— CHAPTER 1 —

Approaches to Mind

Robert G. Burton

I

Willard Quine has observed that the ontological question "can be put in three Anglo-Saxon monosyllables: 'What is there?' It can be answered, moreover, in a word—'Everything' . . ." (1961, 1). And few would deny that *everything* includes minds. But what are minds, and how are they known? How well do we know our own minds? And can we ever know the minds of others as well as we know our own?

The most important general answers to the ontological question that orient contemporary studies of mind can be traced to antiquity. Materialism is the view that reality consists of matter and its modifications; that every property, object, state, event, or process is material. Materialists offer varying definitions of mind depending on the physical level at which they seek to locate and study it. Thus the mind is nothing but the functioning central nervous system, or the mind is a fiction attributed to certain organisms exhibiting intelligent behavior, or the mind is a computational system. Idealism is the view that reality consists of minds and their ideas; that every property, object, state, event, or process is mental. Idealists are difficult to find these days, and none is included in this anthology. Substance dualism is the view that reality consists of two distinct kinds of things, nonmaterial minds and material bodies, and that properties, states, events, and processes may be either mental or material. Dualism remains the most common theory of mind for the general public even though it is rather rare among philosophers and scientists.

This anthology explores six current approaches to the study of mind: the neuroscientific, the behavioral-experimental, the competence approach, the ecological, the phenomenological, and the computational. These approaches differ with respect to the types of mental phenomena investigated and the methods of investigation. It is the organizing thesis of the book that these approaches are interdependent, that they cannot and should not be pursued in provincial isolation from one another. The concluding chapter defends the organizing thesis and develops a systematic explication

of some important forms of interdependence often overlooked and sometimes denied. The prospects for fruitful cooperation among researchers following these diverse approaches will be enhanced if they can be shown exactly how their research strategies are interdependent.

The contributors to this book include philosophers, psychologists, and a computer scientist who is also an adjunct professor of philosophy. It is customary to distinguish philosophy from science. Philosophy is primarily concerned with basic conceptual issues such as the conceptual foundations of science. But it is difficult to draw a sharp line between science and philosophy, and this is particularly true of psychology and the philosophy of mind.

The neuroscientific approach to the study of mind attempts to describe and explain mental phenomena in terms of their underlying neural processes. This is an ambitious undertaking. It has been estimated that a single human brain has roughly 100 billion nerve cells or neurons. Each neuron has roughly three thousand connections to other neurons. Given this enormous number of microscopic elements, and given the complexity of their interconnectedness, it should not be surprising that we have barely begun to understand the mysteries of the human nervous system. Also, moral considerations limit the kind and extent of experimentation on all living humans, including those who have genetic defects, those who have degenerative brain disease, and those who have suffered serious injury to the brain.

Behavioral-experimental psychology attempts to describe and explain mental phenomena in terms of publicly observable behavior. Methodological behaviorists argue that environmental stimuli, rather than conscious mental phenomena, determine what we do. Psychology, if it is to become a legitimate empirical science, should confine itself to investigating the functional relationships between publicly observable behavior and environmental stimuli. The prediction and control of behavior should be its goal. Radical behaviorists attempt to extend the techniques of behavioral analysis to cover the private mental events declared inaccessible by the methodological behaviorists. Psychological behaviorism has a philosophical cousin, sometimes designated logical or analytical behaviorism, that should be mentioned. Logical behaviorism is a theory of the meaning of mentalistic terms. It seeks to analyze such terms as *desire, belief,* or *pain* into terms that mention only behaviors or dispositions to behave in certain ways.

The competence approach to the study of mind employs artificial minds derived from the "idealized speaker-listener" of Chomskyan linguistics (Chomsky 1965). Central to this approach is the distinction between competence and performance. *Performance* refers to actual language use and the cognitive processing behind it. *Competence* refers to

the knowledge, largely tacit, that the speaker-listener must have in order to use language grammatically, intelligibly, and creatively. The use of an idealized speaker-listener in an idealized context enables the theorist to develop and explore formal models of natural language use that are free from the grammatical idiosyncrasies of actual speaker-listeners in natural contexts. And this competence approach to theorizing, using idealized artificial minds, can be adapted to the study of other symbolic-cultural systems in ways that, coincidently, extend our knowledge of cognitive functions.

Ecological psychology is the youngest kid on the block. Like all cognitive sciences, it begins with an analysis of the information available to the perceiving and thinking subject. But it opposes the practice, shared by both behavioral and most other cognitive psychologists, of investigating cognitive phenomena in laboratory settings that are highly artificial and grossly oversimplified in comparison to natural settings. It argues that studies of maze running and bar pressing or the memorization of lists of meaningless symbols will tell us little about such cognitive processes as problem solving and memory in the real world. Instead, it begins with a careful description of the environment and our ordinary activities within it, and it attempts to identify the informational structures provided by the environment before theorizing about how we make use of them.

Phenomenological analysis takes all data of experience as its object. It attempts to investigate and describe the data of consciousness in a disciplined way that identifies biases and minimizes their effect by temporarily suspending such questions as the origin of the phenomena under investigation. The privileged access that each one has to the contents of his or her own consciousness, is, therefore, an important source of knowledge rather than a barrier to be declared out of bounds. Phenomenological analysis not only describes phenomena that become objects for lower-level analysis, but also provides a perspective from which to assess the plausibility of the results of such lower-level analysis.

The computational approach to the study of mind is embodied in the research program that investigates artificial intelligence. The field of artificial intelligence has two goals: (1) to make computers more useful by making them more intelligent, and (2) to understand what makes intelligence as such, including human intelligence, possible (Winston 1984, 1–3). The traditional information processing approach to AI views the mind as a computational system that processes informational input to produce the meaningful phenomena of consciousness and behavioral output. So far, there have been many failed attempts to construct programs and computers that can duplicate such ordinary mental functions as natural language parsing. They have "worked" only in a very limited sense, for very

restricted domains. But some believe that what is now computer simulation (at best) will one day be computer duplication.

Let us now take a brief look at some of the scientific developments and philosophical arguments behind our six different approaches to the study of mind. This will enable us better to place, understand, and assess the work presented in the subsequent chapters.

II

In *A History of Experimental Psychology,* E. G. Boring (1950) observed that biological science grew out of the medical science of antiquity—a mixture of surgery, folk medicine, magic, and anatomy. The nearly universal taboo on the dissection of human cadavers delayed the acquisition of knowledge of human anatomy and the subsequent development of physiology. Long after Hippocrates (ca. 460–370 B.C.), Galen (ca. A.D. 129–199) localized the mind in the brain, distinguished between sensory and motor nerves, and by experimental transection of the spinal cord localized some of its motor functions.

But what does it mean to localize the mind in the brain? And what is the nature of that which is localized? To localize the mind in the brain is not equivalent to proving that the brain, or some part of the functioning brain, is the subject of consciousness. Furthermore, there are many modes of consciousness, many different phenomena that only beings capable of consciousness experience. Perhaps, in the light of the enormous diversity of mental phenomena, it should not be surprising that we still do not know what consciousness is. Perhaps consciousness is no *one* thing but only the various forms of awareness (including self-awareness), involving differing neural processes. On this view, the minimal methodological assumption of neurophysiological psychology is that every aspect of one's conscious experience is correlated with some aspect of one's brain activity. This is not to say that there must be isomorphism between aspects of conscious experience and correlated aspects of brain activity. But it does mean that any change in one's conscious experience will involve a correlated change in one's brain processes. Therefore neuropsychology begins with the attempt to correlate conscious experience with brain processes.

But what are the phenomena of conscious experience that are to be correlated with brain processes, and how is one to get at them? First of all, there are sensations with their unique qualitative natures or qualia. There are also conscious states, such as desires, beliefs, wants, and fears, that exhibit the feature of intentionality; that is, they are directed toward objects, events, or states of affairs beyond themselves. For example, one

desires some definite object or that some possible state of affairs should come to pass or one believes that some state of affairs has come to pass. Other mental phenomena such as moods of depression or anxiety are not directed outward toward any definite object, event, or state of affairs. These are the sorts of phenomena cited in our everyday, commonsense descriptions and explanations of human experience. But how do we get at them?

The traditional answer is that each of us has privileged access to the contents of his or her own consciousness. I know my own mental states either immediately or on reflection in a way that no one else can know them. Others may guess what is on my mind or infer it from my behavior or take my word for it should I choose to talk about it. But their access is always indirect, whereas mine is direct. This fact of privileged epistemic access is at the core of our concept of mind. It does not entail that one always know one's own mind infallibly, indubitably, or incorrigibly. Nor does it prevent varying degrees of self-deception or incompleteness. These difficulties are serious but not insurmountable (see Alston 1971). The discipline of phenomenological analysis attempts to minimize these difficulties. But for those who assume that only publicly accessible objects are possible candidates for scientific investigation, the phenomena of consciousness would seem to be beyond the reach of science unless they could be explained in terms of that which *is* publicly accessible.

Neuropsychology begins with the attempt to correlate conscious experience with brain processes, but it does not end there. It seeks to explain the phenomena of conscious experience in neurophysiological terms. But this suggests that careful phenomenological description of that which is to be explained would be essential to the neuropsychological enterprise. And the history of experimental psychology bears this out. A case in point is the Purkinje effect. In two volumes of excellent visual phenomenology published in 1824–25, the Czech physiologist Purkinje described the phenomenon wherein colors toward the long-wavelength end of the visible spectrum in a multicolor display lose their perceived brightness more rapidly as illumination is reduced than do colors toward the short-wavelength end. But it was seventy years before the Purkinje effect was explained in terms of the fact that the retina contains rods as well as cones and that the rods, which function only in reduced illumination, have greater overall sensitivity than the cones and are also maximally sensitive to short wavelengths. This case illustrates that experimental psychology began as an attempt to develop an analogue of morphological description in biology so as to create a morphology of mind, a phenomenological taxonomy of consciousness (Boring 1950, 20–21). But assuming that phenomena of consciousness have been identified and described with

as little bias as possible, how are they to be explained in neurophysiological terms?

There seem to be three main possibilities: (1) reduction, (2) elimination, or (3) some mixture of both. It is the thesis of reductionism that social, psychological, and biological phenomena are nothing over and above physicochemical phenomena. Reduction has been analyzed in terms of definability and derivability. Therefore psychological theory is reducible to physicochemical theory if and only if psychological terms are definable in the terms of physicochemical theory and psychological laws are derivable from physicochemical laws.

In the 1950s, U. T. Place (1956) and J. J. C. Smart (1959) advanced the type identity theory according to which mental states are nothing but brain states and mental processes are nothing but brain processes. Each type of mental state or process is identical with a type of physical state or process in the brain. For example, the experience of having a toothache is identical with a certain type of brain process just as lightning is identical with a particular type of electrical discharge. This is a very strong claim that goes far beyond the minimal methodological assumption of neurophysiological psychology. After all, a thing cannot be correlated with itself, and the discovery of the countless ways in which mental phenomena are dependent on neural phenomena does not by itself establish their identity. The identity theorist claims that neuroscience will eventually develop a taxonomy of neural states and processes that are in one-to-one correspondence with the mental states and processes of our commonsense psychology. The one-to-one correspondence holds because the corresponding entities are really identical. But be that as it may, the only empirical evidence for type identity theory is the actual correlations and causal dependencies that neuroscience is able to discover. Nevertheless, it does have considerations of simplicity as well as the impressive success of the physical and biological sciences to recommend it.

But type identity theory has fallen out of fashion for a number of reasons. One class of objections has caused type identity theorists themselves to become eliminativists with respect to certain mental phenomena. According to Leibniz's formulation of the principle of the identity of indiscernibles, A is identical with B if and only if every property of A is a property of B. But consider the following apparent counterexample. One's afterimage is yellowish-orange and private, whereas brain processes have no color and are public. J. J. C. Smart's reply to this challenge is to declare that there are no such things as afterimages or sense data. The experience of having an afterimage is identical with a certain type of brain process (1959, 151). Therefore not all mental phenomena that figure in our

commonsense psychology are candidates for reduction; some are candidates for elimination. But how does one distinguish between those that are candidates for reduction and those that are candidates for elimination?

Daniel Dennett has offered one sort of answer in the course of considering "Why you can't make a computer that feels pain" (1981, 190–229). He argues that there is an irredeemable incoherence in our ordinary concept of pain which means that *it* cannot be used in a true theory of pain that a computer or robot might be made to instantiate. We must wait for neuroscience to explain how the various phenomena of pain are produced. Only if brain scientists are able to produce a good subpersonal theory of pain will we be able in principle to construct a robot that instantiates it.

It is important to note that Dennett is not denying the reality of the phenomena of pain. The phenomena of pain are candidates for reduction, not elimination. It is the internally inconsistent concept of pain that is a candidate for elimination. Of course, many would not agree that our commonsense concept of pain is hopelessly incoherent. If it were hopelessly incoherent, how would one explain its persistence and utility in our everyday, commonsense, "folk psychology," which Dennett (1987) recognizes to be a "powerful predictive craft" even though he argues that it is not a very good theory?

There is probably no simple algorithm for distinguishing candidates for elimination from candidates for reduction. Avoidance of internal inconsistency is one criterion. A second criterion is avoidance of indiscriminate reification: that is, treating too many theoretical constructs as though they designate things in the world. As an example of this, Dennett has imagined a society of beings very much like ourselves except for the fact that they speak of being beset by "fatigues" whenever they are tired. The analogy to our speaking of having pains whenever we are in pain is unavoidable. But according to Dennett, neither "fatigues" nor pains should be considered to be things in the world (1988, xix–xx).

Returning to objections to type identity theory, we should consider a group of arguments from multiple realizability. The source of these objections is functionalism. Functionalism takes many forms, but in the most general terms it is the thesis that mental states and processes are the embodiments of abstract functional states and processes that mediate between environmental inputs and behavioral outputs. Thus the mind is a computational system, and psychological explanation consists of developing a functional analysis of mental processes on the model of a computer program that would enable a suitably designed computer to process the environmental inputs to achieve the behavioral outputs in question. But a given program may be realized in an indefinite variety of computers so that

there is no reason to suppose that only living organisms, "carbon-based computers," and, most impressively, the higher mammalian types are capable of mental states and processes.

Now, it is important to observe that the argument from multiple realizability is not only an argument against type identity theory but also an argument against type functionalism. It is true that many different neurophysiological processes might eventuate in the same output, and it is no less true that many different programs might yield the same output. Therefore the only plausible form of these theories is the weaker token form. Token theory correlates specific instances or tokens of mental phenomena with tokens of brain process or function. For token identity theory each specific mental event is identical with some brain event or other, and for token functionalism each specific mental event is some functional event or other. And, needless to say, a given mental event could be described as either a brain event or a functional event or both.

But some mental phenomena, for example qualia or consciousness, seem resistant to information-flow analysis. To bring this out, let us paraphrase Thomas Nagel's phenomenological question (1974). What is it like to be a chess-playing digital computer? Have we any reason to think that any such device knows or cares that it is playing chess? Could a robot experience the thrill of victory or the agony of defeat? Ned Block (1978) has argued convincingly that we would be question-begging liberals to answer yes and question-begging human chauvinists to answer no. Perhaps we will simply have to wait and see what computers of the future can and cannot do.

The older forms of functionalism, exemplified in the "rules and representations" approach to the study of artificial intelligence, now have rivals in connectionism (see Horgan and Tienson 1988) and Gerald Edelman's theory of neuronal group selection (1987). Both of these make use of computational modeling, but they are also strongly influenced by neurobiology. They reject the pure functionalism of traditional AI. They argue that embodiment makes a difference, and they look to neuroscientific insights into the embodiment of natural intelligence for suggestions concerning design of the artificial.

Traditional AI has been much better than we are at such tasks as number crunching, but much worse than we are at such tasks as pattern recognition, understanding speech, recalling and recognizing relevant information, and learning in general (see Tienson 1988, 1–16). Human brains evidently do not do things the way digital computers do them. Neurons are approximately one million times slower in transmitting impulses than are electronic computers. It would be impossible to do what we do, as quickly as we do it, if we had to process all information in the

serial fashion of information flow analysis. Furthermore, neuronal networks degrade gradually; that is, they can suffer damage and still continue to function reasonably well. By contrast, traditional AI systems are "brittle"; that is, when even minor things go wrong they tend to break down completely. This has led to the development of the new connectionist models that simulate neural networks processing inputs in a massively parallel way. They are capable of being "trained up," they may be able to solve such problems as pattern recognition, and they degrade gradually. Edelman's paradigm is more heavily constrained by biological considerations, but, despite his claims to the contrary, it seems to be at least a cousin to the new connectionism.

At this point, we should consider some developments in behaviorism, the form of experimental psychology that became dominant in Anglo-American psychology from the 1920s to the mid-1950s and exercised a strong influence on philosophy as well (see Cohen 1987).

At the turn of the century, when John B. Watson (1914, 1919) was a student, psychology was the most recent science child to have broken away from philosophy. Much effort was spent in the attempt to develop an experimental method of introspection, modeled after physics and chemistry, that would enable psychologists to penetrate the mysteries of consciousness. Introspective observers were trained to report on what was going on in consciousness when, for example, they were given a display of dots or when they were told to wait for a tone. They were asked to distinguish and quantify states of clearness in attention. The ultimate goal of this research was to discover the "atoms" of consciousness. But Watson hated serving as a subject in these stuffy, artificial settings, and noticing that different subjects invariably gave very different reports, he came to believe that he could learn more by studying observable behavior in humans even as it is studied in lower animals. Psychology thus became a part of biology. Watson, following the classical conditioning model of I. P. Pavlov, claimed that most human behavior is conditioned reflexes. Mental phenomena were dismissed as unimportant epiphenomena; that is, phenomena that do no causal work. This position eventually came to be called *methodological behaviorism.*

B. F. Skinner (1938, 1945, 1953, 1957, 1971, 1974, 1977) seemed to share Watson's epiphenomenalism for the most part, although at times he seemed closer to an identity theorist (see Burton 1984, and Creel 1974, 1980). Skinner claimed that "mentalism" is the chief obstacle to a science of human behavior. The mentalistic explanations of cognitive psychologists and philosophers divert inquiry from the role of the environment. In describing *how* people behave, they allay curiosity and bring inquiry to a halt without ever explaining *why* people behave as they do. Methodological

behaviorism was a step in the right direction. But in ruling that private events have no place in science because there can be no public agreement about their validity, the methodological behaviorists unwittingly perpetuated dualism. They ignored the mentalistic link in the chain between environmental input and behavioral output. And thus, according to Skinner, they left the domain of the private to the "mentalists" who continue their profitless investigations. Skinner claimed that his own radical behaviorism does not rule private events out of bounds; it restores introspection but rejects the traditional mentalistic account.

Skinner gave two contrasting accounts of private events (Burton 1984). The most interesting and systematic was introduced in an early paper, "The Operational Analysis of Psychological Terms" (1945), and developed in *Verbal Behavior* (1957). In this work Skinner tried to discover and describe the processes by which we acquire and maintain the linguistic capacity to describe private events. But to pursue this research would have involved using phenomenological data, and this Skinner was reluctant to do. Instead, he became preoccupied with the technology of behavioral control and the attack on mentalism. His later, popular works (1971, 1974) contain a treatment of private events that is superficial, cavalier, and unconvincing. But this should not blind us to the importance of Skinner's contribution to the development of experimental psychology.

Skinner discovered that the consequences of behavior are even more significant than its antecedents in determining the probability that a behavior will be repeated. He introduced the term *reinforcer,* empirically defined as any event that strengthens a behavior on which it is contingent. He termed the classical reflex behavior investigated by Pavlov "respondent" behavior to capture the fact that it is elicited by antecedent conditions. The type of behavior that we are said to emit voluntarily, the type with which Skinner is most concerned, he termed *operant* behavior to capture the fact that it is our means of operating on the environment. Using an experimental environment, Skinner and his associates investigated the effects of various amounts and schedules of reinforcement on the response rates of such diverse organisms as pigeons, rats, and humans for simple operants such as bar pressing. From these investigations certain lawlike regularities have been discovered. Operants that have been strengthened through reinforcement may be weakened and eventually extinguished if they are permitted to occur without reinforcement. Operants may be temporarily weakened by allowing the subject to become satiated on the reinforcement. Aversive stimuli may serve either as punishment that weakens behavior or as negative reinforcement when their removal strengthens behavior. And finally, operants may be weakened through the conditioning of antagonistic operants.

One is tempted to say that, apart from precise laboratory data on the operant behavior of subjects whose histories of conditioning are carefully recorded, there is little news in all of this. Most of these lawlike principles are well entrenched in our commonsense, everyday, "folk psychology." And our ordinary language is adequate to the task of expressing it. Indeed, ordinary language philosophy attempts to mine the implicit resources of our ordinary linguistic practices. Two important versions of ordinary language philosophy are the later work of Ludwig Wittgenstein (1953) and Gilbert Ryle's logical behaviorism as set forth in *The Concept of Mind* (1949).

In his *Philosophical Investigations* (1953), published posthumously, Wittgenstein repudiated referential theories of meaning, including his own former theory (1961), according to which the meaning of a word, for example a typical noun or adjective, is the object to which the word refers, that is, the object or property of an object denoted by the word (see Bechtel 1988). He noted and described an enormous variety of ways in which we use language. Stating facts and referring to objects are not the only "language games" we play; others include giving and obeying orders, describing appearances, constructing, reporting, speculating, forming and testing hypotheses, making jokes, solving problems, and translating from one language to another. Words are used differently in different language games; they are not always used to refer. It is pointless to ask for *the* meaning of a given word apart from its use in the context of a particular language game. Most philosophical problems are really pseudoproblems caused by the failure to distinguish the various language games and their unique "grammars," the implicit rules that govern them. A case in point is the assumption that the word *pain* in sentences such as "I have a pain," is used to refer to some private thing. This leads us to search for evidence that someone has a private thing called a *pain*. Wittgenstein claims that if we attend to the actual circumstances in which we say such things as "I have a pain," we will discover that we are simply expressing our pain; we are not reporting the existence of a private object.

In *The Concept of Mind* (1949), Ryle argued that the problem of the relation between mind and body is a pseudoproblem involving a "category mistake," the mistake of treating mentalistic terms as if they denoted entities analogous to parts of the body. Such terms as *desire, belief,* or *pain* ought, instead, to be analyzed in terms of behaviors and dispositions to behave in certain ways. Ryle's distinction between different linguistic categories is in some respects like Wittgenstein's distinction between different language games. Both reject referential theories of the meaning of mental terms. But logical behaviorism has fallen out of fashion. In the first place, there is an indefinitely long list of behavioral dispositions and

possible behaviors that might constitute any particular mental event. And in the second place, any of the behaviors might occur in the absence of the mental event in question, and the mental event might occur without any of the behaviors on the list. One can have a toothache without saying anything, such as "I have a toothache," or "My tooth aches," or doing anything else about it. And one can engage in all sorts of behaviors leading others to think that one has a toothache when one does not, whether or not one intends to deceive.

We should say more about intentionality, a feature of such conscious states as desires, beliefs, hopes, fears, and wants; namely, the fact that they are *about* something beyond themselves (see Dennett and Haugeland 1987). Franz Brentano (1973 [1874]) claimed that intentionality is the irreducible feature of all mental phenomena. He seems to have gone too far in this; moods, for example, do not seem to be *about* anything. Nevertheless, the problem of developing an adequate analysis of intentionality is shared by philosophers and psychologists alike.

Intentionality, the feature of *aboutness,* should not be confused with the more familiar use of the word *intentional* in which we say that an act was intentional, meaning that it was done deliberately or on purpose. The confusion is natural enough, for both kinds of intentionality are included in ordinary explanations of rational action. For example, someone asks, "Where is Mary?" We answer, "She went to the post office to mail a package." We are describing a purposive action. In this sense, Mary went to the post office intentionally. But our answer also contains an explanation in terms of implicit desire and belief. And desire and belief are intentional in Brentano's technical sense. The object of Mary's desire is to mail a package, and the content of her belief is that she can do so by taking the package to the post office.

One of the most striking things about intentionality is the fact that intentional objects need not exist. Brentano called this fact the *intentional inexistence* of intentional objects. For example, one can want things or believe in things that do not exist. But what, then, is the nature of the relation between an intentional state and its object? And what is the status of the object? The object cannot be simply an idea in the mind. If one wants a drink of water, one wants water, not an idea in the mind. One may already have an idea of water in the mind, but the idea will not quench one's thirst, and the idea is not what one wants. Ordinary relations hold between things that exist or have existed. But how is one to characterize the relation between an intentional state and its object in case its object does not exist? We have yet to reach agreement on how to characterize the relation when the object does exist.

Bertrand Russell (1940) suggested that we represent intentional states

as "propositional attitudes" (see Bechtel 1988). This format represents intentional states with a verb and a proposition. Thus, one can be said to desire, or believe, or hope, or fear that a given proposition is the case. One can have the differing attitudes of desire, belief, hope and fear toward the same proposition, or one can have the same attitude toward differing propositions. Propositions seem to be accessible in a way that intentional objects are not, and they seem to provide a way to specify the contents of mental states. Functionalists have tried to exploit these advantages. In fact, Russell's proposal to move to the logical analysis of the language we use to talk about intentional states has been widely adopted in Anglo-American philosophy. But there are serious problems. What is a proposition? It is neither a particular sentence nor a speaker's particular speech act. It is sometimes said to be that which is expressed in statements, namely, the meaning of the statement, and the bearer of truth. But what is meaning in this context, and how can it bear truth? There is no agreed answer to these questions. The representation of intentional states as propositional attitudes does not explain intentionality. And there is danger that, in following Russell's proposal, we may forget that intentional states are about such things as events in the world—for example, having a drink—and not about the propositions that may be used to represent such intentional objects. To avoid these problems, phenomenologists have usually tried to analyze intentional phenomena directly rather than move to the level of the language that may be used to represent them. Others, most notably P. S. Churchland (1986), P. M. Churchland (1979, 1981, 1984), and Stephen Stich (1983), have argued that intentional phenomena such as beliefs and desires, which are of central importance in our commonsense, "folk psychology," will not find a place in the cognitive science of the future. In fact, "folk psychology" is doomed to go the way of "folk chemistry" and "folk physics."

The label *folk psychology* is obviously loaded. It suggests a primitive and stagnant psychological theory whose central theoretical concepts are as ill-founded as the concepts "phlogiston" and "caloric" proved to be. Folk psychology may be a "powerful predictive craft" in ordinary situations wherein its tacit assumption of the agent's rationality is justified. But we are not always rational. And there are cases of extreme irrationality. Folk psychology is worse than useless in the diagnosis and treatment of mental illness.

Another argument against the survival of folk psychology is the fact that we want to attribute cognitive processes to subhuman species and human infants when it would make no sense to attribute to them such intentional states as definite beliefs and desires. But this would suggest that intentional states are not basic and, therefore, will not be factors in basic

cognitive theory. Be that as it may, it is difficult to imagine what human discourse, including discourse about cognitive science, would be like in a language devoid of intentional idioms. Neither the social sciences nor such social practices as politics and law, as we know them today, would be possible (see Graham and Horgan 1988).

With these background considerations in hand, let us now preview the individual contributions to this anthology.

III

Paul M. Churchland presents a neuroscientific approach to the study of mind in "On The Nature of Theories: A Neurocomputational Perspective" (1990). He begins with a sketch and critique of the classical view of theories as a set of propositions, expressible in the first-order predicate calculus. He considers a number of alternatives before adopting a neurofunctional framework. This framework is based on recent insight into the functional significance of the brain's microstructure for the tasks of representation and computation. According to this perspective, propositional representation is not the most important form used by cognitive creatures. This seems to support the argument sketched earlier against the survival of folk psychological concepts or "propositional attitudes" in basic cognitive theory, at least at the neuroscientific level. Churchland examines AI models of elementary brainlike networks, considering how they might achieve representation and learning. He argues that the functional properties of such systems suggest a conception of "knowledge" or "understanding" that owes nothing to the sentential categories of folk psychology. He addresses the question of how faithfully these networks model the brain and concludes with an extremely optimistic prognosis for computational neuroscience.

William Bechtel and A. A. Abrahamsen defend folk psychology in "Connectionism and the Future of Folk Psychology." They answer arguments of eliminative materialists such as Churchland. They argue that contemporary work on neurally inspired connectionist or parallel distributed processing (PDP) models of cognition does not suggest a possible source of replacement for folk psychology, even though there are good reasons for rejecting the "propositional attitude" form of folk psychology that is a philosophical construction resting on very doubtful assumptions. Neither the PDP models nor any other accounts of internal processing threaten folk models, and even if such models should succeed in giving the correct account of internal processing, folk psychology will still be required to describe the informational capacities of cognitive agents interacting with

one another in real environments. Such descriptions provide an account of what internal processing models must explain.

Roger Thomas describes behavioral research on the conceptual abilities of squirrel monkeys in "Squirrel Monkeys, Concepts, and Logic." He employs a hierarchical scale that equates intelligence with learning ability. He reviews research on the animals' ability to "use" both absolute and relative class concepts and relational concepts involving logical operations. There is well-established precedent for using the basic operations of symbolic logic to define the structures of relational concepts in research on human concept learning. Higher levels in the hierarchy—for example, the level of the conditional—are defined in terms of lower levels, for example, conjunction and negation. Thomas uses an operational definition of conceptual behavior designed to assure that the behavior being studied actually demonstrates conceptual ability and not just simple forms of operant conditioning. He does this by using a wide variety of discriminative stimuli and avoiding the reuse of specific stimuli that have been associated with reinforcement. For example, a tree must be chosen because it is an exemplar of the concept "tree" and not because it is the same stimulus associated with reinforcement on an earlier trial. The subjects are also conditioned to "use" concepts such as conjunction, disjunction, and the conditional. But Thomas notes a potential impasse. He has been unable, so far, to design an experiment that shows that the behavior that supposedly demonstrates the use of conditional and biconditional concepts really does so, because it can also be interpreted in terms of simple conjunctions.

Robert McCauley and Thomas Lawson employ the concept of an artificial mind, the "idealized speaker-listener" of Chomskyan linguistics, in "Connecting the Cognitive and the Cultural." They begin with discussion of a number of problems in the study of sociocultural systems. To mention a few, such systems are difficult to perceive, individuate, and comprehend because of their magnitude, complexity, and constant flux. They cannot be studied in laboratories where experimental controls are most readily established. And finally, how is one to handle the idiosyncracies and intentionality of the individuals who participate in these systems? McCauley and Lawson advocate a stratagem for "neutralizing the sting of intentionality and idiosyncracy"; namely, the construction of idealized artificial minds. They explore alternative ways of constructing such artificial minds before adopting Chomsky's competence approach, which was originally developed as a method of theorizing in linguistics. They argue that the competence approach to theorizing can be adapted to the study of other symbolic-cultural systems in a variety of interesting ways that, coincidently, extend our knowledge of cognitive functions.

Ulric Neisser, who reviewed the state of the art in the classic

Cognitive Psychology (1967), sees cognitive science dividing into two alternative approaches: the computational or information-processing approach and the ecological approach. "Without Perception, There Is No Knowledge: Implications for Artificial Intelligence" is a critique of the former and a defense of the latter. He begins with an assessment of the progress of AI, noting its striking failures in the area of perception. AI has developed computers that play chess and expert systems that give advice, but such computers can neither see or hear, and in a very important sense, they do not really *know* anything. Neisser argues that perception provides the foundation for all other knowledge and that without perception other knowledge is not possible. Focusing on human vision, he distinguishes two distinct functions, *orientation* and *recognition.* Following J. J. Gibson (1979), he argues that orientation is "direct perception" and almost always veridical. Unlike recognition, orientation requires no symbolic representations, and therefore it cannot be understood within the perspective of mentalistic theories of cognition that assume that it does require them. In fact, the failed attempts to produce machine vision based on representational theories constitutes a plausibility argument for the alternative ecological theory. Neisser describes and illustrates Gibson's three levels of analysis in the study of perception: (1) the level of the environment, (2) the level of stimulus information, and (3) the level of the perceiver. He then develops a theory of direct visual orientation based on the pick up of coherent stimulus information by a perceiver whose visual system has "tuned" itself to "resonate" coherently to that optical structure. He argues against the widely held view that all perception is theory laden; his view is Darwinian, not Kantian. He concludes with the suggestion that such AI problems as machine vision will prove insoluble unless the problem of machine orientation is addressed and solved.

Edward Casey's "On the Phenomenology of Remembering: The Neglected Case of Place Memory" is inspired by Neisser's work on memory (1982) calling for the investigation of memory in natural contexts. Casey begins with an overview of six leading forms of nonrepresentational memory: reminding, reminiscing, recognizing, body memory, place memory, and commemoration. He then turns to the analysis of place memory, observing that it has been neglected far more than any of the other five forms and explaining why that is so. Casey's analysis of place memory involves the exploration of six factors: containment, boundary, horizon, protuberances, depth, and world. He then considers the question of computer simulation of human memory in general and the nonrepresentational forms in particular. He allows that there is a high probability of cogent machine simulation of the first three forms: reminding, recognizing, and reminiscing. He argues that the case

of body memory is moot, whereas the case of commemoration, involving such cultural phenomena as ritual, is too complex to try out even as a thought experiment at present. He focuses attention on place memory and argues that its primary parameters of horizon and world are neither determinate nor determinable to the degree necessary for computer simulation. This "argument from the indefeasible indeterminacy of horizon and world" is further elaborated in the course of ten concluding comments. The argument should be construed as a challenge to AI, psychology, and phenomenology alike rather than as an in-principle proof of impossibility. After all, we do manage to represent the indeterminate by the determinate in everyday discourse. Therefore, the indeterminate phenomena of place memory may provide a point of convergence for workers in all three fields.

Brian Cantwell Smith's "The Owl and the Electric Encyclopedia" is written as a review of a paper by Douglas Lenat and Edward Feigenbaum (L&F), "On the Thresholds of Knowledge" (1991). He begins by challenging their radical claim that "just a million frames, massaged by already-understood control structures, could intelligently manifest the sum total of human knowledge" (Smith 1991, 252, and this volume, p. 187). This provides the opportunity for Smith to develop a profile of his alternative view, which he calls *embedded computation* (EC). He does this by considering a dozen foundational questions from three different perspectives: that of traditional formal logic, that of L&F, and that of EC. Smith is in agreement with the other two perspectives on only one item, the central importance of reasoning and inference; EC differs from one or the other, and usually both of the other perspectives, on the remaining eleven foundational questions. The picture of EC that emerges is a view of computation that is not primarily focused on explicit representation. Instead, EC is physically embodied; it participates and acts in the world along with us; it has content that is situated in a context in which meaning is dependent on use; and it can support "original" semantics.

Robert Burton in the concluding chapter, "Reduction, Elimination, and Strategic Interdependence," defends the thesis that the differing approaches to the study of mind are interdependent, that they cannot and should not be pursued in provincial isolation from one another. He begins his argument by criticizing classical reduction theory and then by presenting William Wimsatt's alternative theory of emergence and reduction (1976; see also McCauley 1986, and Abrahamsen 1987). Wimsatt's theory provides the foundation for Burton's thesis of strategic interdependence. He then develops a systematic explication of some important forms of interdependence including some that are so commonplace they are not noticed but simply taken for granted, as well as others that are perceived to be so

threatening they are sometimes denied. These forms of interdependence are exemplified in four different research strategies that may be characterized as follows: (1) the explanation of higher level phenomena in terms of the lower, (2) the explanation of lower level phenomena in terms of the higher, (3) the investigation of phenomena at an intermediate level constrained by considerations from both higher and lower levels, and (4) the investigation of phenomena using models or methods borrowed from other disciplines at the same level. Burton exposes the various forms of interdependence that are implicit in these four strategies as they are exemplified in the contributions to this book. By showing exactly how the research strategies of the differing approaches to the study of mind are interdependent, Burton's analysis eliminates some of the sources of fear and hostility and opens the way for more fruitful cooperation.

References

Abrahamsen, A. A. 1987. "Bridging Boundaries versus Breaking Boundaries: Psycholinguistics in Perspective." *Synthese* 72:355–88.

Alston, W. 1971. "Varieties of Privileged Access." *American Philosophical Quarterly* 8 (no. 3): 223–41.

Bechtel, W. 1988. *Philosophy of Mind: An Overview for Cognitive Science.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Block, N. 1978. "Troubles with Functionalism." In *Perception and Cognition: Issues in the Foundations of Psychology,* edited by C. W. Savage, 261–325. Minneapolis: University of Minnesota Press.

Boring, E. G. 1950. *A History of Experimental Psychology.* New York: Appleton-Century-Crofts.

Brentano, F. 1973 [1874]. *Psychology from an Empirical Standpoint*, trans. A. C. Pancurello, D. B. Terrell, L. L. McAlister. New York: Humanities Press.

Burton, R. G. 1984. "B. F. Skinner's Accounts of Private Events: A Critique." *Journal for the Theory of Social Behaviour* 14 (no. 1): 125–40.

Chomsky, N. 1965. *Aspects of a Theory of Syntax.* Cambridge: MIT Press.

Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind.* Cambridge: Cambridge University Press.

———. 1981. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78:67–90.

———. 1984. *Matter and Consciousness.* Cambridge: MIT Press.

———. 1990. "On the Nature of Theories: A Neurocomputational Perspective." In *Minnesota Studies in the Philosophy of Science*, vol. 14: *Scientific Theories*, edited by C. W. Savage, 51–101. Minneapolis: University of Minnesota Press.

Churchland, P. S. 1986. *Neurophilosophy.* Cambridge: MIT Press.

Cohen, D. 1987. "Behaviourism." In *The Oxford Companion to the Mind,* edited by R. L. Gregory, 71–74. Oxford: Oxford University Press.

Creel, R. E. 1974. "Skinner's Copernican Revolution." *Journal for the Theory of Social Behaviour* 4 (no. 2): 131–45.

——. 1980. "Radical Epiphenomenalism: B. F. Skinner's Account of Private Events." *Behaviorism* 8 (no. 1): 31–53.

Dennett, D. C. 1981. "Why You Can't Make a Computer That Feels Pain." In *Brainstorms,* 190–229. Cambridge: MIT Press.

——. 1987. *The Intentional Stance.* Cambridge: MIT Press.

Dennett, D. C., and J. C. Haugeland. 1987. "Intentionality." In *The Oxford Companion to the Mind,* edited by R. L. Gregory, 383–86. Oxford: Oxford University Press.

Edelman, G. M. 1987. *Neural Darwinism,* New York: Basic Books.

Gibson, J. J. 1979. *The Ecological Approach to Perception.* Boston: Houghton Mifflin.

Graham, G., and T. Horgan. 1988. "How to Be Realistic About Folk Psychology." *Philosophical Psychology* 1:69–81.

Horgan, T., and J. Tienson. 1988. *Spindel Conference 1987: Connectionism and the Philosophy of Mind.* Supplement to *Southern Journal of Philosophy* 26.

Lenat, D. B., and Feigenbaum, E. A. 1991. "On the Thresholds of Knowledge." *Artificial Intelligence* 47 (nos. 1–3): 185–250.

McCauley, R. N. 1986. "Intertheoretic Relations and the Future of Psychology." *Philosophy of Science* 53 (no. 2): 179–99.

Nagel, T. 1974. "What Is It Like to Be a Bat?" *Philosophical Review* 83:435–50.

Neisser, U. 1967. *Cognitive Psychology.* Englewood Cliffs, N.J.: Prentice-Hall.

——. 1982. *Memory Observed.* San Francisco: W. H. Freeman.

Place, U. T. 1956. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47: 44–50. (Reprinted in *The Philosophy of Mind,* edited by V. C. Chappell, 101–9. Englewood Cliffs, N.J.: Prentice-Hall, 1962.)

Quine. W. V. O. 1961. "On What There Is." In *From a Logical Point of View,* 2d ed., 1–19. New York: Harper and Row.

Russell, B. 1940. *An Inquiry Into Meaning and Truth.* London: George Allen & Unwin.

Ryle, G. 1949. *The Concept of Mind.* New York: Barnes and Noble.

Skinner, B. F. 1938. *The Behavior of Organisms: An Experimental Analysis.* New York: Appleton-Century-Crofts.

——. 1945. "The Operational Analysis of Psychological Terms." *Psychological Review* 52:270–77, 291–94.

——. 1953. *Science and Human Behavior.* New York: Macmillan.

——. 1957. *Verbal Behavior.* New York: Appleton-Century-Crofts.

——. 1971. *Beyond Freedom and Dignity.* New York: Alfred A. Knopf.

——. 1974. *About Behaviorism.* New York: Alfred A. Knopf.

——. 1977. "Why I Am Not a Cognitive Psychologist." *Behaviorism* 5 (no. 2): 1–10.

Smart, J. J. C. 1959. "Sensations and Brain Processes." *Philosophical Review* 68: 141–156. (Reprinted in *The Philosophy of Mind,* edited by V. C. Chappell, 160–72. Englewood Cliffs, N.J.: Prentice-Hall, 1962.)

Smith, B. C. 1991. "The Owl and the Electric Encyclopedia," *Artificial Intelligence* 47 (nos. 1–3): 251–88.

Stich, S. P. 1983. *From Folk Psychology to Cognitive Science.* Cambridge: MIT Press.

Tienson, J. 1988. "An Introduction to Connectionism." *Spindel Conference 1987: Connectionism and the Philosophy of Mind.* Supplement to *Southern Journal of Philosophy* 26:1–16.

Watson, J. B. 1914. *Behavior: An Introduction to Comparative Psychology.* New York: Holt.

——. 1919. *Psychology from the Standpoint of a Behaviorist.* Philadelphia: Lippincott.

Wiener, N. 1961. *Cybernetics.* Cambridge: MIT Press.

Wimsatt, W. C. 1976. "Reductionism, Levels of Organization, and the Mind-Body Problem." In *Consciousness and the Brain,* edited by G. Globus, G. Maxwell, and I. Savodnik, 205–67. New York: Plenum Press.

Winston, P. H. 1984. *Artificial Intelligence.* 2d ed. Reading, Mass.: Addison-Wesley.

Wittgenstein, L. 1961 [1921]. *Tractatus Logico-philosophicus.* Translated by D. F. Pears and B. F. McGuiness. London: Routledge and Kegan Paul.

——. 1953 [1945–49]. *Philosophical Investigations.* Translated by G. E. M. Anscombe. New York: Macmillan.