# 1 Introduction

## 1.1 Some Background

In the past decade and a half important new developments in instrumentation capable of studying the functioning brain have appeared. These devices, most notably positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) but also now including such exotic tools as magnetoencephalography, now unquestionably make it possible to study the anatomy and physiology of the brain (among other organ systems) better than ever before. There is no question that brain imaging devices represent one of the most important diagnostic and scientific developments of all time. Human suffering has been alleviated to a degree because of these devices in a way that is comparable only to the introduction of anesthesia or the purification of public water supplies. MRI machines are capable of tracking brain transmitter distribution, spotting potential weaknesses in circulation, defining the tracts connecting distant regions, and, to an as yet unknown degree, determining regions of heightened metabolic activity that may be associated with cognitive activity.

Despite this abundant progress, it must be clearly understood that anatomic and physiological images are not direct measurements or indicators of cognitive processes. Indeed, their meaning as correlations of our mental life is open to a wide variety of disputes, empirical inconsistencies, and internal uncertainties.

Nevertheless many researchers in the field of cognitive neuroscience argue that brain images can be used to study the neural foundations of our mental activities in a way that had hitherto been beyond the hopes of even the most imaginative researchers. Supplanting the older techniques of the electroencephalograph (EEG) and the event-related potential (ERP), these new techniques promised to provide a means of studying the function of the brain as it carries out its adaptive cognitive processes. However, many of us are beginning to believe that it is a promise yet to be fulfilled. In this book I critically examine just what has happened and what we have learned from the astonishingly large corpus of published experiments in which brain images are compared to cognitive processes.

The argument that brain imaging techniques will revolutionize cognitive science is based on the idea that they are direct measurements of salient brain activity during controlled cognitive activities. Many applications of brain imaging have been proposed, some of which are thought to offer alternative, if not better, means of measuring cognitive states and processes than those provided by the traditional psychological methods. Many of these suggestions thus promise what are considered to be objective measures of what had traditionally been limited to subjective measures of mental states.

However, there are others who believe that there are substantial unresolved problems with and limitations of this approach that suggest that some of the initial expectations may be unachievable not only in the short run but also in the distant future. The problems arise in many guises and include some that are conceptual, some that are technical, and some that are empirical. This gradually emerging awareness has led to a somewhat belated critical examination of the plausibility of the assumption that brain imaging techniques will permit us to "read the mind" of a human being or even to find adequately correlated biological markers for particular cognitive states.

The current book considers the role that brain imaging has made or might make to cognitive neuroscience. It is a new embodiment of what had hitherto been known as physiological psychology. The complexity and variability of human behavior have made progress in this field difficult to evaluate and even to conceptualize. Clearly, any novel method of evaluating, predicting, and controlling behavior would be of extreme interest—if it could be shown that these methods work or are likely to work in the future. This is the crux of the problem faced by modern cognitive neuroscience—what is the likelihood that brain imaging techniques will be able to bring added value to the existing behavioral science research? How deep should be our science's commitment to techniques that many researchers believe are deeply flawed and, despite their popularity, are neither theoretically nor empirically seminal nor even, in some cases, possible?

An important goal of this book, therefore, is critically to evaluate the extent to which brain imaging and other recording techniques have informed scientific psychology. This is not just an empirical problem; there is a profound theoretical question lurking in the background—what is the likelihood that we will be able to add to the fundamental theory of the mind-brain question using these powerful methods? In other words, do these new technologies offer us an expedited pathway to the great question of how the brain makes the mind?

Although the ultimate answer to the possibility of supplementing, if not substituting, brain research for behavioral research is going to be primarily empirical (will it work?), it must also be appreciated that there are major philosophical and logical issues raised whenever one has the audacity to compare mental and neural activities. There is no denying that this is a task of universal interest and monumental implications, but at least a few scholars agree that at present there has been limited conceptual progress despite the great diversity and number of empirical studies. We can no more ignore some of the imponderable foundation philosophical concepts that are involved than we can the limits of the technology. Given the current state of our knowledge, it may indeed be that our choice of the "correct" level of analysis, whether it is behavioral or neural, may be instrumental in interpreting the meaning as well as the applicability of what are complex and indirect experiments. Clearly, this is a problem of extreme complexity, and resolving it will be instrumental in the future development of scientific psychology as well as neuroscience.

The primary question—how does the brain make the mind?—cannot be studied in isolation. Cognitive neuroscience may have a short history, but it evolved not only from a century or so of physiological psychology but also from a longterm concern with the basic question. Therefore, other tasks in this book will be to review and evaluate the history of the cognitive neurosciences that preceded the invention of the imaging devices. The observation that much of this earlier work is also deeply flawed adds some depth of understanding to why modern imaging techniques have so far failed to achieve some of their most extravagant claims.

## 1.2 The Great Question—The World Knot

The greatest scientific question of all time, the one to which most human attention has been directed over the millennia, is—how are our minds and our brains related? The profundity of the question has led to its being referred to by Arthur Schopenhauer (1788–1860) as the "world knot."

Although there is considerable debate about the reality of the mind (see for example the article by Schlinger, 2005), to deny its reality or to declare it merely epiphenomenal would be to make human existence meaningless. Furthermore, there is at least one piece of solid evidence that the mental processes are real. That singular piece of evidence is that each of us is endowed with a personal awareness, a process that has come under many names. Whatever the term used—mind, mentality, soul, ego, self, intellect, consciousness, awareness, sentience, psyche, or cognition—we all have first-hand knowledge of what it is that we are talking about when we use any one of these words. There is no way that we could deny the reality of the mind because proof positive exists within each of us—our own sentience. We could not do so without destroying the meaningfulness of our ability to converse and interact at many different social levels.

However, as much as I am convinced that my mind exists, I have long ago resigned myself to the fact that defining the mind is an unachievable goal. These days I look upon it as a process of the brain analogous to rotations being a feature of material devices called wheels—albeit infinitely more complicated. In other words mind is neither nothing more nor nothing less than a function of the material brain.

Despite this near universal appreciation of the reality of mental activities, the description of the mind and the explanation of its neural origins have proven to be extremely difficult challenges for the science that has grown over the centuries to study them. That science is psychology, not the psychology of the therapeutic couch or inferred, but inaccessible, cognitive entities but the hard science of observable behaviors. If the interests of psychological science are combined with neurophysiological science, we refer to this science as physiological psychology, psychobiology, or, most recently, cognitive neuroscience. A major question that is implicit in this discussion is—can (or should) psychology exist without its neural co-studies? More precisely, what do psychology and neuroscience offer each other?

Psychology, confronted as it is by enormous obstructions and difficulties in constructing explanations of mental phenomena, has fractionated into a number of subsciences that have taken many different and often idiosyncratic directions over the years. Schools of thought have proliferated, and over time, strategic collaborations with other sciences have repeatedly formed.

At the root of cognitive psychology, however, has been the assumption that the nature of the mind (or its effects) can be studied experimentally. A further special assumption of modern cognitive neuroscience is that we will also be able to determine the neural conditions that lead to the mind. On the basis of this premise it is argued that, in principle, we should be able to understand the neural mechanisms that account for mental and behavioral activities. It is not yet clear whether or not this goal can or will be achieved; only time will tell. What we can discern now are the intellectual and philosophical roots that underlie the neuroreductionist goal of explaining mind in terms of the brain and the many obstacles that prevent us from achieving that goal.

The most fundamental root of all of these questions lies not in the laboratory but in speculative ontology—a major division of metaphysical philosophy. Ontology is that branch of metaphysics that deals with the philosophy of reality, of the nature of existence itself. The ontology of cognitive neuroscience is especially complex for two reasons: first, we have no direct access to or empirical evidence of the mind (Uttal, 2007); we have only indirect evidence from which we must infer its nature and construct hypotheses concerning its function. Second, mental activity is not sufficiently constrained by behavioral observations so that a robust analysis can be made of it into modular elements: in other words, all of our cognitivist-reductionist theories of mind are underdetermined.

Many questions for which we have no current answers, therefore, lay solely within the confines of the speculative philosophy that we call ontology. In the place of specific empirical answers to some of these most profound questions, philosophers have over the centuries tried to establish certain beliefs about the nature of reality that are based on whatever relevant knowledge is available and rational and logical arguments and derivations that may make these beliefs plausible, even if pure speculation cannot confirm them.

In cognitive neuroscience there is a major ontological assumption that, however controversial, guides the day-by-day activities of laboratory researchers as well as those who conjure up new theories of the relation between the mind and the brain. That basic assumption is that, however inexplicable it may be at the moment, the brain makes the mind. Although we do not know how, it is widely accepted that a complete neural explanation is, in principle, possible. Those who labor in the laboratory rarely make this monistic assumption explicit, and yet few cognitive neuroscientists would challenge this fundamental idea.[1]

Nevertheless, the assumption of mind-brain equivalence is without any compelling empirical foundation; none of the required tests of necessity and sufficiency have ever been carried out to confirm it generally or specifically. However likely it may seem, there is no evidence other than plausibility and reason to support this foundation assumption.

This profound foundation assumption comes in two parts (box 1.1). The first part is a general hypothesis, implicitly honored by all cognitive neuroscientists. It asserts that any mental or cognitive activities and processes as well as all of those that control behavior are the functions, the outcomes, or the results of the activities of the nervous system. Herein is the foundation assumption of what ontologists would refer to as monism or physicalism or mind-brain neuroreductionism.

Only those who believe in some kind of dualism would deny this part of the basic ontological postulate. (See Uttal, 2004, for a more complete discussion of the impact of dualistic thinking throughout history on theology, philosophy, and psychology.) This assumption links the worlds of the mind and the nervous system into a single inseparable reality; one part is structure, and the other is function. We can no more conceptually separate the two than we can separate the circular motion of a wheel from the wheel itself. This does not mean, however, that the two sciences—psychology and neuroscience—are inseparable empirically. Despite the ontological, in principle, inseparability, practical considerations (e.g., complexity) may keep these two scientific paths separate. Examining this issue is also a part of the challenge faced in this book.

The essential point of the first part of the basic ontological postulate is that the function cannot exist without some kind of equivalent physical structure. Our minds are products of our nervous system, and any idea of the consciousness or mind existing after the deterioration of the brain is without merit. Indeed, without this kind of mind-brain[3] monism the whole cognitive neuroscience enterprise would be meaningless and pointless; we could never be sure that our studies were not contaminated by other forces that were totally out of our control and totally unaccounted for in our experimental protocols.

Beyond the general mind-brain, monistic postulate just described lies the second part—one that is much more specific. It is the hypothesis that our minds are not just functions

**Box 1.1**
The Two Parts of the Basic Ontological Postulate

1. All mental processes are the outcome of neural activity.
2. All mental processes are the outcome of the microscopic interactions and actions of the great neuronal networks of the brain.[2] This is the proper level of analysis of the mind-brain problem.

of our material nervous system (the first part) but that they are the specific result of the cumulative integration and interaction of complex and innumerable neuronal activities that go on in the brain as opposed to other levels of neural activity.

It is this complex and intricate pattern of neuronal activity and interactions that cognitive neuroscientists assert becomes or *is* mind; it is in the complex network of neurons that memories are stored, that decisions are made, that personalities are forged, and that behavior is controlled. It is there that the physiological actions are transmuted in some mysterious way into all of the many kinds of mental states, processes, feelings, and faculties that grace human existence. The mind, according to this postulate, arises out of the complex interactions of billions of component parts in ways that we do not now know and possibly may never to be able to know.[4] The relation between the brain and the mind, cognitive neuroscientists agree, is something akin to the Sherrington's (1940/1963) "enchanted loom":

> The brain is waking and with it the mind is returning. It is as if the Milky Way entered upon some cosmic dance. Swiftly the head-mass becomes an enchanted loom where millions of flashing shuttles weave a dissolving pattern, always a meaningful pattern, though never an abiding one; a shifting harmony of sub-patterns. (p. 178)

This is beautiful poetry but hardly a rigorous scientific finding; it is simply a vague metaphor for the point being made by the second part of the basic ontological postulate.

This piece of poetry by Sherrington aside for the moment, the general principle expressed in the second postulate is widely held by contemporary psychologists and neuroscientists. The modern version of the idea was probably first expressed by McCulloch and Pitts (1943) and Pitts and McCulloch (1947) in their pioneering work on the logic of networks and then in a follow-up on form recognition by such networks. However, the first specifically neuroscientific expression of the second postulate was published by Hebb (1949). In it he suggested specific patterns of neural interaction as the basis of cognitive activities. His theoretical neurophysiology was based on his elaboration of what had originally been a psychological principle—Thorndike's (1931) "Law of Effect."[5] This purely psychological observation was that repeated practice led to enhanced behavioral strength. Hebb argued that this law must also have a neural equivalent and in 1949 presciently formulated the following neural equivalent of it:

> When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes takes [*sic*] place in one or both cells such that A's efficiency, as one of the cells firing B, is increased. (Hebb, 1949, p. 62)

Hebb then went on to suggest that the "growth process" was the increased growth of synapses as they were exercised. This idea—that it is the change in synaptic conductivity that accounts for the changes in the neural network—is the basis of most physiological theories of learning and memory. Synaptic conductivity changes can account for short-term memory by invoking reverberating circuits that fade as the temporary synaptic changes lose the transient "potentiation." Long-term memories are accounted for by permanent changes in conductivity so that the information in the synaptic patterns becomes locked in. Martin, Grimwood, and Morris (2000) present a compelling argument that such synaptic changes are necessary but that "little data currently support the notion of sufficiency" (p. 649).

Hebb, nevertheless, made some remarkable speculative leaps from this foundation idea of neural networks. He postulated the existence of "cell assemblies"—a "diffuse structure" of neurons in the brain that was created as a result of use and one that could encode complex responses. To this construct he added the notion of the "phase sequence"—a series of cell assemblies that actually was the level at which "thoughts" became extant. It is important to reiterate that the empirical evidence for these hypothetical neuronal net structures is as nonexistent now as it was then.

Many contemporary scholars also followed in Hebb's footprints. In a recent debate in a popular magazine Koch and Greenfield (2007)[6] argued from two opposed speculative points of view in attempting to answer the question—"How does consciousness happen?" Their answers were almost the same but differ in one main way; Koch argues that consciousness occurs when a specific set of neurons in a specific part of the brain fires in a specific manner. Greenfield argues that the brain produces consciousness when neurons in all parts of "the brain are synchronized into coordinated assemblies, and then disband."

Both hypotheses share a common principle, what I have referred to as the second part of the basic ontological postulate—the plausible, but unsubstantiated, idea that it is the arrangement of the great neuronal networks in the brain that accounts for consciousness, their term for the mind.

The distinctive anatomical attribute that distinguishes between their two theories, on the other hand, is the degree to which the neural network of the brain producing consciousness is localized or distributed. However, they do not differ with regard to the level of analysis; both assume that it is based on the detailed state and interactions of the neuronal network of the brain.

Despite the disclaimer in Koch and Greenfield's joint paper that neither one is "is attempting to explain how consciousness arises" (p. 83), in fact both are actually operating at pretty much the same level at which Hebb was at the beginning of the neural network days. Both are proposing ingenious, but nonspecific and untestable, hypotheses that closely conform to the ontological postulates presented earlier. Both make the same foundation assumptions, and both suggest ways that such assumptions might be implemented. However, neither makes any specific statements about the details of how the neural networks produce the mind. Their speculative contributions are plausible and reasonable; however, they are without any empirical support. Koch is very explicit about this in their joint article when he says:

> Neuroscience does not yet understand enough about the brain's inner workings to spell out exactly how consciousness arises from the electrical and chemical activity of neurons. Thus, the first big step is to determine the best neural correlates of consciousness (NCC)—the brain activity that matches up with specific conscious experiences. (pp. 76–77)

In this manner he retreats back to confront the traditional problems faced by all cognitive scientists. First, all of the problems faced by correlation methods are once again brought to our attention; second, the brain measures—the NCCs—to which he alludes are generally drawn from irrelevant levels of analysis such as brain images, EEGs, and other cumulative methods; and,third there is no direct access to the conscious experiences that permits us to directly compare mental and neural events. The details of the neural networks, however gracefully and eloquently expressed, are totally finessed. Indeed, with considerable justification, we may conclude that Koch and Greenfield's ideas are, perhaps, less specific than were those of Hebb!

It is in this context that the greatest misunderstanding of the current brain imaging approach becomes crystal clear. In a recent paper (Posner & Rothbart, 2007) extolling the importance of the Hebbian tradition dealing with neuronal networks, it was suggested that brain imaging "also probes neural networks that underlie all aspects of human thought, feeling and behavior" (p. 5). Unfortunately, there is a disconnect here between the Hebb *neuronal* network model and the *neural* network of brain regions at which brain imaging techniques operate: the Hebbian network is a network of *microscopic* neurons. The networks that are studied with, for example, an fMRI system, are combinations of *macroscopic* brain regions that actually tell us very little about how the brain makes the mind. Nor do brain images tell us anything about the details of the network of neurons. In fact all of the salient details of their function are lost by the processes of accumulation and summation that characterize such techniques as fMRIs and EEGs. It is entirely possible for two totally different neuronal network states to produce the same fMRI response. Thus, there is no functional relation between an fMRI image and the activity of the critical and essential network of neurons that Hebb so presciently pointed out must be the psychoneural equivalent of a cognitive process. In short, the macroscopic neural networks studied with fMRI systems are not the same as the microscopic neuronal networks that cannot (because of their complexity) be studied at all. To link them together is nothing other than a neuroscientific pun.

It should be clear now that the second part of the ontology postulate is widely, but not universally, held among cognitive neuroscientists. For reasons that have more to do with available technology than with either philosophy, logic, or empirical findings, many other theorists place the essential transformation process between brain and mind at other levels of activity such as single neurons, wavelike fields of activity, or activated chunks of the brain. These alternative hypotheses must be appreciated to be temporary surrogates for the impenetrable neuronal net hypothesis. It is the information processing by highly complex microscopic neural networks, rather than any of these alternative measures cum theories, that is the core of the foundation premise of modern cognitive neuroscience.

The main difficulty that makes the neuronal network hypothesis into a postulate (as opposed to a robust empirically observed fact) is that the combinatorial complexity of the brain's neuronal network is so extreme that it cannot be studied directly. The true psychoneural identity level of activity—the idiosyncratic and exceedingly complex interaction of so many neurons—at which brain becomes mind is computationally intractable. As a result, alternative theoretical approaches such as single-cell or field measurements are called into play simply because they are measurable with available measuring instruments. Ethologists refer to this kind of behavior as displacement activity—one does what one can when one cannot do what one should do!

Complexity being what it is, this is not necessarily a criticism of the state of our science. It is no more a problem than the relativistic limit on the speed of light or the second law of thermodynamics' prohibition of perpetual motion machines. Complexity is becoming better understood and increasingly appreciated as being supremely frustrating to the kind of scientific analysis to which cognitive neuroscientists aspire. However, it is important that we realize the implications of the inaccessibility and noncomputability of the neuronal network. These constraints have profound implications for theory and practice in cognitive neuroscience, the way it has developed, and how it will continue to do so in the future.

However widespread is the acceptance among cognitive neuroscientists of this second part of the ontological postulate—the mind is an emergent factor from the interactions among the vast number of neurons[7] that make up the brain—it must also be reiterated that there is no proof of it, and it has to be considered as an unprovable assumption rather than a provable fact. There is no empirical proof in which necessity and sufficiency of the network have been proven; nor is there any computer simulation that exhibits any of the properties of consciousness.[8] Although we cannot prove the second part of the ontological postulate, there is no plausible alternative explanation available at the present time, only details of how neurons might interact at local levels or speculations about the overall nature of the network in the style proposed by Koch and Greenfield (2007). This is why it is presented here as an ontological postulate or presumption, plausible and reasonable, but not proven and probably not provable.

This second part of the basic ontological postulate is critical (along with the practical limits of what we can do) in determining not only the nature of our theories but our day-to-day activities in the laboratory. The elusiveness of empirical answers to the question of the essential level of analysis is the basic reason that the mind-brain problem (how does the brain make the mind?) remains unanswered and why there is such an abundance of questionable theoretical speculation and flawed empirical research in this field.

Unfortunately, the presumed level of brain activity (the interactions among a vast number of neurons) at which we believe the salient information processes are carried out that become sentience, consciousness, and mental activities of all kinds is exactly the level at which our research techniques are least adequate; the most fundamental reason, as noted, being the extraordinary complexity and numerousness of the involved neurons and the idiosyncratic nature of every neuron-to-neuron interaction.

As a result, neuroscientists have turned to other techniques to provide grist for their theoretical mill. All of the most frequently used methods epitomized by the fMRI or the

EEG, however, share a common difficulty—they pool the responses from the many neurons that make up the brain into cumulative, global measures. In doing so, they lose all of the critical information about the neuron-to-neuron interactions that make up the crucial activities of the brain. All, therefore, are not operating at the level of analysis at which the predominant current opinion suggests is the one at which mind and brain activity are most closely associated—the details of the interactions between myriad individual neurons of the brain. It is a practical problem—there are just too many of them—that restricts this approach.

Another main technological alternative is the use of microelectrodes to study the action of individual neurons. This method has been a powerful tool in helping us understand the nature of the components of the neural networks of the brain (the individual neurons) but from the very narrow perspective of the one-micron-wide point of a microelectrode. However microelectrodes offer little information about the interactive organization of the great numbers of neurons that are involved in even the simplest thought. Indeed, the ability to record from a single electrode has driven a major theoretical tradition based on the idea that single neurons can encode complex cognitive processes. This theoretical hypothesis seems also to be based on a flimsy empirical foundation.

In summary, these two ideas—the general first part of the ontological postulate stating that the mind is a function of the brain and the second more specific part that it is the detailed pattern of neuronal interactions that represents or encodes mental activities and processes—with all of their uncertainties seem to be our best current answers to the mind-brain problem.

It should not be inferred that these two postulates are merely topics for philosophers to mull: they exert an enormous influence on the development of theory and the choice of experimental protocols. By so specifying the relevant level of analysis, we can see that two corollaries immediately emerge. First, assigning mind to very complex neural interactions suggests that almost all of the work that has been done and can be done in the future using brain images is aimed at the wrong level of analysis. Because the old phrenological idea of localized cognitive process encoding modular cognitive processes also seems to be on its last legs, it can be expected that attempts to correlate cognition[9] with brain images will also diminish just as the enthusiasm for the EEG as an entrée into the mind has moderated over the years. Rather than these cumulative, pooled, and integrated signals, we should be attending, if we could, to the detailed patterns of activity of a myriad of neuronal interactions.[10]

Second, the computational and combinatorial aspects of the neuronal net hypothesis suggest that the problem as posed by the ontological postulates is intractable. The best that can be hoped for is that there will be some neural correlates of cognitive processes observed with these integrated global measures that may serve as useful biological markers in certain restricted situations (box 1.2).

To summarize, the main point made here is that a priori no macroscopic brain imaging or electrical recording activity, no matter how direct it may seem to be in recording the activity of the brain, can *in principle* provide solutions to the mind-brain problem. The basic

**Box 1.2**
Two Corollaries of the Basic Ontological Postulate

1. Brain imaging techniques are formulated at the wrong level of analysis and thus cannot provide answers to the mind-brain problem.
2. The neuronal network approach is computationally intractable and thus cannot provide answers to the mind-brain problem.

reason for this conceptual barrier is that the current technology used by these methods is operating at the wrong level of analysis. Whereas brain imaging produces cumulative signals, the essence of the mind is more likely to be found in the microdetails of idiosyncratic neuronal interactions. This clash is between the innumerable states of a discrete network and a single cumulative state in which all of these microscopic activities have been pooled at the cost of great information loss.

In the section that follows, I show how these basic postulates can have a far-reaching effect on our thinking about the relation between the mind and the brain.

## 1.3 Implications of the Ontological Postulate

This section is concerned with the implications of the two parts of the ontological postulate; first the brain is the organ of mind, and, second, the level of analysis at which that equivalence is manifested is to be found in the details of the great network of interacting neurons. Of the first part, little more need be said. One is either a monistic physical materialist or one is a nonmaterialist dualist.[11] If one denies the idea expressed in this part of the postulate, then the whole enterprise of cognitive psychology is a meaningless and uncontrollable mess. For purely theoretical reasons, the second part of the ontological postulate should be the focus of the entire cognitive neuroscience enterprise. That it is not is due to the practical problems dealing with its great complexity.

### 1.3.1 Implication for Philosophy

Dualism comes in many guises: one can resort to theology or such traditional philosophical concepts as Descartes's substance dualism, Geulincx's and Malebranche's occasionalism, or Leibniz's parallelism, on the one hand, or turn to more modern ideas such as Eccles's tripartite reality, Chalmers's naturalistic, or Kripke's versions of dualism, as well as certain interpretations of Davidson's supervenience.[12]

The adoption of any of these dualistic stances by some philosophers, although this point is likely to be disputed, is inconsistent with the whole motivation behind cognitive neuroscience. It seems completely illogical to expect that one could carry out experiments attempting to study the mind-brain relation guided by the presumption that they represented two different kinds of reality. To suggest that mind and brain, in fact, are not causally or otherwise intimately related to the degree of identity or equivalence would invalidate the very essence of cognitive neuroscience.

This then brings us to the second part of the ontological postulates. Having accepted the proposition that the mind and the brain are two parts of the same basic reality, what more can be said about the specific nature of that relation? The answer to this query, the second part of the ontological postulate asserts, is that they are related in the way that a mechanism and its function are related. In this particular case the ontological assertion is that the salient mechanism is the intricate interconnection pattern of the myriad of neurons that make up the great networks of the brain. This level of analysis is complex and is probably beyond analysis and specific explanation for the reasons I have already mentioned. Indeed, it may represent an intractable problem that neither new measuring devices nor computational engines can ever begin to unravel. There are too many uncertainties, too many neurons, too many idiosyncratic interconnections (e.g., the brain is not neatly organized as is a simple crystalline structure) for us to ever be able to understand its detailed organization and how, specifically, this complex information pattern produces the reality we call mind.

This approach has profound implications for understanding what cognitive neuroscience has achieved and what still remains mysterious and unknown. If we are to accept the second part of the ontological postulate, then almost all of the other approaches to studying the relation between the mind and the brain are being carried out at the wrong level of analysis.

For compelling practical and historical reasons neuroscientists have turned to other more global measures such as the EEG, the ERP, and most recently the brain imaging procedures such as fMRI and PET. All of these methods, however, share a common difficulty—they pool the responses from the many neurons that make up the brain into cumulative, global measures. In doing so, they lose all of the critical information that makes up the salient activities of the brain. All, therefore, are not operating at the level of analysis at which the predominant current opinion suggests is the one at which mind and brain activity are most closely associated.

If this analysis is correct, then all of the work using the molar, integrated, cumulative measures of brain activity is misdirected, and the resulting findings must be considered to be irrelevant in the search for solutions to the mind-brain problem. This is as serious an impediment to scientific understanding as acceptance of dualism would be.

There is implicit in these comments another important and essential point—the fact that it is the pattern of information transactions, not the biochemistry of individual neurons, that accounts for the emergence of consciousness or mind. For example, although we know a lot about the biochemistry of the neuron and of the synapse and can explain the details of the transfer of information along an axon or from cell to cell, these are properties of the microscopic components of the nervous system; the particular technology is not essential to an information-processing system's function. By themselves, therefore, these properties tell us little about the mind-brain problem—only about the details of the particular biochemical mechanisms of the components of which the brain is made. Just as a material from which a transistor is made tells us nothing about the program that is being run on a computer, even the most detailed knowledge of the biochemistry of a neuron tells us nothing about how the overall system is representing mental processes. In some unlikely ideal world, should we be able to build a brain-like structure capable of mimicking all of the complex informational processes and interactions, albeit with a completely different technology than the sodium-potassium-chloride chemistry our brains use, such a system would presumably be able to "think" or "perceive" as well as we do and may even be "conscious." The point is that it does not matter what component technological units are being used; only the arrangement and interactions of those elements are of consequence in representing "mind."

This, then brings us to the next step in this preliminary philosophical study of the nature of mind-brain reality—that of the epistemology of cognitive neuroscience. That is, how can we know (i.e., what strategies can we use to learn about) the critical mind-brain interactions. Epistemologists have different goals and consider different topics then do ontologists. Rather than contemplating what is, they are concerned with the limits on our ability to know what is; that is, given postulates such as those proposed by the ontologists, what does this mean to our ability to study the mind-brain problem (among many others) and to understand, to learn, to describe, or to explain the properties of the mind and the brain. In other words, what is it logically possible for psychologists to do given the ontological postulates as starting points?

Few practicing experimental psychologists or cognitive neuroscientists struggle with such questions. They go about their various projects questioning only how data can be gathered or explained. Their epistemology is an empirical one; decisions are made on the basis of what works or what appears to work—their epistemology is an unredeemably practical one. Nevertheless, they, too, are bound by the epistemological constraints. More or less implicitly, without overt awareness, all cognitive neuroscientists and psychologists constantly make certain practical assumptions within this context of unspoken and implicit epistemological conundrums about what they can learn from their experiments.

In the following pages I distinguish between those epistemological issues that guide psychology and those that guide the neuroreductionist efforts of cognitive neuroscientists.

### 1.3.2 Implications for Psychology

To understand how the two ontological postulates affect the course of scientific psychology, it is important to express a major epistemological principle (box 1.3). Before dealing with the impact of this postulate on psychology, I must consider a very special idea—accessibility and its antithesis, inaccessibility. The issue being dealt with here is how much access do we have to mental processes? Can we gain access to (i.e., measure) the nature of mental

**Box 1.3**
The First Epistemological Postulate for Psychology

> Mental processes are private and are not accessible to any form of measurement, either experimental or introspective.

processes, or are we forever constrained from any direct measurement of the mind? Let me now spell out in detail some of the arguments for both of the two interpretations.

### *Arguments for Inaccessibility*

The argument for the inaccessibility of the mind revolves around the fact that the mind, by definition, is a private personal experience. Since there is no detailed explanation (beyond the second part of the ontological postulate) of the relation between the physical structure and processes of the brain and our individual and private experience of being, sentience, or awareness, there is no instrument that we can attach to or scan the head that will directly measure mental processes. As many psychologists have written previously, mind is an *intrapersonal* and not an *interpersonal* phenomenon. The only possible way that we can begin to get any, however defective, insight into the mind of another person is by means of that individual's introspective reports or by observing that person's behavior.

However, both introspection and drawing inferences from behavioral observation are well known to be deeply flawed methodologies. Researchers such as Nisbett and Wilson (1977) had shown four decades ago that people are not aware of their own logical processes and mental strategies. The reasons for this introspective blindness are manifold but include false memory construction (Loftus, 1996) and the automatic (i.e., unconscious or preconscious) nature of many behavioral processes (Bargh, 1997). Whatever the reasons, and there are many others, it is clear that people cannot always accurately report the logic or reasons they used to arrive at a decision. For some phenomena, for example those associated with the basic qualitative state of such experiences as color or pitch, it is not possible to reconstruct with words what it is that a person is experiencing. Introspection, therefore, must be ruled out as an effective means of accessing mental states.

The main alternative means of accessing mental states is to draw indirect inferences from publically observable behavior. However this strategy, too, is deeply flawed. The reasons behind this assertion are equally numerous and include these:

• There is a well-known engineering principle (that holds for human behavior as well as pieces of electronic equipment) that the inner workings of an unopened "black box" cannot be determined by comparing the box's input to its output. The relation between input and output cannot in principle tell us anything definitive about the functional changes that may be occurring in the box. This is well known to engineers and should be to psychologists.

• The reason for this generalization is that the mechanism inside either a piece of hardware or the human mind is underdetermined by the behavioral observation. That is, there is not enough information in behavior to precisely and uniquely determine inner mechanisms. There are many possible (and far too many plausible) explanations for each behavioral observation. No convergence of observations can lead to an answer to the problem of internal structure. Indeed, additional observation often leads to even more plausible and possible explanations than had been contemplated originally.
• Human behavior is characterized by enormous variability when compared to other sciences. Individual results are not sufficiently repeatable. Therefore, there remains a serious question concerning the reliability as well as the validity of many psychological measurements.
• Human behavior is not bound by robust, stable, universal laws of time, space, and number in the way physical phenomena are. There is, therefore, no way that an entirely external observation can be associated with an inaccessible experience. Physics can generally do this, but its success depends on the assumption that the laws of physics are the same everywhere—in the internal microscopic world as well as distant macroscopic universes.[13]
• Behavior, as expressed in the literature of experimental psychology is not adequately linked to the associated mental activities. People can intentionally or unintentionally display behavior that is quite contrary to what they are really thinking. Questionnaires, stage plays, and the courtroom all present examples that illustrate how separated one's thoughts can be from one's utterances. Even the best experimental protocols do not provide robust constraints or necessarily even plausible links between behavior and the underlying thoughts.
• Mental faculties and components are "hypothetical constructs" (MacCorquodale & Meehl, 1948) created by psychologists to describe behavior and, therefore, may not exist in some physical or psychobiological sense.
• Finally, the complexities of both behavior and the neuronal mechanism are so great that there is no computational way in which they can be linked. There is no one-to-one correspondence among measurable behavior, mental activity, and brain responses.

### *Arguments for Accessibility*

The honorable epistemological opposition argues that these arguments are too stringent and demanding and that much is lost when we limit ourselves by assuming mental inaccessibility. Although the words may differ in the many arguments, the ubiquitous core argument for accessibility is that consciousness exists, and without assuming accessibility we would be denied any hope of measuring and explaining it. By denying accessibility, they argue, we lose one of the main raisons d'être of psychological science as well as basic matters of our own humanity. In any event proponents of accessibility argue that the links between mind and behavior are solid enough for us to draw good inferences.

Herein lies the core of the greatest debate in psychology—the one between behaviorism and cognitive mentalism. It is here that the empirical and theoretical strategies of these two great schools of psychological thought diverge. My argument in the present context is that they so diverge primarily because of their differing stances on the epistemological question—are the process and activities of the mind accessible?

The issue has been debated for many years, and the opposing answers to it are based on beliefs and assumptions more comparable to ontological and epistemological speculation than on any empirical evidence. On one side of the debate are those who traditionally have been called mentalists and more recently cognitive mentalists. Mentalism is based on an initial epistemological assumption asserting that mind is sufficiently directly accessible to be studied by introspective or experimental assay techniques. Behaviorists, in opposition, accept that the mind is not directly accessible and, therefore, that we can only observe and measure the final outcome of mental activity—behavior. From that point the respective strategies of the two approaches to the study of the problems of interest to psychology are set in a conceptual concrete.

The arguments for and against behaviorism and mentalism, respectively, can be summed up in the following brief lists abstracted from my earlier work (Uttal, 2000).

The Essential Arguments against Mentalism

1. There is a lack of public availability, objectivity, and repeatability for metaphysical or mental processes.
2. Mentalism leads to homunculus or infinite regression arguments.
3. Mentalism produces unprovable hypothetical constructs.
4. The empirical data argue against the accessibility of mental processes.
5. Mentalism requires complex experimental designs and unprovable assumptions that produce fragile data.
6. Mentalism arises because of the vested interests of its humanist, theological, and personal protagonists or from the professional needs of psychotherapists.

The Essential Arguments against Behaviorism

1. There is only a limited range of behaviorist psychology.
2. Behaviorism dehumanizes humans.
3. Behaviorism is too “mechanical” or is “not sufficiently mechanical.”
4. Behaviorism is not a step forward.
5. Behaviorism overemphasizes the environment and underemphasizes heredity as a source of behavior.
6. Behaviorism is nothing more than common sense.
7. Behaviorism is antidemocratic.
8. Behaviorism is antireligious.

Clearly, none of these arguments is compelling by itself. They all depend in large part on an original decision to accept or reject accessibility. Having said that, it is important not to try to finesse this issue, but, instead, to take the bull by the horns and make a value judgment. In my opinion most of the arguments against behaviorism and its attendant inaccessibility are wishes and hopes rather than scientific arguments. They dote on humanistic judgments about the desirability of understanding the human condition.

The arguments against mentalism, on the other hand, have a greater degree of scientific robustness and support (or are supported by) the idea of inaccessibility. Without any question, however, it is the acceptance or denial of the epistemological postulate of accessibility that directs and guides one to either behaviorism or mentalism. My opinion is that psychology would be better off scientifically being behaviorist rather then mentalist. I go so far as to make it an epistemological postulate asserting my preference for behaviorism (box 1.4).

There is another way in which these two postulates impact on thinking in psychology. Because, according to the second part of the ontological postulate and the first epistemological postulate, the brain level at which mind is embodied is that of the great and unanalyzable neural network and that mental processes are not directly accessible, there are few constraints on how we might assume the mind-brain to be structured. Coupled with the great complexity of the system, this means that we are relatively unfettered in making certain further assumptions concerning the nature of the organization of the mind-brain. Two of these assumptions stand out in the history of psychology—separability and analyzability. Because we cannot deal with the whole complex system with all of its interacting parts and variables at once, we fall back on Descartes's admonition to break the system into parts. This leads to two extremely potent, but highly questionable, governing assumptions. The first is that the mind is modular; that it is made up of quasi-independent units—the faculties and processes that are explored in conventional psychological experiments. The second is that the neural equivalents of these mental modules are located in particular parts of the brain.

Throughout the history of cognitive psychology and its predecessors, up to and including the early days of the brain imaging movement, experiments have been based on these two assumptions. The search was on—find the areas of the brain that were activated by such cognitive processes as "solve a problem," "decide which candidate you prefer," "think about a loved one," or "think about a cow." As the discussion in this book progresses it will become clear that modularization and localization are no longer tenable interpretations. In their place two alternative statements must be substituted. First, mental components cannot be

**Box 1.4**
The Second Epistemological Postulate for Psychology

> Psychology is better served by a behaviorist approach that dotes on the observable parameters of human activity rather than the inferences of a reductionist mentalism.

analyzed into independent and separable cognitive modules; to do so in a Cartesian sense belies the interactive complexity of our thoughts. It is likely that we will begin to realize the mental modules represent a convenient organizing principle but do not necessarily reflect the actual nature of our mental activities. Because compelling evidence for neither modularization nor holism is yet available, I also characterize this idea as an epistemological postulate (box 1.5).[14]

### 1.3.3 Implications for Neuroscience

Just as the choice of one's theoretical psychological stance depends on certain assumptions about the accessibility and modularity of the mind, the choice of one's neuroreductionist stance depends on the second part of the ontological postulate—that the instantiation of the mind is to be found in the actions and interactions of the many neurons of the great networks in the brain.

There is rapidly accumulating empirical evidence that the range of brain regions involved in even the simplest thought is widely distributed throughout the brain as summarized as The First Epistemological Postulate for Neuroscience (box 1.6). This assertion, however, is much less speculative and represents the first of the postulates driving neuroscientific research.

As these holist ideas (the mind must be treated more as a whole than as a system of separable modules, and the brain activities associated with a thought are widely distributed) have increasingly begun to percolate into experiment and theory, the epistemological situation has gotten much worse. A diffusely distributed system is not conceptually simpler than a system of discrete nodes. Furthermore, distribution complicates the search for an objective neural correlate of any behavioral activity. Since multiple regions are involved, distribution has led to the use of complex pattern recognition analysis methods that were far more challenging and the results of which were far less certain than those based on the

**Box 1.5**
The Third Epistemological Postulate for Psychology

> Although convenient as a means of experimental protocol simplification, mental processes are not modular and cannot be divided up into quasi-independent entities.

**Box 1.6**
The First Epistemological Postulate for Neuroscience

> Brain activity associated with mental activity is broadly distributed on and in the brain. The idea of phrenological localization must be rejected and replaced with a theory of broadly distributed neural systems accounting for our mental activity.[15]

simple modular and localizationist ideas of the past. Some investigators (e.g., Hilgetag, O'Neil, & Young, 1996) have suggested that the situation is actually much worse than just being "more complicated." They argued that the analysis of a heavily interconnected system of cooperating and interacting regions in the manner cognitive neuroscience may require might not be possible in fundamental principle. Instead, they argue that the situation would get increasingly complicated (not less so) as more and more experiments are carried out.

Brain images, it must be clearly understood, still provide us only with the capability to search for highly variable locales of activation that may be associated with vaguely defined cognitive processes. It is important to remember that no matter how complex the analysis, brain images essentially search only for answers to the "where" question. The essence of the mind-brain problem, however, is still the "how" question, and it is not yet clear just what the "where" question tells us about the mind-brain problem.

Just as it is necessary to invoke an additional epistemological postulate in order to understand the roots of psychology, it is also necessary to invoke another basic assumption to define one's approach to neuroscience. In the case of psychology, as I noted earlier, the additional epistemological assumption concerns the inaccessibility of the intrapersonal events we designate as mental. Depending on one's choice, it was a more or less logical progression from the respective assumptions of accessibility or inaccessibility to the kind of empirical and theoretical research to be pursued. If one accepted the intrinsic arguments for inaccessibility, the challenges to cognitive psychology were not just practical but of deep principle; inaccessibility denied even the hope of a remote future in which we might find some way to even indirectly measure the attributes of the mind.

The comparable supplemental epistemological assumption underlying modern cognitive neuroscience, however, is a practical one rather than one of deep principle. It is the respective answer to the query—is it possible to measure or examine the details of the neural network that is the basic psychoneural equivalent of mental activity? This supplementary epistemological postulate for neuroscience can be formalized as shown in box 1.7. Although this postulate may well run counter to the current Zeitgeist, a strong argument supporting this postulate can be made based on combinatoric arguments.

Inherent in any such postulate, of course, is the possibility that, at some unforeseeable future time, unexpected developments may make possible what is currently impossible in

**Box 1.7**

The Second Epistemological Postulate for Neuroscience

> Because of their great complexity and number, it is not possible for us to analyze the great neuronal networks of the brain in a way that would permit us to identify the neural equivalent of any kind of mental activity at this microscopic level of analysis.

practice. Nevertheless, the reality today is that there is no practical way to analyze such a complex and irregular network, and some mathematical arguments concerning combinatorics and computability strongly suggest that this is a reasonable working rule for the near and perhaps even for the far-distant future.

What are these arguments supporting the second epistemological postulate for neuroscience? Some of them are these:

- The kinds of network problems that are observed in the brain are known to be computationally intractable *in practice*. They are not infinitely complex (which would introduce an *in principle* constraint) but merely so consuming of any conceivable computational power that they could never be solved. This is equivalent to what complexity theorists call an NP complete problem, a problem that cannot be solved in any determined amount of time.
- Few of our neural network simulations scale up. That is, the simple models we are able to program onto a computer typically fall apart or saturate in one way or another when we try to increase the number of interacting simulated neurons beyond a few hundred.
- Efforts to simplify the difficulties inherent in these problems (e.g., by assuming regularity, adding additional nonbiological constraints, or by breaking it up into smaller parts) do not work.[16]

There are really no good counterarguments to these practical constraints on understanding the neuronal basis of the mind. What actually happens is that investigators implicitly accept the limitations and then turn to alternative experimental and theoretical strategies. They implicitly accept the monumental barriers to the direct evaluations of these complex neuronal networks and utilize whatever measuring devices are available (e.g., the EEG or the fMRI), whether or not these devices are operating at the appropriate level of analysis. In so doing, often without realizing it, they are accepting the unanalyzability of the neuronal network—the second epistemological postulate for neuroscience—and opt for some alternate, but questionable, strategy that does not even promise to answer the essential mind-brain problem. It is important to point out again that this does not imply an "in principle" rejection of the second ontological assumption; instead it is simply a practical and necessary response to the fact that studying the microdetails of the neuronal network is not an effective strategy.

## 1.4 Some Relevant Conceptual Issues

### 1.4.1 The Seductive Attractiveness of Brain Images

A major issue in cognitive neuroscience concerning the use of brain images such as the fMRI is that their impact on our science may be far more than they deserve. That is, we are seduced by the pretty pictures and the seeming "face validity" that these images seem to offer. For example it is now established (McCabe & Castel, 2008) that people are more likely to accept the credibility of a published report when a brain image rather than an informationally equivalent graph or table is used. Roskies (2008) referred to brain images as perpetuating an "illusion of inferential proximity" that makes us feel we know something about something that, in fact, actually remains inscrutable.[17]

Weisberg, Keil, Goodstein, Rawsdon, and Gray (2008), arguing in a similar vein, suggested that neuroscientific information itself, even if it is not relevant, made a theory more acceptable than when that kind of information was withheld. This effect was maximum when the judging subjects were not expert in the field beings discussed. These authors warned especially against the problems posed for nonexperts in evaluating neuroscience information added as decorations to scientific story.

The point is that the attractiveness and the seeming, but illusory, directness of these images give them a conceptual and scientific impact that they may not entirely deserve. Their charm, their novelty, and their pictorial splendor tend to overwhelm critical consideration of the serious epistemological issues revolving around the limits of what these images can actually tell us. It is only in recent years that the empirical facts have begun to raise further questions about some of the facile misinterpretations of their meaning.

It can be argued that the widespread and uncritical acceptance of the brain image as a measuring tool of cognitive processes is based on a widespread misunderstanding of the actual progress that has been made in linking results from the two fields. Although cognitive neuroscience journals have been flooded with publication of what are often very preliminary reports, the neural basis of cognition and the neural activity depicted by brain images operate at vastly different conceptual levels. We do not yet understand what either of these differences means or how we might link them together.

In short, there is no theory or putative explanation that yet explains how mental processes emerge from neural ones. There is, instead, an emerging corpus of scientific opinion that the mind-brain problem is intractable due to the complexity of the neural interconnections that actually lie at its core. The misunderstanding that we have made more progress on this fundamental issue than we have is also exacerbated by the hyperbolic and exaggerated popularization of very preliminary or unsubstantiated scientific findings by the press and the lay community.

### 1.4.2 The Problem of Defining Mental Processes

One of the most serious impediments to unraveling the mind-brain problem is that mental states are very difficult to precisely define. In fact many of the cognitive processes that we wish to correlate with either surgical interventions or brain images are merely neologisms for experimental results or hypothetical constructs used to flesh out some speculative psychological theory. To compare the objective neurological data with such poorly defined, and often arbitrary, mental entities stretches logical analysis to its limits. The actual connection is so loose that it is all too easy to carry out what are, in retrospect, misleading comparisons. It is also possible in systems as complex as this to find empirical support for almost any theory.

The problem is that the history of psychology is filled with a huge vocabulary for a large number of different psychological faculties, traits, or modules. Yet there has been no clear development of a coherent taxonomy or classification system for psychology comparable to the Linnaean one in biology or Mendeleev's in chemistry. Instead, idiosyncratic and obscure terms come and go as psychologists suggest new hypothetical entities, study them until they are no longer of interest, and then move on to some other topic. Words like "acquisitiveness" or "ego" have now been replaced by such equally vague concepts as "consciousness" or "attention." Only operationally defined terms such as "reaction time" or "percent correct" tell us anything and then only about the behavior of the organism.

The issue of definition becomes extremely vexing when a comparison is made between a mental activity and a brain response, especially if it is uncertain just what mental processes are being invoked and how such obscure processes as "attention" and "perception" actually are different or independent of each other (if they are). The point is that it becomes difficult to locate in the material brain what are little more than hypothetical constructs or tags attached to experimental protocols.

This issue raises serious practical problems of interpretation in any mental process-brain response comparison. For example, a study purportedly of people's preferences for a cola or a political candidate may end up measuring a brain response that has nothing directly to do with our preferences per se but may be measuring some subtle aspect of a general emotional response or of some previous experience. This misdirection to an irrelevant aspect of the cognitive state plays havoc with any attempt to use a brain image as an indicator of mental activity as well as any effort to develop a coherent theory of mind-brain relations.

In a more general sense, it raises questions about the validity of any purported neural measure of any cognitive process. If there is always the possibility (because of poor definitions) that we are measuring something other than what we thought we were, no matter how reliable the findings, those findings may be theoretically meaningless. In short, poor definition of mental entities degrades the validity of any neural correlations with those nebulous cognitive processes. To note that it also makes independent stimulus control more difficult is simply to restate the obvious.

### 1.4.3 The One-to-Many Issue

It is becoming increasingly clear that many different cognitive processes can activate the same area or system of areas of the brain. (For example, see the work of Culham & Kanwisher, 2001.) Thus, if the available findings are limited to answers to questions about "where" a response is occurring, it is theoretically impossible to exclusively associate any particular brain activation site or pattern of activation with any particular cognitive state. In the words of Poldrack (2006), it is extremely difficult because the putative location of a cognitive module is not unique, to use "reverse inference" to assign specific mental meanings to even the most discrete and reliable brain activations. Any attempt to do so, according to Poldrack, is "deductively invalid." He goes on to say that it "still can provide some information," but this depends on the empirical "selectivity of activation" (p. 59).

This is a very important, but largely overlooked, point. It raises severe limitations for any attempt to "read a person's mind" by measuring brain responses. First, multiple functionality of single brain regions disassociates specific brain responses from particular cognitive processes as a matter of principle. That is, activation in any particular brain region cannot be solely assigned to any particular cognitive process when that brain area is involved in representing many different cognitive processes.

Second, the one-to-many problem adds to the practical difficulty of assigning either qualitative or quantitative significance to what are often only modest correlations in brain image-cognitive process comparisons. No matter how carefully an experimenter controls the salient experimental variables,[18] there must necessarily always be other forces operating to modulate the response of a given brain region. Efforts to use brain imaging in legal proceedings as a "lie detector" to mitigate culpability are invalidated from the outset by this principle.

### 1.4.4 The Many-to-One Issue

Just as the fact that many psychological tasks and stimuli can simultaneously activate a single brain region (and, therefore, we cannot in principle say that any particular neural activity or place is a unique indicator of any particular kind of mental activity), it must also be remembered that it is likely that many behaviors or cognitive processes may be instantiated by a number of different and redundant brain mechanisms. We have little knowledge about the full range of brain regions that may be equivalent or substitutable for each other. However, we do have plenty of evidence that many different regions of the brain are activated during any kind of cognitive task. Furthermore, we also know that under extreme conditions (such as damage due to ischemic stroke) some regions of the brain are capable of taking over functions of damaged regions. Whether or not this redundancy under the extreme conditions of a stroke is also implicitly or explicitly present under normal conditions remains an important question for cognitive neuroscience.

The potential for redundant representation strongly suggests that cognitive processes need not be encoded by the same neural mechanisms in different people. Just as there are different cognitive strategies to solve a particular problem, it is probably the case that many different brain regions or clusters of brain regions may account for a particular behavioral outcome. This is what we refer to as the many-to-one principle. This is also what is meant by the general underdeterminative nature of behavioral responses—behavior cannot tell us what brain mechanism is active just as activations of brain regions do not tell us which mental process is active. Behavior by itself is neutral with regard to underlying mechanisms.

This limitation on our understanding is also known, as I have discussed earlier, as the "black box problem." To know precisely what mechanism is inside the black box, one must open the box. Unfortunately, for mind-brain theorists, even after imaging the brain, the complexity of the system and the variability of the responses are so great that our path to understanding is blocked by another kind of virtual “closeness”—complexity.

### 1.4.5 The Sign-Code Distinction

Over the years (starting with Uttal, 1967) I have repeatedly pointed out that there are two possible meanings—signs and codes—of the correlated neurophysiological responses that are obtained when one compares brain activity to cognitive processes. A “sign” is a correlate of brain activity that indicates that something is happening neurophysiologically, but it is only a candidate to be the “psychoneural equivalent” of the associated mental activity. That is, a sign may be correlated neural activity in response to stimulus or mental task. However, it has not been established that it is *the* neural activity that specifically results in or *is* the cognitive experience itself. It does not encode, represent, or in any way is it the equivalent of the mental experience. All that a sign does is to tell us that there is a recordable brain response to some stimulus or cognitive state.

A sign may be used in powerful ways to measure some property of brain activity and possibly even someday serve as a biomarker of some dysfunctional cognitive activity. However, it does not necessarily explain or represent the mechanism by means of which brain activity is transmuted into mental activity. In short, the concept of a sign reminds us that not all neural responses recorded from the brain are psychobiologically relevant.

On the other hand I have designated a correlated neurological response that *is* the psychoneural equivalent of some mental activity as a “code.” A code is a measure of neural activity that is the actual mechanism of whatever cognitive process is being manipulated. It is the necessary and sufficient mechanism, not merely a concomitant or correlated sign, of some mental activity. In short, it is the neural activity whose activity *is* the mental activity.

Distinguishing between a sign and a code is not an easy task given that either may correlate highly with brain activity. To determine that something is a code requires that we prove both its necessity and its sufficiency, an empirical task of considerable difficulty. Clearly, because the requirements for a code are so high, there are very few that have been robustly identified in the cognitive neuroscience literature beyond the transmission codes of the sensory and motor systems. The study of higher-level cognitive processes remains virtually untouched by such progress.

It now seems clear that most of the molar,correlated brain responses, whether they are EEGs or fMRIs, are signs. In fact, as some scholars have pointed out, brain images tell us little more than that there is some brain activity when our minds are active, an idea that is hardly surprising given the ontological postulates discussed earlier. However, as far as specifying the specific neural processes that are the coded equivalents of mind, virtually all such cumulative measures of brain activity are bankrupt. The persisting question is—what does knowing what part of the brain is activated by some stimulus or task tell us about how that part might encode mental activity?

## 1.5 Some Relevant Technical Issues

The conceptual, occasionally philosophical, issues discussed in the previous sections raise serious questions about the applicability of brain imaging devices to the measurement of cognitive processes. There are, in addition, a number of purely technical issues that complicate the matter further. These issues and challenges arise not because of any subtle logical or epistemological uncertainties but because of well-documented and tangible issues with the day-to-day details of using brain imaging devices to conduct research on cognitive processes. The technical issues collectively also provide serious challenges to any optimism to the use of brain imaging as a means of evaluating such poorly defined psychological faculties as learning, attention, perception, personality, thinking, intelligence, level of learning, decision making, or other complex, high-level cognitive states. The following paragraphs discuss some of the technical issues that still bedevil research efforts to correlate brain images and cognitive states.

### 1.5.1 Cumbersome Procedures

However beautiful a colorful brain image may eventually turn out to be, it is the final result of a massive investment in time and money. The PET system, for example, not only requires the detector system itself, but also a radioisotope- generating capability to produce the injectable radioactive materials and a computer facility to process the raw data from that detector. (The medical and ethical issues of using such an invasive procedure, furthermore, should not be minimized.)

The complexity and expense of MRI systems are also well appreciated, but the major issue of invasiveness associated with the PET procedure is largely overcome by MRI systems. Functional MRIs are totally noninvasive; no one has ever shown any deleterious physical effects from the large magnetic fields used to orient the protons of the body's atoms other than being hit by an errant piece of metal attracted by the powerful magnetic fields surrounding the device. Nevertheless, there are a number of practical issues in their use that also make the process cumbersome, complicated, and expensive.

It takes an extended period of time to produce a single fMRI image with most current techniques. Furthermore, subjects must cooperate to an extreme degree including remaining motionless and attending to a single cognitive theme for the duration of the measurement in what can be an acoustically noisy and highly constricted environment. Even the slightest head or respiratory movements can distort the final image (Raz et al., 2005). Furthermore, because of the extended time required to collect the data in an fMRI-based experiment, the number of subjects is usually relatively low compared to behavioral evaluations.