TOWARD TECHNICAL UNDERSTANDING*

Part 1. Brain Structure and Function

J.M. HAILE
Clemson University • Clemson, SC 29634-0909

> *One must acquire many different ways to understand.*
>
> Minsky[1]

"You know, Prof, my grades on the quizzes don't really reflect my understanding of the material." . . . "When you talk to Clarence about the material, it's evident he understands a lot about it, so why can't he do the homework?" . . . "I know the class understands this concept, we've been through it many times, so why can't they apply it when they need it? Why can't students access what they know?"

These kinds of comments, from students and colleagues, are familiar to any of us who have spent time in education. They signal frustration in various guises, and often a voiced frustration is but a symptom of deeper dissatisfaction and perplexity. In pondering such comments, I've concluded that many of them spring from a common basis: confusion and misconception about what we mean by an understanding of technical material. Such confusion should not be dismissed lightly, for it can hamper our attempts to help others learn; and so it seems worthwhile to try to clarify what it means to understand. But to unravel such confusion is no small task. The word *understanding* is itself obscured by a vagueness that approaches the enigmatic. For once, Webster fails us, merely offering as synonyms "perception," "comprehension," "appreciation," and "mental grasp." These move us no closer to the root of the matter.

The ambiguity arises because there is not one understanding, or even just a few. There are, in fact, many—many kinds of understanding and many ways to reach them. It is one thing to recognize you have a problem, another thing to articulate the problem, yet another to identify what is needed to solve it, still another to carry out the solution, and even another to appreciate what the solution means. Given those many goals and the many paths to each, it is no wonder we have difficulty articulating general rules—or even rules of thumb—that will consistently lead us to an understanding. But what is difficult in general may be manageable in particular. Perhaps by restricting our attention to particular realms of knowledge, such as those embodied by engineering and the physical sciences, we can clarify what it means to understand—at least for those restricted realms. That is the thesis for the papers in this series.

Having recognized that there are many ways to understand technical material, we then ask how those ways can be organized. One appealing organization is a hierarchy because hierarchies identify levels, and this usage coincides with commonly used, but ill-defined, ideas concerning levels of understanding. In addition, a hierarchy provides a systematic progression that can serve as the basis for helping people learn. For example, a hierarchy of understanding can

J.M. Haile, *a professor of chemical engineering at Clemson University, is the author of* Molecular Simulation, *Published by John Wiley & Sons in 1992.*

---

* *EDITOR'S NOTE: This is the first of three installments. The second installment, "Elementary Levels," will appear in our Fall 1997 issue, and the third installment, "Advanced Levels," will be published in our Winter 1998 issue.*

help us identify the current stage in a student's study of a topic, it can help us show the student what must be done to reach the next stage, and it can help us determine when a transition between stages has been successful. This series of papers is primarily concerned with presenting and discussing a hierarchy for understanding technical material.

But the job of fostering understanding can also be clarified if we know something about how people learn—that is, how the human mind assimilates new information and integrates it with old information. Over the past ten years we have seen significant progress in neuroscience, especially in neurobiology, psychology, and artificial intelligence. As educators, we should take advantage of that progress, recognizing that the next ten years will bring still more progress. By clarifying how the brain functions, we can obtain clues as to how to improve learning. We therefore will use the rest of this paper to review, in an elementary way, relevant aspects of brain structure and function. These discussions will support the hierarchy of understanding presented in the second and third papers in this series.

Since learning creates new structures in the brain by modifying existing structures, learning can only begin from things the student already knows. This has implications as to whether a topic should be taught top-down (deductively) or bottom-up (inductively).

BRAIN STRUCTURE AND FUNCTION

The human brain is not a single entity, but rather a composite of several brains. The top of the spinal cord forms the *medulla,* which supervises basic motor functions, including heart beat, respiration, and digestion. Behind the medulla lies the *cerebellum,* which coordinates body position, movement, and balance. Above the cerebellum we find the *limbic system,* which includes the pituitary gland, the hypothalamus, the hippocampus, and other structures. The pituitary makes hormones that control the function of most other glands in the body; its action is controlled by the hypothalamus. More generally, the hypothalamus regulates all life-support functions, including heart rate, body temperature, chemical balances in the body, hunger, thirst, and emotional responses to threats for survival. The hippocampus apparently participates in the formation of long-term memories; this will be discussed in the third of this series of papers. Atop the limbic systems sits the *cerebrum,* which is devoted to all higher mental activities, including language, conscious awareness, and abstract thought.

Of all these structures, the cerebrum is by far the largest, yet most high-level mental activities are confined to its surface—the *cerebral cortex*—a layer only 2mm thick and convoluted into folds to increase its surface area within a confined volume. The cortex contains a significant portion of the brain's gray matter—the little gray cells so favored by Hercule Poirot. The volume enclosed by the cerebral cortex is filled largely with white matter: the strands and filaments that connect brain cells. That is, much of the human brain is mere wiring.

The following sections of this paper present an elementary overview of the functioning of the neuron and of the huge number of neurons that form the cerebral cortex. These functions allow us to draw certain conclusions about the nature of learning. For a more detailed introduction to brain structure and function, see references 2-6; the illustrations by Macaulay[6] are particularly instructive.

THE NEURON

The basic unit of mental activity is a single nerve cell—the *neuron.* Functionally, a neuron in the cortex collects signals from other neurons, integrates them into a single signal, and then either suppresses the signal or forwards it to other neurons. Structurally, a neuron is composed of three principal parts: a *cell body,* which contains the nucleus and performs the life-support functions common to any biological cell; a tree-like array of branches called *dendrites* that carry signals from other neurons to the cell body; and an *axon,* a single strand that carries the signal from the cell body to other neurons.

More generally, neurons are the primary functional elements of the nervous system; for example, a nerve is a bundle of axons. Axons vary in length from millimeters in the cortex to about a meter in the case of the axons that connect the toes to the spinal cord. Variations in geometry provide a means for classifying neurons by structure;[5] those in the cerebral cortex are called *pyramidal neurons* because of the distinctive shape of the cell body.

The function of an individual neuron is illustrated schematically in Figure 1 (next page). From contacts with other neurons, a dendrite carries a signal, as a voltage difference, to the cell body. At the cell body, signals from all dendrites are combined into a single voltage that propagates to the head of the axon. If this output voltage exceeds a certain threshold, the neuron is said to fire, and a pulse voltage propagates down the axon. The end of the axon divides into branches, providing hundreds of terminals to other neurons.

Each axonal terminal is separated from another neuron by a microscopic gap called a *synapse.* When a voltage pulse reaches an axonal terminal, it causes *vesicles* in the terminal region to fuse onto the wall (the presynaptic membrane) of the neural cell. This, in turn, causes the vesicles to open, releasing a few thousand molecules of a *neurotransmitter* into the synapse. These molecules diffuse across the synapse to a dendrite or cell body of another neuron. If the neurotransmitter can find an appropriate *receptor*—a protein—embedded in the postsynaptic membrane, then the signal is successfully passed from one neuron to the other.

The voltage difference propagating from dendrite through cell body and down an axon is not carried electrically, but chemically; that is, it is not carried by free electrons but by sodium ions. Therefore, the speed of signal propagation is of the order of milliseconds, which is slow relative to the speed of electrical conduction in metal wires.

Further, an axon does not conduct the signal by a current propagating axially, as in a wire. Instead, a local voltage difference in one part of an axon relative to a neighboring part induces opening and closing of molecular gates on channels within the cell membrane; these channels allow flow of ions between the interior and exterior of the neuron, changing the voltage in one region of the neuron relative to an adjacent region. Thus, sequential radial flow of ions through cell walls produces the effect of a voltage propagating axially.

The activity induced by the voltage reaching a synapse amounts to a key-lock-gate scenario. If a neurotransmitter (the key) can find its receptor (the lock), then a molecular gate opens, allowing ions to enter the dendrite, creating a local voltage difference.

A few dozen neurotransmitters have been identified, and more probably remain to be discovered.[5] The common ones include glutamate, dopamine, acetylcholine, and γ-aminobutyric acid (GABA). Certain drugs, including the opiates, nicotine, and the antipsychotics, are known to either mimic or block the actions of certain neurotransmitters.[5,6] This is possible because neurotransmitters activate receptors by matching physical structures, so any molecular fragment that matches the receptor structure might activate that receptor; not only will a key open a lock, but so too will a skeleton key.

Synaptic connections are of two general types. *Excitatory* synapses tend to promote firing of the neuron by activating receptors that allow sodium ions to enter the neuron through the postsynaptic membrane. These connections typically occur on dendritic branches, with the common excitatory neurotransmitter being glutamate. In contrast, *inhibitory* synapses tend to suppress firing by activating receptors that allow chlorine ions to enter through the postsynaptic membrane. These typically occur directly on the cell body, with the common inhibitory neurotransmitter being GABA.

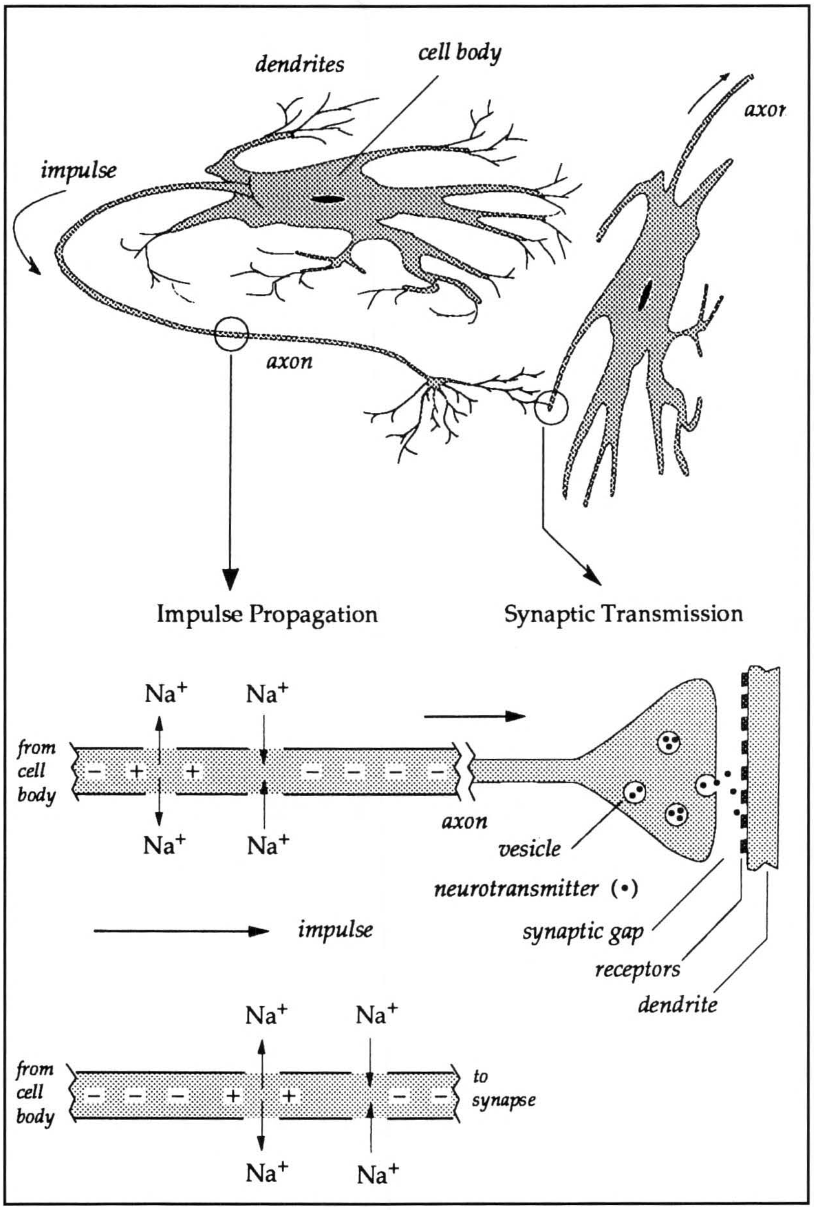

Figure 1. *(top) Principal parts of a single neuron, including synaptic connection to the dendrite of another neuron. (lower left) Exploded view of an axonal segment; radial flow of sodium ions between axon and extracellular fluid propagates a nerve impulse from the cell body to the synapse. (lower right) Exploded view of an axonal terminal and synapse. At the terminal, a nerve impulse stimulates vesicles to fuse with the presynaptic membrane, releasing neurotransmitters into the synaptic gap. The bottom views are adaptations of drawings by Macaulay.*[6]

THE NEURAL NETWORK

Although the functioning of a single neuron is a fascinating electrochemical process, the really astonishing functions occur not at the molecular or cellular levels, but in the collective behavior of large numbers of connected neurons. The number of neurons in the human brain is estimated at between 20 and 100 billion. Further, the average pyramidal neuron makes roughly 1000 connections to other neurons; so the total number of connections may well be about 10^{14}.[3,6] Most connections are among neighboring neurons, but many axons connect neurons that lie in very different regions of the cortex. In principle, any neuron can influence the firing of any other neuron.

By itself, the firing of a single neuron is essentially meaningless. Meaning only arises when a pattern is established by the simultaneous or sequential firings of many neurons. We do not know how firing patterns encode meanings (that is part of the puzzle), but the following metaphor may capture the essence. Imagine an array of lights forming a scoreboard. The array itself has no meaning; if none of the lights are activated, we have no meaning, and if all of the lights are activated, we still have no meaning. Informative meaning occurs only when some lights are activated while others are not. Moreover, the meaning is in the pattern, not in any particular lights; that is, meaning is encoded in the spatial and temporal relations among the lights that are activated and those that are not. For example, meaning is preserved even when the pattern scrolls across the array. Now reread the previous four sentences, everywhere replacing "light" with "neuron" and "array" with "neural net." This metaphor suggests one reason for having both excitatory and inhibitory synapses for in this way, not only can any one neuron participate in any pattern, but in addition, when the same pattern is replayed at different times, a neuron can participate by sometimes being activated and other times being quiescent.

Neural activity in the cerebral cortex distinguishes brain from mind; that is, brain is the structure and mind is the function. As Minsky has written,[1] "Minds are what brains do." But what is it that minds do? In particular, what does the cerebral cortex do?

There is always at least a baseline of neural activity in the cortex. It can be seen crudely on an electroencephalogram (EEG). But that minimal activity can be driven to more active modes by stimulation, either from the external world through the senses or from the internal world through other parts of the brain. The response appears to be a search for meaning—an attempt to find a pattern that interprets or makes sense of the stimulus. In other words, the baseline firing of a huge number of interconnected neurons amounts to a chaotic dynamics—not random, but apparently random with some underlying order.[7-9] Such dynamics are, by definition, sensitive to small disturbances, so even a small stimulation of the cortex can produce a qualitative change in the character of the firing pattern. Some changes in the dynamics take the form of convergence to a local *attractor*—a firing pattern that is recreated and sustained whenever the firing trajectory passes sufficiently "close" to neurons that activate the pattern. Such attractors constitute meaning to the organism. This is illustrated in Figure 2.

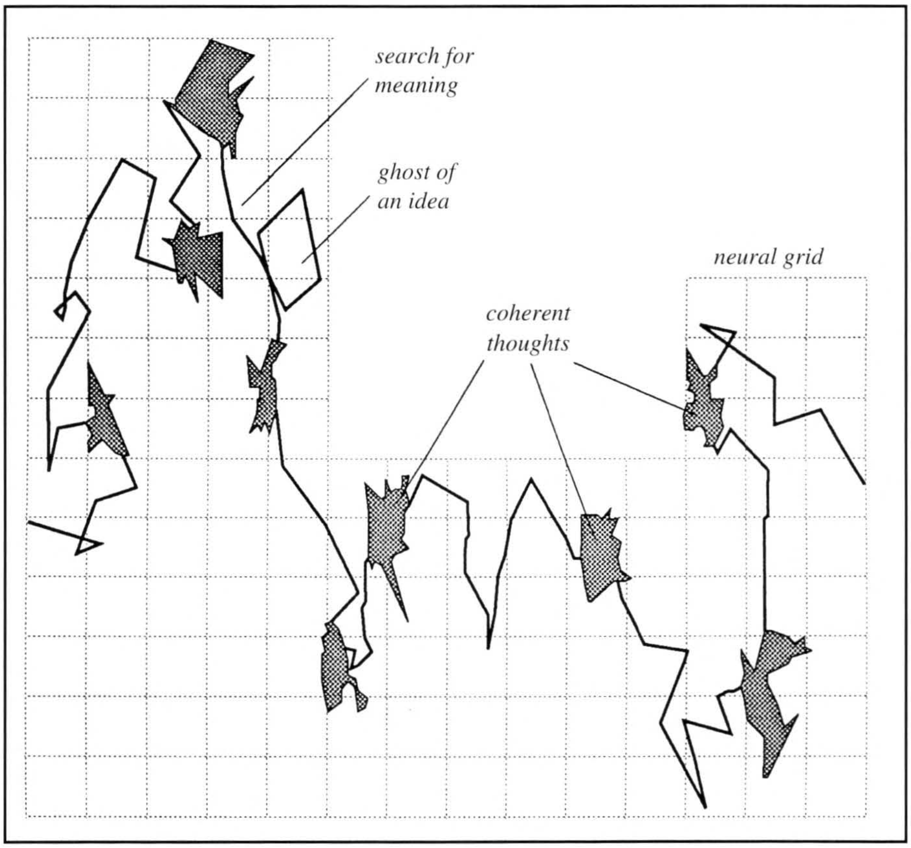

Figure 2. *Schematic representation of sequences of neural firings in the cerebral cortex. Grid represents an array of neurons; path represents the firing sequence. Shaded areas represent repeated firings of patterns of neurons that have meaning, such as recognition of a sight or formation of a coherent thought. Path segments between shaded areas represent the search for meaning. In studies of dynamic systems, such as in process control and statistical mechanics, the grid is interpreted as a phase space and the path is a trajectory of the system. Then the shaded regions are local attractors that order the trajectory into patterns. Based on a figure in Calvin[7] and a phase-space plot in Freeman[8] from EEG data taken from a rat's olfactory bulb.*

We may become conscious of attractors—that is, conscious of thought patterns—when we encounter ambiguity. An ambiguous stimulus causes neural firing patterns to bifurcate into conflicts or competition between two or more attractors; the result is mental confusion. In such situations, the mind contrives more than one pattern that is consistent with the data, and the conflict can only be resolved with more data.

For example, consider the object shown in Figure 3, which presents a conflict between foreground and background. If you focus to bring the shading to the foreground of the figure, you see the letter E. But if you shift your focus slightly, the shading can be pushed to the background, and you see the characters L and 3.

Your recognition of the E is an attractor produced by one assembly of neurons, while recognition of the L and the 3 is a second attractor produced by another assembly of neurons. Both attractors are consistent with the data and additional visual cues would be needed for one attractor to dominate.

The interpretation of thought as a dynamic process driven to local attractors, as shown schematically in Figure 2, is appealing, but it is likely an oversimplification for at least a couple of reasons. First, the coding of meaningful patterns is probably not just in the relative positions of firing neurons; it may also involve firing rates and sequences. That is, meaning may involve both positional codes and temporal codes.[10] Second, the action of the cortex appears to involve distributed processes, in which multiple subprocesses are performed simultaneously.[1,10] For example, visual recognition of an object involves perception of contours, depth, and color—three activities that are performed simultaneously but in separate regions of the cortex.

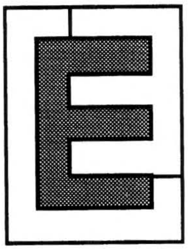

***Figure 3.** Mental confusion can arise when the brain devises two conflicting patterns from the same information. This figure can be ambiguous because it does not provide enough cues to allow one to distinguish foreground from background. Focusing to bring the shaded area into the foreground reveals one character, but focusing to push the shaded area into the background reveals two different characters.*

More complex functions appear to progress through hierarchies of distributed processes, which may explain why we like to use hierarchies for organizing societies, institutions, and problem-solving tasks.[1] Distributed processes make efficient use of neural networks because the same assemblies of neurons can be used for the same kinds of tasks in different situations.

Although we do not yet know the details for *how* meanings are assigned to patterns, we at least know *what* is being done: minds are what brains do, and the search for meaning is what the cortex does.

MODIFICATION OF BRAIN STRUCTURE: LEARNING

The association of meanings with particular stimuli constitutes one aspect of learning. For example, if the stimulus is a right triangle, then part of the associated meaning would be the Pythagorean theorem. But meaning is connected to a stimulus through neural firing patterns created in the cerebral cortex, so to learn new meanings, we must create new firing patterns. For these new patterns to be accessible over long times, the mind must bias connections among neurons so that the new pattern is recreated whenever an appropriate stimulus is encountered. That is, to make long-lasting changes in function (the mind), we must make changes in structure (the brain).

We do not yet know much about how learning modifies the brain, though many observations are suggestive and the more obvious possibilities are listed below. But we do know one mechanism that is *not* used—the brain does not modify itself by growing new neurons. The number of neurons is constant through adolescence, and then neurons begin to degenerate and die over the remaining lifetime. Estimates usually put the average total loss at 10% of the original number.[3] During youth, the brain grows in complexity by forming and pruning dendritic trees and axonal branches, that is, by increasing and refining connections among existing neurons.[3,6]

Here are some ways by which learning may change brain structure and function:[5]

- Changes in the size and shape of axon terminals coupled with changes in the number of presynaptic vesicles (which hold the neurotransmitter) to increase or decrease the amounts of neurotransmitter released
- Changes in the size and shape of postsynaptic receptors and channels to change the level of activation voltages created when a receptor opens
- Increases and decreases in the number of receptors at certain synapses
- Changes in the number and location of synapses
- Sprouting of new axonal terminals
- Remolding of terminal bulbs on dendrites.

IMPLICATIONS FOR EDUCATION

The brain, then, is a self-modifying neural network. The processes carried out by that network constitute the mind, and the function of a part of that mind—the cerebral cortex—is to ascribe meanings or interpretations to both external and internal stimuli. These observations lead to certain implications about the nature of education, including

- Learning is a natural activity of the human mind.
- Learning is not storage and retrieval of information; the brain does not store information.[8,11] It only develops a propensity to reproduce neural firing patterns that have been found beneficial; that is, what we call memory is actually a *re-creation* of information. To be able to use what they know, students must learn cues that re-create useful patterns.
- Since learning creates new structures in the brain by modifying existing structures, learning can only begin from things the student already knows. This has implications as to whether a topic should be taught top-down (deductively) or bottom-up (inductively).
- The brain apparently modifies neural connections as part of its response in those neurons that are activated to form a pattern. Thus, learning new things amounts to a perturbation of things already known; but if the perturbation is too large, then no related neural firing pattern can be created and no learning takes place. Thus, students must be led to new knowledge in small chunks of information that allow the brain to modify existing neural networks. Repetition is then needed to strengthen new neural connections. The importance of repetition is addressed in the third of this series of papers.
- Experts in a topic have highly interconnected constellations of neurons that can be activated by stimulating any of many different nodes.[1] These elaborate networks allow experts to quickly learn new things because their vast networks offer numerous nodes that can be easily modified to assimilate new information. In contrast, students generally have few networks related to technical material; the networks they do have tend to be meager and largely fragmented. The assimilation of new information into those networks is often a laborious task because a small addition to a small structure can require a large change in the structure.

We often witness instances at which disjointed neural assemblies finally become fused into a coherent network. It happens to those students who struggle with a topic for several weeks, laboriously piecing together several disjoined networks. Then, about midterm or shortly thereafter, one more piece of new information perturbs the entire system sufficiently that those several disjointed networks become united—revelation! The student understands.

- Recognition is easier than recall.[1,7] *Recognition* forms a meaningful pattern in response to an external stimulus, while *recall* forms a meaningful pattern in response to an internal stimulus. To test recognition, we might pose a question such as "What quantity is defined by $Re = u\rho d/\mu$?" But to test recall, we would ask, "What equation defines the Reynolds number?" Recall is more difficult because we must not only create the pattern, but we must also generate the stimulus that produces the pattern. Because of this difference in difficulty, students generally prefer certain kinds of quizzes over others.
- Learning is easier than unlearning. Unlearning refers to correcting misunderstandings from earlier learning. During learning, we modify a neural network to create a new net; but during unlearning, we not only create a new network, but we must also suppress formation of the old erroneous pattern.
- To learn, students must actively participate in their own education. Only the individual can modify its own synapses, dendritic trees, and axonal terminals. No instructor can do this for the student.
- Quickness of mind (a commonly used indicator of intelligence) decouples from the ability to think.[12] A quick mind is one that moves immediately and decisively to a local attractor. But the Latin root for *intelligent* is *inter + legere,* which means *to select.* And to select implies a consideration of alternatives; that is, intelligent thinking involves the identification of alternative attractors and choosing from among them. This cannot necessarily be done quickly.

REFERENCES

1. Minsky, M., *The Society of Mind,* Simon and Schuster, New York, NY (1986)
2. Restak, R., *The Brain,* Bantam Books, New York, NY (1984)
3. Restak, R., *Brainscapes,* Hyperion, New York, NY (1995)
4. Calvin, W.H., and G.A. Ojemann, *Conversations with Neil's Brain,* Addison-Wesley, Reading, MA (1994)
5. Fischbach, G.D., "Mind and Brain," *Sci. Am.,* **267**(3), 48 (1992)
6. Ornstein, R., and R.F. Thompson, *The Amazing Brain,* illustrated by D. Macaulay, Houghton Mifflin, Boston, MA (1984)
7. Calvin, W.H., *The Cerebral Code,* MIT Press, Cambridge, MA (1996)
8. Freeman, W.J., "Role of Chaotic Dynamics in Neural Plasicity," *Progress in Brain Research,* **102**, 319 (1994)
9. Kauffman, S.A., *The Origins of Order,* Oxford University Press, New York, NY (1993)
10. Singer, W., "Synchronization of Cortical Activity and Its Putative Role in Information Processing and Learning," *Ann. Rev. Physiol.,* **55**, 349 (1993)
11. Searle, J., *The Rediscovery of the Mind,* MIT Press, Cambridge, MA (1992)
12. de Bono, E., *de Bono's Thinking Course,* Facts on File Publications, New York, NY (1985) ❐