3. SYMPOSIUM ON BRAIN IMAGING METHODOLOGY FOR THE STUDY OF COGNITIVE PSYCHOLOGY

Chaired by Walter Schneider, *University of Pittsburgh*

Functional magnetic resonance imaging: Overview and methods for psychological research

JONATHAN D. COHEN

Carnegie Mellon University and University of Pittsburgh, Pittsburgh, Pennsylvania

and

DOUGLAS C. NOLL and WALTER SCHNEIDER

University of Pittsburgh, Pittsburgh, Pennsylvania

Functional magnetic resonance imaging (fMRI) allows noninvasive imaging of hemodynamic changes related to neural activity. This technique can be used in single-subject designs and can provide millimeter spatial resolution and temporal resolution in the range of 5–10 sec. This paper provides a brief introduction to MRI techniques and their application to functional neuroimaging, focusing on methodological issues that are of particular concern to psychologists, including methods for presenting computerized stimuli to subjects without disrupting the scanner, experimental design issues, and statistical analysis and image processing procedures. To illustrate methodological issues, recent results from a series of studies looking at the topographic organization of visual cortex are presented. General issues concerning limitations in this technique, future directions in its development, its relationship to other neuroimaging techniques, and the role of functional neuroimaging in psychological research are addressed in the Discussion.

Methodological developments often drive both the direction and pace of scientific research. Neuroimaging techniques are clearly playing this role with regard to the development of cognitive neuroscience as a scientific discipline. Functional neuroimaging is the measurement of brain activity during task performance. Recent studies of visual and language processing using positron emission tomography (PET; Fox, Raichle, Mintun, & Dence, 1988; Petersen, Fox, Posner, Mintun, & Raichle, 1989) and of attention and language processing using event-related potentials (ERPs; e.g., Mangun, Hillyard, & Luck, 1992) have

played an important role in establishing cognitive neuroscience as a rapidly developing new field of research.

The strength of these techniques is that they provide information about brain function from awake human beings while they are performing controlled psychological tasks. However, in their present form, each technique is still limited by restricted spatial or temporal resolution. In the best of circumstances, PET can provide good spatial resolution (3 mm). In most cases, however, this is compromised by the need to average data across multiple subjects. Cortical anatomy varies markedly across individuals, compromising the ability to use PET to detect relatively small areas of activation, or to locate these reliably at the level of individual cortical gyri (the effective resolution of PET in between-subject designs is about 15 mm). Furthermore, the temporal resolution of PET is poor (on the order of 1 min). ERP offers excellent temporal resolution (1 msec); however, localizing areas of activity using this technique has proved to be a serious challenge. Magnetic resonance imaging (MRI; also referred to as nuclear magnetic resonance [NMR] imaging) provides the highest spatial resolution (in some cases, .1 mm; see Clark, Courchesne, & Grafe,

This work was supported in part by the Pittsburgh NMR Institute, University of Pittsburgh, Western Psychiatric Institute and Clinic, Carnegie Mellon University, NIMH Grant MH 47073, Office of Naval Research Contract N00014-91-J-1708, and Pew Foundation Grant T89-01245-14. The authors would like to thank Emmanuel Kanal and Bruce Rosen for their encouragement and guidance of the work reported in this paper, as well as B. J. Casey for her comments and suggestions concerning this manuscript. D. C. Noll is in the Department of Radiology at the University of Pittsburgh. Correspondence should be addressed to J. D. Cohen, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213 (e-mail: jc5e+@andrew.cmu.edu).

1992). Traditionally, however, this technique has provided only structural information.

In the last year, new methods, referred to collectively as functional MRI (fMRI), have been developed for using MRI to detect changes in brain activity. The millimeter spatial resolution provided by fMRI for each subject obviates the need for radioactivity or injected agents. This development has the potential to revolutionize functional neuroimaging and to have a dramatic impact on cognitive neuroscience and its parent disciplines. In this paper, we describe current fMRI techniques, compare them with existing techniques, and discuss considerations specific to their application in cognitive psychological research.

OVERVIEW OF fMRI

Functional brain imaging refers, in general, to the monitoring of local changes in brain activity using noninvasive techniques. When a local network of neurons fire, there are many changes that occur, including increases in electrical and metabolic activity, as well as hemodynamic changes such as in blood volume and flow. ERP and magneto-encephalography (MEG) measure electrical activity. PET and fMRI can be used to measure metabolic and hemodynamic changes (see Raichle & Mellinckrodt, 1988). It is helpful to understand some of the basics of MRI before considering how this technique can be used to produce functional images.

The Basics of MRI

MRI relies on the energy released by protons in atomic nuclei as they return to their resting state after having been temporarily perturbed. Protons precess (wobble) as they return to rest, which produces a characteristic energy signal. The production of a signal strong enough to be recorded requires that a large number of protons return to their resting state in a coherent fashion—that is, that they precess in unison. When this occurs, the individual signals sum and produce an aggregate signal that can be detected by the magnetic resonance (MR) scanning device. The large magnet used in MR scanners produces a strong ambient magnetic field, which ensures that a large number of protons within range of the field are in a similar resting state (e.g., sitting upright). To generate an image, the scanner produces a brief energy pulse (at radio frequency) within the plane to be imaged, which momentarily perturbs (tilts) the protons within that plane. It then records the energy that is released as those protons precess back to their original state.

There are two time constants associated with the decay in the MR signal: the T1 and T2 relaxation times. T1 corresponds to the decay in the signal that results from the protons gradually returning to their initial state (which can be thought of as the longitudinal component, since the protons are “tilting” back up to their “upright” state). T2 corresponds to the decay in the MR signal that occurs as the protons gradually fall out of phase in their precession (this can be thought of as a transverse component, since it is in this plane that the motion of the protons is

becoming nonuniform). T2 decay results from inherent, random variations in the precession frequency of individual protons or from local interactions between nearby nuclei. However, larger scale inhomogeneities in the magnetic field can also produce differences in precession frequency, resulting in dephasing and signal reduction. When T2 dephasing is attributable to one or more localizable sources it is referred to as T2*. This will be relevant in our discussion of fMRI below (see Young, 1988, for a more technical description of MRI techniques).

Protons belonging to different types of molecules release energy at different frequencies as they recover, providing characteristic signatures for each molecular type. Standard MRI applications measure the energy released from protons in water, since this is the most abundant molecule in biological tissues and therefore provides the strongest signal. This signal is useful, since different tissues have different water contents. Differences in signal strength, therefore, can be used to generate high-resolution structural images. However, since the water content of neural tissue does not vary as a function of activity, this approach does not yield functional information.

Recording Physiological Activity with MRI

There are a number of approaches that have been taken in the effort to use MRI to measure brain activity. One of these is MR spectroscopy. This technique attempts to measure the energy produced by protons in molecules other than water. Most efforts have focused on phosphorus, which is contained in adenosine-triphosphate (ATP) and therefore could provide an index of metabolic activity (Keshavan, Kapur, & Pettegrew, 1991). However, because the number of phosphorus molecules is so much smaller than the number of water molecules, efforts to measure the MR signal with this technique have suffered from poor signal-to-noise ratio (SNR).

Recently, a number of groups have taken a different approach to recording physiological activity using MRI. Belliveau et al. (1992), Detre, Leigh, Williams, and Koretsky (1992), Kwong et al. (1992), Ogawa, Lee, Nayak, and Glynn (1990), Ogawa et al. (1992), Rosen, Belliveau, and Chien (1989), Rosen, Belliveau, Vivea, and Brady (1990), Thulborn, Waterton, Matthews, and Radda (1982), Turner et al. (1992), and Turner, Le Bihan, Moonen, Despres, and Frank (1991) have all contributed to the development of MRI techniques for recording localized changes in blood volume or blood flow. All of these take advantage of two basic phenomena: the ability of paramagnetic agents to produce contrast in the MR signal, and the fact that regional changes in brain activity are associated with local hemodynamic changes. An elementary understanding of each of these is necessary for understanding how MRI can be used to record regional changes in brain activity.

Paramagnetism is the ability of an otherwise nonmagnetic material to exhibit magnetic properties in the presence of a magnetic field (e.g., the ability of a paper clip to attract others when it is near a magnet). The initial efforts to record physiological activity with MRI used

paramagnetic contrast agents (such as Gadolinium-DTPA), injected into the blood, to measure regional blood flow (Rosen et al., 1990). In the presence of the MR magnet, the molecules in these agents produce local magnetic fields. Recall that MRI requires that large numbers of protons recover their initial state in phase so that their signals sum to produce one aggregate signal that is detectable by the imaging apparatus. The larger the number of molecules that are in phase, the larger the signal. The molecules in the contrast agent introduce local inhomogeneities in the magnetic field, which reduce the coherence of, or "dephase," the signal generated by the protons at that location. This dephasing reduces the size of the signal detected by the scanner. As a result, brain regions that are receiving greater blood flow produce a weaker MR signal than do other regions. The difference in size of signal caused by local dephasing effects can be detected in images acquired by using MR parameters that are sensitive to T2* decay. Therefore, blood-borne contrast agents, in conjunction with T2*-weighted imaging, can be used to produce an image of brain perfusion. By comparing perfusion in activated and nonactivated states, areas of relative brain activity can be identified (Belliveau et al., 1992). One obvious limitation of this approach, however, is the need to use exogenous contrast agents to produce the image.

Recently, several groups have developed methods that do not require the use of injected contrast agents (Ogawa et al., 1990; Turner et al., 1991; Kwong et al., 1992). These are based on the observation that hemoglobin (Hb) becomes highly paramagnetic in its deoxygenated state. In other words, it appears that deoxygenated Hb can be used as a naturally occurring contrast agent: Highly oxygenated areas should produce a larger MR signal than should less well-oxygenated regions. These changes in signal intensity related to the oxygenation of Hb should therefore be detectable in T2*-weighted images. A large number of MR studies, including the ones reported below, are consistent with this reasoning.

Exploiting the oxygen sensitivity of the MR signal for functional brain imaging relies on another physiological phenomenon. A number of studies (e.g., Fox et al., 1988; Raichle & Mellinckrodt, 1988) have demonstrated that at least some brain areas increase their blood flow disproportionately to metabolic need when they become active, resulting in a net *increase* in tissue oxygenation. As a result, the MR signal in those areas increases relative to others, and this contrast can be used to produce MR images of the activated region. This method has already begun to produce reliable results. Bandettini, Wong, Hinks, Tikofsky, and Hyde (1992), Kwong et al. (1992), and Ogawa et al. (1992) have all shown activity-related changes in occipital cortex in response to visual stimuli comparable to those that have been demonstrated using paramagnetic contrast media to image changes in blood flow directly (Belliveau et al., 1992).

All of the initial efforts to image oxygen-related changes directly (without the use of contrast agents) have used specialized MRI equipment. Bandettini et al. (1992) and

Kwong et al. (1992) used echo planar imaging (EPI) techniques, which provide a means of acquiring MRI images much faster than can be acquired with ordinary equipment. Ogawa et al. (1992) and Turner et al. (1992) have reported success using high-field (4-T) magnets. Unfortunately, high-field magnets and the equipment necessary for EPI are of limited availability. Recently, however, other laboratories have begun to report similar results using conventional MRI equipment (Frahm, Bruhn, Merboldt, Hanicke, & Math, 1992; Frahm, Merboldt, & Hanicke, 1993; Gore et al., 1992). In the section that follows, we describe an example of work that we have recently conducted, using standard, "off-the-shelf" hardware and software, that is available at any clinical MRI facility. We then turn to special considerations that must be taken into account when using this technique in conjunction with behavioral research, developments to watch for in this technique, and how the technique compares with the other neuroimaging methods that are currently available.

APPLICATION OF fMRI

We have used a conventional, unmodified 1.5-T scanner (Signa, 4.7 software release, GE Medical Systems, Milwaukee, WI) to study brain activity under a variety of behavioral conditions, including visual stimulation, simple motor activity, and more complex cognitive processing (involving attention and memory tasks). Our initial work has focused primarily on attempts to activate occipital cortex in response to visually presented stimuli. The goal of these studies was to validate this neuroimaging technique by corroborating the results of invasive neurophysiological studies in primates and PET studies in humans concerning the topographic organization of visual cortex. We have reported the results of our studies in detail elsewhere (Schneider, Noll, & Cohen, 1992). In the section that follows, we discuss aspects of the methods that are most relevant to psychologists (and their collaborators) interested in pursuing similar types of studies. Following this, we summarize some of the findings from our visual studies.

Method

Behavioral Techniques

Stimulus presentation apparatus. MRI poses unique challenges for the experimental psychologist regarding the equipment that is used as well as experimental design. The magnet itself places constraints on the physical characteristics of the hardware that can be used, and the image-acquisition process, which is highly sensitive to radio-frequency (RF) signals, places constraints on the electronics. Ferromagnetic materials cannot be brought into close proximity of the scanner, and electronic devices must not produce a significant amount of RF noise, or must be shielded (see discussion below). We have had success using an active matrix LCD display panel (Sharp 1050A) and a high-luminance overhead projector, placed approximately 5 m from the subject's head (about 1-2 m from the foot of the subject table; see Figure 1). The subject observes the screen through a set of prism glasses. The display panel is connected to a microcomputer—which is usually kept in the control room—via a standard serial cable. The image is projected onto a rear-projection screen, the frame of which is made of plastic or

Figure 1. Schematic drawing of the arrangement of the stimulus display apparatus in the magnetic resonance scanner and control rooms.

aluminum, and stands at the foot of the subject's table. We have found that this provides a stable high-contrast and high-resolution visual image without producing any significant degradation in the MR signal. Other groups have also reported success using specially designed video projector systems either placed in the scanner room or projecting through the control room window.

The connections between the experimental control computer and the display must be made carefully. The typical MRI scanner room has copper-mesh shielding in the walls to minimize RF noise that can arise from radio station transmissions or nearby electronic devices such as computers. A standard scanner records data at 62.5 MHz. It is critical that any cables entering the room not carry signals of this frequency. High-speed computers and display devices in the 60-MHz range are therefore problematic. Noise produced by such sources will show up as broad lines that obscure the MR image. We have had success running both a 16-MHz IBM-compatible 386 computer with VGA output and a 25-MHz Macintosh IIci with direct video output. Ideally, it is best to use a 1-MHz bandpass filter on the video cable, to run the cable through a torus, and to keep the computer in the control room. However, not all facilities may have a convenient way of running cables from the scanner to the control room. Under such circumstances, we have found that, with a 1.5-T magnet, it is possible to place the computer in the scanner room and use floppy disks as long as the equipment remains 5 m or more from the scanner.

It is also possible to present auditory stimuli for feedback and/or auditory studies. This can be done using the pneumatic audio apparatus with which many scanners are equipped. However, the noise produced by the scanner during the scanning cycle places serious constraints on the bandwidth and fidelity of the auditory signal that can be presented to the subject. This also precludes the use of ear-plugs, which are usually used to reduce the amount of noise that the subject must endure.

For studies in which subject responses must be recorded, we have used a specially constructed, handheld, fiber-optic response box. This connects, via a fiber-optic cable, to a transducer which, in turn, is connected to the microcomputer via a serial input device. This arrangement eliminates the presence of any alternating electrical signal within the scanner that might produce RF interference.

Stimuli. In our visual studies, we presented subjects with stimuli composed of alternating checkerboard patterns (8-Hz frequency of alternation). Alternating checkerboards were used in order to

replicate the positive findings from experiments with PET in which such stimuli were used (Fox et al., 1988; Raichle & Mellinckrodt, 1988). The shape of the stimuli was manipulated to activate specific regions of visual cortex. For example, a half-circle was presented alternately to the left and right hemifields in order to produce activation of visual cortex alternating between the two cerebral hemispheres. More complex patterns were presented in order to track the activity of individual cortical gyri and to identify distinct topographic areas. This is discussed below.

Scanning Techniques

Each study involved the acquisition of two sets of images: activation images and structural images. One of the primary advantages of using MRI is that functional information can be acquired at the same time as, and directly related to, anatomic data so that areas of activity can be accurately localized. Other techniques, such as PET, require that the functional and anatomic data be acquired at different times, using different apparatuses, so that image registration becomes a problem for localizing areas of activation.

Activation images were acquired using a T2*-weighted, gradient echo pulse sequence similar to the one used by Kwong et al. (1992) and a 5-in. surface coil centered directly over the occiput. A typical set of MR parameters was a TR of 100 msec, a TE of 40 msec, a flip angle of 30°, a field of view of 24 mm, 1 NEX, an in-plane resolution of .93 mm × 1.86 mm per pixel (256 × 128 pixels/image), and a slice thickness of 5–10 mm. With these parameters, images can be acquired at a rate of approximately one every 7 to 10 sec, with an additional 15 sec per set required by the scanner. In other words, four images can be acquired in about 45 sec, and six images can be acquired in about 1 min. A typical study involved the presentation of each stimulus condition (e.g., left field vs. right field stimulation) for approximately 1 min, during which images were acquired at 4 to 6 slice locations. Each condition was repeated 10–15 times, alternating between conditions, for a total time of approximately 20–30 min per study. We were usually able to conduct two or three studies per 2-h session. Structural images were acquired immediately after the termination of each study at the same locations as the activation images for that study.

Because of the high spatial resolution of the MR image and its consequent sensitivity to movement artifact, it is important to restrain the subject's head during scanning. We have used an inflatable surgical pillow for this purpose. Other groups have used bite-

bars and head clamps. We anticipate improvements in this area that will play an important role in improving the resolution of functional MR images.

Analysis Techniques

Identifying the location and extent of activation requires a variety of statistical procedures. Hemodynamic changes produce changes in the MR signal of 2%–8%. This falls within the range of inherent variation in the MR signal, thus activity-related changes are not usually visible in the raw T2*-weighted images acquired using standard equipment. To make these apparent, we construct difference images (this is similar to procedures used in PET studies; see Friston, 1990).¹ Percent difference images are constructed by taking the mean value for each pixel in one condition, subtracting it from its mean value in the other condition, and then dividing this difference by its average value for the two conditions. These values provide an index of the effect size, which can be compared with information obtained from other methods, such as PET. In addition

to percent change, we use statistical procedures to determine the reliability of effects. We wrote a simple program that performs a *t* test on the value of each pixel across images obtained in the two stimulation conditions. The output of this program is a new file that contains a difference image, each pixel of which is the *t* value for the difference in signal intensity of that pixel across the two conditions. This file can then be loaded into an image-processing program that allows visualization of the image and analysis of the patterns of activation. *Image* is one program that can be used for this purpose; it runs on Macintosh computers and is available for free from the NIH.²

The information in statistical difference images can be used to address two types of questions: (1) where, anatomically, do significant differences in activation occur, and (2) what kind of changes occur in anatomically prespecified areas. Answering each of these questions requires a different approach.

Where did activation occur? Although it would seem to be easy to answer this question simply by looking at the difference image

Figure 2. Activation of occipital cortex in the right and left hemispheres in response to visual stimulation of right and left hemifields. The insets at the upper left show the stimulus pattern presented to the subject in each condition. The inset at the upper right shows the position (white lines) at which sagittal activation images were obtained. The large sagittal images at the bottom of the figure illustrate areas of activation (white pixels) during left hemifield stimulation and right hemifield stimulation superimposed on structural images of the right and left hemispheres. Areas of activation represent regions of significant difference in signal intensity in the T2*-weighted images (≥ 10 contiguous pixels with $t \geq 1.7$, $df = 19$) in the activation versus control condition. Note that there are only very small regions of significant difference in the hemisphere ipsilateral to the stimulus.

and comparing it with a corresponding structural image, this approach is subject to two opposing concerns. On the one hand, given the large number of t tests that are performed (32,768, for a 256×128 pixel image), it can be expected that a large number of pixels will reach significance by chance (roughly 1,600 for $p < .05$). On the other hand, performing a standard correction for repeated measures (e.g., dividing the significance value by the number of tests) is overly conservative. This is because we expect areas of activation to span a number of contiguous pixels. The conditional probability that contiguous pixels will have a significant t value is substantially less than the probability that each alone would be significant. We are not aware of any formal statistical corrections that address this set of concerns. As such, we have adopted the following procedure. We identify all regions made up of a minimum number of contiguous pixels (usually 10) whose t value is above some threshold (usually $t = 1.7$ or above), and then report the area of this region (in pixels) and the average and maximum t values for pixels within it. The region is then overlaid on the corresponding structural image in order to identify the anatomic area involved. The probability that a given set of 10 pixels would all randomly reach a t value of 2 is very small ($p < 10^{-16}$). In control studies, in which we subtracted two sets of images obtained while the subject was at rest, we have found that the application of our criteria (> 10 con-

iguous pixels with $t > 2$) does not usually produce false-positive results. There are, however, conditions under which false-positive results can occur. These include motion artifact and banding artifact, which are discussed below (see Technical Issues in the Discussion).

Has a particular area shown activation? The approach to this question is quite straightforward: Identify the anatomic region of interest and then examine the difference values for that region. Below, we provide an example of a case in which we traced a particular gyrus of visual cortex in the structural image and then plotted activation values for the locations obtained from the difference image, which allowed us to identify topographically distinct areas of functional significance.

Temporal analyses. Even taking the effects of repeated measures into account, there are still other ways in which significant difference values can arise artifactually. A high t value could be caused by spuriously high values in one or two scans within a given condition, or, in simple alternating designs (e.g., ABAB . . .), it could be caused by signal drift. A monotonically changing baseline will produce an overall difference in the first condition compared with the second condition. These concerns can be addressed qualitatively by looking at the raw signal values across time. If significant difference values are due to consistent, functionally relevant physiological changes, then signal intensity should alternate regularly and

Figure 3. Areas of significant activation (shown as white pixels) in a right sagittal slice during upper versus lower hemifield stimulation. Note that the total area activated in the two conditions is the same as the area shown in the right sagittal slice of Figure 1 (full left field stimulation) and that the two areas are divided by the calcarine fissure.

Figure 4. Pattern of activation along the calcarine fissure under different stimulus conditions. The inset at the upper left shows a tracing (white line) of the calcarine fissure in a coronal image of occipital cortex. The inset at the upper right shows a schematic of two of the stimulus conditions in which this subject was studied: alternating left-right hemifield stimulation and alternating upper-lower hemifield stimulation (differences in shading are used in the figure only to distinguish the stimuli presented in each condition; all actual stimuli were equiluminant). The gray-scale graphs at the bottom of the figure illustrate the t values at each point along the calcarine fissure for each pair of stimulus conditions (white represents a large positive t value, and black, a large negative t value). The regions labeled with letters correspond to areas that have distinct functional topography. For example, region "a" shows patterns of activity that are consistent with primary visual cortex: It activates uniformly in response to full hemifield stimulation, and it divides along the calcarine fissure for upper versus lower field stimulation. Elsewhere (Schneider, Noll, & Cohen), we have proposed that the four labeled regions (a-d) correspond to visual areas V1, V2, V3, and V4, respectively.

consistently across the behavioral conditions of the task. Isolated spikes in one condition, or baseline drift, suggest artifactual causes for the difference values observed. Of course, counterbalancing the design and using paired *t* tests will also help ameliorate these concerns. There are also quantitative approaches to analysis of the data in the temporal dimension (e.g., nonparametric tests and wave-form analyses) that should, and no doubt will, be exploited.

Summary of Results

Topographic Mapping of Visual Cortex

By using the techniques described above, we have been able to elicit areas of cortical activation that are strikingly faithful to local cerebral anatomy. Figures 2-4 show the results obtained from studies in which visual stimuli were presented alternately to the left and right hemifields, to the upper and lower hemifields, and in the form of alternating radial wedges. As expected, left/right hemifield stimulation produces activation of occipital cortex in the contralateral hemisphere (left stimulation activating right visual cortex, and vice versa; see Figure 2). Upper/lower hemifield stimulation activates the lingual gyrus (upper field stimulation) and the cuneus (lower field stimulation), dividing the areas of activation seen in each of the left/right images along the calcarine fissure that separates the lingual and cuneal gyri (see Figure 3). We have also followed the migration of activation anteriorly along the corresponding gyri as the visual stimulus is made more eccentric.

Figure 4 shows the pattern of activation along the right calcarine fissure, in response to alternating hemifields (left and right, and upper and lower). The ribbon of gray matter lying along this fissure (the inferior surface of the lingual gyrus and the superior surface of the cuneus) was traced using a modified Bezier function in a high-resolution coronal structural image. The corresponding points were then identified in the *t*-test image, and their values were laid out linearly and represented as gray-scale values. Left-field stimulation produced a solid area of activation along the upper and lower banks of the right calcarine fissure (region labeled a). Stimuli presented to the upper and lower visual fields each activated one-half of this region, corresponding to the opposing bank of the calcarine fissure. We have used this technique to plot activation observed in response to wedge-shaped stimuli, which produced alternating bands of activation corresponding appropriately to the distinct regions of the stimulus (Schneider, Noll, & Cohen, 1992). In several cases, such measurements demonstrate changes in activation from one extreme to the other within one or two pixels of the border between regions, suggesting that the technique has the capability for 1-2-mm resolution. This approach has also allowed us to define functionally distinct topographic regions of activation. We have proposed that the region described above corresponds to primary visual cortex (V1) and that the other labeled regions in Figure 4 correspond

Figure 5. Sagittal views of the occipital cortex for 3 subjects showing the pattern of activation obtained from each using the identical stimulus conditions (upper vs. lower hemifield stimulation). Note the marked variations of anatomy, and the patterns of activation that correspond to them, across the 3 subjects.

to visual areas V2, V3, and V4, respectively (see Schneider et al., 1992).

Between-Subject Variability

The studies described above have been replicated in 6 different subjects. In all cases, we were able to observe significant patterns of activation that predictably corresponded to the stimulation condition and were remarkably faithful to anatomic boundaries. However, the exact location and configuration of these boundaries varied markedly from subject to subject. Figure 5 shows the structural anatomy and pattern of activation obtained during the upper/lower conditions from 3 subjects. Not only does the shape of the calcarine fissure differ strikingly between these subjects, but so does the distribution of activation along it.

This tremendous variability in cortical anatomy has been noted by other investigators as well (Clarke & Miklosy, 1990). We will return to this issue in the Discussion below.

DISCUSSION

The work we have described in this paper, along with that of others that is beginning to appear in the literature (Frahm et al., 1992; Frahm et al., 1993; Gore et al., 1992), indicate that standard MRI equipment can be used to conduct functional studies. This technique is capable of producing functional images that are close to, if not equal to, the resolution that standard MRI provides for structural anatomy. Given the wide availability of the basic equipment, the modest investment needed to adapt this equipment for functional imaging, and the safety of the procedure, fMRI promises to quickly become an important new tool in psychological research.

As exciting and important as this new technique is, like any other technique, it is also associated with its limitations. In the remainder of this paper, we consider some of the important issues concerning fMRI as it relates to other methods that are currently available. We conclude by discussing some of the directions that are being pursued for its continued development.

Safety, Cost, and Spatial Resolution

The two primary advantages of fMRI are its safety and its high spatial resolution. As we discussed above, the fact that the procedure is noninvasive and nontoxic means that multiple scans can be obtained from a single subject. Furthermore, MRI is significantly less costly, per image, than PET and SPECT, which are the only methods currently available that provide spatial resolution that is close to that of MRI. These factors make it both possible and practical to conduct within-subject averaging of images, which is extremely valuable given the degree of between-subject variability that has been observed in cortical anatomy and activity. Most PET and SPECT studies require that images be averaged across subjects in order to obtain significant effects. This is because only a limited number of scans can be performed on a single subject (due to restrictions in the maximum allowable exposure to the radio-

active tracer agents that are involved). As we have seen, however, anatomy can vary dramatically between subjects, significantly reducing both the resolution and accuracy of the results that are obtained when data are averaged across subjects. Because there are no exposure restrictions associated with fMRI, image averaging can be performed within a single subject, and areas of activity can then be localized in structural images obtained from the same subject.

Nevertheless, the anatomic diversity observed across subjects poses a serious challenge to the functional neuroimaging enterprise as a whole, especially as it begins to take on higher cognition. The problem is that as we begin to focus on such questions, the relevant anatomy has not been prespecified (as it has for primary sensory and motor function). Furthermore, it seems plausible to assume that as we move into more cognitive domains, the mapping of function onto structure may become increasingly complex and idiosyncratic to each subject. Suppose, for example, that in a study of memory function we see consistent patterns of activation within a given subject but that the areas involved vary substantially across subjects. Does this represent anatomic diversity or differences in the way subjects are performing the task? These types of concerns are not new to psychology. Similar problems concerning individual differences arise in reaction time studies, protocol analyses, and so forth. These challenges do, however, point out the increasingly important role that well-trained psychologists will play in experimental design in this rapidly growing new field.

Temporal Resolution

The temporal resolution provided by fMRI is better than that of PET or SPECT but substantially less than that of ERP or MEG. At present, specially equipped MR scanners can acquire images as rapidly as 4 per sec (100-msec acquisition time, 150 msec between images). However, temporal resolution is also constrained by the time course of the physiological processes that are generating the signal. Current methods rely on hemodynamic responses, which appear to take place over seconds rather than milliseconds (see below). This means that, at present, behavioral conditions must produce a steady state that lasts several seconds. This is an improvement on PET and SPECT, which require that conditions last a minute or more. However, it places current fMRI techniques out of the range of being able to measure many cognitive processes, which can occur in 100 msec or less. At present, ERP and MEG are the only noninvasive neuroimaging techniques that offer this level of temporal resolution. In the future, it is likely that these techniques will be used to complement one another (see discussion below, under Current and Future Directions).

Physiological and Anatomic Limitations

Ultimately, fMRI is limited by the physiological processes it is used to measure. Current methods appear to be detecting changes in blood oxygenation that result from the brain's local regulation of its blood supply in response

to local neural activity. Animal studies suggest that this response can begin as rapidly as 500 msec after the onset of neural firing. However, even with high-speed MRI equipment (with temporal resolutions of 250 msec), researchers have not seen changes in the MR signal until about 2–4 sec after stimulation. Furthermore, these do not appear to peak until 4–6 sec (Bandettini et al., 1992; Blamire et al., 1992; DeYoe, Neitz, Bandettini, Wong, & Hyde, 1992). It may be possible to improve on this with gated acquisition of images (see below) and signal averaging; however, such methods have yet to be explored.

Another potential limitation is that some brain areas may be more efficient in regulating their blood supply than are the primary sensory and motor areas that have been studied most intensively so far. This is suggested by the growing PET literature: It has been much more difficult to generate detectable levels of activation in association cortex than it has been in primary sensory or motor cortex. This may prove to be even more difficult using fMRI (the sensitivity of fMRI relative to PET has not yet been established). The use of magnetic contrast agents, which provide additional sensitivity, may be necessary for studying subtle effects in these areas. In any event, both PET and fMRI will benefit by efforts to better our understanding of the basic physiological processes that regulate cerebral blood flow and its relationship to blood oxygenation and metabolism.

Finally, there are anatomic constraints on the use of fMRI. Some brain areas lend themselves better to MR imaging than others. In particular, it is difficult to obtain good images from areas that lie near large cavities such as the frontal sinuses. The differences in magnetic susceptibility that occur along the boundaries of these areas produce a distortion of the ambient magnetic field, which degrades the MR signal (the effect is analogous to the optical distortions produced by a glass bottle submerged in a pool of water). It is possible to compensate for these distortions by introducing appropriate additional fields (this process is called “shimming,” and is analogous to placing a set of prisms around the bottle to “smooth” out the image). However, this process is time consuming and technically demanding.

Technical Issues

There are also still technical problems associated with fMRI. As we noted earlier, the high resolution of the MR image makes it particularly susceptible to motion artifact. This can manifest in a variety of ways, the most notable of which is the appearance of gray-white-matter boundaries in difference images that can compromise the fidelity and significance of activation-related changes. These arise as structures of different magnetic susceptibility traverse the field of view of a given set of pixels. If this happens in a gradual or regular way across the conditions of the experiment, it can represent a confounding source of difference in the MR signal, which will show up in the difference image. A number of approaches have been taken to physically immobilize the head. In the long term, however, the most promising approach will be the develop-

ment of computer algorithms that will permit post hoc correction of images.

A second source of artifact is transient scanner effects. On occasion, the MR scanner can produce an aberrant image. Since activity-related changes are small ($< 5\%$), a single bad image out of 10 or 20, with large discrepant values, can significantly bias analysis results. It is important, therefore, to examine all images and eliminate ones of poor quality.

A third type of artifact has been observed by a number of groups working with standard MRI equipment. This artifact appears as large, low-spatial-frequency variations (bands) in difference images. The source of this artifact has not yet been identified. It appears to be related to transient shifts in the strength of the MR signal during image acquisition. One possible source of this effect may be diffuse changes in cerebral blood flow related to the cardiac cycle. If these changes or some other transient change in signal intensity is the cause, it should be possible to eliminate this effect by appropriately normalizing the data for each acquisition period. This approach is currently being explored.

Blood flow can also produce another form of artifact. A spatial-frequency echo is produced in areas in which a large volume of blood is moving during the scan (see Huk, Gademann, & Friedmann, 1986). T2*-weighted images may be particularly sensitive to venous blood flow. In a few scans, we have seen spatial echo from a vein, which appears in areas of gray matter.

These sources of artifact require that images be carefully examined and screened before it is determined that t test images reflect functional brain activity. Fortunately, the first three forms of artifact produce patterns in t test images that are distinct from patterns produced by functional activity. Of course, t test images should always be compared with structural images to verify that putative areas of activation lie within gray matter.

Finally, in fMRI studies, as in any other type of study, replication is the best measure of reliability. Replications are best performed on the same subject in the same session. The location of a given anatomic area can vary between subjects by as much as a centimeter or more, even after correction for overall brain size; even with the same subject, it is difficult to ensure similar head placement across sessions. Therefore, it is good practice to repeat the same set of conditions in the same session with a given subject.

Current and Future Directions

High-field magnets. The strength of the MR signal is directly related to the strength of the ambient field used to align the protons in their resting state. Most scanners that are available in clinical settings are 1.5-T. However, commercial scanners are now available at 3-T, and there are a number of 4-T research scanners that are in use. The benefit of increased field strength, in terms of SNR, appears to fall somewhere between a linear and an exponential improvement (Ogawa et al., 1992; Turner et al., 1992). Clearly, improvements in SNR will provide

improved spatial, and possibly improved temporal, resolution. Spectroscopy (e.g., the ability to image phosphorus) will also benefit by increased magnetic field strength. However, as field strength increases, so, too, does the energy of the RF pulse needed to generate a signal. At present, this energy is well below levels that are considered to be safe for human subjects. However, the safety limit, as well as the technical problems associated with building larger magnets, are likely to significantly constrain developments in this direction.

High-speed imaging. Although larger magnets are not likely to become widely available in the immediate future, faster imaging techniques are. There have been a number of recent developments in MR technology aimed at reducing image-acquisition time, such as EPI (e.g., Bandettini et al., 1992; Kwong et al., 1992), spiral scan pulse sequences (Noll, Meyer, Cohen, & Schneider, 1993), and others. Some of them require additional apparatuses, and all rely on special software. The technical details of these developments are beyond the scope of this paper. What they offer, however, is much more rapid imaging times: Images can be acquired as quickly as 4/sec rather than once every 7 sec. One or more of these techniques are likely to become more widely available within the next few years. They will open up a number of ways for improving functional studies: (1) A given number of images can be acquired over a much shorter period of time, reducing subject fatigue, minimizing motion artifact, and so forth; (2) a greater number of images can be acquired for the same amount of time, permitting better averaging, more detailed time course data, and so forth; and (3) the duration of each behavioral condition can be reduced to more closely approximate mixed designs. As noted above, the temporal resolution of MR is currently limited by the time course of local hemodynamic responses in the brain. However, as new measurement techniques that more closely measure neural activity are developed, the actual temporal resolution of MR may improve.

Gated acquisition. As the temporal resolution of MRI improves, it will become increasingly valuable to synchronize image acquisition with stimulus presentation (or response generation). This will permit image averaging to be done in a manner similar to that done with ERP and MEG. This approach might allow detection of physiological responses closer to their onset than have been observed to date—possibly as early as 500 msec after behavioral stimulation. The primary constraint on this approach is that MR images must be acquired at regular intervals. The time between images determines the degree to which protons have returned to their resting state, which in turn determines the magnitude of the signal recovered in subsequent images. Uneven intervals will produce confounding variations in signal strength.

Use of multiple imaging modalities. The greatest promise for fMRI, as for all of the other approaches to neuroimaging, is its use in conjunction with complementary techniques. Each method has its own strengths, potentials, and areas of weakness. Tremendous spatial resolution is offered by fMRI. However, it does not have the temporal

resolution of ERP or MEG. Conversely, it has been very difficult to localize activity identified by the use of ERP and MEG. By applying these techniques together, however, it may be possible to identify an area of activity with fMRI and then trace the details of its time course by using ERP and/or MEG. PET and SPECT, in turn, can provide neurochemical specificity not currently possible with other techniques. They can be used to trace the underlying neurotransmitter systems involved in functions whose anatomic location and time course have been identified with the use of a combination of MRI, ERP, and MEG. The next major phase of growth in neuroimaging research, no doubt, will be the development of techniques for coordinating the use of multiple modalities. This has important conceptual significance, which we turn to in the final section.

Conceptual Issues

There are two fundamental conceptual issues that must be addressed when considering any neuroimaging technique: the role of subtractive methodology and the theoretical significance of the data that are generated. The problems associated with subtractive methodology are not new to psychology. Donders (1868/1969) wrote about these problems long ago. Nevertheless, every study that reports activation in a particular area of the brain that is associated with a particular function has used subtractive methodology to establish that relationship: Signals obtained while the subject is at rest—or performing a control task—are subtracted from signals obtained while the subject is performing the experimental task. The concern, of course, is whether all of the other “irrelevant” functions—those that are supposed to be subtracted out—are being performed in the same way or in the same locations during the two conditions. There is no question that careful task analysis, accompanied by sophisticated experimental design, can help ameliorate these concerns. Ultimately, however, data acquired with a particular neuroimaging technique must be compared with data obtained using other methods—other neuroimaging techniques, human behavioral studies (both with normal and clinical populations), and animal studies. Only with the use of convergent methodologies will it be possible to establish clear and reliable relationships between anatomic structures and their function(s).

However, even when a particular cognitive function has been reliably identified with a particular anatomic structure, it is important to ask what has been learned. Is it interesting that vision occurs in the back of the brain? To the surgeon, perhaps, but clearly not to the psychologist. What is of interest to the psychologist are the mechanisms that subserve vision—and how they might participate in other processes (e.g., mental imagery)—not simply their geographic location. The data provided by neuroimaging techniques alone are not sufficient to answer questions about mechanisms. Once again, this requires the application of a rich array of approaches—from detailed, invasive neurophysiological studies of target anatomic areas to sophisticated behavioral studies that identify the func-

tions they subserve. Only when neuroimaging techniques are used to provide a link between these two traditional endeavors will they realize their potential as the powerful tools that they currently promise to be.

SUMMARY AND CONCLUSION

Functional MRI is a promising new addition to the array of functional neuroimaging techniques that are becoming increasingly available to cognitive neuroscientists and psychologists. It provides a noninvasive means for imaging hemodynamic changes that appear to be closely related to functional brain activity. Major advantages of fMRI are that it does not require the injection of contrast agents and does not pose any currently known risk to the subject. As such, it permits within-subject designs, which can be used to dramatically improve spatial resolution and to test the replicability of findings. Compared with other neuroimaging techniques, it offers the best available spatial resolution, with temporal resolution that surpasses PET and SPECT but is inferior to ERP and MEG. The current limitations of this technique—from the point of view of the psychologist—are its supramillisecond temporal resolution and sources of artifact (some of which have not yet been identified) that can obscure results. There is little doubt that the future of this technique lies in its coordinated use with complementary techniques. The increasing use of neuroimaging techniques to study the relationship between brain activity and behavior guarantees psychologists an important role in experimental design and interpretation of results.

REFERENCES

- BANDETTINI, P. A., WONG, E. C., HINKS, R. S., TIKOFKY, R. S., & HYDE, J. S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, **25**, 390-397.
- BELLIVEAU, J. W., KENNEDY, D. N., MCKINSTRY, R. C., BUCHBINDER, B. R., WEISSKOFF, R. M., COHEN, M. S., VEVEA, J. M., BRADY, T. J., & ROSEN, B. R. (1992). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, **254**, 716-719.
- BLAMIRE, A. M., OGAWA, S., UGURBIL, K., ROTHMAN, D., MCCARTHY, G., ELLERMAN, J., HYDER, F., RATTNER, Z., & SCHULMAN, R. G. (1992). Echo-planar imaging of the activated human visual cortex shows a time delay between stimulus and activation. *Proceedings of the 11th Annual Scientific Meeting and Exhibition of the Society of Magnetic Resonance in Medicine* (p. 1823).
- CLARK, V. P., COURCHESNE, E., & GRAFE, M. (1992). In vivo myeloarchitectonic analysis of human striate and extrastriate cortex using magnetic resonance imaging. *Cerebral Cortex*, **2**, 417-424.
- CLARKE, S., & MIKLOSSY, J. (1990). Occipital cortex in man: Organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *Journal of Comparative Neurology*, **298**, 188.
- DETRE, J. A., LEIGH, J. S., WILLIAMS, D. S., & KORETSKY, A. P. (1992). Perfusion imaging. *Magnetic Resonance in Medicine*, **23**, 37-45.
- DEYOE, E. A., NEITZ, J., BANDETTINI, P. A., WONG, E. C., & HYDE, J. S. (1992). Time course of event-related MR signal enhancement in visual and motor cortex. *Proceedings of the 11th Annual Scientific Meeting and Exhibition of the Society of Magnetic Resonance in Medicine* (p. 1824).
- DONDERS, F. C. (1969). Over de snelheid van psychische processen. Onderzoekingen gedaan in het Psychologisch Laboratorium der Utrechtsche Hoogeschool: 1868-1869. *Acta Psychologica*, **30**, 412-431. (Original work published 1868)
- FOX, P. T., RAICHEL, M. E., MINTUN, M. A., & DENCE, C. (1988). Nonoxidative glucose consumption during focal physiologic neural activity. *Science*, **241**, 1445-1448.
- FRAHM, J., BRUHN, H., MERBOLDT, K. D., HANICKE, W., & MATH, D. (1992). Dynamic MR imaging of human brain oxygenation during rest and photic stimulation. *Journal of Magnetic Resonance Imaging*, **2**, 501-505.
- FRAHM, J., MERBOLDT, K. D., & HANICKE, W. (1993). Functional MRI of human brain activation at high spatial resolution. *Magnetic Resonance in Medicine*, **29**, 139-144.
- GORE, J. C., MCCARTHY, G., CONSTABLE, R. T., ANDERSON, A. W., KENNAN, R. P., RATTNER, Z., & ZHONG, J. (1992). Imaging regional brain activation at 1.5T using conventional imaging techniques. *Proceedings of the 11th Annual Scientific Meeting and Exhibition of the Society of Magnetic Resonance in Medicine* (p. 1826), Book of Abstracts, Society of Magnetic Resonance in Medicine.
- HUK, W. J., GADEMANN, G., & FRIEDMANN, G. (1986). *Magnetic resonance imaging of central nervous system diseases*. New York: Springer-Verlag.
- KESHAVAN, M. S., KAPUR, S., & PETTEGREW, J. W. (1991). Magnetic resonance spectroscopy in psychiatry: Potential, pitfalls, and promise. *American Journal of Psychiatry*, **148**, 976-985.
- KWONG, K. K., BELLIVEAU, J. W., CHESLER, D. A., GOLDBERG, I. E., WEISSKOFF, R. M., PONCELET, B. P., KENNEDY, D. N., HOPPEL, B. E., COHEN, M. S., TURNER, R., CHENG, H.-M., BRADY, T. J., & ROSEN, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of the Sciences (USA)*, **89**, 5675-5679.
- MANGUN, G. R., HILLYARD, S., & LUCK, S. J. (1992). Electrocortical substrates of visual selective attention. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV*. Cambridge, MA: MIT Press.
- NOLL, D. C., MEYER, H. C., COHEN, J. D., & SCHNEIDER, W. (1993). Spiral scan imaging of cortical activation. *Journal of Magnetic Resonance Imaging*, **3**, 44.
- OGAWA, S., LEE, T. S., NAYAK, A. S., & GLYNN, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, **14**, 68-78.
- OGAWA, S., TANK, D. W., MENON, D. W., ELLERMAN, J. M., KIM, S., MERKLE, H., & UGURBIL, K. (1992). *Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping using MRI*. Manuscript submitted for publication.
- PETERSEN, S. E., FOX, P. T., POSNER, M. I., MINTUN, M., & RAICHEL, M. E. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, **1**, 153-170.
- RAICHEL, M. E., & MELLINCKRODT, E. (1988). In V. B. Mountcastle, F. Plum, & S. R. Geiger (Eds.), *Handbook of physiology: The nervous system*. Bethesda, MD: American Psychological Society.
- ROSEN, B. R., BELLIVEAU, J. W., & CHIEN, D. (1989). Perfusion mapping by nuclear magnetic resonance. *Magnetic Resonance Quarterly*, **5**, 263-281.
- ROSEN, B. R., BELLIVEAU, J. W., VIVEA, J. M., & BRADY, T. J. (1990). Perfusion imaging with NMR contrast agents. *Magnetic Resonance in Medicine*, **14**, 249-265.
- SCHNEIDER, W., NOLL, D. C., & COHEN, J. C. (1992). *Functional topographic mapping of human visual cortex using conventional MRI*. Manuscript submitted for publication.
- THULBORN, K. R., WATERTON, J. C., MATTHEWS, P. M., & RADDA, G. K. (1982). Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochimica et Biophysica Acta*, **714**, 265-270.

- TURNER, R., JEZZARD, P., WEN, H., KWONG, K., LE BIHAN, D., & BALABAN, R. (1992). Functional mapping of the human visual cortex at 4 tesla using deoxygenation contrast EPI. *Proceedings of the 11th Annual Scientific Meeting and Exhibition of the Society of Magnetic Resonance in Medicine* (p. 304).
- TURNER, R., LE BIHAN, D., MOONEN, C. T. W., DESPRES, D., & FRANK, J. (1991). Echo-planar time course MRI of cat brain oxygenation changes. *Magnetic Resonance in Medicine*, **22**, 159-166.
- YOUNG, S. W. (1988). *Magnetic resonance imaging: Basic principles*. New York: Raven Press.

NOTES

1. It is a simple matter to write computer programs to analyze MR images. Each image is stored as a binary file that contains a header and then a sequence of values—usually 16-bit signed integers—that are the MR signal intensity for each of the pixels in the image.

2. *Image* can be obtained free via anonymous ftp at zippy.nimh.nih.gov[128.231.98.32] in the /pub/image directory, or by contacting its author, Wayne Rasband, at wayne@helix.nih.gov.