Mind

Introduction to Cognitive Science

second edition

Paul Thagard

9 Brains

You probably know that your ability to think depends on your brain, but understanding of how processes in the brain contribute to thought is relatively new. Some ancient Greek philosophers believed that the brain is one of three organs of thought, along with the heart and liver. But Aristotle argued that the brain is merely a cooling device for the heart, which he took to be the center of intellectual and perceptual functions (Finger 1994). This chapter reviews the main methods for investigating the nature of brains and how they contribute to thought, ranging from looking at the effects of brain damage to using machines to scan brain activities. It then discusses how discoveries about brain processes have enriched our understanding of the representations and computations that produce thinking. The chapter then considers the relevance of molecular processes involving neurotransmitters for understanding the relation between thought and brain, and discusses the practical applicability of knowledge about the brain, especially in the treatment of mental illness.

How Brains Are Studied

Brain Structure and Lesions

The first important method for studying brains was dissection, in which brains were carefully cut apart to reveal their anatomical structure. In the second century A.D., the Roman physician Galen described many brain structures based on the dissection of animals such as cows and baboons. The results of dissections on human brains were not reported until the sixteenth century when Vesalius provided detailed anatomical descriptions. He thought that the ventricles, which are open spaces in the brain, are crucial for thought because they produce "animal spirit" that is distributed

to the nerves. Studying the anatomy of the brain does not by itself reveal much about its physiology—how it works.

Insights into the physiology of the brain and its relation to thought came about by the method of studying *lesions*, which are injuries to specific parts of the body. Lesions of the brain can occur naturally because of tumors, blood clots, or accidents; or they can be produced by cutting or burning. In the eighteenth century, lesion experiments on dogs revealed that breathing depends on a brain area called the *medulla*: damage to the area causes severe breathing problems. Early recognition that a human cognitive function depends on a specific brain area came in the 1860s when Paul Broca attributed a patient's inability to use language to a specific part of the brain's frontal lobes now called Broca's area. Since then, the contributions of many specific parts of the brain to particular cognitive functions have been discovered, as illustrated in figure 9.1. For example, people with damage to the hippocampus have difficulty forming new memories, and damage to the amygdala can cause inability to feel fear and other emotions.

Electrical Recording and Stimulation

In 1875, Richard Caton reported the existence of electric currents in the brain that vary with different stimuli presented. This made it possible to
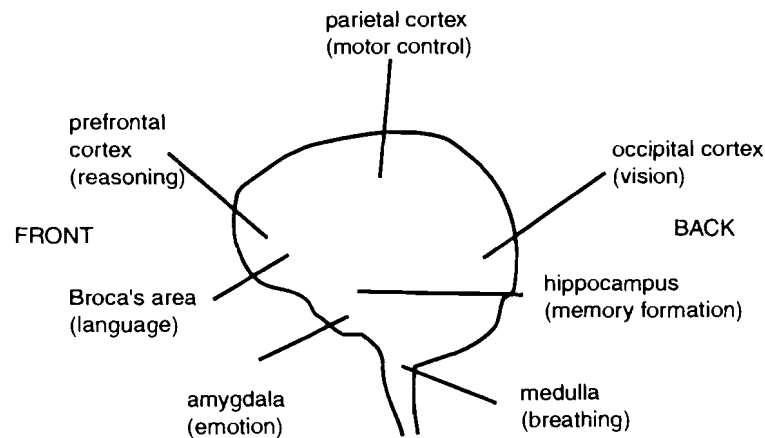

Figure 9.1
Some important brain areas, with their associated functions. For much more detailed maps, see the web sites listed at the end of this chapter.

determine what parts of the brain are active without having to rely on lesions. Today, electrical activity in the brain can be recorded by a machine called an *electroencephalograph*, or EEG for short. With an EEG, electrodes attached to the head detect electrical activity in different parts of the brain, and a history of this activity is transferred to a computer. EEGs can be used to monitor cognitive functions, and to diagnose brain disorders such as epilepsy, which is brought about by an abnormal electrical discharge of brain cells.

EEGs identify electrical activity in large regions of the brain, but the electrical activity of particular neurons can be identified using the technique of *single cell recording*. Electrodes are inserted into the brain to record the firing activity of specific neurons. For example, there are neurons in monkeys that respond to the texture of fat in the mouth, with especially high firing rates when the monkeys are given heavy cream (Rolls 1999, 34). Single cell recording is too invasive to be used routinely on humans, but is sometimes used during brain surgery to identify the contribution of particular neurons to specific cognitive functions.

Electrical activity in the brain can be stimulated as well as recorded—for example, during surgery by applying current to exposed parts of the brain. Less invasively, *transcranial electronic stimulation* is performed when electrodes are placed on the head to make current flow through the brain. Alternatively, *transcranial magnetic stimulation* applies powerful electromagnets to stimulate or disrupt brain activity.

Brain Scans

Whereas EEGs can identify activity in large brain regions and single cell recording applies to individual neurons, neither method reveals much about small brain areas. In contrast, modern brain imaging technologies can identify activity in areas a few millimeters (less than 1/8 of an inch) across, comprising a few million out of the 100 billion or so neurons in the brain. The two technologies that are currently most useful are *positron emission tomography* (*PET*) and *magnetic resonance imaging* (*MRI*).

In a PET scan, the subject is first injected with a radioactive material that spreads through the bloodstream. The most active parts of the body, including particular brain regions, require increased blood flow to nourish the cells that are producing a given activity. The PET scanner detects increases in radioactivity due to an increase in blood flow, thereby

identifying groups of cells such as neurons that are most active. When you are given a particular mental task to perform, the PET scan is able to detect what parts of your brain are used to perform it. For example, if you are given a picture to look at, the PET scan will show increased blood flow to the primary visual cortex located at the back of your brain.

PET scanning has a number of limitations, including the use of radioactive material and the inability to localize activity to regions smaller than a cubic centimeter. Hence, it is now less commonly used than magnetic resonance imaging, which was originally developed to display the structure of parts of the body. With an MRI, the subject is inserted into a large magnet that generates signals from the hydrogen nuclei of water molecules in the body. The MRI machine detects these signals and uses computers to distinguish physical structures based on the different signals that they generate. For example, MRIs are often used to diagnose sports injuries by detecting changes to joints and other structures.

To investigate brain processes, specific magnetic pulses can be generated that enable the detection of changes in blood flow in a person performing a mental task; this is called *functional MRI*, or *fMRI*. Hence fMRI, like PET scans, can be used to identify brain regions with increased blood flow corresponding to increased neuronal activity. Images of the brain can be produced in a few seconds with a spatial resolution of a few millimeters. Unfortunately, fMRI does not have the temporal resolution of an electrical recording technique such as the EEG, which can detect changes in brain activity that take much less than a second. Nevertheless, fMRI studies have become crucial for helping to identify the specific brain regions involved in various kinds of thinking. For a history and review of techniques for brain mapping and imaging, see Savoy 2001 and Posner and Raichle 1994.

Now that techniques are available for identifying activity of brain regions and even of single neurons, do we still need the computational-representational understanding of mind? Why not explain thinking directly in terms of neuronal activity without talking about rules, concepts, and other representations? Why not focus on the kinds of physical processes found in the brain rather than on computational processes? To answer these questions, we need to examine how brains exhibit representational and computational properties.

How Brains Represent

Spiking Neurons

A representation is a structure that stands for something by virtue of relations such as similarity, causal history, and connections with other representations. For example, a photograph of you is a representation of you because it looks like you and because photography causally links it with you. The word "cat" is not similar to cats, but there is a causal link between utterances of this word and the presence of cats, as well as relations between the concept *cat* and other concepts. Let us look at how individual neurons and especially groups of neurons can serve as representations.

The artificial neurons (units) discussed in chapter 7 represent aspects of the world by means of numbers called activations that correspond roughly to the firing rates of real neurons. A typical neuron may fire hundreds of times per second, and we can think of it as representing a degree of presence or absence of what it represents. For example, if a unit represents the concept *cat*, then its firing many times per second signifies the presence of a cat. However, all natural and most artificial neural networks use distributed representations in which concepts are encoded by a population of neurons: a group of neurons represents a concept by virtue of a pattern of firing rates in all of the neurons. Thus a group of neurons, each with its own firing rate, can encode a large number of aspects of the world.

Focusing on firing rates, however, seriously underestimates the representational capacity of neurons and groups of neurons. The *spike train* of a neuron is its pattern of firing or not firing over a period of time. We can represent a spike train by a sequence of 1s (firing) and 0s (not firing). The spike train 10100 and 00011 both involve a neuron with a firing rate of 2 times out of 5, but they are different patterns. There are far more different spike trains than firing rates (see the notes at the end of this chapter). Thus, a group of neurons with varying spike trains has the capacity to encode an enormously large number of features of the world. See Maass and Bishop 1999 for analysis of the representational and computational capacities of spiking neurons, and Eliasmith and Anderson 2003 for an elegant analysis of neural representation.

We have seen that a single neuron can represent a feature of the world as the result of being tuned to fire more rapidly when that feature is

presented. More powerful neural representations arise if the neuron can encode more possibilities by using the temporal properties of different spike trains, and if the neuron is part of a population of neurons that work together to represent many features. In sum, a representation in the brain is a population of neurons whose firing patterns encode information by virtue of having acquired regular responses to particular kinds of input.

Brain Maps

The brain does not try to use all of its billions of neurons to represent everything; different brain regions represent different kinds of sensory stimuli. For example, the visual cortex at the back of the brain has neurons that respond to different visual inputs. There are neuronal groups whose firing patterns correspond spatially to the structure of the input—for example, when a column of neurons fires together to represent the fact that a line is part of the visual stimulus. Thus, different parts of the brain have groups of neurons that fire when different kinds of visual, olfactory (smell), taste, auditory, and tactile stimuli are presented. The human brain can do a lot more than just represent stimuli presented to it, because a group of neurons can respond to inputs from many groups of neurons. This can produce a combined representation of what the input neurons represent. For example, there are regions in the frontal cortex of monkeys where the sensory modalities of taste, vision, and smell converge, enabling the representation of fruits and their key properties (Rolls 1999). It is clear, therefore, that the brain is a superb representational device.

How Brains Compute

Transformations

But is the brain a computer? The most familiar kinds of computation involve rules for transforming symbols—for example, calculating that 2 + 2 = 4 and inferring q from p and if p then q. Such computations are transformations of representations. The brain can also be viewed as performing transformations of representations encoded by the firing patterns of neurons. In general, a physical system is a computational system "when its physical states can be seen as representing states of some other systems, where transitions between its states can be explained as operations on the representations" (Churchland and Sejnowski 1992, 62). To put it in a

slogan: *No computation without representation*. From this perspective, digital computers and brains are two different kinds of computational system.

Consider, for example, the operations of your visual system. At the back of your eye is the *retina*, millions of cells that are sensitive to light. Retinal cells respond to light reflected from objects into your eye, and send signals through a series of layers to neurons in the visual cortex. Successive layers in the cortex detect more and more complex aspects of the objects that originally sent light into the eye, as neurons in each layer abstract and transform the firing patterns of neurons in the preceding layer. Thus, the visual cortex progressively constructs representations of lines, patterns in two dimensions, and finally three-dimensional colored objects.

The brain transforms neuronal representations into new ones by means of synaptic connections. As we saw in chapter 7, the firing of one neuron can excite or inhibit the firing of another neuron. Hence, one group of neurons with its patterns of firing can alter the patterns of firing of another group of neurons to which it is connected by means of synapses between pairs of neurons. There can also be feedback connections from one group of neurons to another, enabling them to influence each other. The brain contains many such feedback influences. In general, computation in the brain consists of interactions between groups of neurons that produce transformations of firing patterns.

Integration

The brain's operation is much more complicated than simply taking sensory input and transforming it. In order to eat a banana, a monkey needs to combine visual, tactile, and other kinds of information about it, and then use this integrated information to guide actions such as ingesting it. Hence, much of what the brain does involves operations in central brain areas that combine information from multiple other areas. At the level of individual neurons, we can describe computation in terms of the ideas about activation, excitation, and inhibition presented in chapter 7, but a full understanding of the computational accomplishments of brains requires attention to the higher level operations of transformation and integration just described. Figure 9.2 depicts some of the interconnections of the prefrontal cortex (the front of the front of the cortex) with many other brain areas. Chapter 10 describes a computational model of emotion according to which the brain makes emotional judgments by combining
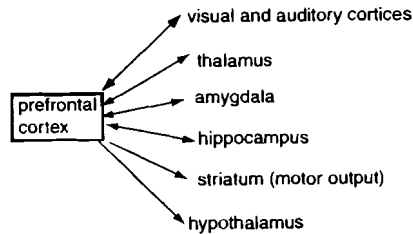
Figure 9.2
Inputs and outputs from prefrontal cortex. Based on Groenewegen and Uylings (2000). Note that connections usually go in both directions. Many connections between the other areas are not shown. See figure 9.1 for a better picture of the spatial organization of the brain.

information from the frontal cortex (high-level thought), the amygdala (bodily information), and the hippocampus (memory).

Learning

One of the most impressive computational accomplishments of neural networks is learning, in which changes in the synaptic weights between neurons produce major improvements in the performance of the network. However, networks trained by the backpropagation algorithm discussed in chapter 7 exhibit a problem called *catastrophic interference* (McCloskey and Cohen 1989). This happens when a network is trained to perform one task, such as forming associations between words, and then trained to perform another similar task. When people undergo such retraining, they usually experience only some loss of the ability to perform the first task, but artificial neural networks can suffer a dramatic drop in performance on the first task when they learn the second.

McClelland, McNaughton, and O'Reilly (1995) argue that the brain's solutions to this problem is to have two complementary learning systems in two different brain regions, the hippocampus and the neocortex (the most recently evolved part of the cortex). The hippocampal system permits rapid learning of new items, whereas the neocortex learns slowly by small adjustments of synaptic strengths through something like the backpropagation algorithm. As illustrated in figure 9.3, initial storage of most information takes place in the hippocampus, and is only gradually consolidated in the neocortical system. Catastrophic interference is avoided because new
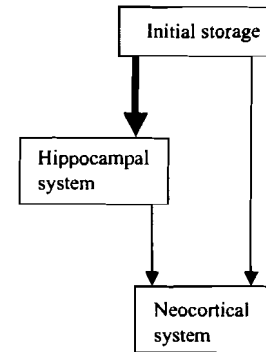
Figure 9.3
A two-system model of memory storage. Adapted with permission from McClelland, McNaughton, and O'Reilly 1995, p. 444.

information coming into the hippocampus has only a small and gradual effect on the neocortex, which retains most of what it already knew. Because learning depends on two different interconnected systems operating with different learning procedures, understanding of neural computation requires specification of the operations of multiple regions.

Thus, it is not metaphorical to say that the brain represents and computes, even though its computations are done differently from the kinds most familiar to us in modern digital computers. Understanding of how brains work requires attention to the roles played by particular brain regions such as the hippocampus and prefrontal cortex. Hence, we should not think of the brain as one big connectionist network of the sort described in chapter 7, but rather as a highly organized and interconnected system of specialized neural networks. This section has only begun to sketch the computational activities of brains; see Churchland and Sejnowski 1992 and Eliasmith and Anderson 2003 for much more detailed discussions.

How Molecules Matter

If you have you ever had a cup of coffee or an alcoholic drink, then you have experienced the effects of chemicals on the brain. Coffee contains caffeine, which blocks the action of the molecule adenosine, which makes

people drowsy by inhibiting the firing of some neurons. Hence, caffeine increases neuronal activity and keeps you awake. In contrast, alcohol can disrupt mental functioning by inhibiting the action of the molecule glutamate, which excites neurons. Both caffeine and alcohol also increase activity of the chemical dopamine, which produces feelings of pleasure. Adenosine, glutamate, and dopamine are all neurotransmitters, molecules that enable one neuron to influence another.

All the models of neurons and brains described earlier in this book are based on electrical activity: neurons fire and provide electrical inputs to other neurons. However, as caffeine and alcohol illustrate, the direct effects of real neurons on each other are chemical as well as electrical, in that molecules are emitted from one neuron and then passed over to another neuron, where they initiate chemical reactions that generate the electrical activity of the stimulated neuron. Figure 9.4 depicts how neurotransmitters are passed from the axon of one neuron to the dendrite of another neuron.

There are dozens of neurotransmitters operating in the human brain, some with excitatory and others with inhibitory effects. This operation is consistent with the general connectionist ideas described in chapter 7, which assumed excitatory and inhibitory links between neurons. But broader chemical effects on neural computation are produced by hormones such as estrogen and testosterone, which can affect the firing of neurons independent of direct connections. A neuron in one part of the brain such as the hypothalamus may fire and release a hormone that travels to a part of the body such as the adrenal glands, which stimulates the release of other hormones that then travel back to the brain and influence the firing of different neurons. Complex feedback loops can result, involving interactions between the neurotransmitter control of hormone
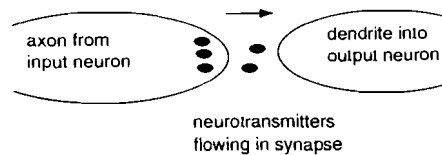

neurotransmitters
flowing in synapse

Figure 9.4
Neurotransmitter molecules flowing from one neuron into another.

release and the hormonal regulation of neurotransmitter release. These feedback loops can also involve the immune system, because brain cells also have receptors for cytokines, which are protein messengers produced by immune system cells such as macrophages. Thus, attention to the neurochemistry of hormones shows an important limitation to connectionist models, in that whether a neuron fires is not just a function of neurons that have synaptic inputs to it (see Thagard 2002 for further discussion).

Practical Applicability

Motivations for studying brains are both theoretical—how does it work?—and practical—how can we help it to work better? Many advances have been made in understanding how mental problems can arise from defects in the functioning of different brain areas and neurotransmitters. For example, children with attention deficit disorder have difficulty concentrating and often fall behind in school. They are generally treated with Ritalin, which stimulates areas of the brain involved in filtering information by increasing the activity of neurotransmitters such as dopamine.

The growth of different brain areas can help to explain variations in behavior as people age. In the past decade, brain scans of children and teenagers have revealed that the brain undergoes remarkable growth at roughly the ages of 1–2 and 11–12 years. The latter growth spurt was a big surprise to researchers, and is generating new explanations of why many teenagers display difficult and risky behavior (Strauch 2003). Areas of the prefrontal cortex are not fully developed until the early twenties, so that teenagers' decisions are often heavily driven by emotional information in the amygdala, rather than by reasoning about potential risks. They may, for example, take drugs such as ecstasy and cocaine, which provide short-term pleasure by intensely stimulating production of dopamine, but lead to addiction as the result of depletion of dopamine receptors that produces cravings for higher and higher doses.

Neurotransmitters and brain areas are also relevant to explaining many mental illnesses. Schizophrenia, in which people lose touch with reality because of bizarre beliefs and hallucinations, is associated with excess dopamine activity in the prefrontal cortex. Drugs that block dopamine reduce the symptoms of schizophrenia. On the other hand, lack of dopamine can produce the problems with motor control found in

Parkinson's disease. Depression is often treated with Prozac and other drugs that increase the availability of the excitatory neurotransmitter serotonin by decreasing its reuptake in synapses.

Summary

The early decades of cognitive science, and even the connectionist models of the 1980s, largely ignored how brains produce thinking. But since the 1990s the brain challenge has been increasingly answered by experimental and computational investigations of how brains work. Brain scanning techniques such as PET and fMRI have provided a huge amount of information concerning how different brain regions contribute to various cognitive functions. Computational models of the brain have become biologically richer, both with respect to employing more realistic neurons such as ones that spike, and with respect to simulating the interactions between different areas of the brain such as the hippocampus and the cortex. These models are not strictly an alternative to computational accounts in terms of logic, concepts, rules, images, and connections, but should mesh with them and show how mental functioning performs at the neural level. The remarks on neurological plausibility in chapters 2–6 show that such meshing is rapidly progressing.

Moreover, there is increasing understanding of the chemical functioning of brains, in particular how different neurotransmitters and hormones affect neuronal firing. These advances do not require abandonment of the general view of thinking as representation and computation, but they do show the need to expand and supplement earlier cognitive theories.

Understanding of brain mechanisms is invaluable for explaining and treating mental illness. The explanation schema for kinds of mental illness such as schizophrenia is as follows:

Explanation target

Why do people have a particular kind of mental illness?

Explanatory pattern

People usually have normal brain function that involve identifiable brain structures and neurotransmitters.

Defects to these brain structures and chemical processes can disrupt normal functioning.

These defects produce the symptoms of the mental illness.

Many mental illnesses are treated by means of drugs that help to restore or approximate the normal functioning of chemical-electrical processes in the brain. The dramatic advances in neuroscience of recent decades do not by themselves solve the mind–body problem, since a skeptic can always maintain that there are aspects of mind such as consciousness that will never succumb to scientific explanation. But we will see in the next two chapters that even questions about the nature of emotions and consciousness are increasingly yielding to neurological investigation.

Discussion Questions

1. How do different methods reveal different aspects of brain structure and function?

2. What is the relation between brain structures and the kinds of representation discussed in chapters 2–7?

3. Is brain processing really computation?

4. To understand brains, is it necessary to move down to the molecular and chemical level?

5. What aspects of thinking seem to be hardest to explain in terms of brain structures and processes?

Further Reading

See Finger 1994 for a readable history of neuroscience. Kandel, Schwartz, and Jessell 2000 is a standard neuroscience textbook. Allman 1999 discusses brain evolution. O'Reilly and Munakata 2000 describes computational models for cognitive neuroscience. Churchland 2002 and Bechtel et al. 2001 provide philosophical discussions of neuroscience.

Web Sites

Explore the brain and spinal cord: http://faculty.washington.edu/chudler/introb.html

Science Daily (mind and brain news): http://www.sciencedaily.com/news/mind_summaries.php

The Whole Brain Atlas: http://www.med.harvard.edu/AANLIB/home.html

Notes

If a neuron fires up to 100 times per second, then there are 101 possible firing rates it can have, including 0. The average firing rate of the neuron might be, for example, 25 or 50 times per second. A group of 1000 neurons can then represent a huge number of possibilities, 100^{1000}. For spiking neurons, there are 2^{100} different possible spike trains compared to only 101 different firing rates. The number of different possible combinations of spike trains in a group of 1000 neurons is therefore astronomical: $(2^{100})^{1000}$.

Chapter 3 mentions the debate between the view that many rules and concepts are innate and the alternative view that emphasizes learning. A related debate concerns the extent to which brain structures have been selected by biological evolution for specific functions. Evolutionary psychologists claim that the brain contains a large number of evolved computational devices that are specialized in function, such as a face recognition system, a language acquisition device, navigation specializations, and a routine for detecting cheaters in social situations (Cosmides and Tooby 1999). In contrast, Quartz and Sejnowski (2002) argue that the brain has evolved to make possible flexible learning: the main representational features of cortex are built from the dynamic interaction between neural growth mechanisms and environmentally derived neural activity.

10 Emotions

How do you feel about the following items? *Group A*: death, cancer, poison, traffic tickets, insults, vomit. *Group B*: lottery winnings, fine restaurants, sex, victories, babies, parties. For most people, the things in group A are associated with negative emotions such as sadness, fear, and anger, whereas the things in group B are associated with happiness and pleasure. If you think of the main events of your day so far, you will probably be able to recall the emotions that accompanied them—for example, the joy you felt when your sports team won, or the worry you felt when you realized that exams are coming soon.

Traditionally, cognitive science ignored the study of emotions, seeing it as a side issue to the more central study of cognition. Philosophers as far back as Plato have tended to view emotion as a distraction or impediment to effective thought. In the past decade, however, there has been a dramatic increase in the appreciation of the relevance of emotion to cognition, particularly with respect to decision making. On the old view, decisions can be made either rationally or emotionally, and cognitive science is mainly concerned with the rational ones. In contrast, the view now emerging is that emotions are an inherent part of even rational decision making. This chapter will describe how emotions contribute to both representation and computation. But first we need to discuss the nature of emotion.

What Are Emotions?

Everyone is familiar with basic emotions such as happiness, sadness, fear, anger, disgust, and surprise. But there is much disagreement among cognitive scientists concerning the nature of emotions. Favored theories fall

Synapse Space in which a signal passes from one *neuron* to another.

Target analog Set of objects, properties, and relations that can be learned about by comparison to a source analog.

Theory Set of hypotheses that explain observations.

Thought experiment Use of the imagination to investigate nature.

Ventromedial prefrontal cortex The bottom-middle part of the *prefrontal cortex*.

Whorf hypothesis Conjecture that language determines how we perceive and think about the world.

References

Abraham, R. H., and C. D. Shaw. 1992. *Dynamics: The geometry of behavior.* 2nd ed. Redwood City, Calif.: Addison-Wesley.

Adams, M. J. 1990. *Beginning to read.* Cambridge, Mass.: The MIT Press.

Aitchison, J. 1987. *Words in the mind: An introduction to the mental lexicon.* Oxford: Basil Blackwell.

Akmajian, A., R. A. Demers, A. K. Farmer, and R. M. Harnish. 2001. *Linguistics: An introduction to language and communication.* 5th ed. Cambridge, Mass.: The MIT Press.

Allen, R. H., ed. 1992. *Expert systems for civil engineers: Knowledge representation.* New York: American Society of Civil Engineers.

Allman, J. M. 1999. *Evolving brains.* New York: Scientific American Library.

Anderson, J. A., and E. Rosenfeld, eds. 1988. *Neurocomputing.* Cambridge, Mass.: The MIT Press.

Anderson, J. A., and E. Rosenfeld, eds. 1998. *Talking nets: An oral history of neural networks.* Cambridge, Mass.: The MIT Press.

Anderson, J. R. 1983. *The architecture of cognition.* Cambridge, Mass.: Harvard University Press.

Anderson, J. R. 1993. *Rules of the mind.* Hillsdale, N.J.: Erlbaum.

Anderson, J. R. 2000. *Cognitive psychology and its implications.* 5th ed. New York: Worth.

Anderson, J. R., Y. Qin, V. A. Stenger, and C. S. Carter. Forthcoming. The relationship of three cortical regions to an information-processing model. *Cognitive Neuroscience* 16, 637–653.

Asada, M. et al. 2003. An overview of RoboCup-2002 Fukuoka/Busan. *AI Magazine* 24 (2, summer), 21–40.

Ashby, F. G., and E. Walrdron. 2000. The neuropsychological bases of category learning. *Current Directions in Psychological Science* 9, 10–14.

Barnes, A., and P. Thagard. 1997. Empathy and analogy. *Dialogue: Canadian Philosophical Review* 36, 705–720.

Barrett, R. A. 1991. *Culture and conduct: An excursion in anthropology.* 2nd ed. Belmont, Calif.: Wadsworth.

Barsalou, L. W. 1983. Ad hoc categories. *Memory and Cognition* 11, 211–227.

Barsalou, L. W., W. K. Simmons, A. K. Barbey, and C. D. Wilson. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 7, 84–91.

Bartlett, F. C. 1932. *Remembering.* Cambridge: Cambridge University Press.

Barton, S. 1994. Chaos, self-organization, and psychology. *American Psychologist* 49, 5–14.

Bates, E. A., and J. L. Elman. 2002. Connectionism and the study of change. In M. H. Johnson, Y. Munakata, and R. O. Gilmore, eds., *Brain development and cognition,* 2nd ed., 420–440. Oxford: Blackwell.

Bechtel, W., and A. A. Abrahamsen. 2002. *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks.* 2nd ed. Oxford: Basil Blackwell.

Bechtel, W., and G. Graham, eds. 1998. *A companion to cognitive science.* Malden, Mass.: Blackwell.

Bechtel, W., P. Mandik, J. Mundale, and R. S. Stufflebeam, eds. 2001. *Philosophy and the neurosciences: A reader.* Malden, Mass.: Blackwell.

Benington, J. H., and H. C. Heller. 1995. Restoration of brain energy metabolism as the function of sleep. *Progress in neurobiology* 45, 347–360.

Bergmann, M., J. H. Moor, and J. Nelson. 2003. *The logic book.* 4th ed. Columbus Ohio: McGraw-Hill.

Blackmore, S. 2004. *Consciousness: An introduction.* Oxford: Oxford University Press.

Block, N. 1978. Troubles with functionalism. In C. W. Savage, ed. *Perception and cognition.* Minneapolis, Minn.: University of Minnesota Press.

Block, N., O. Flanagan, and G. Güzeldere, eds. 1997. *The nature of consciousness: Philosophical debates.* Cambridge, Mass.: The MIT Press.

Boroditsky, L. 2003. Linguistic relativity. In L. Nadel, ed., *Encylopedia of cognitive science.* Vol. 2, 917–921. London: Nature Publishing Group.

Boroojerdi, B., M. Phipps, L. Kopylev, C. M. Wharton, L. G. Cohen, and J. Grafman. 2001. Enhancing analogic reasoning with rTMS over the left prefrontal cortex. *Neurology* 56, 526–528.

Braine, M. D. S. 1978. On the relation between the natural logic of reasoning and standard logic. *Psychological Review* 85, 1–21.

Braine, M. D. S., and D. P. O'Brien. 1998. *Mental logic.* Mahwah, N.J.: Erlbaum.

Brewer, W. F., and J. C. Treyens. 1981. Role of schemata in memory for places. *Cognitive Psychology* 13, 207–230.

Brooks, R. A. 1991. Intelligence without representation. *Artificial Intelligence* 47, 139–159.

Brooks, R. A. 2002. *Flesh and machines: How robots will change us.* New York: Pantheon.

Bruer, J. T. 1993. *Schools for thought: A science of learning in the classroom.* Cambridge, Mass.: The MIT Press.

Bruner, J. S. 1990. *Acts of meaning.* Cambridge, Mass.: Harvard University Press.

Bruner, J. S., J. J. Goodnow, and G. A. Austin. 1956. *A study of thinking.* New York: Wiley.

Buchanan, B., and E. Shortliffe, eds. 1984. *Rule-based expert systems.* Reading, Mass.: Addison-Wesley.

Busemeyer, J. R., and J. T. Townsend. 1993. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Revew* 100, 432–459.

Cacioppo, J. T. et al., eds. 2002. *Foundations in social neuroscience.* Cambridge, Mass.: The MIT Press.

Card, S. K., T. P. Moran, and A. Newell. 1983. *The psychology of computer-human interaction.* Hillsdale, N.J.: Erlbaum.

Carruthers, P. 2000. *Phenomenal consciousness: A naturalistic theory.* Cambridge: Cambridge University Press.

Chalmers, D. J. 1996. *The conscious mind.* Oxford: Oxford University Press.

Cheng, P. W., and K. J. Holyoak. 1985. Pragmatic reasoning schemas. In *Cognitive Psychology* 17, 391–416.

Chi, M. 1992. Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere, ed., *Cognitive models of science,* 129–186. Minneapolis, Minn.: University of Minnesota Press.

Chomsky, N. 1957. *Syntactic structures.* The Hague: Mouton.

Chomksy, N. 1959. Review of *Verbal behavior,* by B. F. Skinner. *Language* 35, 26–58.

Chomsky, N. 1972. *Language and mind.* 2nd ed. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1980. *Rules and representations*. New York: Columbia University Press.

Chomsky, N. 1988. *Language and problems of knowledge*. Cambridge, Mass.: The MIT Press.

Chomsky, N. 2002. *On nature and language*. Cambridge: Cambridge University Press.

Christoff, K., V. Pabhakaran, J. Dorfman, Z. Zhao, J. K. Kroger, K. J. Holyoak, and J. D. K. Gabrieli. 2001. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage* 14, 1136–1149.

Churchland, P. M. 1989. *A neurocomputational perspective*. Cambridge, Mass.: The MIT Press.

Churchland, P. M. 1995. *The engine of reason: The seat of the soul*. Cambridge, Mass.: The MIT Press.

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

Churchland, P. S. 2002. *Brain-wise: Studies in neurophilosophy*. Cambridge, Mass.: The MIT Press.

Churchland, P. S., and T. Sejnowski. 1992. *The computational brain*. Cambridge, Mass.: The MIT Press.

Clark, A. 1997. *Being there: Putting brain, body, and world together again*. Cambridge, Mass.: The MIT Press.

Clark, A. 1999. An embodied cognitive science. *Trends in Cognitive Sciences* 3, 345–351.

Clark, A. 2001. *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press.

Conati, C., A. Gertner, and K. Vanlehn. 2002. Using Bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Instruction* 12, 317–417.

Cooper, L. A., and R. N. Shepard. 1973. Chronometric studies of the rotation of mental images. In W. G. Chase, ed., *Visual information processing*, 75–176. New York: Academic Press.

Cosmides, L., and J. Tooby. 1999. Evolutionary psychology. In R. A. Wilson and F. C. Keil, eds., *The MIT encyclopedia of the cognitive sciences*, 295–298. Cambridge, Mass.: The MIT Press.

Costello, F. J., and M. T. Keane. 2000. Efficient creativity: constraint-guided conceptual combination. *Cognitive Science* 24, 299–349.

Crick, F. 1994. *The astonishing hypothesis: The scientific search for the soul*. London: Simon and Schuster.

Crick, F., and C. Koch. 1998. Consciousness and neuroscience. *Cerebral Cortex* 8, 97–107.

Croft, D., and P. Thagard. 2002. Dynamic imagery: A computational model of motion and visual analogy. In L. Magnani, ed., *Model-based reasoning: Scientific discovery, technological innovation, values*, 259–274. New York: Kluwer/Plenum.

Crowley, K., and R. S. Siegler. 1993. Flexible strategy use in children's tic-tac-toe. *Cognitive Science* 17, 531–561.

Damasio, A. R. 1994. *Descartes' error*. New York: Putnam.

Damasio, A. R. 1999. *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace Jovanovich.

D'Andrade, R. G. 1995. *The development of cognitive anthropology*. Cambridge: Cambridge University Press.

Davidson, R. J., D. Pizzagalli, J. B. Nitschke, and K. Putnam. 2002. Depression: Perspectives from affective neuroscience. *Annual Review of Psychology* 53, 545–574.

Davidson, R. J., K. R. Scherer, and H. H. Goldsmith, eds. 2003. *Handbook of affective sciences*. New York: Oxford University Press.

Dawson, M. R. W. 1998. *Understanding cognitive science*. Oxford: Blackwell.

Dean, T. L., and M. P. Wellman. 1991. *Planning and control*. San Mateo, Calif.: Morgan Kaufmann.

Dehaene, S., E. Spelke, P. Pinel, R. Stanescu, and S. Tsivkin. 1999. Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science* 284, 970–974.

Dennett, D. 1991. *Consciousness explained*. Boston: Little, Brown.

Dietrich, E., ed. 1994. *Thinking computers and virtual persons: Essays on the intentionality of machines*. San Diego, Calif.: Academic Press.

Dourish, P. 2001. *Where the action is: The foundations of embodied interaction*. Cambridge, Mass.: The MIT Press.

Dreyfus, H. L. 1991. *Being-in-the-world*. Cambridge, Mass.: The MIT Press.

Dreyfus, H. L. 1992. *What computers still can't do* 3rd ed. Cambridge, Mass.: The MIT Press.

Dym, C. L., and R. E. Levitt. 1991. *Knowledge-based systems in engineering*. New York: McGraw-Hill.

Edelman, G. M., and G. Tononi. 2000. *A universe of consciousness: How matter becomes imagination*. New York: Basic Books.

Ekman, P. 2003. *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. New York: Henry Holt.

Eliasmith, C. 2003. Moving beyond metaphors: Understanding the mind for what it is. *Journal of Philosophy* 100, 493–520.

Eliasmith, C., and C. H. Anderson. 2003. *Neural engineering: Computation, representation and dynamics in neurobiological systems*. Cambridge, Mass.: The MIT Press.

Eliasmith, C., and P. Thagard. 2001. Integrating structure and meaning: A distributed model of analogical mapping. *Cognitive Science* 25, 245–286.

Elman, J. L., E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, Mass.: The MIT Press.

Evans, T. 1968. A program for the solution of a class of geometric analogy intelligence test questions. In M. Minsky, ed., *Semantic information processing*, 271–353. Cambridge Mass.: The MIT Press.

Fazio, R. H. 2001. On the automatic activation of associated evaluations: An overview. *Cognition and Emotion* 15, 115–141.

Feigenbaum, E., P. McCorduck, and H. Nii. 1988. *The rise of the expert company*. New York: Vintage.

Feldman Barrett, L., and J. A. Russell. 1998. Independence and bipolarity in the structure of affect. *Journal of Personality and Social Psychology* 74, 967–984.

Feldman, J. A. 1981. A connectionist model of visual memory. In G. E. Hinton and J. A. Anderson, eds., *Parallel models of associative memory*, 49–81. Hillsdale, N.J.: Erlbaum.

Fellbaum, C., ed. 1998. *WordNet: An electronic lexical database*. Cambridge, Mass.: The MIT Press.

Fikes, R., and N. Nilsson. 1971. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2, 189–208.

Fineman, S., ed. 2000. *Emotion in organizations*. London: Sage.

Finger, S. 1994. *Origins of neuroscience: A history of explorations into brain function*. New York: Oxford University Press.

Finke, R. 1989. *Principles of mental imagery*. Cambridge, Mass.: the MIT Press.

Finke, R., S. Pinker, and M. Farah. 1989. Reinterpreting visual patterns in mental imagery. *Cognitive Science* 13, 51–78.

Finke, R., T. B. Ward, and S. M. Smith. 1992. *Creative cognition: Theory, research and applications*. Cambridge, Mass.: The MIT Press/Bradford Books.

Finucane, M. L., E. Peters, and P. Slovic. 2003. Judgment and decision making: The dance of affect and reason. In S. L. Schneider and J. Shanteau, eds., *Emerging*

perspectives on judgment and decision research, 327–364. Cambridge: Cambridge University Press.

Flanagan, O. 1992. *Consciousness reconsidered*. Cambridge, Mass.: The MIT Press.

Flanagan, O. J. 2000. *Dreaming souls: Sleep, dreams, and the evolution of the conscious mind*. Oxford: Oxford University Press.

Fodor, J. 1975. *The language of thought*. New York: Crowell.

Forbus, K. D., Gentner, and K. Law. 1995. MAC/FAC: A model of similarity-based retrieval. *Cognitive Science* 19, 144–205.

Forbus, K., P. Nielsen, and B. Faltings. 1991. Qualitative spatial reasoning: The CLOCK project. *Artifical Intelligence* 51, 417–472.

Forbus, K. 2001. Exploring analogy in the large. In D. Gentner, K. H. Holyoak, and B. K. Kokinov, eds., *The analogical mind: Perspectives from cognitive science*, 23–58. Cambridge, Mass.: The MIT Press.

Foss, J. E. 1995. Materialism, reduction, replacement, and the place of consciousness in science. *Journal of philosophy* 92, 401–429.

Frege, G. 1960. *Translations from the philosophical writings of Gottlob Frege*. Trans. by P. Geach and M. Black. Oxford: Basil Blackwell.

French, R. M. 2002. The computational modeling of analogy-making. *Trends in Cognitive Sciences* 6, 200–205.

Frijda, N. H. 1986. *The emotions*. Cambridge: Cambridge University Press.

Funt, B. 1980. Problem solving with diagrammatic representations. *Artificial Intelligence* 13, 201–230.

Galegher, J., R. E. Kraut, and C. Egido, eds. 1990. *Intellectual teamwork: Social and technological foundations of cooperative work*. Hillsdale, N.J.: Erlbaum.

Gardner, H. 1985. *The mind's new science*. New York: Basic Books.

Genesereth, M. R., and N. J. Nilsson. 1987. *Logical foundations of artificial intelligence*. Los Altos, Calif.: Morgan Kaufmann.

Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7, 155–170.

Gentner, D. 1989. The mechanisms of analogical learning. In S. Vosniadou and A. Ortony, eds., *Similarity and analogical reasoning*, 199–241. Cambridge: Cambridge University Press.

Gentner, D., K. H. Holyoak, and B. K. Kokinov, eds. 2001. *The analogical mind: Perspectives from cognitive science*. Cambridge, Mass.: The MIT Press.

Hirschfeld, L. A., and S. A. Gelman, eds. 1994. *Mapping the mind: Domain specificity in cognition and culture*. Cambridge: Cambridge University Press.

Hoffman, M. L. 2000. *Empathy and moral development: Implications for caring and justice*. Cambridge: Cambridge University Press.

Hofstadter, D. 1995. *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. New York: Basic Books.

Hollan, J., E. Hutchins, and D. Kirsh. 2000. Distributed cognition: Toward a new foundation for human-computer interaction. *ACM Transactions on Computer-Human Interaction* 7, 174–196.

Holland, J. H., K. J. Holyoak, R. E. Nisbett, and P. R. Thagard. 1986. *Induction: Processes of inference, learning, and discovery*. Cambridge, Mass.: The MIT Press.

Holtzman, S. 1989. *Intelligent decision systems*. Reading, Mass.: Addison-Wesley.

Holyoak, K. J., and P. Thagard. 1995. *Mental leaps: Analogy in creative thought*. Cambridge, Mass.: The MIT Press/Bradford Books.

Howson, C., and P. Urbach. 1989. *Scientific reasoning: The Bayesian tradition*. Lasalle, Ill.: Open Court.

Hummel, J. E., and K. J. Holyoak. 1997. Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review* 104, 427–466.

Hummel, J. E., and K. J. Holyoak. 2003. A symbolic-connectionist theory of relational inference and generalization. *Psychological Review* 110, 220–264.

Hutchins, E. 1995. *Cognition in the wild*. Cambridge, Mass.: The MIT Press.

James, W. 1884. What is an emotion? *Mind* 9, 188–205.

Johnson, M. 1987. *The body in the mind*. Chicago: University of Chicago Press.

Johnson-Laird, P. N. 1983. *Mental models*. Cambridge, Mass.: Harvard University Press.

Johnson-Laird, P. N. 1988. *The computer and the mind*. Cambridge, Mass.: Harvard University Press.

Johnson-Laird, P. N., and R. M. Byrne. 1991. *Deduction*. Hillsdale, N.J.: Erlbaum.

Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kandel, E. R., J. H. Schwartz, and T. M. Jessell. 2000. *Principles of neural science*. 4th ed. New York: McGraw-Hill.

Kant, I. 1965. *Critique of pure reason* Trans. by N. Kemp Smith. 2nd ed. London: MacMillan.

Keil, F. 1989. *Concepts, kinds, and cognitive development*. Cambridge, Mass.: The MIT Press/Bradford Books.

Keysar, B. 1990. On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language* 28, 375–385.

Kintsch, W. 1988. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review* 95, 163–182.

Kintsch, W. 1998. *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kintsch, W., D. Welsch, F. Schmalhofer, and S. Zimny. 1990. Sentence memory: A theoretical analysis. *Journal of memory and language* 29, 133–159.

Kirshner, D., and J. A. Whitson. 1997. *Situated cognition: social, semiotic, and psychological perspectives*. Mahwah, N.J.: Erlbaum.

Kitayama, S., and H. R. Markus, eds. 1994. *Emotion and culture: Empirical studies of mutual influence*. Hyattsville, Md.: American Psychological Association.

Kitcher, P. 1993. *The advancement of science*. Oxford: Oxford University Press.

Kolodner, J. 1993. *Case-based reasoning*. San Mateo, Calif.: Morgan Kaufmann.

Konolige, K. 1992. Abduction versus closure in causal theories. *Artificial Intelligence* 53, 255–272.

Kosslyn, S. M. 1980. *Image and mind*. Cambridge, Mass.: Harvard University Press.

Kosslyn, S. M. 1994a. *Elements of graph design*. New York: W. H. Freeman.

Kosslyn, S. M. 1994b. *Image and brain: The resolution of the imagery debate*. Cambridge, Mass.: The MIT Press.

Kosslyn, S. M., G. Ganis, and W. L. Thompson. 2001. Neural foundations of imagery. *Nature Reviews Neuroscience* 2, 635–642.

Kosslyn, S. M., G. Ganis, and W. L. Thompson. 2003. Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Sciences* 7, 109–111.

Kosslyn, S. M., and O. Koenig. 1992. *Wet mind: The new cognitive neuroscience*. New York: Free Press.

Kosslyn, S. M., and S. P. Shwartz. 1977. A simulation of visual imagery. *Cognitive Science* 1, 265–295.

Koza, J. R. 1992. *Genetic programming*. Cambridge, Mass.: The MIT Press.

Kramer, P. D. 1993. *Listening to Prozac*. New York: Viking.

Krauss, L. M. 1996. *The physics of Star Trek*. New York: HarperPerennial.

Nussbaum, M. 2001. *Upheavals of thought*. Cambridge: Cambridge University Press.

Oatley, K. 1992. *Best laid schemes: The psychology of emotions*. Cambridge: Cambridge University Press.

Oatley, K., and J. M. Jenkins. 1996. *Understanding emotions*. Oxford: Blackwell.

O'Brien, D. P., M. D. S. Braine, and Y. Yang. 1994. Propositional reasoning by mental models? Simple to refute in principle and in practice. *Psychological Review* 101, 711–724.

O'Hare, G. M. P., and N. Jennings, eds. 1996. *Foundations of distributed artificial intelligence*. New York: Wiley.

O'Reilly, R. C., and Y. Munakata. 2000. *Computational explorations in cognitive neuroscience*. Cambridge, Mass.: The MIT Press.

Paivio, A. 1971. *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Panksepp, J. 1998. *Affective neuroscience: The foundations of human and animal emotions*. Oxford: Oxford University Press.

Pearl, J. 1988. *Probabilistic reasoning in intelligent systems*. San Mateo, Calif.: Morgan Kaufman.

Pearl, J. 2000. *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Peirce, C. S. 1992. *Reasoning and the logic of things*. Cambridge, Mass.: Harvard University Press.

Penrose, R. 1994. *Shadows of the mind: A search for the missing science of consciousness*. Oxford: Oxford University Press.

Perry, E., H. Ashton, and A. Young, eds. 2002. *Neurochemistry of consciousness*. Amsterdam: John Benjamins.

Piaget, J., and B. Inhelder. 1969. *The psychology of the child*. Trans. by H. Weaver. New York: Basic Books.

Picard, R. W. 1997. *Affective computing*. Cambridge, Mass.: The MIT Press.

Pinker, S. 1994. *The language instinct: How the mind creates language*. New York: William Morrow.

Pinker, S. 1999. *Words and rules: The ingredients of language*. New York: HarperCollins.

Pinker, S. 2002. *The blank slate: The modern denial of human nature*. New York: Viking.

Pinker, S., and A. Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193.

Pinker, S., and M. T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6, 456–464.

Polk, T. A., and C. M. Seifert, eds. 2002. *Cognitive modeling*. Cambridge, Mass.: The MIT Press.

Pollack, J. B. 1991. The induction of dynamical recognizers. *Machine Learning* 7, 227–252.

Pollock, J. L. 1995. *Cognitive carpentry: A blueprint for how to build a person*. Cambridge, Mass.: The MIT Press.

Port, R., and T. van Gelder, eds. 1995. *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, Mass.: The MIT Press.

Posner, M. I., and S. W. Keele. 1970. Retention of abstract ideas. *Journal of Experimental Psychology* 83, 304–308.

Posner, M. I., and M. E. Raichle. 1994. *Images of mind*. New York: Freeman.

Prietula, M. J., K. M. Carley, and L. Gasser, eds. 1998. *Simulating organizations: Computational models of institutions and groups*. Menlo Park, Calif.: AAAI Press.

Prior, A. N. 1967. Logic, history of. In P. Edwards, ed., *Encylopedia of philosophy*, vol. 4, 513–571. New York: Macmillan.

Putnam, H. 1975. *Mind, language, and reality*. Cambridge: Cambridge University Press.

Pylyshyn, Z. 1984. *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, Mass.: The MIT Press.

Pylyshyn, Z. 2002. Mental imagery: In search of a theory. *Behavioral and Brain Sciences* 25, 157–237.

Quartz, S. R., and T. J. Sejnowski. 2002. *Liars, lovers, and heroes: What the new brain science reveals about how we become who we are*. New York: William Morrow.

Quinlan, J. R. 1983. Learning efficient classification procedures and their application to chess end games. In R. Michalski, J. Carbonell, and T. Mitchell, eds., *Machine learning: An artificial intelligence approach*. Palo Alto, Calif.: Tioga.

Read, S., and A. Marcus-Newhall. 1993. The role of explanatory coherence in the construction of social explanations. *Journal of Personality and Social Psychology* 65, 429–447.

Resnick, L., J. Levine, and S. Behrend, eds. 1991. *Socially shared cognitions*. Hillsdale, N.J.: Erlbaum.

Richards, D. 1990. Is strategic decision making chaotic? *Behavioral Science* 35, 219–232.