Society of Experimental Biology, 4, 454-482. Cambridge University Press, New York.

Lauterbur P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance, Journal of Magnetic Resonance, 81, 43-56.

Le Bihan D. (1995). Molecular Diffusion, tissue microdynamics and microstructure. NMR in Biomedicine, 8, 375-386 .

Macovski A. (1983). Medical Imaging Systems, Prentice-Hall, Inc., Englewood Cliffs, NJ.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences of the United States of America, 97, 4398-403.

Mansfield P., Pykett I.L. (1978). Biological and Medical Imaging by NMR. Journal of Magnetic Resonance, 29, 355-373.

Marshuetz, C., Salthouse, T. A., Ombao, H., Smith, E. E., Jonides, J., Diego, L. R., Bates, J. E., Chenevert, T. L., Krishnan, S., & Betley, A. T. (1999). A single subject statistical approach to modeling the componenets of verbal working memory. Paper presented at the Annual Meeting of the Society for Neuroscience, Miami Beach.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. Science, 270, 102-5.

- Maudsley A. A., Hilal S. K., Perman W. H., Simon H. E. (1983). Spatially resolved high resolution spectroscopy by four dimensional NMR. Journal of Magnetic Resonance, *51*, 147-152.
- McKeown, M.J., Jung, T.P., Makeig, S., Brown, G., Kindermann, S.S., Lee, T.W., Sejnowski, T.J. (1998a). Spatially independent activity patterns in functional MRI data during the stroop color-naming task. Proceedings of the National Academy of Sciences of the United States of America. *95*(3):803-10.
- McKewon, M.J., Makeig, S., Brown, G.B., Jung, T.P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J. (1998b). Analysis of fMRI Data by Blind Separation into Independent Spatial Components, Human Brain Mapping, *6*, 160-188.
- Milner, B., Corkin, S., Teuber, H. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow up study of H.M. Neuropsychologia, *6*, 215-234.
- Moseley M. E., Cohen Y., Mintorovitch J., Chileuitt L., Shimizu H., Kucharczyck J., Wendland M. F., Weinstein P. R. (1990). Early Detection of Regional Cerebral Ischemia in cats: comparison of Diffusion- and T2- weighted MRI and Spectroscopy. Magnetic Resonance in Medicine, *14*, 330-346.
- Nishimura D. G. (1996). Principles of Magnetic Resonance Imaging, Stanford University.
- Ogawa S., Lee T. M., Kay A. R., Tank D. W.(1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proceedings of the National Academy of Sciences of the United States of America. *87*, 9868-72.
- Peled, S., Gudbjartsson, H., Westin, C. F., Kikinis, R., & Jolesz, F. A. (1998). Magnetic resonance imaging shows orientation and asymmetry of white matter fiber tracts. Brain Resonance, *780*, 27-33.

- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single- word processing. Nature, *331*, 585-9.
- Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. Proceedings of the National Academy of Sciences of the United States of America, *95*, 853-60.
- Petersson K. M., Nichols T. E., Poline J. B., Holmes A. P. (1999). Statistical Limitations in functional Neuroimaging (I and II). Philosophical Transactions of the Royal Society of London - Series B, Biological Sciences, *354*, 1239-81.
- Poline J. B., Worsley K. J., Holmes A. P., Frackowiak R. S. J., Friston K. J. (1995). Estimating Smoothness in Statistical Parameter Maps: Variability of p-values. Journal of Computer Assisted Tomography, *19*, 788-796.
- Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. Science, *240*, 1627-31.
- Postle, B. R., & D'Esposito, M. (1999). "What"-Then-"Where" in visual working memory: an event-related fMRI study. Journal of Cognitive Neuroscience, *11*, 585-97.
- Purcell, Torrey H. C., Pound R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. Physics Review, *69*, 37-38.
- Ramsey, N. F., Kirkby, B. S., Van Gelderen, P., Berman, K. F., Duyn, J. H., Frank, J. A., Mattay, V. S., Van Horn, J. D., Esposito, G., Moonen, C. T., & Weinberger, D. R. (1996). Functional mapping of human sensorimotor cortex with 3D BOLD fMRI correlates highly with H₂(¹⁵O) PET rCBF. Journal of

Cerebral Blood Flow and Metabolism, 16, 755-64.

Raz, N. Aging of the brain in its impact on cognitive performance. In F. I. M. Craik & T. A. Salthouse (Eds.),

Handbook of aging and cognition - II. . Mahweh, NJ: Erlbaum.

Reiman, E. M., Fusselman, M. J., Fox, P. T., & Raichle, M. E. (1989). Neuroanatomical correlates of anticipatory anxiety [published erratum appears in Science 1992 Jun 19;256(5064):1696]. Science, 243, 1071-4.

Reuter-Lorenz, P.A., Jonides, J., Smith, E., Hartley, A., Miller, A., Marshuetz, C., & Koeppel, R. (2000) Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. Journal of Cognitive Neuroscience, 12, 174-187.

Roy, C. S., & Sherrington, C. S. (1890). On the regulation of the blood supply of the brain. Journal of Physiology, 11, 85-108.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M. P., Dold, G., & Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects [see comments]. Nature, 380, 526-8.

Sandler et al. (1995). Diagnostic Nuclear Medicine, Williams and Wilkins, Inc., Baltimore, MD. Chapter 12.

Schacter, D.L. and Buckner, R.L. (1998). Priming and the brain. Neuron, 20, 185-195.

Shulman, R. G., & Rothman, D. L. (1998). Interpreting functional imaging studies in terms of neurotransmitter cycling. Proceedings of the National Academy of Sciences of the United States of America, 95, 11993-8.

- Smith, E.E., & Jonides, J. (1995) Working memory in humans: Neuropsychological evidence. In M. Gazzaniga (Ed.) The Cognitive Neurosciences, 1009-1020 Cambridge, MA:MIT Press.
- Smith, E.E., Jonides, J., and Koeppel, R.A. (1996) Dissociating verbal and spatial working memory using PET. Cerebral Cortex, 6, 11-20.
- Strang G. (1998). Linear Algebra and its Applications. Third edition. Harcourt Brace Jovanovich, San Diego.
- Stegman, L. D., Rehemtulla, A., Beattie, B., Kievit, E., Lawrence, T. S., Blasberg, R. G., Tjuvajev, J. G., & Ross, B. D. (1999). Noninvasive quantitation of cytosine deaminase transgene expression in human tumor xenografts with in vivo magnetic resonance spectroscopy. Proceedings of the National Academy of Sciences of the United States of America, 96, 9821-6.
- Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. American Science, 57, 421-57.
- Turner, R., & Friston, K. Functional MRI, unpublished.
- Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In D.J. Engle, M.A. Goodale, and R.J. Mansfield (Eds.), Analysis of visual behavior, 549-586. Cambridge, MA: MIT Press.
- Vazquez, A. L., Noll, D. C. (1998). Non-linear Aspects of the Blood Oxygenation Response in Functional MRI, NeuroImage 8,108-118.
- Williams D. S., Detre J. A., Leigh J. S., Koretsky A. P. (1992). Magnetic Resonance Imaging of Perfusion using Spin Inversion of Arterial Water. Proceedings of the National Academy of Sciences of the United States of America, 89, 212-216.

- Wong E. C., Buxton R. B., Frank L. R.. (1997). Implementation of Quantitative Perfusion Imaging Techniques for Functional Brain Mapping using Pulsed Arterial Labeling. NMR in Biomedicine, 10, 237-249.
- Worsley K. J., Evans A. C., Marrett S., Neelin P. (1992). A Three-Dimensional Statistical Analysis for CBF Activation Studies in Human Brain. Journal of Cerebral Blood Flow and Metabolism 12, 900-918.
- Worsley K. J., Marret S., Neelin P., Vandal A. C., Friston K. J., Evans A. C. (1996). A Unified Statistical Approach for Determining Significant Signals in Images of Cerebral Activation. Human Brain Mapping, 4, 58-73.
- Zarahn, E., Aguirre, G., & D'Esposito, M. (1997). A trial-based experimental design for fMRI. Neuroimage, 6, 122-38.

Table 1

Summary of PET and fMRI Methods

What is imaged	PET	fMRI
Brain structure		Structural T1 and T2 scans
Regional brain activation	Blood flow (^{15}O) Glucose metabolism (^{18}F FDG) Oxygen consumption	BOLD (T2*) Arterial spin tagging (AST) FAIR
Anatomical connectivity		Diffusion tensor imaging
Receptor binding and regional chemical distribution	Benzodiazapines, dopamine, acetylcholine, many others Kinetic modeling	MR spectroscopy
Gene expression	Various radiolabeling compounds	MR spectroscopy with kinetic modeling

Table 2

Relative advantages of PET and fMRI

PET	fMRI
<ul style="list-style-type: none"> • Mapping of receptors and other neuroactive agents • Direct measurement of glucose metabolism • No magnetic susceptibility artifacts • Quiet environment for auditory tasks • Imaging near fluid spaces • Easily combined with ERP and other measurements because there is no magnetic field 	<ul style="list-style-type: none"> • Repeated scanning • Single subject analyses possible • Higher spatial resolution • Higher temporal resolution • Single trial designs • Estimation of hemodynamic response and separation of stimulus and task set related variables • Lower cost

Figure Captions

Figure 1. A graph showing the results of a search of the Medline database for articles with the words functional Magnetic Resonance Imaging (fMRI) in the title.

Figure 2. A spatial and a verbal task used to study item-recognition performance in working memory. Note that the two tasks are similar in structure except for the material that must be retained and retrieved.

Figure 3. Lateral and superior images revealing activations in spatial and verbal working memory tasks. In each row, three views of the brain in grey-scale renderings of a composite MRI have superimposed on them activations from a PET experiment, where the activations are shown in a color scale with blue the least active and red the most active.

Figure 4. A schematic of a task used to study processes required to switch between two tasks. The figure shows that the task entails two types of switches, between different internal counters or between different operations on the contents of those counters.

Figure 5. Behavioral data from the dual-switching task. Note that there are main effects of both types of switch, and that there is no interaction between these two separate effects.

Figure 6. One contrast in brain activations between the two types of switches in the dual-switching task. The top panel shows activation in a ventromedial prefrontal site and the bottom panel shows activation in a lateral prefrontal site for each type of switch. Note the double dissociation in patterns of activations in these two sites for the two types of switch.

Figure 7. Measured rCBF responses in three areas across six conditions, one rest condition and five levels of increasing difficulty in the Tower

of London task. These areas showed linear increases in rCBF with increasing difficulty, whereas other areas (such as visual cortex) showed a response to task versus rest but no changes among the five difficulty levels. Reproduced from Dagher et al. (1999).

Figure 8. A schematic diagram of the main components of a PET scanner.

Figure 9. The PET Scanner. Each detector counts the number of annihilation events that take place in a "column" of tissue. The column can be subdivided into smaller units that represent the image pixels. The detector counts the sum of the events in each of the elements in the column.

Figure 10. PET image reconstruction. The raw data are a set of projections (sums) at different angles as shown in A. "Backprojecting" the raw data onto the image means adding the numbers of counts in the projection to the pixels that are aligned with each point in the projection, as shown in B. An image can be obtained after the data from all the projections has been added as shown in C.

Figure 11. Representation of the proton's magnetization

Figure 12. The Spin Ensemble

Figure 13. Tipping the magnetization vector from the z-axis onto the xy-plane. The duration and strength of the B1 field determine how far the vector is tipped (ie - the "flip angle")

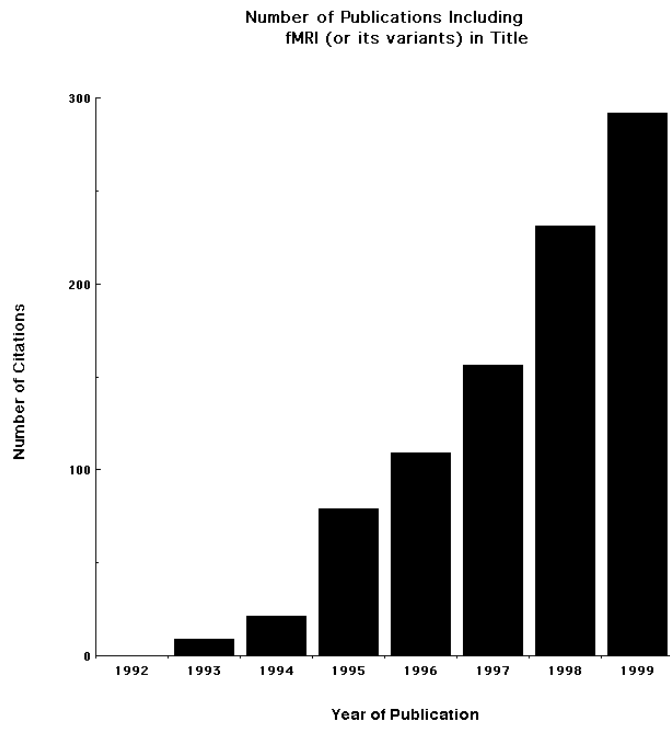
Figure 14. The dephasing process occurs because all the spins in the ensemble do not precess at the exact same rate. Some of them get ahead, and some of them lag behind. The net effect is that they start canceling each other out, shortening the length of the magnetization vector.

Figure 15. The same slice of brain tissue can appear very different, depending on which relaxation mechanism is emphasized as the source of the contrast in the pulse sequence. Using long echo times emphasizes T2 differences between tissues, and shortening the repetition time emphasizes T1 differences in tissue. Left: one slice of a T1 image. Right: the same slice acquired as a T2 image.

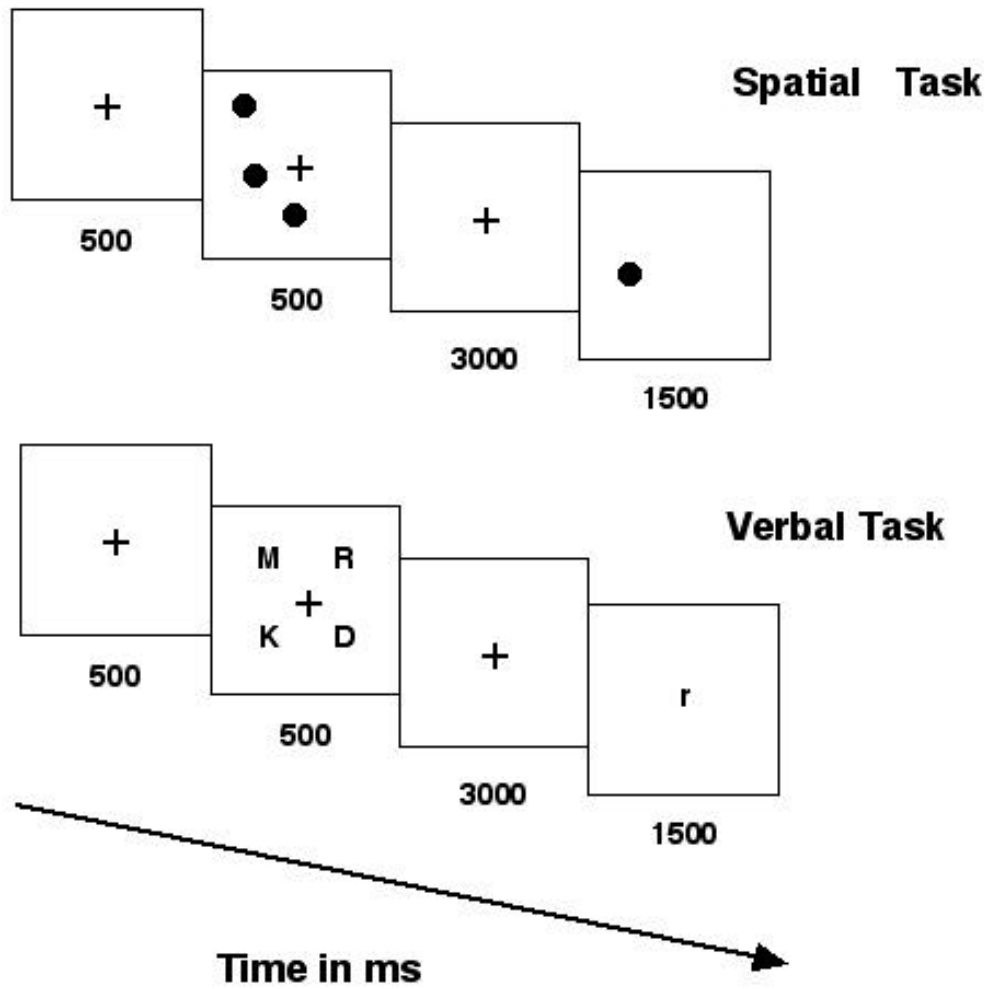
Figure 16. Refocusing of the spins by Gradient Echoes and Spin Echoes.

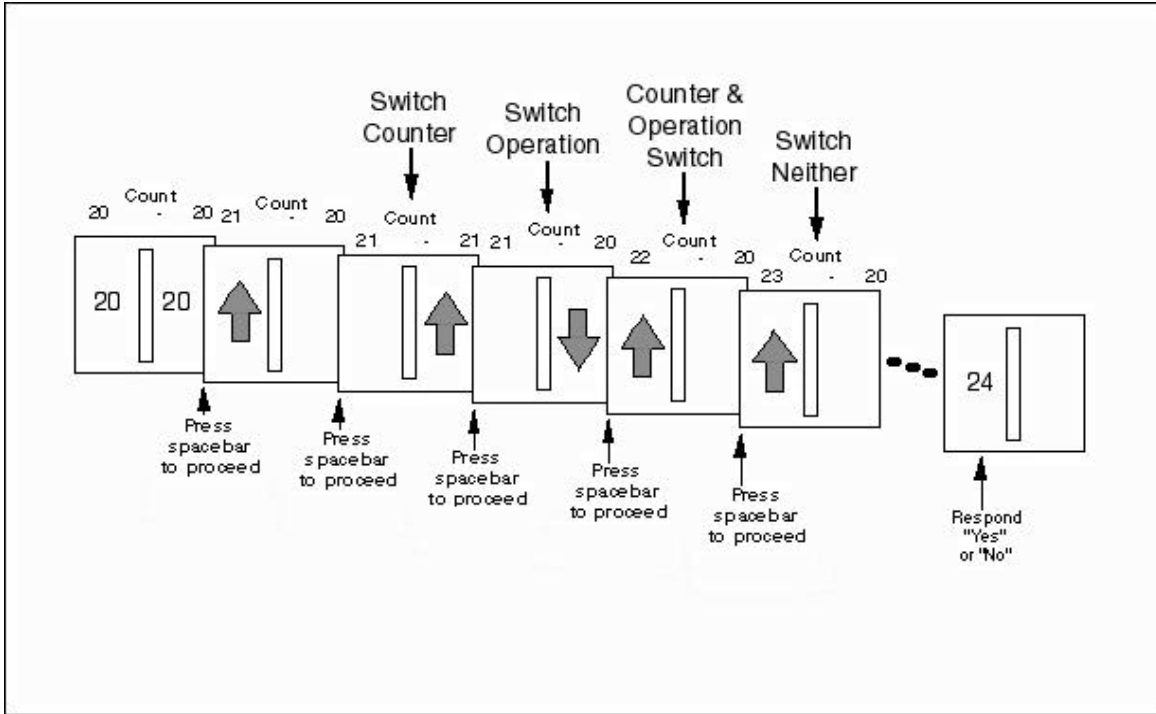
Figure 17. The BOLD response to a single event is shown in the top portion of the figure. This is commonly referred to as the hemodynamic response function, or HRF. A train of events, like the one shown in the middle figure, would produce a BOLD response like the one shown in the bottom part of the figure.

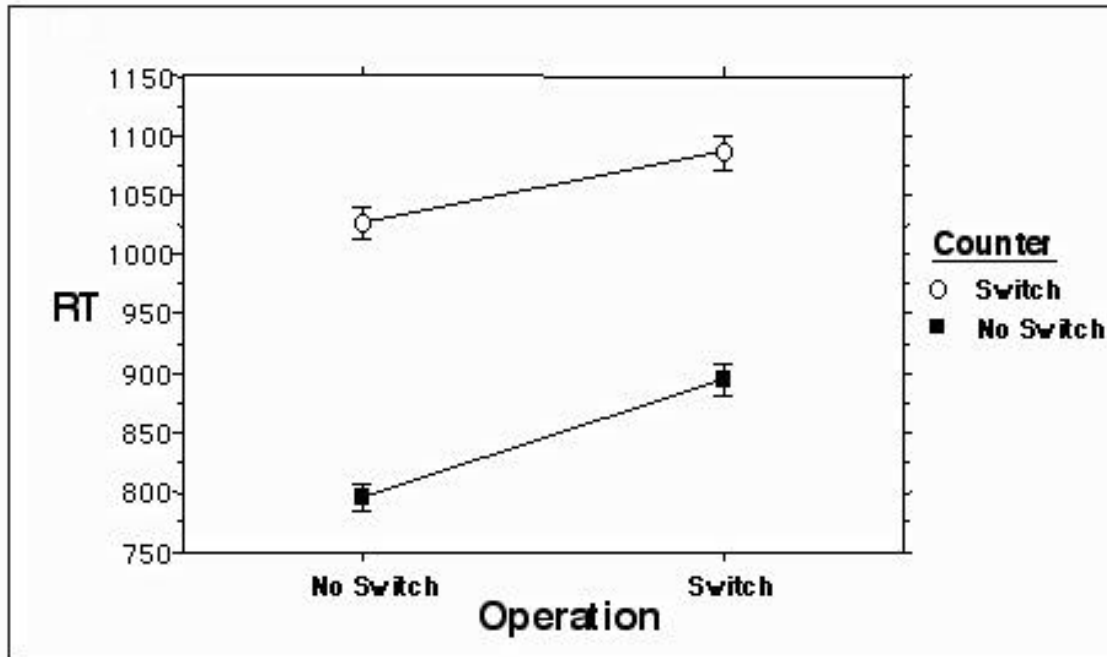
Figure 18. The Design Matrix should include all the significant effects that are present in the experiment. Each effect is represented by a column of data containing the expected time series that one would see if that were the only effect present.

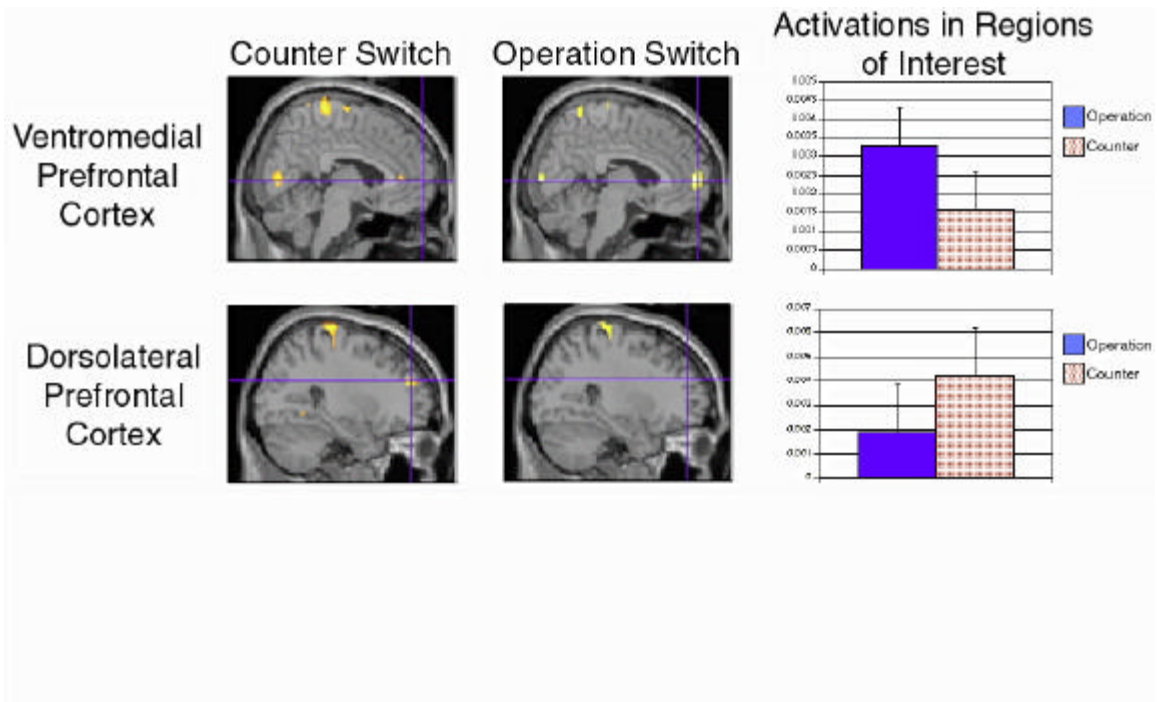


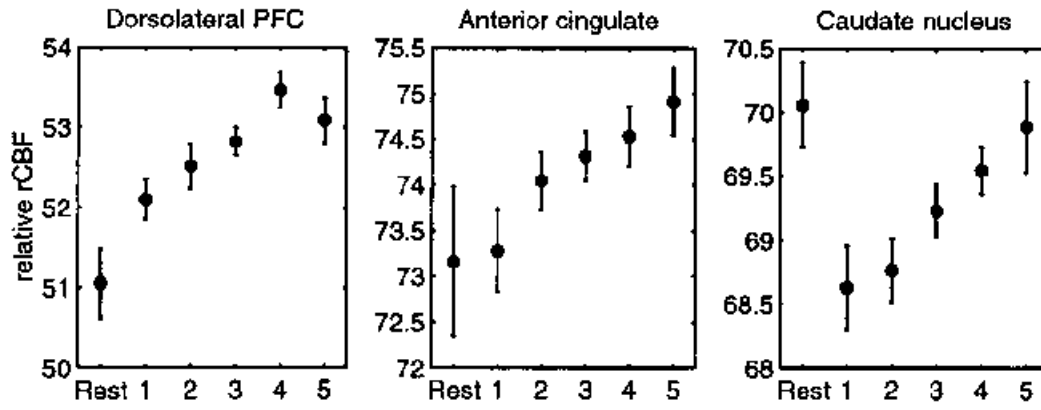
Spatial and Verbal Item-Recognition

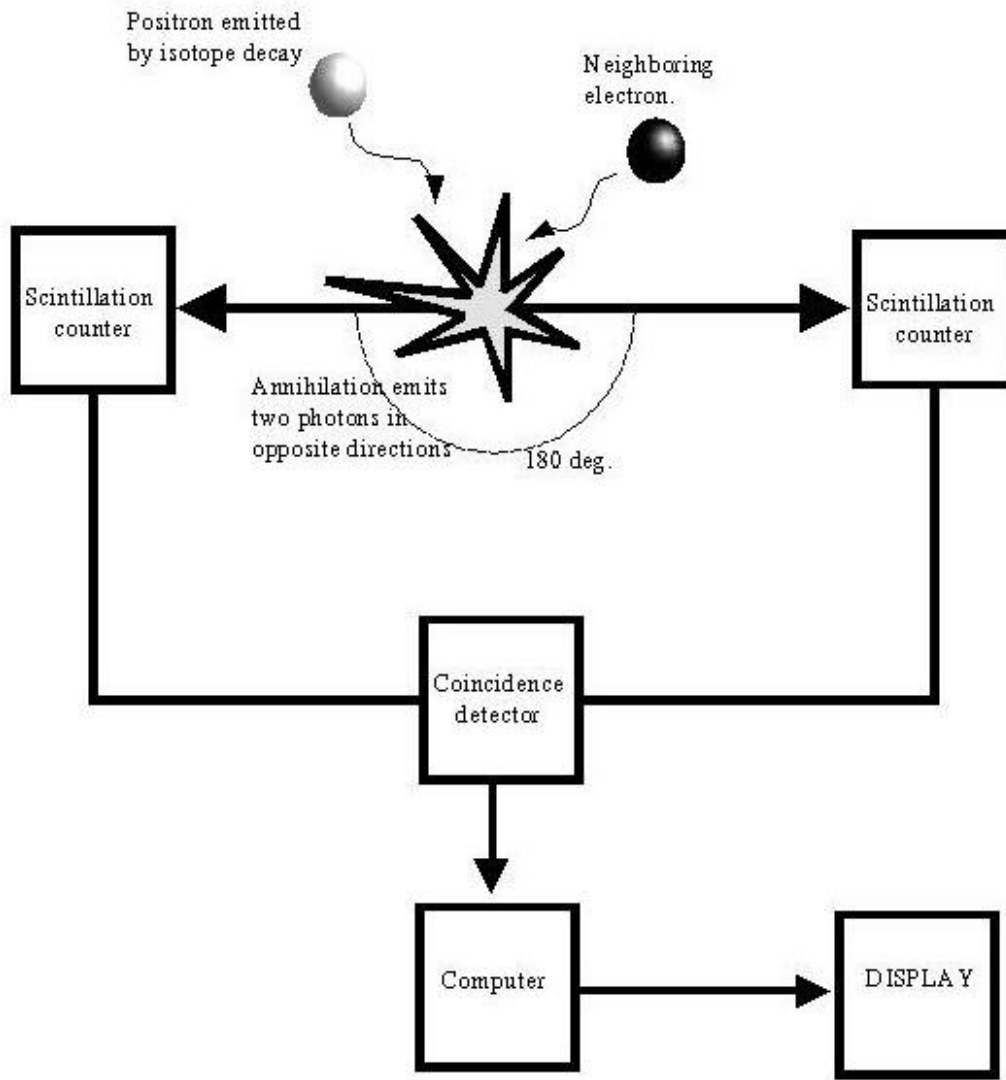


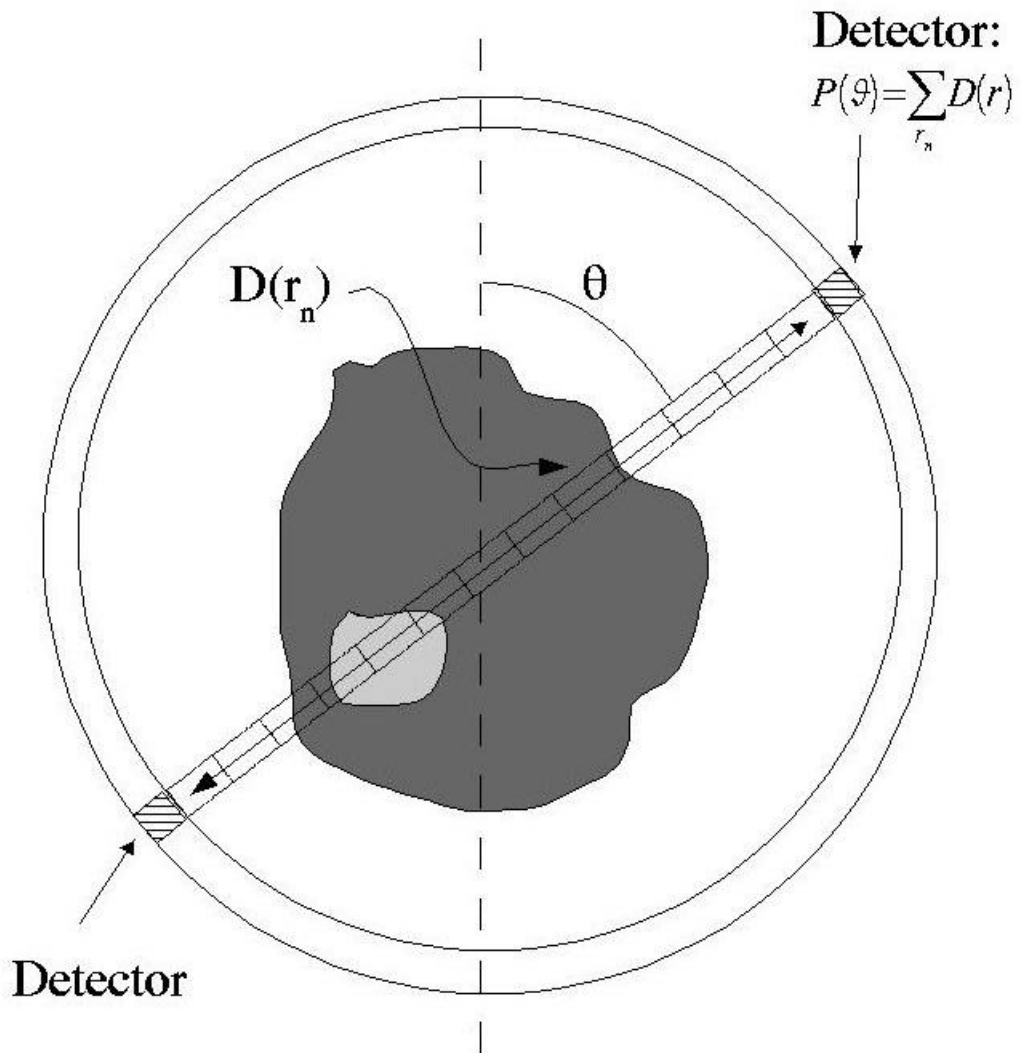


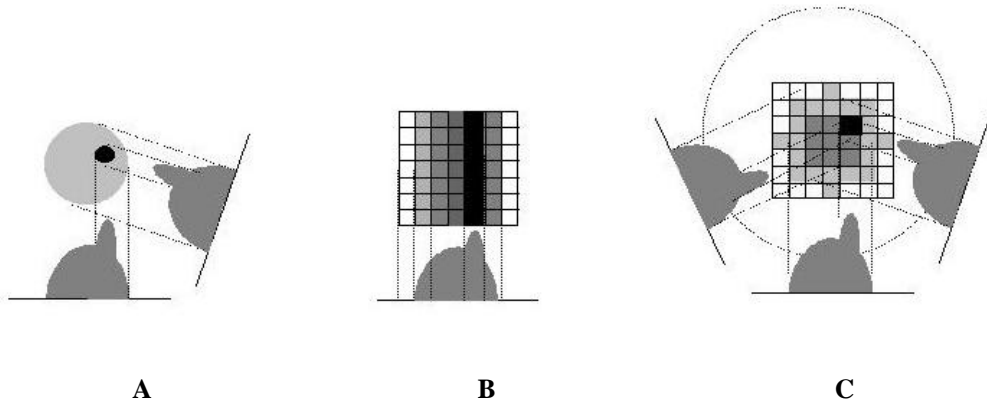


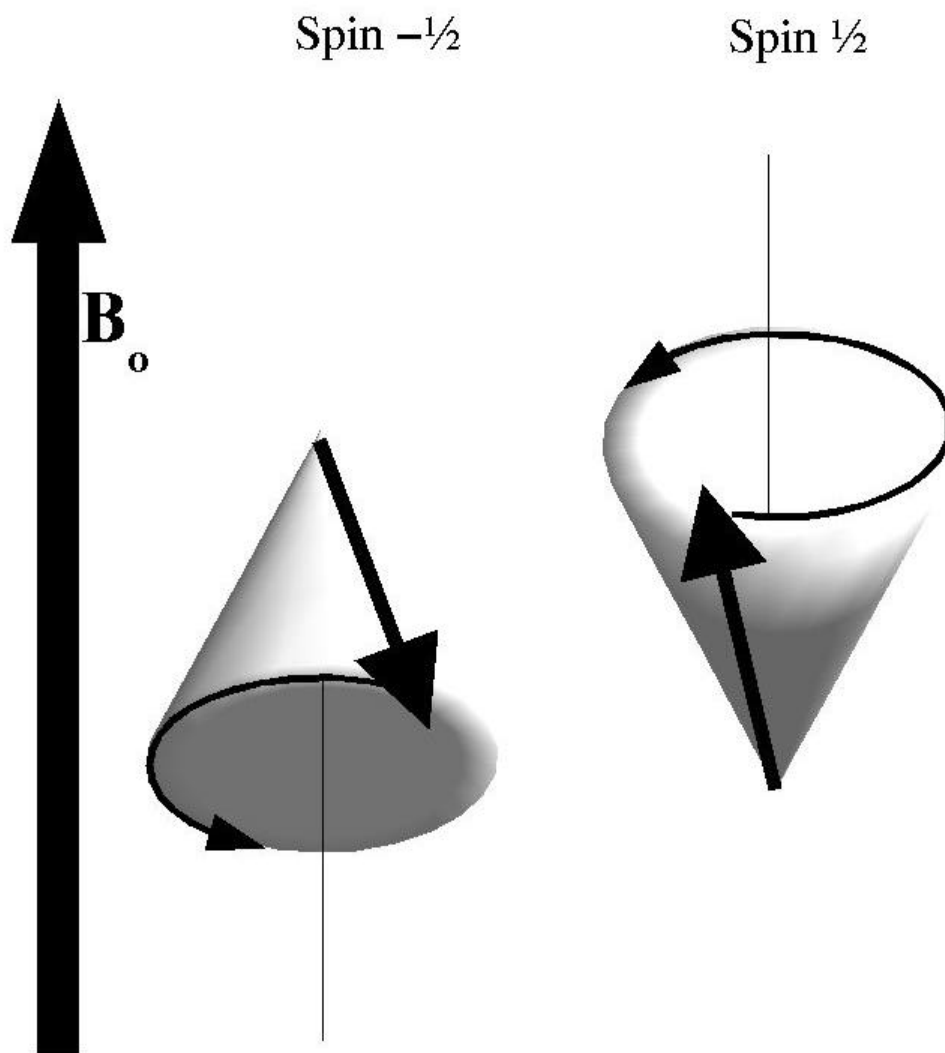




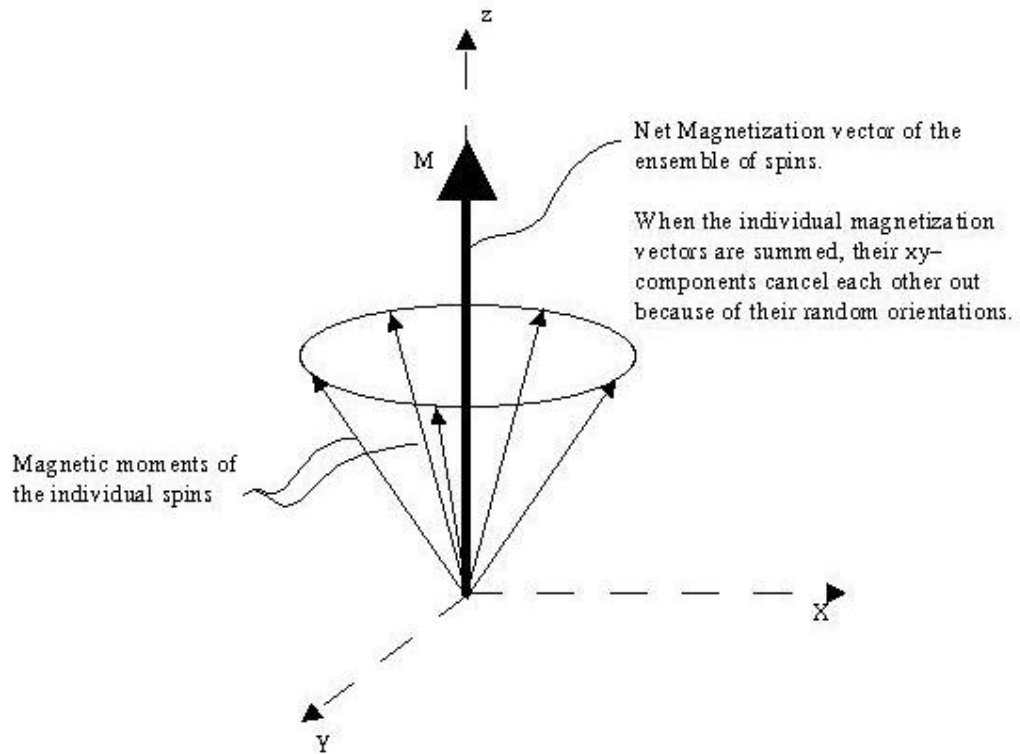


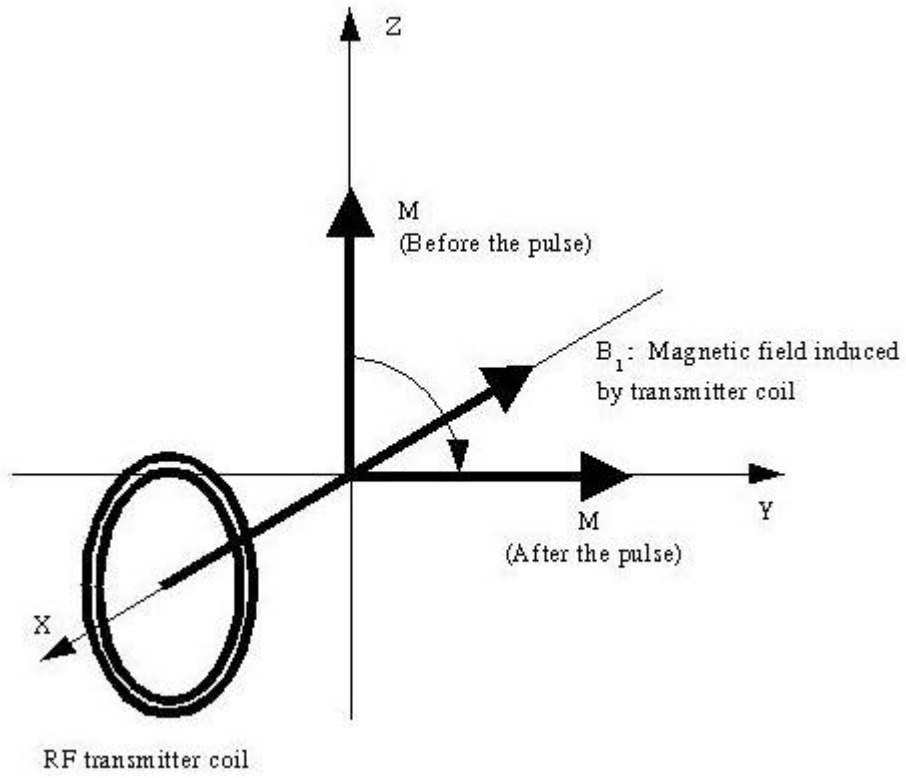


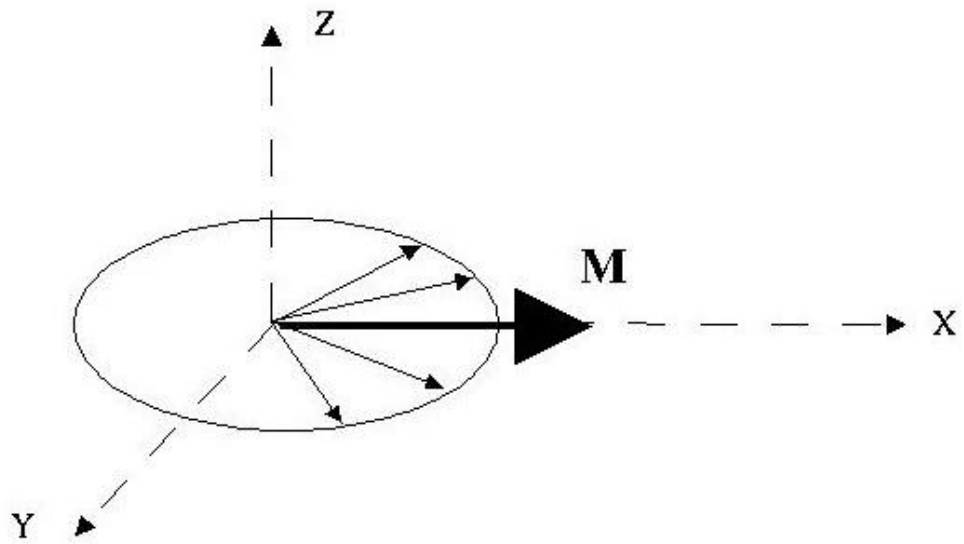
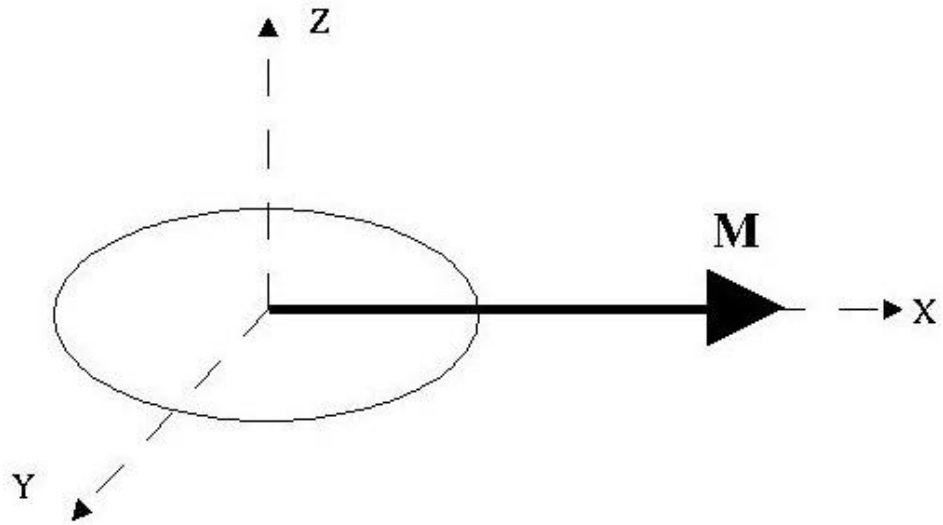


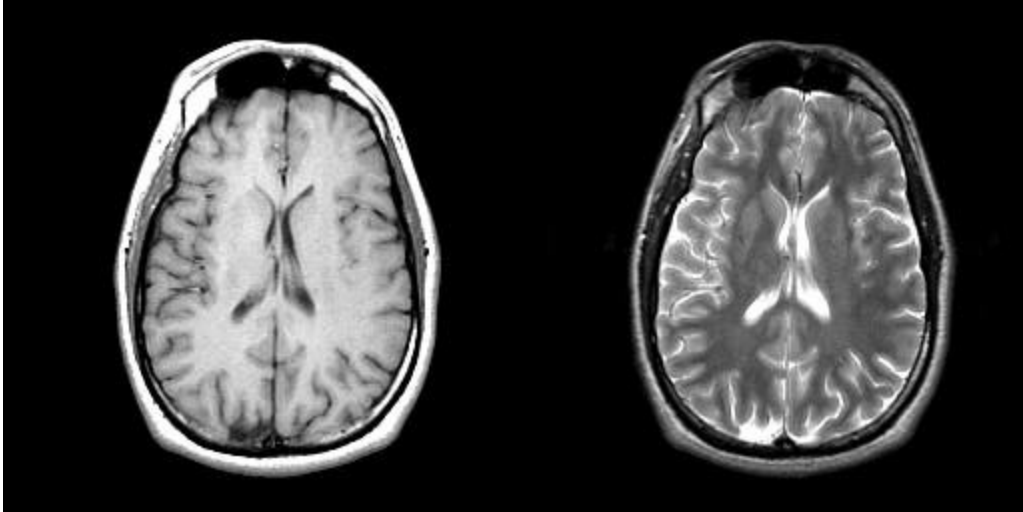


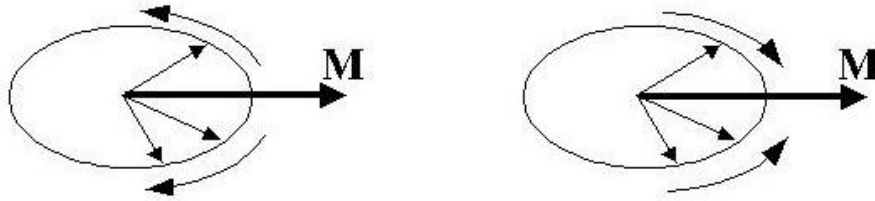
The magnetic moment of the spins precesses around the axis of the main magnetic field (B_0). Its orientation is determined by their quantum state.



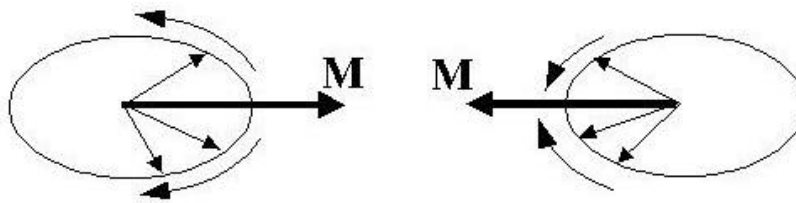








Gradient Echo technique: A gradient in the magnetic field causes the spins to lose phase coherence. Reversal of the gradient causes them to regain it.



Spin Echo technique: spins are dephased by the gradient. After application of a 180 degree pulse, all the spins are rotated about the y-axis, and the application of the same gradient causes the spins to regain coherence along the negative x-axis.

