

*Epistemic Issues in Procuring  
Evidence about the Brain:  
The Importance of Research  
Instruments and Techniques*

*William Bechtel and Robert S. Stufflebeam*

**1 The Epistemic Challenge Posed by Research  
Instruments and Techniques**

According to traditional philosophical accounts of scientific methodology, the evidence for scientific theories stems from observation, especially observation with the naked eye. These accounts portray the testing of scientific theories as a matter of comparing the predictions of the theory with the data generated by these observations, which are taken to be an objective portrayal of reality. One lesson learned by philosophers of science in the last 40 years is that even observation with the naked eye is not as epistemically straightforward as is sometimes assumed. What one is able to see depends upon one's training: a novice looking through a microscope may fail to recognize the neuron and its processes (Hanson, 1958; Kuhn, 1962/1970).<sup>1</sup> But a second lesson is only beginning to be learned: the evidence in science is often not procured through simple observations with the naked eye, but through observations mediated by complex instruments and sophisticated research techniques. In order to acquire evidence about the phenomena under investigation, these instruments must alter it. (For a simple, prosaic example, consider the ordinary thermometer. It requires the transformation of the temperature of the surrounding air into the expansion of a liquid or metal so as to produce a display that we can see [Hacking, 1983].) The fact that evidence consists of altered phenomena then raises a serious question: to what degree is what is taken as evidence just the product of the alteration or in what respects does it reflect the original phenomena for which it is taken to be evidence?

Since most scientific evidence is procured through instruments and research techniques, the question of whether purported evidence is really an artifact is a frequent

one, especially when a new instrument or technique is being introduced. One reason for this is that the purported evidence is often extremely variable. Even when different researchers try to use the same instrument and follow the same procedure, subtle variations in the instrument or technique can produce significantly different results. Another is that often when they are introduced, the ways in which the new instruments and techniques alter the phenomena are not understood, at least in sufficient detail to vindicate directly the purported evidence. A potent example of how purported evidence is produced through techniques not well understood is Golgi's silver nitrate stain which only stains selective neurons and played a critical role in establishing the claim that neurons were discrete individual cells (see chapter 3, this volume); yet 100 years after its introduction, its method of action and the reason why it only stains selective cells are not understood.

If purported evidence is only a reflection of the means of altering the phenomena, then it is only an *artifact*; if it is, in the relevant respects, reflective of the underlying phenomena, then it is genuine evidence. The challenge for scientists is to determine which it is. Scientists themselves are usually very much aware of the problem. Especially at a time when new sources of evidence are being introduced, they often engage in bitter and extended arguments as to whether the purported evidence should be trusted. Typically, these controversies are resolved after a short period of time. Protocols for the use of instruments and performance of the techniques that are accepted as informative are established and routinized; students learn them and treat the resulting evidence as authoritative. As a result, the epistemic issues disappear from sight. As Bruno Latour (1987) graphically describes it, the procedures become *black boxes*.

For those concerned with the epistemic status of science, the means by which new instruments and techniques are evaluated and, if accepted, become black boxes, is of central significance. Initially it is a bit of a mystery how scientists could provide evidence that new instruments and techniques are producing genuine evidence and not artifacts. The approach used to evaluate theoretical knowledge – comparing predictions of theories or models against evidence – is of no use since there is no independent body of evidence against which to evaluate the instruments and techniques.<sup>2</sup> As we will demonstrate later in this chapter in the context of discussing techniques for neuroimaging, scientists generally rely on a variety of indirect measures in their evaluations. These include (1) whether the instrument or technique is producing well-defined or determinate results, (2) the degree to which the results from one instrument or technique agree with results generated in other ways, and (3) the degree to which the purported evidence coheres with theories that are taken to be plausible. Before turning to that case study, we will first describe some of the instruments and techniques on which neuroscience, especially neuroscientific accounts of mental function, have relied and attempt to cultivate an appreciation for how much these instruments and techniques alter the underlying phenomena and raise the prospect of artifact. Our goal is not to promote skepticism about scientific inquiry, but rather an awareness of an epistemic challenge that is central to scientific practice.<sup>3</sup>

## 2 Instruments and Research Techniques Employed in Discovering how the Brain Performs Mental Processes

To study the brain, researchers must make its structure and processes accessible. The functioning brain is usually shrouded in a skull out of sight. Even when the skull is removed, little information about its operation is available through simple observation. One can identify some of its larger landmarks, such as the cerebellum and the major gyri and sulci in convoluted brains such as our own, but these observations do not reveal the relevant structural units in the brain (which tend to be much smaller and not indicated by surface structures; see chapter 3, this volume), let alone the occurrent chemical and electrical processes most significant for maintaining mental life. Even determining that chemical and electrical processes are occurring in the brain requires sophisticated tools of intervention.

In this section we describe a number of the major tools which have been employed in the neuroscience investigations recounted elsewhere in this volume. (We will focus on techniques that are thought to be most informative about the primary neural correlates of cognitive activity, the electrical activities of neurons. One should note, though, that other activities, especially chemical activities, are also highly relevant to cognitive performance and to other mental phenomena such as emotions.) In providing a brief introduction to the instruments and techniques, our emphasis will be on revealing the sort of intervention each requires, the kind of information about the brain each provides, and some of the limitations each encounters.

One general issue that we should note at the outset is that the various techniques for studying brains are employed with members of different species. Most of what we know of the neuroanatomy of the brain has been learned from studies of non-human primates such as the macaque and most of the single-cell electrophysiology has been done on cats and various species of monkeys. Results from these studies are then extrapolated to the human brain. Electroencephalography, PET, and fMRI studies, on the other hand, are generally done on humans, and although some comparative work on other species has begun, such investigations pose serious challenges. Yet, from what is known, it is clear that there are significant differences in the organization of brains across species, rendering the task of making inferences across species challenging.

### Neuroanatomical methods

Neuroanatomy, the characterization of the structure of the brain, both at a macro and a micro level, has provided the foundation for a great deal of understanding of how the brain performs mental functions. The previous chapter described some of the classic discoveries of neuroanatomy – the discovery that neurons are the functional units of the nervous system, and the identification of areas of the brain with different neural composition, patterns of connectivity, etc. – and the procedures by

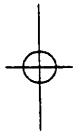
which these were discovered. To determine the physical composition and organization of the brain, neuroanatomists have had to do such things as cut it apart, slice it into thin preparations, treat these with chemical baths of various sorts, and examine the products through an (optical, electron, etc.) microscope. These processes are all disruptive of the normal brain but are pursued because the resulting images are thought to reveal important structures within the non-disrupted brain.

Critical to the functioning of the brain is the manner in which neurons are connected to each other. Consider what is involved in discovering the connectivity pattern in a portion of cortex. You might try viewing a stained preparation of neural tissue, but all you will see is a tangled web of connections. To discover the relevant organization within these complex mazes, neuroscientists have done such things as cut axons from their neurons, causing the axons to die. This results in accumulation of dense granular material along the axons' path which can be seen in histological preparations after the animal dies. Another approach is to use retrograde and anterograde tracers, chemical substances which are taken up by neurons at a given point and transported backwards or forwards along their axons and dendrites. Horseradish peroxidase, for example, is taken up by the axons of cells and transported back to the cell bodies where, over a few days, it oxidizes and takes on vivid colors. This allows the cell bodies connected to particular axons to be readily identified in slice preparation after the animal is killed.

Insofar as neuroanatomical studies depend not just on microscopes to magnify images, but on dissection of the brain and applications of various stains and tracers, these studies clearly involve interventions. While many of these interventions are reasonably well understood (e.g. the process and rate of radioactive decay in radioactive tracers), others are often less well understood (e.g. the process by which various stains bind with substances within the cell). When successful, these procedures can provide rich detail about the structures in the nervous system. But there are also some clear limitations on neuroanatomical approaches. For example, while they may reveal that neurons in one area project on to another, they do not show what kinds of connections are involved (inhibitory or excitatory) and what information is conveyed. A critical example of how such limitations are restricting the emerging understanding of the nervous system is provided by the discovery of enormous numbers of neurons projecting backwards from areas further within the system to areas closer to the sensory periphery. Most neuroscientists think these recurrent connections are extremely important for brain function, but at present it is not clear what information is carried by these connections and hence what information-processing function they perform. Thus, as critical as neuroanatomy is for understanding the brain, understanding function requires techniques that directly intervene in the functioning of the brain and render the functional processes salient.

### Deficits and lesions

One of the oldest approaches to identifying the function of brain components is analysis of the deficits resulting from lesions (localized damage) to those compo-



nents. Lesions can originate either from illness or injury or from neuroscientists actually destroying neural tissue. Whatever the source of the lesion, the goal of this approach is to identify a psychological deficit associated with it and to infer from that what contribution the damaged area made to normal psychological function. Although the deficits to language (e.g. aphasia, dyslexia, agraphemia) are perhaps the best known (and the use of lesions to study language will be illustrated in the contributions to Part II), there is a wide range of mental abilities that can be selectively impaired by lesions to the brain (e.g. face recognition, ability to orient towards the locations of objects, encoding new events into episodic memory).

One challenge in lesion research is determining precisely what areas of the brain are injured. Until the recent introduction of imaging technology, one could only determine what areas in the human brain were damaged after the person died and an autopsy was performed. By then, though, the range of damage may have extended. This required Broca (see chapter 5, this volume), for example, to engage in protracted argument as to what the extent of damage was when the deficit appeared. Perhaps the greatest challenge in using lesions and deficits to understand brain operation is to infer precisely what the damaged component had contributed to normal function. The most general inference is that the damaged area was in some way necessary to the normal performance (e.g. inferring from the fact that lesioning the hippocampus or hippocampal region results in anterograde amnesia, to the conclusion that the region is necessary for encoding new episodic memories). Even this inference is problematic as sometimes an organism can recover or develop an alternative way of performing a function over time after brain injury. The challenge is even greater when one tries to specify just what aspect of the task (e.g. encoding new episodic memories) the damaged part played. It may have been responsible for the whole task, or it may have performed only one contributing function, perhaps even an ancillary one. One can gain an appreciation of the challenge involved by considering how one might go about trying to understand how a radio (or a computer) operates by selectively removing parts and examining the resulting performance. As Richard Gregory (1968) notes, removal of a transistor from a radio may cause it to hum, yet it would be a bad inference to assume that the removed transistor was the hum suppressor.

One strategy that is widely invoked in lesion research is to attempt to dissociate two mental functions by showing that damage to a given brain part may interfere with one but not another. Single dissociations, however, do not show that the damaged brain part is only involved in the impaired function, since it could be that the two functions differ in the demands they make on a component and that with increased damage, the same brain part might interfere with both functions. As a result, researchers often seek double dissociations, where damage to one area causes disruption in one function (while leaving the other largely unaffected), and damage to another area disrupts the other function (while leaving the first largely unimpaired). Double dissociations are often taken as compelling evidence that the two functions are performed separately in the brain (Shallice, 1988). Recent investigations with neural networks and other dynamical systems, however, have shown that

double dissociations can result from differential damaging of a single system where it is known that there are not different subsystems carrying out separate tasks (e.g. applying rules for pronouncing words versus looking up pronunciations in a lexicon – see Hinton and Shallice, 1991; Van Orden et al., in preparation). Thus, double dissociations are not foolproof indicators that there are separate systems responsible for separately impaired functions.

We noted above the problem of determining the precise extent of naturally occurring lesions. Although there remains uncertainty with surgically induced lesions,<sup>4</sup> in general these offer much more control. However, for obvious ethical reasons, permanent lesions are only made in human brains in neurosurgical patients when it is anticipated that removal of a brain area is likely to have a beneficial effect such as reducing epileptic seizures, and in such cases the pre-existing medical problem is likely to complicate any functional interpretation of the consequences of the lesion. This has meant a restriction of experimental lesion studies to non-human animals. Recently, however, new techniques have been pioneered in which researchers can induce temporary lesions in humans. One of these techniques that currently affords great promise is transcranial magnetic stimulation; it involves application of a strong but localized magnetic field so as to disrupt the activity in the affected brain area. Early reports (Walsh and Cowey, 1998) indicate that one can disrupt very specific functions, but the critical question of what the affected area contributed positively to the function remains a challenge, one which typically requires complementing lesion studies with electrical stimulation or recording studies.

### Electrophysiological studies

The discovery of the nature of electricity and that the brain in part operates on electrical principles enabled neuroscientists to study the brain as one would study other electrical systems, probing it with electrical stimuli or recording its electrical activity. It is important to note, though, that the idea that the brain is an electrical system was only formulated in recent times. One of the eighteenth-century discoverers of electricity, Luigi Galvani, proposed that nerve transmission was electrical on the basis of experiments stimulating peripheral nerves and muscles with an electrostatic device. His proposal was only definitively established in the mid-nineteenth century by Emil du Bois-Reymond, who, by using non-polarizable electrodes and a multiplier for nerve current, developed a galvanometer that was sufficiently sensitive to detect electrical currents in nerves. Only after this demonstration did researchers attempt to analyze the electrical processing within the brain either by stimulating it or by recording the electrical currents generated as the brain carried out various tasks.

**Stimulation studies** The strategy in stimulation studies is to inject electrical current into the brain in the attempt to elicit responses, with the assumption that if one can elicit a response from a given area with an exogenous source of electricity, then normal electrical activity in the affected area would also generate the same

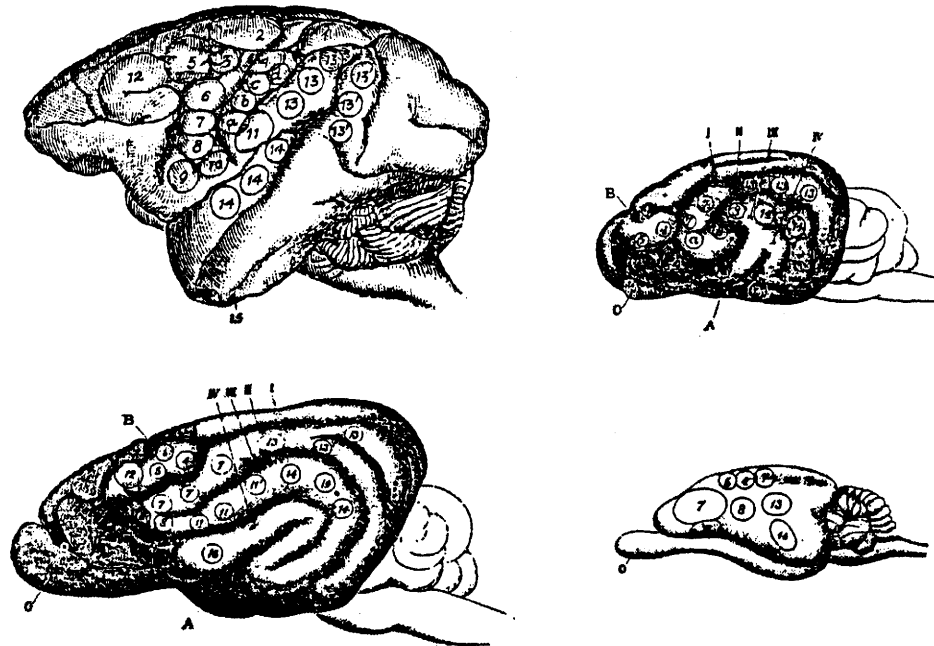


Figure 4.1 Ferrier's (1886) maps of sites on the left hemispheres of monkey (upper left), dog (lower left), cat (upper right) and rabbit (lower right) from which motor responses were elicited with mild electrical stimulation. The same numbering pattern is used for each species and designated a specific motor response.

response. Several nineteenth-century researchers pursued this strategy, but could not discover the right dosage of electricity to elicit a response. By using very mild stimulation, in 1870 Gustav Fritsch and Eduard Hitzig succeeded in eliciting muscle movement in dogs after electrical stimulation of their brains. The approach was generalized by David Ferrier (1876), who elicited responses to electrical stimulations in a variety of areas in the brains of many different species including macaque monkeys, dogs, jackals, cats, rabbits, guinea pigs, rats, pigeons, frogs, and fishes (see figure 4.1). Ferrier argued that many of these loci were not specifically motor, but reflected sensory or other psychological processes leading to motor responses. Ferrier construed stimulation as the natural complement of lesion studies – where deficits resulting from lesions would show what areas were necessary for a function, response after stimulation would show what brain activity was sufficient for a particular response.<sup>5</sup>

Since the pioneering work of Fritsch and Hitzig and Ferrier, electrical stimulation has been widely used in attempts to map brain areas from which particular mental or motor responses could be elicited. Among the major contributors to this work were Walter Rudolf Hess (1949), who applied the technique to subcortical

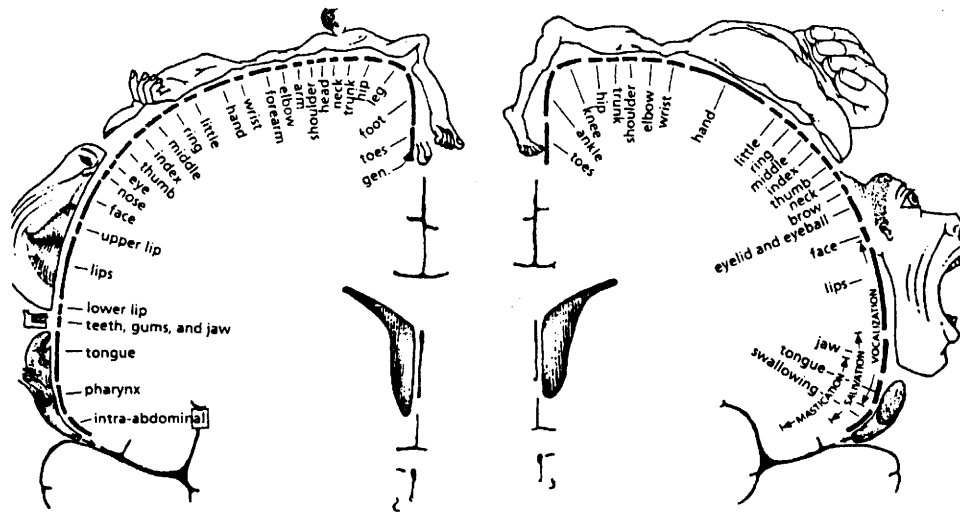


Figure 4.2 Representations of the primary sensory cortex (left) and primary motor cortex (right), using homunculi to show the relative amounts of brain tissue devoted to different portions of the body. Redrawn from Penfield and Rasmussen (1950).

areas, where stimulation seemed to alter emotional behavior in cats, and Wilder Penfield (Penfield and Rasmussen, 1950), who employed it on humans who were candidates for neurosurgery to reduce epileptic symptoms. Penfield's immediate objective was avoiding damage to critical areas such as those responsible for speech, but his research played a major role in mapping cortical areas, including sensory and motor cortices, in humans (see figure 4.2).

Although electrical stimulation has provided a great deal of useful information about the functioning of different brain regions, it relies on some major assumptions. The first is that the electrical stimulus supplied by the electrode induces electrical activity comparable to that which arises within the brain itself. One reason to be cautious is that generally only the joint activity of multiple neurons, transmitted chemically across numerous synapses, is sufficient to elicit a response in another neuron or muscle; to stimulate a neuron artificially, a relatively large burst of electricity must be applied at one location. Such an electrical discharge could easily spread a considerable distance through the cortex and activate areas beyond that which is directly stimulated. Second, just as with lesion studies, there is a serious question of how one should interpret the cognitive contribution of the stimulated area. In the case of the sensory and motor cortices, the idea that these were the projection and major motor command areas in the cortex is supported on other grounds, including neuroanatomy and lesion studies, and what electrical stimulation facilitated was the detailed mapping of, for example, different locations on the primary motor strip which generated motions in different parts of the body. Numerous

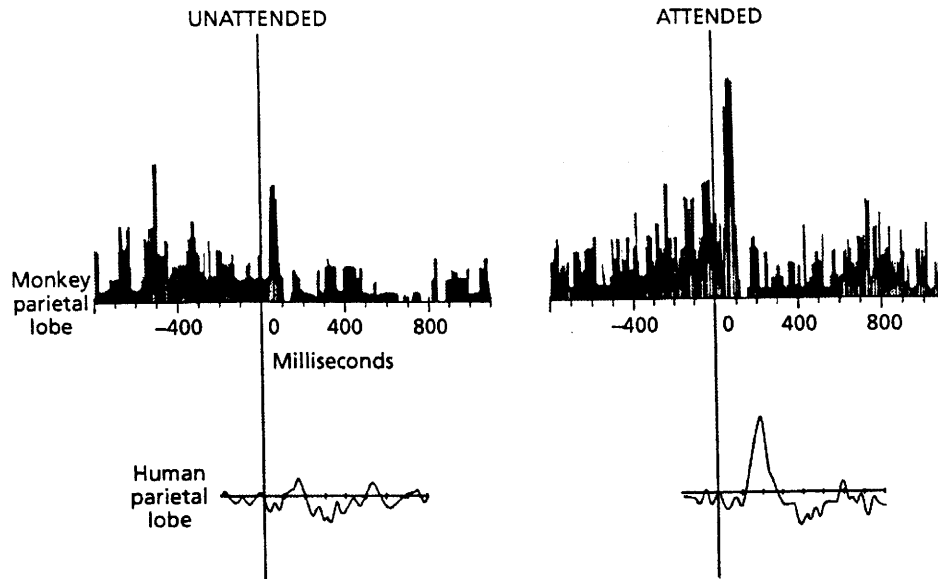


researchers in the 1960s applied the technique of electrical stimulation more widely and drew conclusions that have proven extremely controversial. For example, James Olds claimed to have discovered a pleasure center upon finding that electrical stimulation probably to an area in or near the hypothalamus would cause rats to press a bar repeatedly when an apparatus was configured so as to produce further stimulation (Olds, 1965). Other researchers have contested his conclusion, advancing alternative interpretations of his findings. For example, William Uttal (1978) concluded: "it now appears that many of the hypothalamic effects on feeding and drinking are mediated not by hypothalamic nuclei but rather by interruption of the sensory-motor signals conveyed by fiber tracts that pass in close proximity to these nuclei" (p. 340). Likewise, José Delgado's (1969) interpretation of his finding, that stimulation in the brain of a bull would stop its aggressive charge, as indicating a center for inhibiting aggression was challenged by Elliot Valenstein, who notes that the stimulated bull continually circles in one direction and concludes, "any scientist with knowledge in this field could conclude only that the stimulation had been activating a neural pathway controlling movement" (Valenstein, 1973, p. 98). Just as with lesion studies, the challenge in electrical stimulation studies is to constrain the interpretation of the contribution of the stimulated site to normal mental function.

**Electroencephalogram and evoked response potentials** Electroencephalography involves recording aggregate electrical signals from electrodes placed either on the skull or directly on the cortex. In the late nineteenth century such recordings were made from animals by Richard Caton and Adolph Beck, but it was Hans Berger, a German psychiatrist, who applied the technique to humans by adapting methods designed to record the much stronger electrical signal from the heart muscle. He first recorded electroencephalograms in humans with skull defects, then in 1924, from the skull of his then 15-year-old son. Berger distinguished several different patterns of waves, including what he called alpha waves (large-amplitude, low-frequency waves which appeared when subjects closed their eyes) and beta waves (smaller, higher-frequency waves, which appeared as soon as the subject received sensory stimuli or were asked to solve a problem) (Berger, 1929).

Berger's research provides a clear example of the development of an instrument and technique which seemed to be revealing evidence about the operation of the brain, but which was difficult to interpret. On the one hand, as researchers recorded EEGs from a wide variety of individuals in various different states (e.g. different stages of sleep), a complex classification system for EEG patterns began to develop. But, on the other hand, it was not at all clear either what was the source of the signal or what it was telling researchers about brain processes. Initially researchers assumed it reflected some sort of summation of nerve firing. However, early attempts to record from individual neurons (see below) revealed no correlation between individual neural activity and the EEG (Li et al., 1952). It is now accepted that the EEG signal originates principally with pyramidal cells which are aligned in columns in the cortex; when these cells are stimulated, ion flows into and out of the cell are created, resulting in a dipole. When the cells are aligned spatially and activated synchro-





*Figure 4.3* Examples of both recordings from a single cell and event-related potentials. The upper traces are the electrical activity shown by a single parietal neuron following stimulus, delivered at a time indicated by the long vertical line. The left side was recorded while the monkey was not attending to the location of the stimulus, and the right side was recorded while it was. The lower traces are the average event-related potentials recorded from a scalp electrode over the parietal lobe of a human subject under the same stimulus conditions. Reprinted from M. I. Posner and M. E. Raichle, *Images of Mind*, New York: Scientific American Library, 1994, p. 21.

nously, these ion flows create an electric field strong enough to be recorded at the scalp (Kutas and Dale, 1997).

EEG recordings have turned out to be useful for studying sleep-wake cycles, identifying brain damage, detecting the origin of epileptic tumors, and monitoring the depth of anesthesia. However, the basic EEG has not been useful for studying information processing in the brain, since recorded EEG activity is a combined measure of many different neural processes. To employ EEG for studying information processing, it was necessary to identify a response in the EEG that was specific to a particular stimulus presented to the subject. This was accomplished when G. D. Dawson (1951) adapted a procedure originally designed for detecting lunar tides in the atmosphere by averaging over a large number of cases. Over many cases in which the same stimulus was presented, background noise would likely be randomly distributed whereas any specific response to the stimulus would stand out. Through use of this averaging procedure, researchers have been able to develop evoked response potentials (ERPs) in which a specific electrical response pattern generated in response to a particular stimulus provides a temporal record of neural activity (figure 4.3).

ERPs have turned out to be extremely valuable in relating neural processing and mental activity because of their temporal precision. For example, by studying the changes in the ERP signal under various attention-directing conditions, researchers are able to resolve in part a long-enduring controversy over whether attention operates early or late in the course of processing a stimulus. Attending to the location of a stimulus resulted in increased P100 and N100 waves (P and N stand for positive and negative currents respectively and the numbers refer to the time in milliseconds after stimulus presentation that the effect is produced) (Luck and Ford, 1998). ERP studies have also been informative about many aspects of language processing. For example, sentences with semantic anomalies generate a negative wave beginning around 200 milliseconds after the presentation of the anomalous word. This wave peaks at approximately 400 milliseconds and so is referred to as an N400 wave. By identifying what word in a semantically anomalous sentence elicited the N400, Garnsey and her colleagues provided evidence that people commit to a particular parsing of ambiguous sentences early rather than waiting for disambiguating information (Garnsey et al., 1989).

One of the chief advantages of ERPs is that they can provide very fine-grained information about the timing of neural processes. On the other hand, it is extremely difficult to determine the spatial origin in the brain of the ERP signal recorded on the skull. The problem, known as the *inverse problem*, is to determine what set of generators within the skull could produce the observed pattern of activation. The difficulty is that there are usually multiple solutions to this problem. A further limitation of ERP studies is that while they can show that there is a distinctive electrical activity related to a particular stimulus, they neither reveal the precise nature of the response at the neuronal level nor specifically what information-processing role that underlying brain activity is playing.

**Recording from single cells** One of the most powerful ways of relating brain activity to function is to record electrical activity from individual neurons, either by inserting an electrode into the neuron or by placing it next to the neuron, and relate the activity there to ongoing mental processes (see top half of figure 4.3). The first challenge neurophysiologists faced in recording from single cells was to amplify the electrical activity sufficiently to record it. Success in this effort came from the work of Edgar D. Adrian (1926), who connected an intact sensory nerve in the frog to the input of an amplifier and recorded the resulting spike pattern on a oscillograph with a moving strip of photographic film. The recording technique was further advanced by two Washington University investigators, Joseph Erlanger and Herbert Gasser, who in 1922 introduced a cathode ray oscillograph, which they employed to discover that nerve impulses traveled at different velocities depending on the diameter of the fiber, resulting in the classification of different types of nerve fiber (including the philosophers' favorite example, C-fibres; see chapter 16, this volume).

Once the technology for recording from individual neurons was worked out, the challenge was to interpret the activity by relating it to mental processes. To do so,

researchers require independent access to mental processes. This is done most easily with neurons involved in sensory processing or controlling motor activity. In the course of his research in which he removed the optic nerve from the frog and identified cells in it that responded to stimuli presented at particular locations in the visual field, Haldan Keffer Hartline (1938) introduced the idea that individual neurons had specific *receptive fields*. The first successful attempt to record from cells in living organisms was Stephen Kuffler's recording from retinal ganglion cells; Kuffler discovered that these cells responded most when a stimulus was presented at the center of the receptive field and not in the surround (on center, off surround) or vice versa (Kuffler, 1953).

Many of the pioneering neuroscience discoveries made in the middle of the twentieth century employed cell recording. In a extremely influential paper entitled "What the frog's eye tells the frog's brain," Jerome Lettvin and his colleagues identified retinal ganglion cells which responded to specific stimuli, including some that responded to small moving spots that the researchers characterized as bug detectors (Lettvin et al., 1959). Hubel and Wiesel's pioneering work in the 1950s, which they summarize in chapter 10, below, used cell recording to map out the response characteristics of cells in primary visual cortex. Working in auditory cortex of the cat during the same period, Clinton Woolsey (1960) identified cells at different locations that responded to stimuli at different frequencies. Cell recording has also been used to identify cells engaged in tasks further removed from the sensory and motor periphery; for example, Goldman-Rakic (1987) has identified cells that continue to fire after a stimulus has been removed when the animal must retain that information for a short interval before performing an action.

As useful a technique as cell recording is, it does have limitations. First, since the technique is primarily correlational, it requires identifying a sensory stimulus, motor response, or ongoing cognitive activity that can be correlated with the neural activity. Second, although when successful it allows researchers to identify what stimulus drives the cell, it does not reveal what contribution the cell is making to processing that information. As Marr (1982) argued, this requires an analysis of the task the cognitive system is performing and accounts (pitched at the psychological, information-processing level) of how it is carrying out that task. Third, it assumes that electrical responses of individual cells are the proper correlate of psychological function. Increasingly, researchers are exploring the possibility that the proper correlate may be a pattern distributed over many cells. Procedures for recording from many, possibly hundreds, of cells simultaneously are now being developed, but these pose serious challenges in terms of analyzing the resulting information.

### Structural neuroimaging

As we noted above, until very recently researchers had to wait until they could perform an autopsy to identify structural features of an individual's brain. One of the major recent developments has been the introduction of procedures that allow researchers to image the brain while the person or organism is still alive. Some of

these are principally useful for examining neural structure whereas others have been employed to investigate functions performed by brain structures. Since we will focus on the epistemic issues concerning neuroimaging below, we will defer introducing those issues and simply describe the various neuroimaging techniques.

**Conventional radiographs** X-rays, like other forms of radiation, are absorbed by the material through which they pass, with the density of the matter determining the degree of absorption. Those X-rays passing clear through the material irradiate the photographic plate, producing a white image, whereas a black image results when X-rays have been absorbed by the substrate. Dense materials, such as bone and tumors absorb many X-rays and show up as black in images. The gray and white matter of the cortex absorbs very few X-rays and so appears as white and cannot be distinguished from other soft tissue. Examination of the brain with traditional X-rays therefore produces a two-dimensional image that primarily reveals the cranial structure and tumors; by adding compounds to the blood that absorb X-rays, it is also possible to create X-ray images of the arteries and veins of the brain. For neuroscientists interested in the organization of gray matter, however, X-ray images are of little use.

**Computerized tomography (CT)** In a two-dimensional image such as that produced by X-rays, one cannot determine where in the third dimension an absorbent material might be. By rotating the radiation source (initially, the source was X-rays) and detectors around the object and summing the readings from all of the beams that pass through a given location within the object, researchers developed a means of locating where in the third dimension a structure was located. The Greek word *tomos* means "cut" and this technique is referred to as *tomography* since locations are specified in a single plane cut through the object. Not only do X-ray CTs provide detailed depth information, they also are able to differentiate gray and white matter, blood, and cerebrospinal fluid, thereby providing a much more detailed account of the anatomy of the brain. But the CT technique has also been extended for use with numerous other radiation sources and is used in the last two imaging techniques described below.

**Autoradiography** To image function, one needs to identify a correlate of neural firing which provides a recordable signal. Whereas EEG and ERP uses the electrical fields generated by ion flows in and out of cells, functional neuroimaging relies on signals associated with basic cell metabolism or blood flow. One of the first techniques, autoradiography, relies on labeling metabolically active cells by tagging metabolites such as deoxyglucose with a radioactive element (often  $C^{14}$  or  $F^{18}$ ). After injection, an animal performs a task, then is sacrificed. Slices of its brain are laid on photographic plates to develop images which reveal the areas of the brain most metabolically active. While in some instances this approach can provide stunning images of brain activity (Tootell et al., 1982, for example, demonstrated the topographical layout of visual cortex by having a monkey view a figure resulting in the

reproduction of the pattern on the monkey's visual cortex upon autopsy), it has the obvious disadvantages of requiring the sacrifice of the subject and being able to gather data on only one mental activity per animal.

**Positron emission tomography (PET)** By combining the use of a emissive radioactive element with computerized tomography, PET generates images of metabolic activities in functioning brains. One strategy for PET imaging adapted a technique developed by Louis Sokoloff and his colleagues for labeling glucose metabolism. It employs radioactively labeled 2-deoxyglucose, a close analog of the glucose that figures in basic cell metabolism but which builds up rather than being metabolized in the cell. As the cell requires more energy, more radioactively labeled deoxyglucose-6-phosphate builds up in the cell; as it decays, positrons are ejected from the radioactive atom, travel a short distance until they collide with an electron, whereupon they annihilate in the emission of two gamma rays directed at  $180^\circ$  to each other. The PET scanner contains detectors surrounding the head which record an event only when two gamma rays arrive at different locations simultaneously; sophisticated computational techniques are then employed to determine the site of the annihilation. Although the 2-deoxyglucose strategy is often employed in diagnostic uses of PET, the more common approach in functional neuroimaging is to use labeled  $H_2O$  which has a short half-life (permitting several sequential scans of the same subject); the labeled  $H_2O$  is carried in the bloodstream, so what is being measured is the increased bloodflow that accompanies neural firing. Since PET was developed before fMRI, much of the pioneering research in functional neuroimaging employed it. However, it encounters a number of serious limitations. Since it uses radioactive tracers, for health reasons subjects can only be scanned while performing a limited number of tasks on one occasion in their lifetime. Since the signal is relatively weak, subjects must perform the task repeatedly during the 20-second scan, and data must be averaged over multiple subjects. Increasingly, therefore, fMRI is replacing PET as the neuroimaging technique of choice.

**Magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI)** In a strong magnetic field, the nuclei of elements which have an odd atomic weight (e.g. hydrogen) are induced to align the axes of their spin. A brief pulse of radiowaves can perturb this alignment by tipping the orientation of spin, thereby increasing the energetic state of the nuclei. When the pulse ends, they precess back into their aligned state, releasing energy in the form of radio waves in which the frequency reflects the particular atom and its environment. Since hydrogen atoms in gray and white matter have different relaxation frequencies, MRI can clearly differentiate them and provide detailed structural images of the brain. The ability to find an MRI signal correlated with function rests on the fact that brain activity generates increased blood flow in excess of oxygen utilization, resulting in a task-dependent reduction in deoxyhemoglobin, a paramagnetic molecule (Fox et al., 1988). This gives rise to what is termed the blood oxygen level – dependent (BOLD) contrast between conditions of heightened and less heightened neural activity.

Ogawa et al. (1990) predicted that this would permit BOLD-based MRI, a prediction borne out in processing of simple sensory stimuli in Ogawa et al. (1992). In a very short time, techniques for using fMRI (functional MRI) to study cognition have been dramatically improved, making it now the neuroimaging technique of choice. One indication of the potential for fMRI is the recent development of techniques to relate changes in the MRI signal to individual events (still averaging over multiple such events), thereby avoiding the need to have subjects perform the same cognitive task over 20-second intervals (Rosen et al., 1998).

### 3 Epistemic Evaluation of New Techniques: The Case of Neuroimaging

In our survey of methods for gaining information about brain function related to mental activity, we have emphasized that each involves an indirect measure of the brain's activity, often generated by intervening in the normal activity of the brain. This raises the prospect that the results are artifacts, not data informative about the brain activity normally underlying mental activities. Thus, an extremely important aspect of scientific practice is the process by which scientists evaluate new instruments and research techniques to determine whether results from them should be accepted as data. What makes the issue of evaluating whether a technique is producing data or an artifact challenging is that, unlike the evaluation of theories against independently evaluated data, there is not a prior source of data which one can generally use to evaluate a research technique. In evaluating the reliability of the data, researchers must invoke very different kinds of criteria. We propose that three criteria figure prominently in such evaluations: (1) the definitiveness of the results themselves, (2) the consilience of the results with those generated by other procedures, and (3) the coherence of the results with plausible theoretical accounts.

To make the case that these criteria figure prominently in scientist's actual evaluation of new instruments and techniques, we will focus on the new functional neuroimaging techniques, PET and fMRI. In both cases, the basic physical processes used to produce the image are reasonably well understood, and not the source of any skepticism. Thus, in PET, the processes by which a radioactive isotope emits a positron which, once it collides with an electron, annihilates and produces two gamma rays directed 180° from each other is well understood and not contested. Similarly, the processes by which nuclei align themselves in a magnetic field, are induced to tilt by a pulsed radio wave, then precess back when the pulse ends in fMRI are not at issue. Further, the fact that both reliably measure increased blood flow in brain areas is not a concern for those concerned with artifacts. Rather, much of the concern focuses on the relation between cognitive processes and increased blood flow.

We can differentiate two components of the relation between cognition and blood flow. On the one hand, there is the mechanism responsible for changes in blood flow.

On the other, there is the cognitive interpretation of the activities. Raichle (1998) traces inquiry into the relation of brain activity and blood flow to nineteenth-century researchers such as Angelo Mosso, but shows that the mechanism responsible for increasing blood flow is still not known. The details of the relation between increased neural activity and increase in blood flow was not critical for evaluating PET. For fMRI the issue is more important since, as we noted, the BOLD technique relies on blood flow increasing more than is required by oxygen consumption by neurons. At present there are several proposals to explain the increased blood flow (for a review, see Raichle, 1998), which have differing implications for how the resulting images relate to neural activity.

While not denying the importance of this first issue, we will concentrate on the second – the process of bridging to cognition. We should note first that both PET and fMRI place considerable constraints on the cognitive activities that can be studied. The person has to perform the activities while lying motionless on his or her back within the confines of a scanner. Consequently, the cognitive activities studied in neuroimaging will not be the ordinary ongoing cognitive activities of life, but tasks specifically designed to be carried out under such circumstances (e.g. reading words from a list). Second, the goal of imaging is to localize specific cognitive processes or operations by identifying the increased blood flow associated with them (and, to the degree possible, determine the relation between different processes).<sup>6</sup> Imaging studies thus depend critically on construction of tasks for which researchers already have a plausible cognitive decomposition which they employ to guide interpretation of the imaging results. Information-processing psychology has produced cognitive decompositions of a number of activities into successive stages of processing and neuroimagers have frequently availed themselves of these. But these decompositions are themselves contested, especially by advocates of dynamical systems models who deny that overall behavior is not the result of successive stages of processing, but instead an emergent product of highly distributed dynamical processes (see MacKay, 1998; van Orden and Paap, 1997). Even if one accepts that there is a decomposition of cognitive processes in the brain, there remains the question of exactly what decomposition the brain employs. One point to thus stress is that any imaging study is only as good as the assumption of decomposition of processing components on which it relies.

One of the major ways in which imagers have tried to link the results of imaging with proposed decompositions of tasks into cognitive operations is to image a person while performing two different tasks thought to differ only in that one employs one or more cognitive operations additional to those employed by the other, and then subtracting the second image from the first, generating a difference image. One then identifies the area(s) revealed in the difference image as the locus of the additional operation. For example, in a landmark early imaging study that we will take as our main example in what follows, Petersen et al. (1988, 1989) subtracted the image produced when a subject reads a noun and pronounces it aloud from the image produced when a subject reads nouns, generates a related verb, and pronounces the verb.





They thereby hoped to identify those brain areas required to generate the verb (see chapter 7, this volume).

The procedure just described for relating areas of increased activation in images to cognitive operations is known as the *subtractive method*. It was initially developed by F. C. Donders (1868) for use in chronometric studies of cognitive process: reaction times for one task were subtracted from those for another task and the difference was thought to reflect the time required for the additional processes required in the longer task. In chronometric studies, the subtractive method was broadly criticized in the 1960s. Sternberg (1969) pointed out, for example, that the subtractive method assumed that the additional cognitive activity was a pure insertion into a sequential set of processes, and this assumption might well be false. As a result, he advocated replacing the subtractive method in studies of mental chronometry by techniques which measure whether different tasks interfere with each other (for detailed discussion, see Posner, 1978). Neuro-imagers have returned to the original simple subtraction approach of Donders, a move Raichle (1998) defends by arguing that because they will be able to observe any changes in activation in other brain areas that might arise in the more complex task, researchers will be able to detect failures of pure insertion. There are, however, reasons for skepticism: imaging procedures will only identify statistically significant changes in activation elsewhere in the brain; if there are resulting accommodations elsewhere in the brain, they may fall below this threshold and thus not be noted.

Amongst skeptics, a major source of doubt about the reliability of localization based on imaging studies stems from the variability in results that have been obtained. Through a meta-analysis of PET studies of rhyming, David Poeppel (1996) revealed considerable variability in the areas researchers identified. The five studies he reviewed identified 22 different brain areas, only eight of which were identified in more than one study, and only three appeared in as many as three of the studies. Such variability is not universal in neuroimaging studies (Corbetta, 1998, for example, emphasizes the agreement of different imaging studies in identifying areas with increased activation in tasks involving covert attentional or eye movements to target locations), but when it occurs, it raises questions for researchers and skepticism amongst critics. Poeppel is not himself a skeptic of imaging research, but uses the variability he identified to point to the critical importance of psycholinguistic theories in PET results – differences in the theories used to guide the decomposition can result in very different linkages between brain areas and cognition. An additional factor is the specific way in which the studies are carried out. Often differences in the way the tasks are administered that are not thought to be significant can produce major differences in resulting images. In fact, neuroimagers themselves uncovered one such difference in method that can generate different results – the degree of practice subjects have with the task prior to imaging. After Petersen et al.'s study of verb generation, the researchers decided to repeat the study with subjects who had practiced the task. Performance after even quite short practice resulted

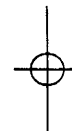
in different patterns of activation when the practiced lists were used (Raichle et al., 1994).

Given the grounds for skepticism, why have so many cognitive neuroscientists quickly adopted neuroimaging? We will focus on three factors that figure in the assessment of neuroimaging as well as other newly introduced techniques: the definiteness of the images themselves, the agreement of the imaging results with those arrived at through other techniques, and finally the ability of the results to support plausible theoretical models.

By *definiteness* of the images, we are referring to the fact that the techniques generate specific, reasonably well delineated images which change under different task conditions. For a lay person, one of the compelling features of the images produced by PET and fMRI is their coloration which makes it appear that particular areas of the brain “light up” on specific tasks. In fact, the results of raw scans or of subtractions are actually numbers, either indicating amount of blood flow or the degree of significance of the increase in blood flow. These numbers are translated into a coloring scheme which was specifically chosen for its suggestiveness – hot colors (reds and yellows) indicating increases in blood flow, cold colors (blues) indicating decrease. The lay impression is, thus, a consequence of the means of presentation researchers have adopted. But what is important is that the images reveal increases and decreases in activation in reasonably well delineated areas of the brain.<sup>7</sup> Moreover, these areas remain roughly constant across trials and subjects (when the tasks remain constant). If the PET or fMRI signal was an artifact not linked to brain processes associated with specific tasks, multiple individual performances would vary and cancel each other out, or yield a pattern in which the pixels whose activations are statistically significant would be distributed randomly over the cortex. The fact that continuous areas are all activated above threshold indicates that there is a brain activity related to the cognitive task that gives rise to the image.

The constancy is revealed most significantly in studies (especially PET studies) which rely on averaging across subjects. Since even the size and shape of the brain varies across individuals, such averaging relies on techniques for mapping individual brains on to a common atlas (such as the Talairach Atlas (Talairach and Tournoux, 1988)). The differences between brains could easily have resulted in diluting the activation so that no areas would show up as having significantly increased activation. The fact that areas of localized increased activations appeared in spite of averaging suggests that the imaging results are robust and not likely to be artifacts.

A second strategy is to demonstrate that the results are consistent with the results of other ways of studying cognitive function in the brain such as lesion studies and single cell recording. Prior to the advent of PET, these tools had provided a modestly rich account of the tasks performed by different areas in primate cortex. These studies had provided good grounds for believing that most visual processing occurs in occipital cortex and surrounding areas of temporal and parietal cortex. Likewise, cortical stimulation studies had revealed the motor and somatosensory cortices in Brodmann’s areas 1–4. As a result, the first studies using PET to study



cognitive performance only sought to demonstrate that simple motor tasks such as finger tapping would produce activation in motor cortex and that looking at visual stimuli would produce activation in primary visual cortex (Fox et al., 1987; Fox et al., 1986).

As researchers have moved to more cognitive tasks, it has become routine to identify similarities between the areas identified in imaging with those identified in lesion and single cell recording research. For example, the stage in the Petersen et al. (1989) study in which they subtracted the activations generated when subjects simply listened to or read words, from those where they pronounced them, yielded bilateral activations in the motor and sensory face areas as well as the cerebellum. They found these results to be highly credible since they cohered with a long history of evidence that the motor and sensory face areas are involved in language production, especially articulatory processing (for example, that lesions in the most focal parts of Broca's area are known to result in speech production deficits). Another example that relates to work discussed elsewhere in this volume is provided by Haxby et al. (1991), who demonstrated that activations on tasks requiring determining the location of visual stimuli versus the identity of the stimulus produced activations in approximately the same areas as Mishkin, Ungerleider, and Macko (chapter 11, this volume) had shown resulted in deficits when lesioned in monkeys.

Insisting on consistency with findings from other techniques creates something of a paradox. The goal of introducing new techniques is to revise and extend our knowledge, but the consistency requirement would seemingly prevent that. One part of the resolution of this paradox is to require that the new technique maintain consistency with established techniques only in the domain of overlap, and on the basis of that overlap to assume that the new technique is also providing correct information when it provides results that extend beyond those techniques. Imaging studies often find activations both in areas which had been associated with a particular cognitive function through lesion or cell-recording studies as well as in new areas not previously expected. For example, Petersen et al. found activation during the verb-generating task in parts of the cerebellum. Traditionally the cerebellum has been viewed as principally engaged in motor activity, but this study and others has contributed to the growing recognition that the cerebellum plays a role in a variety of cognitive tasks (Thach, 1998).

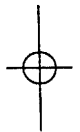
The more challenging situation is when a new technique such as imaging produces results that are at odds with those produced by earlier techniques. Since we cannot be sure that the older techniques were themselves reliable, one would not want to simply dismiss new techniques where they generate conflicting evidence. But clearly there is a greater burden on those advancing the new technique when such discrepancies emerge. Again, the Petersen et al. study with the verb-generating task provides an instructive example since they conceived of this task as adding a semantic component to the pronunciation task. It was expected that the areas of increased activation after subtracting the areas active in the corresponding noun pronunciation task would be those involved in semantic processing. Following the modern reinterpretation of Wernicke's deficit studies by Bradley et al., (1980) and others (see

chapters 7 and 8, this volume), Wernicke's area in the superior temporal lobe has been interpreted as the site of semantic processing. But in the verb-generating task, which requires semantic processing, Petersen et al. found increased activation in the left prefrontal cortex and not in Wernicke's area. They contend that increased activation in Wernicke's area is limited to situations requiring phonological encoding of the input (such as when words are presented auditorily). Their PET results are thus inconsistent with a long history of lesion studies; accordingly, the burden was on Petersen et al. to show that their results were not artifacts.

Our goal is not to endorse the Petersen et al. interpretation, but to examine the strategy of arguing for it. A starting point is to note that lesions only indicate weak points in a processing system (a point through which information is conducted, for example), not necessarily the location where a failed process is performed. Thus, lesion results are not definitive. But Petersen et al. required a more positive defense for their alternative proposal. One strategy they employed was to emphasize consistency with earlier, non-PET blood flow studies. However, since PET also relies on blood flow, that just raises the possibility that both sets of studies rest on artifacts involving blood flow. They also emphasized consistency across PET studies. For example, they report PET scans on five subjects while they carried out a semantic judgment task, in which subjects judged whether words referred to objects in the same category; this study showed increased activation in a very similar area of prefrontal cortex (although these increases did not reach the threshold for statistical significance).

The challenge to Petersen et al. was increased by the results of a PET study by another laboratory, one using related but different tasks that were also expected to tap semantic processing. Thus, Christopher Frith and his colleagues (Frith et al., 1991) found activation in a semantic task in both Wernicke's area and prefrontal areas. Frith et al. present their results as *disconfirmation* of Petersen's results since they did get activation in Wernicke's area, and they attribute the prefrontal activation to "intrinsic generation rather than semantics" (p. 1146). Petersen and Fiez (chapter 7, this volume) respond to this claimed disconfirmation by analyzing the tasks employed in the Frith et al. study. Specifically, they argue that since all of the input conditions in Frith et al.'s study involve auditory input, the activations Frith et al. found in Wernicke's area are due to auditory processing, not semantic processing and that only the activation in prefrontal cortex found in Frith et al.'s study is really a candidate for semantic processing.

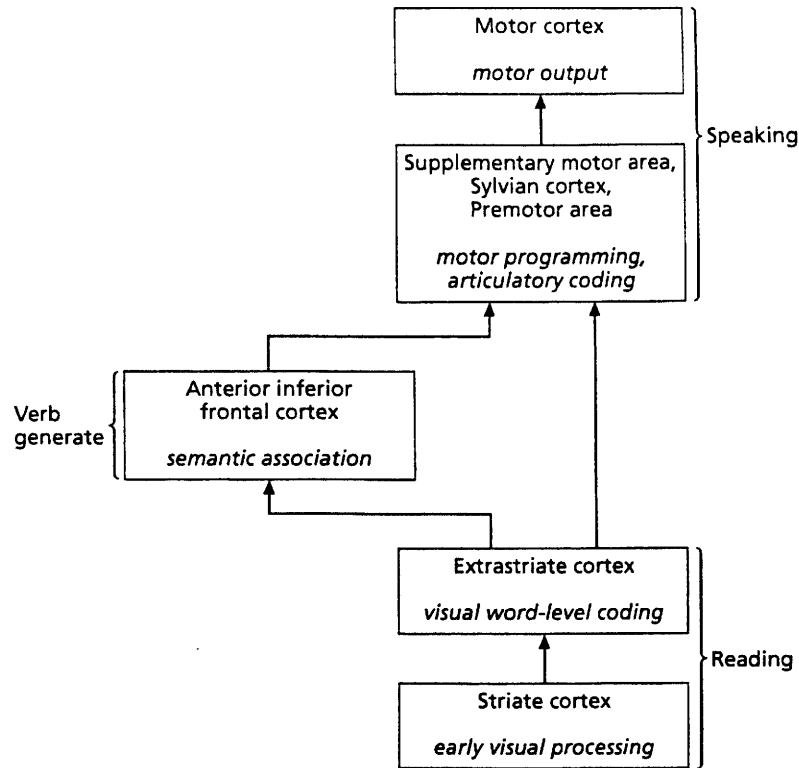
While their critique of the Frith et al. study may suffice to neutralize its challenge to their results, the burden is still primarily on Petersen and his colleagues to show that the prefrontal activations reflect semantic processing since it is their result which is at odds with the history of lesion studies. To make their case, Petersen et al. invoke a third strategy for providing credibility for new techniques, showing that the new evidence fits a compelling theory. The theory claims that the type of processing required for semantics is characteristic of that performed in prefrontal cortex. Particularly important here is the idea that prefrontal cortex is involved in withholding responses to stimuli, a cognitive process that is likely to be important in semantic



processing. Evidence for this proposal is offered by studies by Goldman-Rakic (1987) showing that lesions in anterior prefrontal cortex in monkeys leaves the animals unable to withhold responses to false cues. (This theoretical framework, and its possible relevance to semantic processing, has recently been developed more fully by Deacon, 1997, who argues that what is critical for a semantic system is the ability to establish contrasts between lexical items with different meanings, a capacity Deacon also links with processes in prefrontal cortex.)

In advancing the theoretical perspective in which they situate their PET results, Petersen et al. contrast two theoretical approaches, a serial, single-route framework and a dual-route framework. The serial, single-route framework they associate with Norman Geschwind (1979), who interprets the processing in reading aloud as following a pathway from visual cortex through Wernicke's area in the angular gyrus, to Broca's area, and finally to primary motor cortex. In contrast, the dual-route models distinguish a lexical route (required in order to read words with non-standard pronunciations such as *pint*) and a non-lexical route that utilizes phoneme-to-grapheme correspondence rules (required to read non-words such as *rini*). Dual-route models have received independent support from neuropsychological research that reveals a double-dissociation between patients who can read words with non-standard pronunciations but not non-words, and patients who read non-words, but not words with non-standard pronunciations (Coltheart, 1987); however, they have also been challenged by a variety of investigators (Plaut, 1995). Petersen et al. advocate the dual-route model as indicated in figure 4.4, thereby rejecting the single-route framework which has traditionally assigned a major role in semantic processing to Wernicke's area. It is important to note that what serves to explain away the traditional interpretation of the semantic function of Wernicke's area is not the two routes distinguished by Coltheart (since neither the lexical nor the non-lexical route involves Wernicke's area), but a separate distinction between the processing of visually presented words and of auditorily presented words. The latter necessarily entails phonological encoding of words, while such encoding can be by-passed with visually presented words. The dual-route models of word reading thus provide a framework which Petersen et al. can *expand* upon to incorporate their results indicating lack of semantic processing in Wernicke's area. (Petersen et al. also indicate a variety of pieces of supporting psycholinguistic evidence for this extension of the dual route framework.) With the expanded dual-route framework, they can advance a theoretical model that renders their results plausible, thereby countering the objection that their results are inconsistent with more traditional results indicating a semantic function for Wernicke's area.

From a strongly empiricist point of view, in which evidence or data is the foundation upon which theoretical frameworks are developed and evaluated, the suggestion that researchers use theoretical frameworks to support their experimental results seems seriously misguided. In practice, however, such an approach is rather common (see Bechtel, 2000). Failure to find a plausible theoretical framework in which to understand the results of a technique leads scientists to suspect that the technique is generating artifacts, whereas success in identifying such a framework reduces that



*Figure 4.4* Two routes for processing visual and auditory language inputs (after Petersen et al., 1989). Note that neither the direct route from visual processing to motor output nor that through semantic association employs Wernicke's area.

suspicion and can even counter the threat posed by results generated by more traditional techniques.

#### 4 Conclusions

The attempt to link studies of the brain to psychological processes is critically dependent on a variety of instruments and research techniques, many of which we described briefly in section 2 of this chapter. Because these techniques all rely on indirect measures and generally upon intervening in normal brain processes, each has confronted an epistemic challenge to show that they are generating reliable evidence about normal brains. In addition to identifying some of the reasons for these epistemic questions we have pointed to some of the limitations faced by the various instruments and techniques. Our objective, again, is not to promote skepticism, but

an appreciation of the basis on which scientific claims about brain function are based. In the third section we examined these epistemic issues in greater depth by focusing on the new neuroimaging techniques of PET and fMRI. Particularly important in our analysis is the challenge of drawing psychological interpretations of these measures of blood flow. Using Petersen et al.'s study of single-word processing with PET, we argued that in fact, in evaluating the reliability of such a procedure, scientists rely on the definitiveness of their results, the consilience of their results with those of other techniques, and the plausibility of theoretical models in terms of which they interpret their results.

### Notes

- 1 The idea that perception, the recognition of objects and events, is indirect and rests on "unconscious inferences," was clearly articulated in the nineteenth century by Helmholtz. A number of perceptual theorists (for example, James Gibson) have taken issue with the reference to "inferences" in perception, but the development of a neurobiological understanding of how the visual system operates (see Part III) clearly supports the idea that there are many processing steps between the registration of light on the retina and the recognition of the object or events seen. Accordingly, many of the same issues that arise in evaluating instruments arise in perception itself. Bechtel (2000) argues for parallels between the way we settle the occasional but rare disputes about what is visually perceived and the disputes between scientists over whether an instrument has generated an artifact.
- 2 Sometimes scientists offer theories of how instruments or techniques work, in which case the more standard account of evaluation by comparing these theories against their evidence can be invoked. But although some theoretical ideas often lie behind the development of instruments and techniques, they are often incomplete and incapable of settling controversies that arise over artifact versus evidence. Detailed theoretical knowledge of how the instrument or technique operates is generally not available until long after the instruments and techniques have been employed and the questions about artifacts resolved.
- 3 A second objective is to make philosophers who are inclined to draw upon the results of empirical inquiry more sensitive to the procedures on which such inquiry rests. Responsible philosophical utilization of scientific results requires cognizance of how those results were obtained.
- 4 For example, the focus on the role of the hippocampus in encoding memory resulted from lesions in a patient, H.M. that were made in the attempt to control incurable epilepsy. Although the surgeon thought he had removed the whole hippocampus, more recent MRI studies reveal that much of it remained (although over time it has atrophied).
- 5 Ferrier is rightly construed as one of the major figures in the development of neuroscience, but that does not imply that most of his fundamental claims turned out to be correct. As we shall see in chapter 13, drawing upon both lesion and stimulation studies, he contended that the angular gyrus in the parietal cortex, not the occipital lobe, was the primary site of visual processing in the brain.

- 6 If we focus on whole activities a person performs, the associated brain processes will not be restricted to just one brain area. Rather, performance is generally distributed over numerous areas, each of which does part of what is needed to perform the overall tasks. As noted by Petersen and Fiez (chapter 7, this volume), "a functional area of the brain is not a task area: there is no 'tennis forehand area' to be discovered. . . . Any task or 'function' utilizes a complex and distributed set of brain areas." Accordingly, what must be localized or mapped on to the brain through neuroimaging are "simple operations." The aim is to determine the distinctive contributions, or simple operations, performed in different regions of the brain.
- 7 By referring to specific areas, though, we are not suggesting that these areas are known to correspond to neuroanatomically delineated areas which are thought to be functionally significant. Imagers often try to link areas of activation with anatomically or physiologically identified areas (see Zeki et al., 1991), but this is in general a very difficult task due to the paucity of information about the locus of these neuroanatomical areas in humans. Most of the relevant neuroanatomy has in fact been done on monkeys or other non-human species.

### References

- Adrian, E. D. 1926: The impulses produced by sensory nerve endings, part I. *Journal of Physiology (London)*, 61, 49–72.
- Bechtel, W. 2000: From imaging to believing: Epistemic issues in generating biological data. In R. Creath and J. Maienschein (eds), *Biology and Epistemology*, Cambridge, UK: Cambridge University Press, 138–63.
- Berger, H. 1929: Über das elektroencephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87, 527–70.
- Bradley, D., Garrett, M., and Zurif, E. 1980: Syntactic deficits in Broca's aphasia. In D. Caplan (ed.), *Biological Studies of Mental Processes*, Cambridge, MA: MIT Press, 269–86.
- Coltheart, M. 1987: Cognitive neuropsychology, and the study of reading. In M. I. Posner and O. S. M. Marvin (eds), *Attention and Performance*, vol. 11, Hillsdale, NJ: Lawrence Erlbaum, 3–37.
- Corbetta, M. 1998: Frontoparietal cortical networks for directing attention and the eye to visual locations: identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences, USA*, 95, 831–8.
- Dawson, G. D. 1951: A summation technique for detecting small signals in a large irregular background. *Journal of Physiology*, 115, 2P.
- Deacon, T. W. 1997: *The Symbolic Species*. New York: Norton.
- Delgado, J. 1969: *Physical Control of the Mind*. New York: Harper.
- Donders, F. C. 1868: Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool: 1868–1869. *Tweede Reeks*, 2, 92–120.
- Ferrier, D. 1876: *The Functions of the Brain*. London: Smith, Elder, and Company.
- Fox, P. T., Burton, H., and Raichle, M. E. 1987: Mapping human somatic sensory cortex with positron emission tomography. *Journal of Neurosurgery*, 67, 34–43.



- Fox, P. T., Minton, M. A., Raichle, M. E., Miezin, F. M., Allman, J. M., and van Essen, D. C. 1986: Mapping human visual cortex with positron emission tomography. *Nature*, 323, 806–9.
- Fox, P. T., Raichle, M. E., Mintun, M. A., and Dence, C. 1988: Nonoxidative glucose consumption during focal physiologic neural activity. *Science*, 241, 462–4.
- Frith, C. D., Friston, K. J., Liddle, P. F., and Frackowiak, R. S. J. 1991: A PET study of word finding. *Neuropsychologia*, 29 (12), 1137–48.
- Garnsey, S. M., Tanenhaus, M. K., and Chapman, R. M. 1989: Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, 18, 51–60.
- Geschwind, N. 1979: Specializations of the human brain. *Scientific American*, 238 (3), 158–68.
- Goldman-Rakic, P. S. 1987: Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In J. M. Brookhart, V. B. Mountcastle, and S. R. Geiger (eds), *Handbook of Physiology: The Nervous System*, vol. 5, Bethesda, MD: American Physiological Society, 373–417.
- Gregory, R. L. 1968: Models and the localization of functions in the central nervous system. In C. R. Evans and A. D. J. Robertson (eds), *Key Papers in Cybernetics*, London: Butterworth, 91–102.
- Hacking, I. 1983: *Representing and Intervening*. Cambridge, UK: Cambridge University Press.
- Hanson, N. R. 1958: *Patterns of Discovery*. Cambridge, UK: Cambridge University Press.
- Hartline, H. K. 1938: The response of single optic nerve fibers of the vertebrate retina. *American Journal of Physiology*, 113, 59–60.
- Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B., and Rapoport, S. I. 1991: Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences, USA*, 88, 1621–65.
- Hess, W. R. 1949: The central control of the activity of internal organs, *Nobel Lectures, Physiology or Medicine 1942–1962*. New York: Elsevier.
- Hinton, G. E., and Shallice, T. 1991: Lesioning a connectionist network: Investigations of acquired dyslexia. *Psychological Review*, 98, 74–95.
- Kuffler, S. W. 1953: Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16, 37–68.
- Kuhn, T. S. 1962/1970: *The Structure of Scientific Revolutions*, 2nd edn. Chicago: University of Chicago Press.
- Kutas, M., and Dale, A. 1997: Electrical and magnetic readings of mental function. In M. D. Rugg (ed.), *Cognitive Neuroscience*, Cambridge, MA: MIT Press, 197–242.
- Latour, B. 1987: *Science in Action*. Cambridge, MA: Harvard University Press.
- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., and Pitts, W. H. 1959: What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineers*, 47, 1940–51.
- Li, C.-H., McLennan, H., and Jasper, H. H. 1952: Brain waves and unit discharges in cerebral cortex. *Science*, 116, 656.
- Luck, S. J., and Ford, M. A. 1998: On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences (USA)*, 95, 825–30.
- MacKay, D. G. 1998: Stage theories refuted. In W. Bechtel and G. Graham (eds), *A Companion to Cognitive Science*, Oxford: Basil Blackwell, 671–8.
- Marr, D. C. 1982: *Vision: A Computation Investigation into the Human Representational System and Processing of Visual Information*. San Francisco: Freeman.

- Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W. 1990: Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Science, USA*, 87, 9868–72.
- Ogawa, S., Tank, D. W., Menon, R., Ellerman, J. M., Kim, S., Merkle, H., and Ugurbil, K. 1992: Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Science, USA*, 89, 5951–5.
- Olds, J. 1965: Pleasure centers in the brain. *Scientific American*, 195, 105–16.
- Penfield, W., and Rasmussen, T. 1950: *The Cerebral Cortex in Man: A Clinical Study of Localization of Function*. New York: Macmillan.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. 1988: Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331, 585–8.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. 1989: Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, 1 (2), 153–70.
- Plaut, D. C. 1995: Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 17, 291–321.
- Poeppl, D. 1996: A critical review of PET studies of phonological processing. *Brain and Language*, 55 (3), 317–51.
- Posner, M. I. 1978: *Chronometric Explorations of Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Raichle, M. E. 1998: Behind the scenes of functional brain imaging: A historical and physiological perspective. *Proceedings of the National Academy of Sciences*, 95, 765–72.
- Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, M.-A. K., Pardo, J. V., Fox, P. T., and Peterson, S. E. 1994: Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, 4, 8–26.
- Rosen, B. R., Buckner, R. L., and Dale, A. M. 1998: Event-related functional MRI: Past, present, and future. *Proceedings of the National Academy of Sciences*, 95, 773–80.
- Shallice, T. 1988: *From Neuropsychology to Mental Structure*. New York: Cambridge University Press.
- Sternberg, S. 1969: The discovery of processing stages: Extension of Donders' method. *Acta Psychologica*, 30, 276–315.
- Talairach, J., and Tournoux, P. 1988: *Co-planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers.
- Thach, W. T. 1998: What is the role of the cerebellum in motor learning and cognition. *Trends in Cognitive Science*, 2 (9), 331–7.
- Tootell, R. B., Silverman, M. S., Switkes, E., and DeValois, R. L. 1982: Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218, 902–4.
- Uttal, W. R. 1978: *The Psychobiology of Mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Valenstein, E. 1973: *Brain Control: A Critical Examination of Brain Stimulation and Psychosurgery*. New York: John Wiley and Sons.
- van Orden, G. C., and Paap, K. R. 1997: Functional neural images fail to discover the pieces of the mind in the parts of the brain. *Philosophy of Science*, 64 (4), S85.
- Van Orden, G. C., Pennington, B. F., and Stone, G. O. in preparation: What do double dissociations prove? Inductive methods and isolable systems. Submitted for Publication.
- Walsh, V., and Cowey, A. 1998: Magnetic stimulation studies of visual cognition. *Trends in Cognitive Sciences*, 2, 103–10.

- Woolsey, C. 1960: Organization of the cortical auditory system: a review and a synthesis. In *Neural Mechanisms of the Auditory and Vestibular Systems*, eds. G. L. Rasmussen and W. F. Windle, pp. 165–80. Springfield, Ill.: Charles C. Thomas.
- Zeki, S. M., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C., and Frackowiak, R. S. J. 1991: A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, 11, 641–9.