Introduction

Cognitive neuroscience is a discipline that attempts to determine the neural mechanisms underlying cognitive processes. Specifically, cognitive neuroscientists test hypotheses about brain–behavior relationships organized along two conceptual domains: **functional specialization**—the idea that functional modules exist within the brain, that is, areas of the cerebral cortex that are specialized for a specific cognitive process, and **functional integration**—the idea that a cognitive process can be an emergent property of interactions among a network of brain regions that suggests that a brain region can play a different role across many functions.

Early studies of brain–behavior relationships consisted of careful observation of individuals with neurological injury resulting in focal brain damage. The idea of functional specialization evolved from hypotheses that damage to a particular brain region was responsible for a given behavioral syndrome that was characterized by a precise neurological examination; for instance, the association of nonfluent aphasia with right-sided limb weakness implicated the left hemisphere as the site of language abilities. Moreover, upon the death of a patient with a neurological disorder, clinicopathological correlations provided confirmatory information about the site of damage causing a specific neurobehavioral syndrome such as aphasia; for example, in 1861, Paul Broca’s observations of nonfluent aphasia in the setting of a damaged left inferior frontal gyrus cemented the belief that this brain region was critical for speech output. The introduction of structural brain imaging more than 100 years after Broca’s observations, first with computerized tomography and later with magnetic resonance imaging (MRI), paved the way for more precise anatomical localization in the living patient of the cognitive deficits that develop after brain injury. The superb spatial resolution of structural neuroimaging has reduced the reliance on the infrequently obtained autopsy for making brain–behavior correlations.

Functional neuroimaging, broadly defined as techniques that measure brain activity, has expanded our ability to study the neural
basis of cognitive processes. One such method, functional MRI (fMRI) has emerged as an extremely powerful technique that affords excellent spatial and temporal resolution. Measuring regional brain activity in healthy subjects while they perform behavioral tasks links localized brain activity with specific behaviors; for example, functional neuroimaging studies have demonstrated that the left inferior frontal gyrus is consistently activated during the performance of speech-production tasks in healthy individuals. Such findings from functional neuroimaging are complementary to findings derived from observations of patients with focal brain damage. This chapter focuses on the principles underlying fMRI as a cognitive neuroscience tool for exploring brain–behavior relationships.

**Inference in Functional Neuroimaging Studies of Cognitive Processes**

Insight regarding the link between brain and behavior can be gained through a variety of approaches. It is unlikely that any single neuroscience method is sufficient to investigate fully any particular question regarding the mechanism underlying cognitive function. From a methodological point of view, every method will offer different temporal and spatial resolution. From a conceptual point of view, every method will provide data that will support different types of inferences that can be drawn from it. Thus, data obtained addressing a single question but derived from multiple methods can provide more comprehensive and inferentially sound conclusions.

Functional neuroimaging studies support inferences about the association of a particular brain system with a cognitive process. However, it is difficult to prove in such a study that the observed activity is necessary for an isolated cognitive process because perfect control over a subject’s cognitive processes during a functional neuroimaging experiment is never possible. Even if the task a subject performs is well designed, it is difficult to demonstrate conclusively that he/she is differentially engaging a single identified cognitive process. The subject may engage in unwanted cognitive processes that either have no overt measurable effects, or are perfectly confounded with the process of interest. Consequently, the neural activity measured by the functional neuroimaging technique may result from some confounding neural computation that is itself not necessary for executing the cognitive process seemingly under study. In other words, functional neuroimaging is an observational, correlative method. It is important to note that the inferences that can be drawn from functional neuroimaging studies such as fMRI apply to all methods of physiological measurement [e.g., electroencephalogram (EEG) or magnetoencephalogram.]

The inference of necessity cannot be made without showing that inactivating a brain region disrupts the cognitive process in question. However, unlike precise surgical or neurotoxic lesions in animal models, lesions in patients are often extensive, damaging local neurons and fibers of passage; for example, damage to prominent white matter
tracts can cause cognitive deficits similar to those produced by cortical lesions, such as the amnesia resulting from lesions of the fornix, the main white matter pathway projecting from the hippocampus. In addition, connections from region A may support the continued metabolic function of region B, but region A may not be computationally involved in certain processes undertaken by region B. Thus, damage to region A could impair the function of region B via two possible mechanisms: (1) diaschisis and (2) retrograde transsynaptic degeneration. Consequently, studies of patients with focal lesions cannot conclusively demonstrate that the neurons within a specific region are themselves critical to the computational support of an impaired cognitive process.

Empirical studies using lesion and electrophysiologic methods demonstrate these issues regarding the types of inferences that can logically be drawn from them. In monkeys, single-unit recording reveals neurons in the lateral prefrontal cortex that increase their firing during the delay between the presentation of information to be remembered and a few seconds later when that information must be recalled. These studies are taken as evidence that persistent neural activity in the prefrontal cortex is involved in temporary storage of information, a cognitive process known as working memory. The necessity of prefrontal cortex for working memory was demonstrated in other monkey studies showing that prefrontal lesions impair performance on working memory tasks, but not on tasks that do not require temporarily holding information in memory. Persistent neural activity during working memory tasks are also found in the hippocampus. Hippocampal lesions, however, do not impair performance on most working memory tasks, which suggests that the hippocampus is involved in maintaining information over short periods of time, but is not necessary for this cognitive operation. Observations in humans support this notion. For example, the well-studied patient H.M., with complete bilateral hippocampal damage and the severe inability to learn new information, could nevertheless perform normally on working memory tasks such as digit span. The hippocampus is implicated in long-term memory, especially when relations between multiple items or multiple features of a complex novel item must be retained. Thus, the hippocampus may only be engaged during working memory tasks that requires someone to subsequently remember novel information.

When the results from lesion and functional neuroimaging studies are combined, a stronger level of inference emerges. As in the examples of Broca’s aphasia or working memory, a lesion of a specific brain region causes impairment of a given cognitive process, and when engaged by an intact individual, that cognitive process evokes neural activity in the same brain region. In this type of finding, the inference that this brain region is computationally necessary for the cognitive process is stronger than data derived from each study performed in isolation. Thus, lesion and functional neuroimaging studies are complementary, each providing inferential support that the other lacks.
Other types of inferential failure can occur in the interpretation of functional neuroimaging studies when other common assumptions do not hold true. First, it is assumed that if a cognitive process activates a particular brain region (evoked by a particular task), the neural activity in that brain region must depend on engaging that particular cognitive process; for example, a brain region showing greater activation during the presentation of faces than to other types of stimuli, such as photographs of cars or buildings, is considered to engage face perception processes. However, this region also may support other higher-level cognitive processes such as memory processes, in addition to lower-level perceptual processes. Second, it is assumed that if a particular brain region is activated during the performance of a cognitive task, the subject must have engaged the cognitive process supported by that region during the task; for example, observing activation of the frontal lobes during a mental rotation task, it was proposed that subjects engaged working memory processes to recall the identity of the rotated target. (They derived this assumption from other imaging studies showing activation of the frontal lobes during working memory tasks.) However, in this example, because some other cognitive process supported by the frontal lobes could have activated this region, one cannot be sure that working memory was engaged leading to the activation of the frontal lobes.

In summary, interpretation of the results of functional neuroimaging studies attempting to link brain and behavior rests on numerous assumptions. Familiarity with the types of inferences that can and cannot be drawn from these studies should be helpful for assessing the validity of the findings reported by such studies.

Functional MRI as a Cognitive Neuroscience Tool

Functional MRI has become the predominant functional neuroimaging method for studying the neural basis of cognitive processes in humans. Compared to its predecessor, positron emission tomography (PET) scanning, fMRI offers many advantages; for example, MRI scanners are much more widely available, and imaging costs are less expensive because MRI does not require a cyclotron to produce radioisotopes. Magnetic resonance imaging is also a noninvasive procedure because there is no requirement for injection of a radioisotope into the bloodstream. In addition, given the half-life of available radioisotopes, PET scanning is unable to provide comparable temporal resolution to that of fMRI, which can provide images of behavioral events occurring on the order of seconds rather than the summation of many behavioral events over tens of seconds.

In selected circumstances, however, PET can provide an advantage over fMRI for studying certain questions concerning the neural basis of cognition; for example, at present, fMRI does not adequately image the regions within the orbitofrontal cortex and the anterior or inferior temporal lobe because of the susceptibility artifact near the interface of the brain and sinuses. These artifacts worsen at higher magnetic fields (i.e., 3 or 4 Tesla), and such scanners are becoming commonly available.
and increasingly utilized by cognitive neuroscientists. Improvements in pulse sequences for acquiring fMRI data and development of algorithms for distortion correction of images should eventually eliminate or reduce these artifacts.\textsuperscript{5-12} Currently, however, such sequences and methods are not widely available and implemented. Position emission tomography scanning may remain desirable or necessary when studying certain populations of individuals; for example, amnesic patients resulting from cerebral anoxia often have implanted cardiac pacemakers precluding them from having an MRI scan due to the magnetic field. However, PET scanning is unacceptable for studies of children due to the radiation exposure.

The MRI scanner, compared to a behavioral testing room, is less than ideal for performing most cognitive neuroscience experiments. Experiments are performed in the awkward position of lying on one’s back, often requiring subjects to visualize the presentation of stimuli through a mirror in an acoustically noisy environment. Moreover, most individuals develop some degree of claustrophobia due to the small bore of the MRI scanner and find it difficult to remain completely motionless for a long duration of time that is required for most experiments (e.g., usually 60 to 90 minutes). These constraints of the MRI scanner make it especially difficult to scan children, which has resulted in many fewer fMRI studies involving children than adults.\textsuperscript{20,21} In addition, it has been a technical challenge to develop equipment within the MRI environment that successfully presents different types of stimuli to the individual (e.g., olfactory, tactile), as well as to collect ancillary response or physiological data necessary for a particular experiment.

All sensory systems have been investigated with fMRI, including the visual, auditory, somatosensory, olfactory, and gustatory systems. Each system requires different technologies for successful presentation of relevant stimuli within an MRI environment. At the time of this writing, very few off-the-shelf commercial products exist that are MRI compatible, and most in use today have been engineered locally by individual laboratories. Most published fMRI studies have utilized visual stimuli, although great strides have been made to allow the presentation of other types of stimuli. Details regarding the issues related to presenting visual and auditory stimuli in the MRI environment can be reviewed furthered in a comprehensive chapter on the topic by Savoy and colleagues.\textsuperscript{22} In brief, the most common means of presenting visual stimuli is via a LCD projector system, with the sophistication of the system depending on the quality of image resolution required for the experiment. Several options exist for auditory stimuli, such as piezoelectric or electrostatic headphones. However, the biggest challenge is the acoustically noisy scanner environment. The pulsing of the fMRI gradient coils is the source of such noise, making the study of auditory processes challenging,\textsuperscript{23,24} for example, during echoplanar imaging within a 4 Tesla magnet using a high-performance head gradient set, sound levels can reach 130 decibels. As a reference point, Food and Drug Administration (FDA) safety regulations require no greater than an average of 105 decibels for one hour. With placement
of absorbing materials within the scanner and on the walls of the room, as well as a fiberglass bore liner surrounding the gradient set, we have been able to reduce sound levels by about 25 decibels. For further discussion concerning sound reduction techniques, refer to the chapter on Auditing MRI. One of the biggest technical challenges within an MRI scanner has been the ability to present olfactory stimuli. However, sophisticated MR-compatible olfactometers have been designed and utilized successfully. Such methods use a nasal-mask in which the change from odorant to no-odorant conditions occurs within a few milliseconds.\textsuperscript{25,26}

Acquiring ancillary electrophysiological data such as electromyographic recordings to measure muscle contraction or electrodermal responses to measure autonomic activity enhances many cognitive neuroscience experiments. Devices have been developed that are MR compatible for these types of measurements, as well other physiological measures such as heart rate, electrocardiography, oxygen saturation, and respiratory rate. The recording of eye movements is becoming commonplace in MRI scanners, predominantly with the use of an infrared video camera equipped with long-range optics.\textsuperscript{27,28} Video images of the pupil–corneal reflection can be sampled at 60/120/240 hertz, allowing for the accurate (less than one degree) localization of gaze within 50 horizontal and 40 vertical degrees of visual angle. Although most behavioral tasks used in cognitive neuroscience experiments rely on collecting manual responses, the ability to reliably collect verbal responses without significant artifact being introduced into the data has been demonstrated by several laboratories.\textsuperscript{29–31}

Electroencephalogram recordings also have been performed successfully during MRI scanning.\textsuperscript{32,33} However, the recording of event-related potentials, a signal that is much smaller in amplitude than the signal in EEG, can be more difficult in a magnetic field due to artifacts induced by gradient pulsing and head movement from cardiac pulsation. New monitoring devices and algorithms to remove artifact are being developed, allowing for reliable measurements of event-related potentials during MRI scanning.\textsuperscript{34,35}

In summary, most initial challenges facing performing cognitive and behavioral experiments within the MRI environment have been overcome, creating an environment that is comparable to standard psychophysical testing labs outside of a scanner. Although individual laboratories have achieved most of these advancements, MRI scanners originally designed for clinical use by manufacturers are now being designed with consideration of many of these research-related issues.

Temporal Resolution
Two types of temporal resolution need to be considered for cognitive neuroscience experiments. First, what is the briefest neural event that can be detected as an fMRI signal? Second, how close together can two neural events occur and be resolved as separable fMRI signals?

The time scale on which neural changes occur are quite rapid; for example, neural activity in the lateral intraparietal area of monkeys increases within 100 milliseconds of the visual presentation of a saccade
In contrast, the fMRI signal gradually increases to its peak magnitude within four to six seconds after an experimentally induced brief (less than one second) change in neural activity, and then decays back to baseline after several more seconds.37–39 This slow time course of fMRI signal change in response to such a brief increase in neural activity is informally referred to as the BOLD fMRI hemodynamic response, or simply, the hemodynamic response (see Figure 18.1). Thus, neural dynamics and neurally evoked hemodynamics, as measured with fMRI, are on quite different time scales.

The sluggishness of the hemodynamic response limits the temporal resolution of the fMRI signal to hundreds of milliseconds to seconds as opposed to the millisecond temporal resolution of electrophysiological recordings of neural activity, such as from single-unit recording in monkeys and EEG or magnetoencephalogram in humans. However, it has been clearly demonstrated that brief changes in neural activity can be detected with reasonable statistical power using fMRI; for example, appreciable fMRI signal can be observed in sensorimotor cortex in association with single finger movements40 and in visual cortex during very briefly presented (34 milliseconds) visual stimuli.41 In contrast, the temporal resolution of fMRI limits the detection of sequential changes in neural activity that occurs rapidly with respect to the hemodynamic response. That is, the ability to resolve the changes in the fMRI signal associated with two neural events, often requires the separation of those events by a relatively long period of time compared with the width of the hemodynamic response. This is because two neural events closely spaced in time will produce a hemodynamic response that reflects the accumulation from both neural events, making it difficult to estimate the contribution of each individual neural event. In general, evoked fMRI responses to discrete neural events separated by at least four seconds appear to be within the range of resolution.42 However, provided that the stimuli are presented randomly, studies have shown significant differential functional responses between two events (e.g., flashing visual stimuli) spaced as closely as 500 milliseconds apart.43–45 The effect at fixed and randomized intertrial intervals on the BOLD signal is illustrated in Figure 18.2.

![Figure 18.1. A typical hemodynamic response (i.e., fMRI signal change in response to a brief increase of neural activity) from the primary sensorimotor cortex. The fMRI signal peaked approximately five seconds after the onset of the motor response (at time zero).](image-url)
In some tasks, the order of individual trial events cannot be randomized; for example, in certain types of working memory tasks, the presentation of the information to be remembered during the delay period, and the period when the subject must recall the information, are individual trial events whose order cannot be randomized. In these types of tasks, short time scales (less than four seconds) cannot be temporally resolved. These temporal resolution issues in fMRI have been extensively considered regarding their impact on experimental design.46,47

Spatial Resolution

It has yet to be determined how precisely the measured BOLD fMRI signal, which arises from the vasculature, reflects adjacent neural activity. Thus, the ultimate spatial resolution of BOLD fMRI is unknown. Functional MRI studies in both monkey and man at high field (4 to 4.7 Tesla) have demonstrated that BOLD signal can be obtained with high spatial resolution—approximately 0.75 × 0.75 mm² in-plane resolution.48,49 In monkeys, with novel approaches such as using a small, tissue-compatible, intraostely implanted radiofrequency coil, ultra high spatial resolution of 125 × 125 μm² has been obtained.50 Using this method, Logothetis and colleagues50 demonstrated cortical lamina-specific activation in a task that compared responses to moving stimuli with those elicited by flickering stimuli. This contrast elicited BOLD signal mostly in the granular layers of the striate cortex of the monkey, which are known to have a high concentration of directionally selec-
tive cells. Advances in such methods would allow for imaging of hundreds of neurons per voxel as opposed to hundreds of thousands of neurons per voxel, which is more typical for a human cognitive neuroscience fMRI experiment.

Virtually all fMRI studies model the large BOLD signal increase, which is due to a local low-deoxyhemoglobin state (see Figure 18.1), in order to detect changes correlating with a behavioral task. However, optical imaging studies have demonstrated that preceding this large positive response is an initial negative response reflecting a localized increase in oxygen consumption, causing a high-deoxyhemoglobin state. This early hemodynamic response is called the initial dip and is thought to be more tightly coupled to the actual site of neural activity evoking the BOLD signal as compared to the later positive portion of the BOLD response; for example, Kim and colleagues, scanning cats in a high field scanner, demonstrated that the early negative BOLD response (e.g., initial dip) produced activation maps that were consistent with orientation columns within visual cortex. This finding is quite remarkable given that the average spacing between two adjacent orientation columns in cortex is approximately one millimeter. In contrast, the activation maps produced by the delayed positive BOLD response appeared more diffuse, and cortical columnar organization could not be identified. Thus, empirical evidence suggests that deriving activation maps by correlating behavioral responses with the initial dip may markedly improve spatial resolution. However, it is important to note that observation of the initial dip of the BOLD signal has been inconsistently observed in humans across laboratories for reasons that are still unclear. Several groups, however, were able to detect columnar architecture (in this case, ocular dominance columns) by modeling the positive BOLD response in humans scanning at 4 Tesla. These investigators attributed their success to optimized radiofrequency coils, limiting head motion, optimizing slice orientation, and the enhanced signal-to-noise ratio (SNR) provided by a high magnetic field.

Another unique method for improving spatial resolution has been called functional magnetic resonance–adaptation (fMR-A), which could provide a means for identifying and assessing the functional attributes of sharply defined neuronal populations within a given region of the brain. Even if the spatial resolution of fMRI evolves to the point of being able to resolve a population of a few hundred neurons within a voxel, it is still likely that this small population will contain neurons with very different functional properties that will be averaged together. The adaptation method is based on several basic principles. First, repeated presentation of the same type of stimuli (i.e., a picture of the one object) causes neurons to adapt to those stimuli (i.e., neuronal firing is reduced). Second, if these neurons are then exposed to a different type of stimulus (i.e., a picture of another object) or a change in some property of the stimulus (i.e., the same object in a different orientation), recovery from adaptation can be assessed (i.e., whether or not the BOLD signal returns to its original state). If the signal remains adapted, it implies that the neurons are invariant to the attribute that was changed. If the signal recovers from the adapted
state, it would imply that the neurons are sensitive to that attribute; for example, Grill-Spector and colleagues demonstrated that an area of lateral occipital cortex thought to be important for object recognition was less sensitive to changes in object size and position as compared to changes in illumination and viewpoint. Thus, with this method, it is possible to investigate the functional properties of neuronal populations with a level of spatial resolution that is beyond that obtained from conventional fMRI data analysis methods.

Considering all the neuroscientific methods available today for studying human brain–behavior relationships, fMRI provides an excellent balance of temporal and spatial resolution (see Figure 18.3). Improvements on both fronts will clearly add to the increasing popularity of this method.

**Issues in Functional MRI Experimental Design**

Numerous options exist for designing experiments using fMRI (see Chapter 3 by Aguirre for more in-depth discussion of experimental design). The prototypical fMRI experimental design consists of two behavioral tasks presented in blocks of trials alternating over the course of a scanning session, and the fMRI signal between the two tasks is compared. This is known as a blocked design; for example, a given block might present a series of faces to be viewed passively, which evokes a particular cognitive process, such as face perception. The experimental block alternates with a control block, which is designed to evoke all of the cognitive processes present in the experimental block except for the cognitive process of interest. In this experiment, the control block may comprise a series of objects. In this way, the stimuli used in experimental and control tasks have similar visual attributes, but differ in the attribute of interest (i.e., faces). The inferential frame-
work of cognitive subtraction\textsuperscript{56} attributes differences in neural activity between the two tasks to the specific cognitive process (i.e., face perception). Cognitive subtraction was originally conceived by Donders\textsuperscript{57} in the late 1800s for studying the chronometric substrates of cognitive processes (see Sternberg\textsuperscript{58}) and was a major innovation in imaging.\textsuperscript{56,59}

The assumptions required for cognitive subtraction may not always hold and could produce erroneous interpretation of functional neuroimaging data.\textsuperscript{42} Cognitive subtraction relies on two assumptions: pure insertion and linearity. Pure insertion implies that a cognitive process can be added to a preexisting set of cognitive processes without affecting them. This assumption is difficult to prove because one needs an independent measure of the preexisting processes in the absence and presence of the new process.\textsuperscript{58} If pure insertion fails as an assumption, a difference in the neuroimaging signal between the two tasks might be observed, not because a specific cognitive process was engaged in one task and not the other, but because the added cognitive process and the preexisting cognitive processes interact.

An example of this point is illustrated in working memory studies using delayed-response tasks.\textsuperscript{60} These tasks (for an example, see Jonides and colleagues\textsuperscript{61}) typically present information that the subject must remember (engaging an encoding process), followed by a delay period during which the subject must hold the information in memory over a short period of time (engaging a memory process), followed by a probe that requires the subject to make a decision based on the stored information (engaging a retrieval process). The brain regions engaged by evoking the memory process theoretically are revealed by subtracting the BOLD signal measured by fMRI during a block of trials that the subject performs that do not have a delay period (only engaging the encoding and retrieval processes) from a block of trials with a delay period (engaging the encoding, memory, and retrieval processes). In this example, if the addition or insertion of a delay period between the encoding and retrieval processes affects these other behavioral processes in the task, the result is failure to meet the assumptions of cognitive subtraction. That is, these non-memory processes may differ in delay trials and no-delay trials, resulting in a failure to cancel each other out in the two types of trials that are being compared.

Empirical evidence of such failure exists.\textsuperscript{62} For example, Figure 18.4 demonstrates BOLD signal derived from the prefrontal cortex from a subject performing a delayed response task similar to the tasks described above. The left side of the figure illustrates BOLD signal consistent with delay period activity, whereas the right side of the figure illustrates BOLD signal from another region of prefrontal cortex that did not display sustained activity during the delay, yet showed greater activity in the delay trials as compared to the trials without a delay. In any blocked functional neuroimaging study that compares delay versus no-delay trials with subtraction, such a region would be detected and likely assumed to be a memory region. Thus, this result provides empirical grounds for adopting a healthy doubt regarding the inferences drawn from imaging studies that rely exclusively on cognitive subtraction.
In functional neuroimaging, the transform between the neural signal and the hemodynamic response (measured by fMRI) must also be linear for the cognitive subtractive method to yield valid results. In other words, it is assumed that the BOLD signal being measured is approximately proportional to the local neural activity that evokes it. Surprisingly, although thousands of empirical studies using fMRI to study brain–behavior relationships have been published, only a handful exist that have explored the neurophysiological basis of the BOLD signal (for reviews, see Attwell and Iadecola63 and Heeger and Ress64). In several studies, linearity did not strictly hold for the BOLD fMRI system, but the linear transform model was reasonably consistent with the data; for example, Boynton and colleagues tested whether BOLD signal in response to long duration stimuli can be predicted by summing the responses to shorter duration stimuli.39 Using pulses of flickering checkerboard patterns and measuring within human primary visual cortex, these investigators found that the BOLD signal response to various durations of stimulus presentation (6, 12, or 24 seconds) could be predicted from the responses they obtained from shorter stimulus presentations; for example, the BOLD signal response to a six-second pulse could be predicted from the summation of the BOLD signal response to the three-second pulse with a copy of the same response delayed by three seconds. However, temporal summation did not always hold, and there are clearly nonlinear effects in the

Figure 18.4. Data derived from the performance of a normal subject on a spatial delayed-response task. This task comprised both delay trials (circles), as well as trials without a delay period (no-delay trials; diamonds). (A) Trial-averaged fMRI signal from prefrontal cortex that displayed delay-correlated activity. The gray bar along the x-axis denotes the 12-second delay period during delay trials. The delay trials display a level of fMRI signal greater than baseline throughout the period of time corresponding to the retention delay (taking into account the delay and dispersion of the fMRI signal). The peaks seen in the signal correspond to the encoding and retrieval periods. (B) Trial-averaged fMRI signal from a region in prefrontal cortex that did not display the characteristics of delay-correlated activity. This region displays a significant functional change associated with the no-delay trials, and a significant functional change associated with the encoding and retrieval periods of the delay trials, but not one associated with the retention delay of delay trials. Adapted from Zarahn E, Aguirre GK, D’Esposito M. Temporal isolation of the neural correlates of spatial mnemonic processing with fMRI. Cogn Brain Res. 1999;7:255–268.
transform of neural activity to a hemodynamic response that must be considered. If these nonlinearities lead to saturation of the BOLD effect at a certain stimulus intensity, erroneous interpretation of particular results of fMRI experiments may occur.

Another class of experimental designs, called event-related fMRI, attempt to detect changes associated with individual trials, as opposed to the larger unit of time comprising a block of trials. Each individual trial may be composed of one behavioral event, such as the presentation of a single stimulus (e.g., a face or object to be perceived), or several behavioral events, such as in the delayed-response task described above, (e.g., an item to be remembered, a delay period, and a motor response in a delayed-response task); for example, with an event-related design, activity within the prefrontal cortex has consistently been shown to correlate with the delay period, supporting the role of the PFC in temporarily maintaining information. This finding is consistent with single-neuron recording studies in the PFC of monkeys. Event-related designs offer numerous advantages; for example, it allows for stimulus or trial randomization, avoiding the behavioral confounds of blocked trials. It also permits the separate analysis of functional responses, which are identified only in retrospect (i.e., trials on which the subject made a correct or incorrect response). Of course, an experiment does not have to be limited to either a block or event-related designs—a mixed-type (both event-related and blocked) design, where particular trial types are randomized within a block, is perfectly feasible. In this type of design, both item-related processes (e.g., transient responses to stimuli), as well as state-related processes (processes sustained throughout a block of trials or a task) are perfectly feasible. Overall, much flexibility exists in the type of experimental design that can be utilized in an fMRI experiment, and continued innovation in this area will greatly expand the types of neuroscientific questions that can be addressed.

Issues in Interpretation of fMRI Data

Statistics

Many statistical techniques are used for analyzing fMRI data, but no single method has emerged as the ideal or gold standard (see Chapter 3 by Aguirre for more in-depth discussion of statistical analysis of fMRI data). The analysis of any fMRI experiment designed to contradict the null hypothesis (i.e., there is no difference between experimental conditions) requires inferential statistics. If the difference between two experimental conditions is too large to reasonably be due to chance, then the null hypothesis is rejected in favor of the alternative hypothesis, which typically is the experimenter’s hypothesis (e.g., the fusiform gyrus is activated to a greater extent by viewing faces than objects). Unfortunately, because errors can occur in any statistical test, experimenters will never know when an error is committed, and they can only try to minimize them. Knowledge of several basic statistical issues provides a solid foundation for the correct interpretation of the data derived from functional neuroimaging studies.
Two types of statistical errors can occur. A Type I error is committed when the null hypothesis is falsely rejected when it is true, that is, a difference between experimental conditions is found but a difference does not truly exist. This type of error is also called a false-positive error. In a functional neuroimaging study, a false-positive error would be finding a brain region activated during a cognitive task, when actually it is not. A Type II error is committed when the null hypothesis is accepted when it is false, that is, no difference between experimental conditions exists when a difference does exist. This type of error is called a false-negative error. A false-negative error in a functional neuroimaging study would be failing to find a brain region activated during the performance of a cognitive task when actually it is.

In fMRI experiments, like all experiments, a tolerable probability for Type I error, typically less than five percent, is chosen for adequate control of specificity, that is, control of false-positive rates. Two features of imaging data can cause unacceptable false-positive rates, even with traditional parametric statistical tests. First, there is the problem of multiple comparisons. For the typical resolution of images acquired during fMRI scans, the full extent of single slice (matrix-128, slice-5 mm) of the human brain could comprise 15000 voxels. Thus, with any given statistical comparison of two experimental conditions, there are actually 15000 statistical comparisons being performed. With such a large number of statistical tests, the probability of finding a false-positive activation, that is, committing a Type I error, somewhere in the brain increases. Several methods exist to deal with this problem. One method, a Bonferroni correction, assumes that each statistical test is independent and calculates the probability of Type I error by dividing the chosen probability ($p = 0.05$) by the number of statistical tests performed. Another method is based on Gaussian field theory, and calculates the probability of Type I error when imaging data are spatially smoothed. Many other methods for determining thresholds of statistical maps are proposed and utilized, but unfortunately, no single method has been universally accepted. Nevertheless, all fMRI studies must apply some type of correction for multiple comparisons to control the false-positive rate.

The second feature that might increase the false-positive rate is the noise in fMRI data. Data from BOLD fMRI are temporally autocorrelated, with more noise at some frequencies than at others. The shape of this noise distribution is characterized by a $1/\text{frequency}$ function, with increasing noise at lower frequencies. Traditional parametric and nonparametric statistical tests assume that the noise is not temporally autocorrelated, that is, each observation is independent. Therefore, any statistical test used in fMRI studies must account for the noise structure of fMRI data. If not, the false-positive rates will inflate.

Type II error is rarely considered in functional neuroimaging studies. When a brain map from an fMRI experiment is presented, several areas of activation are typically attributed to some experimental manipulation. The focus of most imaging studies is on brain activation, whereas it is often implicitly assumed that all of the other areas (typically, most
of the brain) were not activated during the experiment. Power as a statistical concept refers to the probability of correctly rejecting the null hypothesis.\(^7\) As the power of a fMRI study to detect changes in brain activity increases, the false-negative rate decreases. Unfortunately, power calculations for particular fMRI experiments are rarely performed, although this methodology is evolving.\(^7\)–\(^8\) Reports that specific brain areas were not active during an experimental manipulation should provide an estimate of the power required for detection of a change in the region. All experiments should be designed to maximize power. Relatively simple strategies can increase power in an fMRI experiment in certain circumstances, such as increasing the amount of imaging data collected or increasing the number of subjects studied. It is also important to note that task designs can affect sensitivity;\(^8\) for example, because BOLD fMRI data are temporally autocorrelated, experiments with fundamental frequencies in the lower range (e.g., a boxcar design with 60-second epochs) will have reduced sensitivity due to the presence of greater noise at these lower frequencies. Finally, in a study that simultaneously measured neural signal via intracortical recording and BOLD signal in a monkey, it was observed that the SNR of the neural signal was, on average, at least one order of magnitude higher than that of the BOLD signal. The investigators of this study concluded that “the statistical and thresholding methods applied to the hemodynamic responses probably underestimate a great deal of actual neural activity related to a stimulus or task.”\(^8\) Thus, the magnitude of Type II error in BOLD fMRI may currently be underestimated and warrants further consideration in the interpretation of almost any cognitive neuroscience experiment.

**Altered Hemodynamic Response**

When comparing changes in BOLD signal levels within the brain of an individual subject across different cognitive tasks and making conclusions regarding changes in neural activity and the pattern of activity, numerous assumptions are made regarding the steps comprising neurovascular coupling (stimulus $\rightarrow$ neural activity $\rightarrow$ hemodynamic response $\rightarrow$ BOLD signal) and the regional variability of the metabolic and vascular parameters influencing the BOLD signal. It should be obvious that fMRI studies of cognition and behavior of individuals with local vascular compromise or diffuse vascular disease (e.g., patients with strokes or normal elderly) are potentially problematic; for example, many fMRI studies have sought to identify age-related changes in the neural substrates of cognitive processes. These studies that directly compare changes in BOLD signal intensity across age groups rely upon the assumption of age-equivalent coupling of neural activity to BOLD signal. However, there is empirical evidence that suggests that this general assumption may not hold true. Extensive research on the aging neurovascular system has revealed that it undergoes significant changes in multiple domains in a continuum throughout the human lifespan, probably as early as the fourth decade (for a review, see Farkas and Luiten\(^8\)). These changes affect the vascular ultrastructure,\(^8\) the resting cerebral blood flow (CBF),\(^8\)–\(^6\) the vascular
responsiveness of the vessels, and the cerebral metabolic rate of oxygen consumption. Aging is also frequently associated with comorbidities such as diabetes, hypertension, and hyperlipidemia, all of which may affect the BOLD signal by affecting CBF and neurovascular coupling. Any one of these age-related differences in the vascular system could conceivably produce age-related differences in BOLD fMRI signal responsiveness, greatly affecting the interpretation of results from such studies.

Our laboratory compared the hemodynamic response function (HRF) characteristics in the sensorimotor cortex of young and older subjects in response to a simple motor reaction-time task. The provisional assumption was made that there was identical neural activity between the two populations based on physiological findings of equivalent movement-related electrical potentials in subjects under similar conditions. Thus, it was presumed that any changes that were observed in BOLD fMRI signal between young and older individuals in motor cortex would be due to vascular, and not neural, activity changes in normal aging. Several important similarities and differences were observed between age groups. Although there was no significant difference in the shape of the hemodynamic response curve or peak amplitude of the signal, a significantly decreased SNR in the BOLD signal was found in older individuals as compared to young individuals. This was attributed to a greater level of noise in the older individuals. A decrease in the spatial extent of the BOLD signal was also observed in older individuals compared to younger individuals in sensorimotor cortex (i.e., the median number of suprathreshold voxels). Similar results have been replicated by two other laboratories. These findings suggest that there is some property of the coupling between neural activity and BOLD signal that changes with age.

The notion that vascular differences among individuals may affect BOLD signal is especially a concern when considering studies of patient populations with known vascular changes such as stroke. A recent fMRI study addressed the issue of the influence of vascular factors on the BOLD signal in a symptomatic stroke population. They analyzed the time course of the BOLD HRF in the sensorimotor cortex of patients with an isolated subcortical lacunar stroke compared to a group of age-matched controls. They found a decrease in the rate of rise and the maximal BOLD HRF to a finger- or hand-tapping task in both the sensorimotor cortex of the hemisphere affected by the stroke and the unaffected hemisphere. These investigators proposed that, given the widespread changes of these BOLD signal differences, the change was unlikely a direct consequence of the subcortical lacunar stroke, but rather a manifestation of preexisting diffuse vascular pathology.

In summary, comparing BOLD signal in two different groups of individuals that may differ in their vascular system should be done with caution; for example, in one scenario, a comparison of activation of young and elderly individuals during a cognitive task may show less activation by elderly (as compared to young subjects) in some brain...
regions, but greater activation in other regions (e.g., see Rypma and colleagues). In this scenario, it is unlikely that regional variations in the hemodynamic coupling of neural activity to imaging signal would account for such age-related differences in patterns of activation. In another scenario, a comparison of young and elderly subjects may show less activation by elderly (as compared to young subjects) in some brain regions, but no evidence of greater activation in any other region. In this case, it is possible that the observed age-related differences are not due to differences in intensity of neural activity, but rather to other non-neuronal contributions to the imaging signal, that is, neurovascular coupling. Several statistical approaches towards the imaging data are being developed that will attempt to address these potential confounds.

Types of Hypotheses Tested Using fMRI

Functional neuroimaging experiments test hypotheses regarding the anatomical specificity for cognitive processes (functional specialization), basic mechanisms of cognition (cognitive theory), and direct or indirect interactions among brain regions (functional integration). The experimental design and statistical analyses chosen will determine the types of questions that can be addressed. Ultimately, the most powerful approach for the testing of theories on brain–behavior relationships is the analysis of converging data from multiple methods.

Functional Specialization

The major focus of fMRI studies of cognition is testing theories on functional specialization. The concept of functional specialization is based on the premise that functional modules exist within the brain, that is, areas of the cerebral cortex are specialized for a specific cognitive process; for example, facial recognition is a critical primary function likely served by a functional module. Prosopagnosia is the selective inability to recognize faces. Patients with prosopagnosia, however, can recognize familiar faces, such as those of relatives, by other means, such as the voice, dress, or body shape. Other types of visual recognition, such as identifying common objects, are normal. Prosopagnosia arises from lesions of the inferomedial temporo-occipital lobe, which usually are due to a stroke within the posterior cerebral artery circulation. No lesion studies have precisely localized the area crucial for facial perception. However, they provide strong evidence that a brain area is specialized for processing faces. Functional imaging studies have provided anatomical specificity for such a module; for example, Kanwisher and colleagues used fMRI to test a group of healthy individuals and found that the fusiform gyrus was significantly more active when the subjects viewed faces than when they viewed assorted common objects. The specificity of a fusiform face area was further demonstrated by the finding that this area also responded significantly more strongly to passive viewing of faces than to scrambled two-tone faces, front-view photographs of houses, and
photographs of human hands. These elegant experiments allowed the investigators to reject alternative functions of the face area, such as visual attention, subordinate-level classification, or general processing of any animate or human forms, demonstrating that this region selectively perceives faces.

Cognitive Theory

An exciting new direction for studies using functional neuroimaging are those that test theories of the underlying mechanisms of cognition; for example, an fMRI study attempted to answer the question, “To what extent does perception depend on attention?” One hypothesis is that unattended stimuli in the environment receive very little processing, but another hypothesis is that the processing load in a relevant task determines the extent to which irrelevant stimuli are processed. These alternative hypotheses were tested by asking normal individuals to perform linguistic tasks of low or high load while ignoring irrelevant visual motion in the periphery of a display. Visual motion was used as the distracting stimulus because it activates a distinct region of the brain (cortical area MT or V5, another functional module in the visual system). Activation of area MT would indicate that irrelevant visual motion was processed. Although task and irrelevant stimuli were unrelated, fMRI of motion-related activity in MT showed a reduction in motion processing during the high-processing load condition in the linguistic task. These findings support the hypothesis that perception of irrelevant environmental information depends on the information processing load that is currently relevant and being attended to. Thus, by the finding that perception depends on attention, this fMRI experiment provides insight regarding underlying cognitive mechanism.

Functional Integration

Functional neuroimaging experiments can also test hypotheses about interactions between brain regions by focusing on covariances of activation levels between regions. These covariances reflect functional connectivity, a concept that was originally developed in reference to temporal interactions among individual neurons. Newer approaches, often using a statistical test called structural equation modeling, attempt to determine whether covariances among brain regions result from direct or indirect interactions, a concept called effective connectivity. Using this method, McIntosh and colleagues found shifting prefrontal and limbic interactions in a working memory task for faces as the retention delay increased (see Figure 18.5). The different interactions between brain regions at short and long delays were interpreted as a functional change; for example, strong corticolimbic interactions were found at short delays, but at longer delays, when the image of the face was more difficult to maintain, strong fronto-cingulate-occipital interactions were found. The investigators postulated that the former finding was due to maintaining an iconic facial representation, and that the latter finding was due to an expanded encoding strategy, resulting...
Integration of Multiple Methods

The most powerful approach toward understanding brain–behavior relationships comes from analyzing converging data from multiple methods. There are several ways in which different methods can provide complementary data; for example, one method can provide superior spatial resolution (e.g., fMRI), whereas the other can provide superior temporal resolution (e.g., event-related potentials). In addition, the data from one method may allow for different conclusions to be drawn from it, such as whether a particular brain region is necessary to implement a cognitive process (i.e., lesion methods) or whether it is only involved during its implementation (i.e., physiological methods). The following sections describe examples of such approaches.

Combined fMRI/Lesion Studies

The combined use of functional neuroimaging and lesions studies can be illustrated with studies of the neural basis of semantic memory, the cognitive system that represents our knowledge of the world. Early studies of patients with focal lesions supported the notion that the temporal lobes mediate the retrieval of semantic knowledge\textsuperscript{107}; for example, patients with temporal lobe lesions may show a disproportionate impairment in the knowledge of living things (e.g., animals) compared with nonliving things. Other patients have a disproportionate deficit in knowledge of nonliving things.\textsuperscript{108} These

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{network_analysis.png}
\caption{Network analysis of fMRI data during performance of a working memory task across three different delay periods.\textsuperscript{105} Areas of correlated increases in activation (solid lines) and areas of correlated decreases in activation (dotted lines) are shown. Note the different pattern of interactions among brain regions at short and long delays. Adapted from McIntosh AR, Grady CL, Haxby JV, Ungerleider LG, Horwitz B. Changes in limbic and prefrontal functional interactions in working memory task for faces. \textit{Cereb Cortex.} 1996;6:571–584.}
\end{figure}
observations led to the notion that the semantic memory system is subdivided into different sensorimotor modalities, that is, living things, compared with nonliving things, are represented by their visual and other sensory attributes (e.g., a banana is yellow), whereas nonliving things are represented by their function (e.g., a hammer is a tool but comes in many different visual forms). The small number of patients with these deficits, and often large lesions, limits precise anatomical–behavioral relationships. However, functional neuroimaging studies in normal subjects can provide spatial resolution that the lesion method lacks.109

These original observations regarding the neural basis of semantic memory conflicted with functional neuroimaging studies consistently showing activation of the left inferior frontal gyrus (IFG) during the retrieval of semantic knowledge; for example, an early cognitive activation PET study revealed IFG activation during a verb-generation task compared with a simple word-repetition task.59 A subsequent fMRI study110 offered a fundamentally different interpretation of the apparent conflict between lesion and functional neuroimaging studies of semantic knowledge: left IFG activity is associated with the need to select some relevant feature of semantic knowledge from competing alternatives, not retrieval of semantic knowledge per se. This interpretation was supported by an fMRI experiment in normal individuals in which selection, but not retrieval, demands were varied across three semantic tasks. In a verb-generation task, in a high-selection condition, subjects generated verbs to nouns with many appropriate associated responses without any clearly dominant response (e.g., wheel), but in a low-selection condition, nouns with few associated responses or with a clear dominant response (e.g., scissors) were used. In this way, all tasks required semantic retrieval, and differed only in the amount of selection required. The fMRI signal within the left IFG increased as the selection demands increased (see Figure 18.6). When the degree of semantic processing varied independently of selection demands, there was no difference in left IFG activity, suggesting that selection, not retrieval, of semantic knowledge drives activity in the left IFG.

To determine if left IFG activity was correlated with, but not necessary for, selecting information from semantic memory, the same task used during the fMRI study was used to examine the ability of patients with focal frontal lesions to generate verbs.111 Supporting the earlier claim regarding left IFG function derived from an fMRI study,110 the overlap of the lesions in patients with deficits on this task corresponded to the site of maximum fMRI activation in healthy young subjects during the verb-generation task (see Figure 18.6). In this example, the approach of using converging evidence from lesion and fMRI studies differs in a subtle but important way from the study described earlier that isolated the face-processing module. Patients with left IFG lesions do not present with an identifiable neurobehavioral syndrome reflecting the nature of the processing in this region. Guided by the fMRI results from healthy young subjects, the investigators studied patients with left IFG lesions to test a hypothesis regarding the necessity of
this region in a specific cognitive process. Coupled with the well-established finding that lesions of the left temporal lobe impair semantic knowledge, these studies further our understanding of the neural network mediating semantic memory.

**Combined fMRI/Transcranial Magnetic Stimulation Studies**

Transcranial magnetic stimulation (TMS) is a noninvasive method that can induce a reversible virtual lesion of the cerebral cortex in a normal human subject. Using both fMRI and TMS provides another means of combining brain activation data with data derived from the lesion method. There are several advantages for using TMS as a lesion method. First, brain injury likely results in brain reorganization after the injury, and studies of patients with lesions assume that the non-lesioned brain areas have not been affected, whereas, in TMS, it is performed on the normal brain. Another advantage for using TMS is that it has excellent spatial resolution and can target specific locations in the brain, whereas lesions in patients with brain injury are markedly variable in location and size across individuals. Such an approach can be illustrated in a recent investigation of the role of the medial frontal cortex in task-switching. In this study, subjects first performed an fMRI study that identified the regions that were active when they...
H stayed on the current task versus when they switched to a new task. It was found that medial frontal cortex is activated when switching between tasks. In order to determine if the medial frontal cortex was necessary for the processes involved in task-switching, the same paradigm was utilized during inactivation of the medial frontal cortex with TMS. Guided by the locations of activation observed in the fMRI study, and using a MRI-guided frameless stereotaxic procedure, it was found that applying a TMS pulse over the medial frontal cortex disrupted performance only during trials during which the subject was required to switch between tasks. Transcranial magnetic stimulation over adjacent brain regions did not show this effect. Additionally, the excellent temporal resolution of TMS allowed the investigators to stimulate during precise periods of the task, determining that the observed effect was during the time when the subjects were presented a cue, indicating they must switch tasks prior to the actual performance of the new task. Thus, combining the results from both fMRI and TMS, it was concluded that medial prefrontal cortex was essential for allowing individuals to intentionally switch to a new task.

**Combined fMRI/Event-Related Potential Studies**

The strength of combining these two methods is coupling the superb spatial resolution of fMRI with the superb temporal resolution of event-related potential recording. An example of such a study was reported by Dehaene and colleagues, who asked the question “Does the human capacity for mathematical intuition depend on linguistic competence or on visuospatial representations?” In this study, subjects performed two addition tasks—one in which they were instructed to select the correct sum from two numerically close numbers (exact condition) and one in which there were instructed to estimate the result and select the closest number (approximate condition). During fMRI scanning, greater bilateral parietal lobe activation was observed in the approximation condition as compared to the exact condition. Because this activation was outside the perisylvian language zone, it was taken as support that visuospatial processes were engaged during the cognitive operations involved in approximate calculation. Greater left lateralized frontal lobe activation was observed to be greater in the exact condition as compared to the approximate condition, which was taken as evidence for language-dependent coding of exact addition facts. In order to consider an alternative explanation of the fMRI findings, the investigators also performed an ERP study. The alternative explanation was that in both the exact and approximate tasks, subjects would compute the exact result using the same representation for numbers, but later processing, when they had to make a decision as to the correct choice, was what led to the differences in brain activation. Because fMRI does not offer adequate temporal resolution to resolve these two behavioral events that occur on brief time scale, event-related potential was the appropriate method to test this hypothesis. In the event-related potential study, it was demonstrated that the evoked neural response during exact and approximate trials already differed significantly during the first 400 milliseconds of a trial before subjects had to make a decision.
Summary

Functional MRI is an extremely valuable tool for studying brain–behavior relationships, as it is widely available, noninvasive, and has superb temporal and spatial resolution. New approaches in fMRI experimental design and data analysis are appearing in the literature at an almost exponential rate, leading to numerous options for testing hypotheses on brain–behavior relationships. Combined with information from other complimentary methods, such as the study of patients with focal lesions, healthy individuals with transcranial magnetic stimulation, or event-related potentials, data from fMRI studies provide new insights regarding the organization of the cerebral cortex, as well as the neural mechanisms underlying cognition.

References


75. Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved assessment of significant activation in functional magnetic


