Consciousness and Cognition **8,** 1–44 (1999)
Article ID ccog.1998.0372, available online at http://www.idealibrary.com on IDEAL®

The Link between Brain Learning, Attention, and Consciousness

Stephen Grossberg

Department of Cognitive and Neural Systems and Center for Adaptive Systems, Boston University

The processes whereby our brains continue to learn about a changing world in a stable fashion throughout life are proposed to lead to conscious experiences. These processes include the learning of top-down expectations, the matching of these expectations against bottom-up data, the focusing of attention upon the expected clusters of information, and the development of resonant states between bottom-up and top-down processes as they reach an attentive consensus between what is expected and what is there in the outside world. It is suggested that all conscious states in the brain are resonant states and that these resonant states trigger learning of sensory and cognitive representations. The models which summarize these concepts are therefore called Adaptive Resonance Theory, or ART, models. Psychophysical and neurobiological data in support of ART are presented from early vision, visual object recognition, auditory streaming, variable-rate speech perception, somatosensory perception, and cognitive–emotional interactions, among others. It is noted that ART mechanisms seem to be operative at all levels of the visual system, and it is proposed how these mechanisms are realized by known laminar circuits of visual cortex. It is predicted that the same circuit realization of ART mechanisms will be found in the laminar circuits of all sensory and cognitive neocortex. Concepts and data are summarized concerning how some visual percepts may be visibly, or modally, perceived, whereas amodal percepts may be consciously recognized even though they are perceptually invisible. It is also suggested that sensory and cognitive processing in the What processing stream of the brain obey top-down matching and learning laws that are often complementary to those used for spatial and motor processing in the brain's Where processing stream. This enables our sensory and cognitive representations to maintain their stability as we learn more about the world, while allowing spatial and motor representations to forget learned maps and gains that are no longer appropriate as our bodies develop and grow from infanthood to adulthood. Procedural memories are proposed to be unconscious because the inhibitory matching process that supports these spatial and motor processes cannot lead to resonance.

Key Words: learning, expectation, attention, adaptive resonance, neural network, procedural memory, consciousness, object recognition, speech perception.

HOW DO WE CONTINUE TO LEARN THROUGHOUT LIFE?

We experience the world as a whole. Although myriad signals relentlessly bombard our senses, we somehow integrate them into unified moments of conscious experience

The author thanks Cynthia E. Bradford and Diana Meyers for their valuable assistance in the preparation of this manuscript. This work was supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-92-J-1309, ONR N00014-95-1-0409, and ONR N00014-95-1-0657).

Address correspondence and reprint requests to Stephen Grossberg, Department of Cognitive and Neural Systems Center for Adaptive Systems, Boston University, Boston, MA 02215.

1053-8100/99 $30.00

that cohere together despite their diversity. Because of the apparent unity and coherence of our awareness, we can develop a sense of self that can gradually mature with our experiences of the world. This capacity lies at the heart of our ability to function as intelligent beings.

The apparent unity and coherence of our experiences is all the more remarkable when we consider several properties of how the brain copes with the environmental events that it processes. First and foremost, these events are highly context sensitive. When we look at a complex picture or scene as a whole, we can often recognize its objects and its meaning at a glance, as in the picture of a familiar face. However, if we process the face piece-by-piece, as through a small aperture, then its significance may be greatly degraded. To cope with this context sensitivity, the brain typically processes pictures and other sense data in parallel, as *patterns* of activation across a large number of feature-sensitive nerve cells, or neurons. The same is true for senses other than vision, such as audition. If the sound of the word GO is altered by clipping off the vowel O, then the consonant G may sound like a chirp, quite unlike its sound as part of GO.

During vision, all the signals from a scene typically reach the photosensitive retinas of the eyes at essentially the same time, so parallel processing of all the scene's parts begins at the retina itself. During audition, each successive sound reaches the ear at a later time. Before an entire pattern of sounds, such as the word GO, can be processed as a whole, it needs to be recoded, at a later processing stage, into a simultaneously available spatial pattern of activation. Such a processing stage is often called a working memory, and the activations that it stores are often called short-term memory (STM) traces. For example, when you hear an unfamiliar telephone number, you can temporarily store it in working memory while you walk over to the telephone and dial the number.

In order to determine which of these patterns represents familiar events and which do not, the brain matches these patterns against stored representations of previous experiences that have been acquired through learning. Unlike the STM traces that are stored in a working memory, the learned experiences are stored in long-term memory (LTM) traces. One difference between STM and LTM traces concerns how they react to distractions. For example, if you are distracted by a loud noise before you dial a new telephone number, its STM representation can be rapidly reset so that you forget it. On the other hand, if you are distracted by a loud noise, you (hopefully) will not forget the LTM representation of your own name.

The problem of learning makes the unity of conscious experience particularly hard to understand, if only because we are able to rapidly learn such enormous amounts of new information, on our own, throughout life. For example, after seeing an exciting movie, we can tell our friends many details about it later on, even though the individual scenes flashed by very quickly. More generally, we can quickly learn about new environments, even if no one tells us how the rules of each environment differ. To a surprising degree, we can rapidly learn new facts without being forced to just as rapidly forget what we already know. As a result, we do not need to avoid going out into the world for fear that, in learning to recognize a new friend's face, we will suddenly forget our parents' faces.

Many contemporary learning algorithms would not be so lucky. Speaking technically, the brain solves a very hard problem that many current approaches to technology have not solved. It is a self-organizing system that is capable of rapid yet stable autonomous learning of huge amounts of data in a nonstationary environment. Discovering the brain's solution to this key problem is as important for understanding ourselves as it is for developing new pattern recognition and prediction applications in technology.

I have called the problem whereby the brain learns quickly and stably without catastrophically forgetting its past knowledge the *stability–plasticity dilemma*. The stability–plasticity dilemma must be solved by every brain system that needs to rapidly and adaptively respond to the flood of signals that subserves even the most ordinary experiences. If the brain's design is parsimonious, then we should expect to find similar design principles operating in all the brain systems that can stably learn an accumulating knowledge base in response to changing conditions throughout life. The discovery of such principles should clarify how the brain unifies diverse sources of information into coherent moments of conscious experience.

This article reviews evidence that the brain does operate in this way. It summarizes several recent brain modeling studies that illustrate, and further develop, a theory called Adaptive Resonance Theory, or ART, that I introduced in 1976 (Grossberg, 1976a,b, 1978, 1980, 1982). In the present article, I briefly summarize results selected from four areas where ART principles have been used to explain challenging behavioral and brain data. These areas are visual perception, visual object recognition, auditory source identification, and variable-rate speech recognition. On first inspection, the behavioral properties of these visual and auditory phenomena may seem to be entirely unrelated. On a deeper computational level, their governing neural circuits are proposed to incorporate a similar set of computational principles.

I should also say right away, however, that ART principles do not seem to be used in all brain learning systems. Whereas ART learning designs help to explain sensory and cognitive processes such as perception, recognition, attention, reinforcement, recall, working memory, and memory search, other types of learning seem to govern spatial and motor processes. In these latter task domains, it *is* adaptive to forget old coordinate transformations as the brain's control systems adjust to a growing body and to other changes in the body's sensory–motor endowment throughout life.

Sensory and cognitive processes are often associated with the What cortical processing stream that passes from the visual cortex through the inferotemporal cortex, whereas spatial and motor processes are associated with the Where (or How) cortical processing stream that passes from the visual cortex through the parietal cortex (Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983; Ungerleider & Mishkin, 1982). Our research over the years has concluded that many processes in the two distinct streams, notably their matching and learning processes, obey different, and even complementary, laws. This fact bears heavily on questions of consciousness and helps to explain why procedural memories are not conscious (Cohen & Squire, 1980; Mishkin, 1982; Scoville & Milner, 1957; Squire & Cohen, 1984). Indeed, a central hypothesis of ART since its inception is

ART Hypothesis: All Conscious States are Resonant States.

As noted in greater detail below, many spatial and motor processes involve a form of inhibitory matching and mismatch-based learning that does not support resonant states. Hence, by the ART Hypothesis, they cannot support a conscious state. Although ART predicts that all conscious states are resonant states, the converse statement, that all resonant states are conscious states, is not yet asserted.

It might be worthwhile to note immediately that various other models of cognitive learning and recognition, such as the popular backpropagation model (Parker, 1982; Rumelhart, Hinton, & Williams, 1986; Werbos, 1974), are based on a form of mismatch-based learning. They cannot, therefore, generate resonant states and, in fact, are well known to experience catastrophic forgetting under real-time learning conditions. A comparative survey of ART vs backpropagation computational properties is provided in Grossberg (1988).

THE THEORETICAL METHOD

Another point worth noting is how one arrives at a psychophysiological theory such as ART which attempts to link behavioral properties to the brain mechanisms which generate them. Such a linkage between brain and behavior is, I believe, crucial in any mature theory of consciousness, since a theory of consciousness that cannot explain behavioral data has failed to deal with the contents of consciousness, and a theory of consciousness that cannot link behaviors to the brain mechanisms from which they emerge must remain, at best, a metaphor.

A particular type of theoretical method has been elaborated over the past 30 years with which to approach such complex behavioral and brain phenomena. The key is to begin with behavioral data, typically scores or even hundreds of parametrically structured behavioral experiments in a particular problem domain. One begins with behavioral data because the brain has evolved in order to achieve *behavioral* success. Any theory that hopes to link brain to behavior thus needs to discover the computational level on which brain dynamics control behavioral success. One works with large amounts of data because otherwise too many seemingly plausible hypotheses cannot be ruled out.

A crucial metatheoretical constraint is to insist upon understanding the behavioral data—which comes to us as static numbers or curves on a page—as the emergent properties of a dynamical process which is taking place moment-by-moment in an individual mind. One also needs to respect the fact that our minds can adapt on their own to changing environmental conditions without being told that these conditions have changed. One thus needs to frontally attack the problem of how an intelligent being can *autonomously adapt to a changing world*. Knowing how to do this is presently an art form. There are no known algorithms with which to point the way.

Whenever we have attempted this task in the past, we have resisted every temptation to use homunculi or else the crucial constraint on *autonomous* adaptation would be violated. The result has regularly been the discovery of new organizational principles and mechanisms, which we have then realized as a minimal model operating according to only locally defined laws that are capable of operating on their own in real time. The remarkable fact is that, when such a model has been written down, it has always been interpretable as a neural network. These neural networks have always included known brain mechanisms. The functional interpretation of these mechanisms has, however, often been novel because of the light thrown upon them by the behavioral analysis. The networks have also typically predicted the existence of unknown neural mechanisms, and many of these predictions have been supported by subsequent neurophysiological, anatomical, and even biochemical experiments over the years. Once this neural connection has been established by a top-down analysis, one can work both top-down from behavior and bottom-up from brain to exert a tremendous amount of conceptual pressure with which to better characterize and refine the model. A fundamental empirical conclusion can be drawn from many experiences of this type; namely, the brain as we know it can be successfully understood as an organ that is designed to achieve successful autonomous adaptation to a changing world. I like to say that, although I am known as one of the founders of the field of neural networks, I have never tried to derive a neural network. They are there because they provide a natural computational framework with which to control autonomous behavioral adaptation to a changing world.

Such a real-time analysis is not easy because it requires that one have knowledge, and even mastery, of several disciplines. For example, it has always proved to be the case that the level of brain organization that computes behavioral success is the network or system level. Does this mean that individual nerve cells, or even smaller components, are unimportant? Not at all. One needs to properly define the individual nerve cells and their interactions in order to correctly define the networks and systems whose interactive, or emergent, properties map onto behavior as we know it. Thus one must be able to freely move between (at least) the three levels of Neuron, Network, and Behavior in order to complete such a theoretical cycle.

Doing this requires that one has a sufficiently powerful theoretical language. The language of mathematics has proved to be the relevant tool, indeed a particular kind of mathematics. All of the self-adapting behavioral and brain systems that I have ever derived are nonlinear feedback systems with large numbers of components operating over multiple spatial and temporal scales. The nonlinearity just means that our minds are not the sum of their parts. The feedback means that interactions occur in both directions within the brain and between the brain and its environment. The multiple temporal scales are there because, for example, processes like STM are faster than the processes of learning and LTM. Multiple spatial scales are there because the brain needs to process parts as well as wholes. All of this is very easy to say intuitively. But when one needs to work within the tough honesty of mathematics, things are not so easy. Most of the difficulties that people seem to have in understanding what is already theoretically known about such systems derives from a literacy problem in which at least one, but often more than one, of the ingredients of neuron, network, behavior, and nonlinear feedback mathematics are not familiar to them.

A second important metatheoretical constraint derives from the fact that no single step of theoretical derivation can derive a whole brain. One needs to have a method that can evolve with the complexity of the environmental challenges that the model is forced to face. This is accomplished as follows. After introducing a dynamic model of a prescribed set of data, one analyzes its behavioral and brain data implications as well as its formal properties. The cycle between intuitive derivation and computational analysis goes on until one finds the most parsimonious and most predictive realization of the organizational principles that one has already discovered. Through this analysis, one can also identify various ‘‘species-specific variations’’ of such a prototypical model and apply them to different types of data. Such a theoretical analysis also discloses the *shape* of the boundary, within the space of data, beyond which the model no longer has explanatory power. The shape of this boundary between the known and the unknown then often clarifies what design principles have been omitted from the previous analyses. The next step is to show how these additional design principles can be incorporated into a more powerful model that can explain even more behavioral and neural data. In this way, the model undergoes a type of evolutionary development, as it tries to cope behaviorally with environmental constraints of ever increasing subtlety and complexity.

The metatheoretical constraint that comes into view here is an *embedding* constraint; in other words, one needs to be able to embed the previous model into the new model. Otherwise expressed, the previous model needs to be ‘‘unlumpable’’ as it evolves into an increasingly complex ‘‘brain.’’ This is a type of *correspondence principle* that places a surprisingly severe test on the adequacy of the previously discovered theoretical principles. Many models regularly fail the embedding constraint. That is why they come and go with surprisingly rapidity and do not get integrated into burgeoning theories of ever greater predictive power.

The crucial importance of being able to derive behavioral mechanisms as emergent properties of real-time brain mechanisms, and being able to embed a previous model into a more mature model that is capable of adapting to more complex environments, led me to the name Embedding Fields for my earliest models of brain and behavior (Grossberg, 1964). The word ‘‘fields’’ is a short-hand for the neural network as a computational unit whose interactions generate behavioral emergent properties; the word ‘‘embedding’’ refers to the unlumpability constraint. Many stages of model evolution have occurred since the mid-1960s and all of them have successfully built a foundation for their progeny. The present article will necessarily omit these modeling cycles and will instead discuss some of its results from the viewpoint of consciousness research.

HOW DO WE PERCEIVE ILLUSORY CONTOURS AND BRIGHTNESS?

Let me start by providing several examples of the diverse phenomena that ART clarifies. Consider the images in Fig. 1. Figure 1A shows an image called an Ehrenstein figure in which some radial black lines are drawn on a uniformly white paper. Remarkably, our minds construct a circular illusory contour that touches each line end at a perpendicular orientation. This illusory contour is a collective, emergent property of all the lines that only occurs when their positions relative to each other are suitable. For example, no illusory contour forms at the line ends in Fig. 1B even though they end at the same positions as the lines in Fig. 1A. Note also that the illusory contour in Fig. 1A surrounds a disk that seems uniformly brighter than its surround. Where does the brightness enhancement come from? It certainly does not always happen when illusory contours from, as can be seen by inspecting Fig. 1C. Here a vertical illusory contour can be recognized as interpolating the two sets of offset horizontal lines, even though neither side of the contour seems brighter than the other. How we can consciously *recognize* something that we cannot *see* and is thus perceptually invisible is a fascinating aspect of our conscious awareness about which quite a bit is now known. Such percepts are known as *amodal* percepts (Mi-

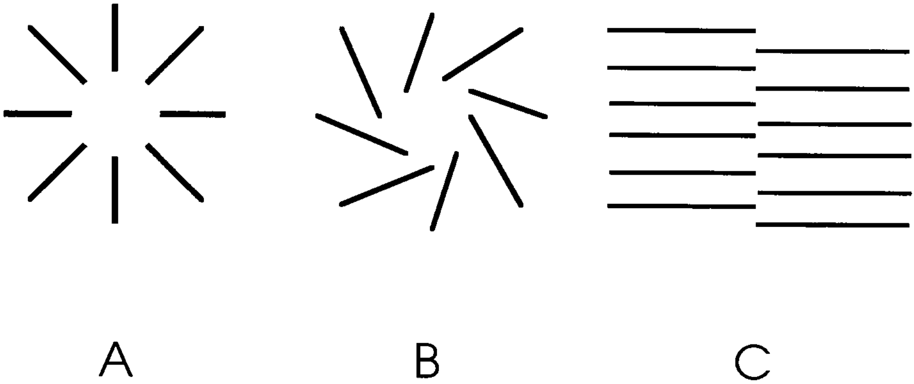


FIG. 1. (A) The Ehrenstein pattern generates a circular illusory contour that encloses a circular disk of enhanced illusory brightness. (B) If the endpoints of the Ehrenstein pattern remain fixed while their orientations are tilted, then both the illusory contour and the brightness vanish. (C) The offset pattern generates a vertical boundary that can be recognized even though it cannot be seen.

chotte, Thines, & Crabbe, 1964) in order to distinguish them from modal, or visible, percepts. Amodal percepts are experienced in response to many naturalistic scenes, notably in response to scenes in which some objects are partially occluded by other objects. How both modal and amodal percepts can occur will be discussed below. Of particular interest from the viewpoint of ART processing is why the Ehrenstein disk looks bright, despite the fact that there are no local contrasts within the image itself that describe a disk-like object.

HOW DO WE LEARN TO RECOGNIZE VISUALLY PERCEIVED OBJECTS?

The Ehrenstein example concerns the process of visual perception. The next example concerns a process that goes on at a higher level of the visual system. It is the process whereby we visually recognize objects. A key part of this process concerns how we learn to categorize specific instances of an object, or set of objects, into a more general concept. For example, how do we learn that many different printed or script letter fonts can all represent the same letter A? Or how do we learn that several different combinations of patient symptoms are all due to the same disease? Moreover, how do we control how general our categories will become? For some purposes, like recognizing a particular face, we need highly specific categories. For others, like knowing that every person has a face, the categories are much more general. Finally, how does our learning and memory break down when something goes wrong in our brain? For example, it is known that lesions to the human hippocampal system can cause a form of amnesia whereby, among other properties, patients find it very hard to learn new information and hard to remember recently learned information, but previously learned information about which their memory has ‘‘consolidated’’ can readily be retrieved. Thus, an amnesic patient can typically carry out a perfectly intelligent conversation about experiences that occurred a significant time before the lesion that caused the amnesia occurred.

What computational properties do the phenomena of bright illusory disks and amnesic memory have in common? I will suggest below that their apparent differences conceal the workings of a general unifying principle.

HOW DO WE SOLVE THE COCKTAIL PARTY PROBLEM?

To continue with our list, let us now consider a different modality entirely; namely, audition. When we talk to a friend in a crowded noisy room, we can usually keep track of our conversation above the hubbub, even though the sounds emitted by the friendly voice may be substantially overlapped by the sounds emitted by other speakers. How do we separate this jumbled mixture of sounds into distinct voices? This is often called the cocktail party problem. The same problem is solved whenever we listen to a symphony or other music wherein overlapping harmonic components are emitted by several instruments. If we could not separate the instruments or voices into distinct sources, or auditory streams, then we could not hear the music as music or intelligently recognize a speaker's sounds. A striking and ubiquitous property of such percepts, and one which has not yet been understood by alternative modeling approaches, is how future events can alter our conscious percepts of past events in a context-sensitive manner.

A simple version of this competence is illustrated by the auditory continuity illusion (Bregman, 1990). Suppose that a steady tone shuts off just as a broadband noise turns on. Suppose, moreover, that the noise shuts off just as the tone turns on once again; see Fig. 2A. When this happens under appropriate conditions, the tone seems to continue right through the noise, which seems to occur in a separate auditory ‘‘stream.’’ This example shows that the auditory system can actively extract those components of the noise that are consistent with the tone and use them to track the ‘‘voice’’ of the tone right through the hubbub of the noise.

In order to appreciate how remarkable this property is, let us compare it with what happens when the tone does not turn on again for a second time, as in Fig. 2B. Then the first tone does not seem to continue through the noise. It is perceived to stop before the noise. How does the brain know that the second tone will turn on after the noise shuts off so that it can continue the tone through the noise, yet not continue the tone through the noise if the second tone does not eventually occur? Does this not seem to require that the brain can operate ‘‘backward in time’’ to alter its decision as to whether to continue a past tone through the noise based on future events?

Many philosophers and scientists have puzzled about this sort of problem. I argue that the process whereby we consciously hear the first tone takes some time to unfold so that by the time we hear it, the second tone has already begun. To make this argument, we need to ask why does conscious audition take so long to occur after the actual sound energy reaches our brain? Just as important, why can the second tone influence the conscious percept so quickly, given that the first tone could not? Finally, I indicate what these auditory phenomena have to do with bright Ehrenstein disks and amnesia.

HOW DO WE CONSCIOUSLY PERCEIVE SPEECH?

The final examples also involve the auditory system, but at a higher level of processing. They concern how we understand speech. In these examples, too, the process

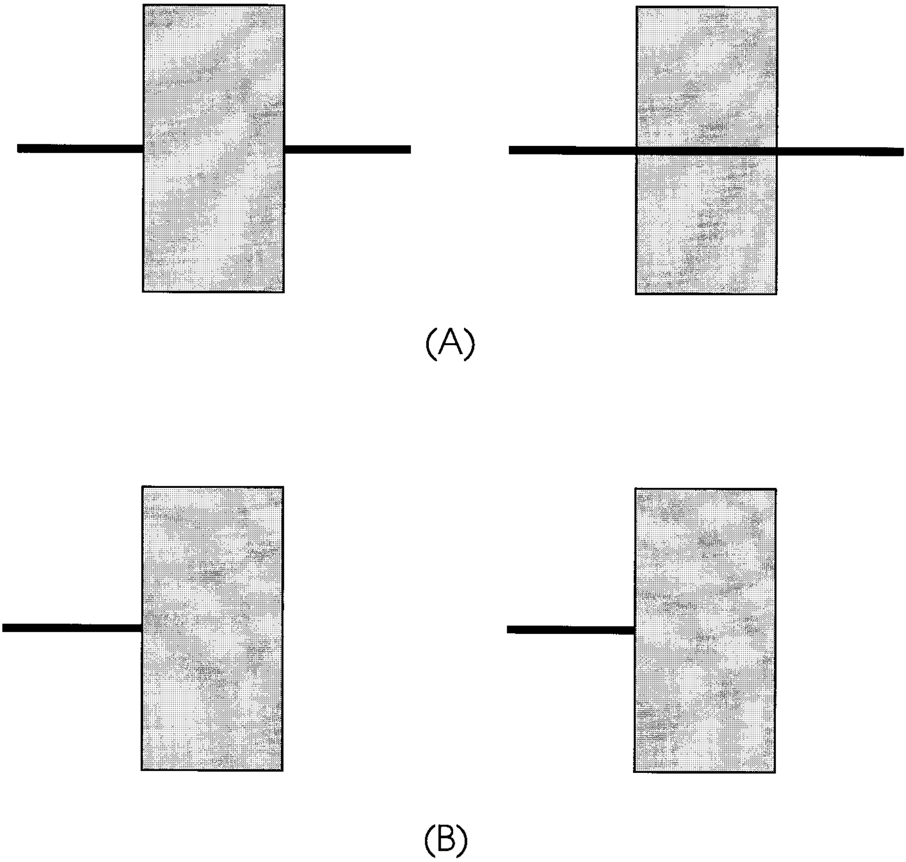


FIG. 2. (A) Auditory continuity illusion: When a steady tone occurs both before and after a burst of noise, then under appropriate temporal and amplitude conditions, the tone is perceived to continue through the noise. (B) This does not occur if the noise is not followed by a tone.

whereby conscious awareness occurs takes a long time, on the order of 100 ms or more. An analysis of these percepts will also give us more clues about the nature of the underlying process. The first example is called phonemic restoration. Suppose that a listener hears a noise followed immediately by the words ‘‘eel is on the . . . .’’ If this string of words is followed by the word ‘‘orange,’’ then ‘‘noise-eel’’ sounds like ‘‘peel.’’ If the word ‘‘wagon’’ completes the sentence, then ‘‘noise-eel’’ sounds like ‘‘wheel.’’ If the final word is ‘‘shoe,’’ then ‘‘noise-eel’’ sounds like ‘‘heel.’’

This marvelous example, which was developed by Richard Warren and his colleagues more than 20 years ago (Warren, 1984; Warren & Sherman, 1974), vividly shows that the bottom-up occurrence of the noise is not sufficient for us to hear it. Somehow the sound that we *expect* to hear based upon our previous language experiences influences what we do hear, at least if the sentence is said quickly enough. As in the auditory continuity illusion, it would appear that the brain is working ‘‘backward in time’’ to allow the meaning imparted by a later word to alter the sounds that we consciously perceive in an earlier word.

I suggest that this happens because as the individual words occur, they are stored temporarily via STM traces in a working memory. As the words are stored, they activate LTM traces which attempt to categorize the stored sound stream into familiar language units like words at a higher processing level. These list categories, in turn, activate learned top-down expectations that are *matched* against the contents of working memory to verify that the information expected from previous learning experiences is really there. This concept of bottom-up activation of learned categories by a working memory, followed by read-out of learned top-down expectations, is illustrated in Fig. 3A.

What is the nature of this matching, or verification, process? Its properties have been clarified by experiments of Arthur Samuel (Samuel, 1981a,b) and others in which the spectral content of the noise was varied. If the noise includes all the formants of the expected sound, then that is what the subject hears, and other spectral components of the noise are suppressed. If some formants of the expected sound are missing from the noise, then only a partial reconstruction is heard. If silence replaces the noise, then only silence is heard. The matching process thus cannot ‘‘create something out of nothing.’’ It can, however, selectively amplify the expected features in the bottom-up signal and suppress the rest, as in Fig. 3B.

The process whereby the top-down expectation selectively amplifies some features while suppressing others helps to ‘‘focus attention’’ upon information that matches our momentary expectations. This focusing process helps to filter out the flood of sensory signals that would otherwise overwhelm us and to prevent them from destabilizing our previously learned memories. Learned top-down expectations hereby help to solve the stability–plasticity dilemma by focusing attention and preventing spurious signals from accidentally eroding our previously learned memories. In fact, Gail Carpenter and I proved mathematically in 1987 that such an ART matching rule assures stable learning of an ART model in response to rapidly changing environments wherein learning becomes unstable if the matching rule is removed (Carpenter and Grossberg, 1987a).

What does all this have to do with our conscious percepts of speech? This can be seen by asking: If top-down expectations can select consistent bottom-up signals, then what keeps the selected bottom-up signals from reactivating their top-down expectations in a continuing cycle of bottom-up and top-down feedback? Nothing does. In fact, this reciprocal feedback process takes awhile to equilibrate, and when it does, the bottom-up and top-down signals lock the STM activity patterns of the interacting levels into a resonant state that lasts much longer and is more energetic than any individual activation. ART hereby suggests how only resonant states of the brain can achieve consciousness and that the time needed for a bottom-up/top-down resonance to develop helps to explain why a conscious percept of an event takes so long to occur after its bottom-up input is delivered.

The example of phonemic restoration also clarifies another key point about the conscious perception of speech. If noise precedes ‘‘eel is on the shoe,’’ we hear and understand the meaning of the sentence ‘‘heel is on the shoe.’’ If, however, noise is replaced by silence, we hear and understand the meaning of the sentence ‘‘eel is on the shoe’’ which has a quite different, and rather disgusting, meaning. This example shows that the process of resonance binds together information about both meaning

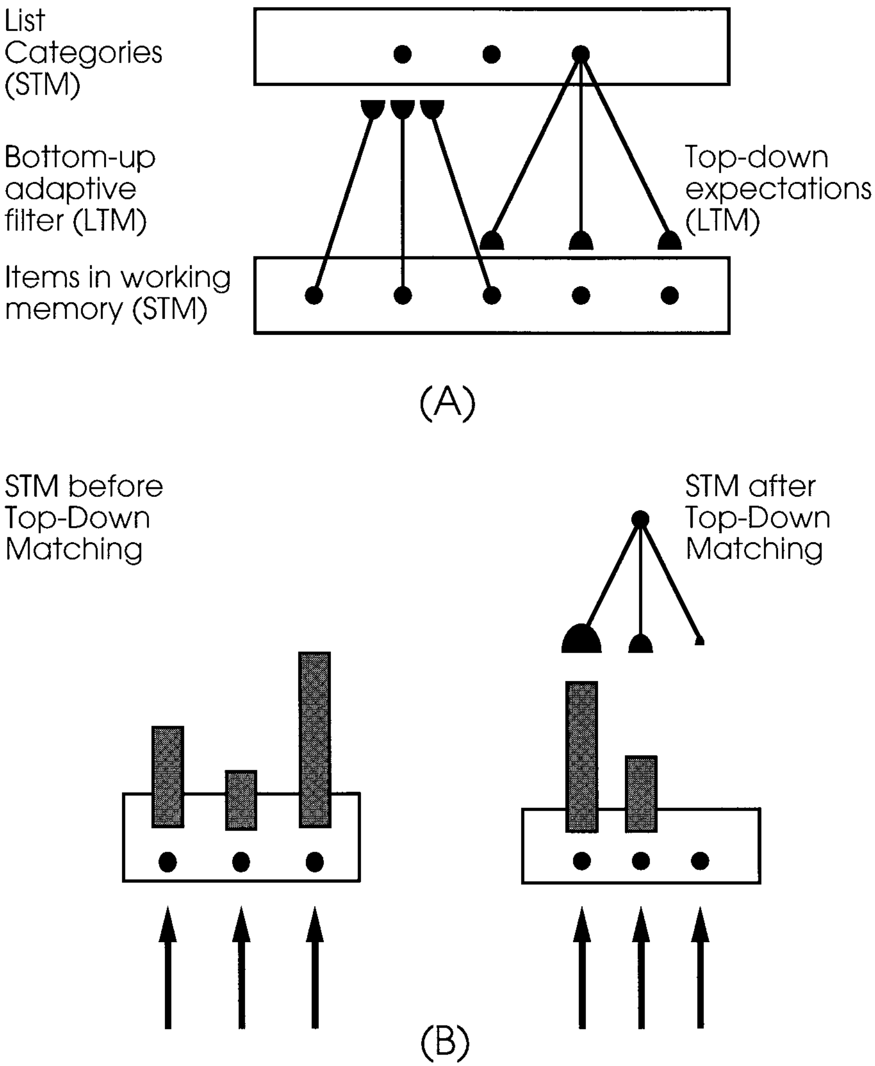


FIG. 3. (A) Auditory items activate STM traces in a working memory, which send bottom-up signals toward a level at which list categories, or chunks, are activated in STM. These bottom-up signals are multiplied by learned LTM traces which influence the selection of the list categories that are stored in STM. The list categories, in turn, activate LTM-modulated top-down expectation signals that are matched against the active STM pattern in working memory. (B) This matching process confirms and amplifies STM activations that are supported by contiguous LTM traces and suppresses those that are not.

and phonetics. Meaning is not some higher-order process that is processed independent from the process of conscious phonetic hearing. Meaning and phonetics are bound together via resonant feedback into a global emergent state in which the phonetics that we hear are linked to the meaning that we understand.

ART MATCHING AND RESONANCE: THE LINK BETWEEN ATTENTION, INTENTION, AND CONSCIOUSNESS

Adaptive resonance theory claims that in order to solve the stability–plasticity dilemma only resonant states can drive new learning. That is why the theory is called *adaptive* resonance theory. I explain how this works more completely below. Before doing so, let me emphasize some implications of the previous discussion that are worth reflecting about. The first implication provides a novel answer as to why, as philosophers have asked for many years, humans are ‘‘intentional’’ beings who are always anticipating or planning their next behaviors and their expected consequences. ART suggests that ‘‘stability implies intentionality.’’ That is, stable learning requires that we have expectations about the world that are continually matched against world data. Otherwise expressed, without stable learning, we could learn very little about the world. Having an active top-down matching mechanism greatly amplifies the amount of information that we can stably learn about the world. Thus the mechanisms which enable us to know a changing external world, through the use of learned expectations, set the stage for achieving internal self-awareness.

It should be noted here that the word ‘‘intentionality’’ is being used, at once, in two different senses. One sense concerns the role of expectations in the anticipation of events that may or may not occur. The second sense concerns the ability of expectations to read-out planned sequences of behaviors aimed at achieving definite behavioral goals. The former sense will be emphasized first; the latter toward the end of the article. My main point in lumping them together is that ART provides a unified mechanistic perspective with which to understand both uses of the word.

The second implication is that ‘‘intention implies attention and consciousness.’’ That is, expectations start to focus attention on data worthy of learning, and these attentional foci are confirmed when the system as a whole incorporates them into resonant states that include (I claim) conscious states of mind.

Implicit in the concept of intentionality is the idea that we can get *ready* to experience an expected event so that when it finally occurs we can react to it more quickly and vigorously, and until it occurs, we are able to ignore other, less desired, events. This property is called *priming*. It implies that, when a top-down expectation is read-out in the absence of a bottom-up input, it can subliminally sensitize the cells that would ordinarily respond to the bottom-up input, but not actually fire them, while it suppresses cells whose activity is not expected. Correspondingly, the ART matching rule computationally realizes the following properties at any processing level where bottom-up and top-down signals are matched: (1) *bottom-up automatic activation:* A cell, or node, can become active enough to generate output signals if it receives a large enough bottom-up input, other things being equal; (2) *top-down priming:* A cell can become sensitized, or subliminally active, and thus cannot generate output signals if it receives only a large top-down expectation input. Such a top-down priming signal prepares a cell to react more quickly and vigorously to subsequent bottom-up input that matches the top-down prime; (3) *match:* A cell can become active if it receives large convergent bottom-up and top-down inputs. Such a matching process can generate enhanced activation as resonance takes hold; (4) *mismatch:* A cell is suppressed even if it receives a large bottom-up input if it also receives only a small, or zero, top-down expectation input.

I claim that this ART matching rule and the resonance rule that it implies operate in all the examples that I have previously sketched and do so to solve the stability–plasticity dilemma. All the examples are proposed to illustrate how we can continue to learn rapidly and stably about new experiences throughout life by matching bottom-up signal patterns from more peripheral to more central brain processing stages against top-down signal patterns from more central to more peripheral processing stages. These top-down signals represent the brain's learned expectations of what the bottom-up signal patterns should be based upon past experience. The matching process is designed to reinforce and amplify those combinations of features in the bottom-up pattern that are consistent with the top-down expectations and to suppress those features that are inconsistent. This top-down matching step initiates the process whereby the brain selectively pays attention to experiences that it expects, binds them into coherent internal representations through resonant states, and incorporates them through learning into its knowledge about the world.

Given that such a resonant matching process occurs in the brain, how does the brain react when there is a mismatch situation? The ART matching rule suggests that a big enough mismatch between a bottom-up input and a top-down expectation can rapidly attenuate activity at the matching level. This collapse of bottom-up activation can initiate a rapid *reset* of activity at both the matching level itself and at the subsequent levels that it feeds, thereby initiating a memory search for a more appropriate recognition category or creating a new one.

RESONANT DYNAMICS DURING SPEECH CATEGORIZATION

Many examples of such a reset event occur during variable-rate speech perception. As one example, consider how people hear combinations of vowels (V) and consonants (C) in VC–CV sequences. Bruno Repp at Haskins Laboratories has studied perception of the sequences [ib]–[ga] and [ib]–[ba] when the silence interval between the initial VC syllable and the terminal CV syllable is varied (Repp, 1980). If the silence interval is short enough, then [ib]–[ga] sounds like [iga] and [ib]–[ba] sounds like [iba]. Repp ran a number of conditions, leading to the several data curves displayed in Fig. 4. The main point for present purposes is that the transition from a percept of [iba] to one of [ib]–[ba] occurs after 100–150 ms more silence than the transition from [iga] to [ib]–[ga]. One hundred milliseconds is a very long time relative to the time scale at which individual neurons can be activated. Why is this shift so large?

My colleagues Ian Boardman and Michael Cohen and I have quantitatively simulated these data using a model, called the ARTPHONE model, of how a resonant wave develops due to bottom-up and top-down signal exchanges between a working memory that represents the individual speech items and a list categorization network

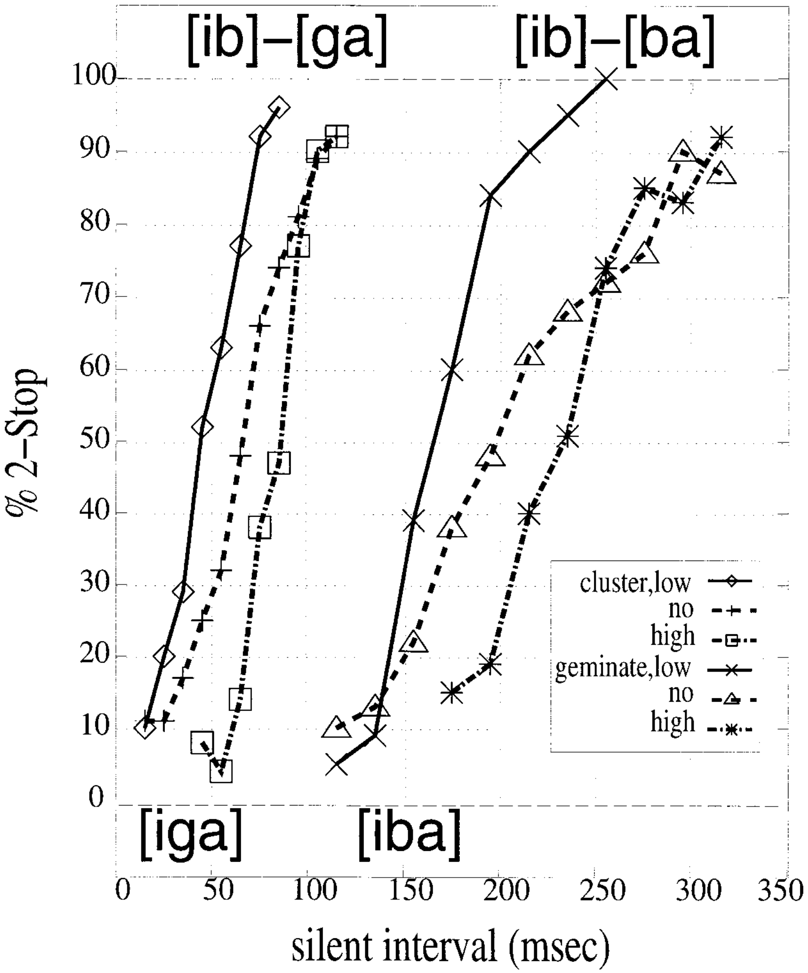


FIG. 4. The left-hand curves represent the probability, under several experimental conditions, that the subject will hear [ib]-[ga] rather than [iga]. The right-hand curves do the same for [ib]-[ba] rather than the fused percept [iba]. Note that the perception of [iba] can occur at a silence interval between [ib] and [ba] that is up to 150 ms longer than the one that leads to the percept [iga] instead of [ib]-[ga]. (Data are reprinted with permission from B.H. Repp (1980), Haskins Laboratories Status Report on Speech Research, SR-61, 151–165.)

that groups them together into learned language units, or chunks (Grossberg, Boardman, & Cohen, 1997). We have shown how a mismatch between [g] and [b] rapidly resets the working memory if the silence between them is short enough, thereby preventing the [b] sound from reaching resonance and consciousness, as in Fig. 5. We have also shown how the development of a previous resonance involving [b] can resonantly fuse with a subsequent [b] sound to greatly extend the perceived duration of [iba] across a silence interval between [ib] and [ba]. Figure 6A illustrates this property by suggesting how the second presentation of [b] can quickly reactivate the

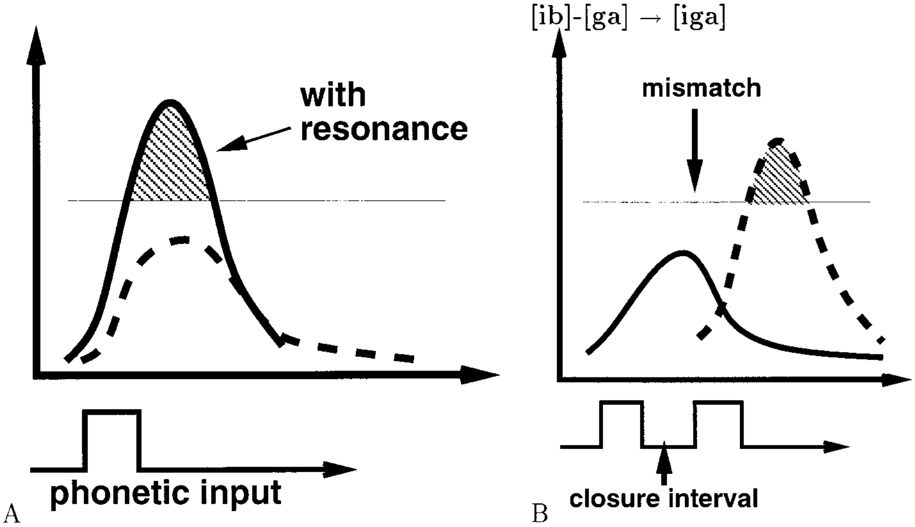


FIG. 5. (A) Response to a single stop, such as [b] or [g], with and without resonance. Suprathreshold activation is shaded. (B) Reset due to phonologic mismatch between [ib] and [ga].

resonance in response to the first presentation of [b] before the resonance stops. This phenomenon uses the property that it takes longer for the first presentation of [b] to reach resonance than it does for the second presentation of [b] to influence the maintenance of this resonance.

If, however, [ib] can fuse across time with [ba], then how do we ever hear distinct [ib]–[ba] sounds when the silence gets long enough? Much evidence suggests that after a resonance fully develops, it spontaneously collapses after awhile due to a habituative process that goes on in the pathways that maintain the resonance via

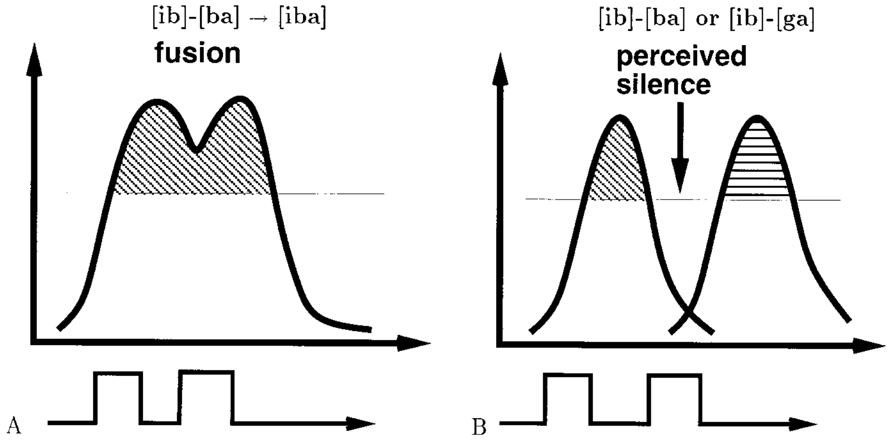


FIG. 6. (A) Fusion in response to proximal similar phones. (B) Perceptual silence allows a two-stop percept.

bottom-up and top-down signals. Thus, if the silence is long enough for resonant collapse of [ib] to occur, then a distinguishable [ba] resonance can subsequently develop and be heard, as in Fig. 6B.

Such a habituative process has also been used to explain many other data about perception, learning, and recognition, notably data about the reset of visual, cognitive, or motor representations in response to rapidly changing events. Relevant visual data include properties of light adaptation, visual persistence, aftereffects, residual traces, and apparent motion (Carpenter & Grossberg, 1981; Francis & Grossberg, 1996a,b; Francis, Grossberg, & Mingolla, 1994). Abbott *et al.* (1997) have recently reported data from the visual cortex that they modeled using the same habituative law that was used in all of these applications. At bottom, such a habituative law is predicted to be found so ubiquitously across brain systems because it helps to rapidly adapt, reset and rebalance neural circuits in response to rapidly changing input conditions, notably as part of an opponent process (Grossberg, 1980).

The Repp (1980) data illustrate the important fact that the duration of a consciously perceived interval of silence is sensitive to the phonetic context into which the silence is placed. These data show that the phonetic context can generate a conscious percept of continuous sound across 150 ms of silence—that can be heard as silence in a different phonetic context. Our explanation of these data in terms of the maintenance of resonance in one case, but its rapid reset in another, is consistent with a simple, but revolutionary, definition of silence: Silence is a temporal discontinuity in the rate with which the auditory resonance evolves in time. Various other models of speech perception, having no concept like resonance on which to build, cannot begin to explain data of this type. Several such models are reviewed in Grossberg, Boardman, and Cohen (1997).

RESONANT DYNAMICS DURING AUDITORY STREAMING

A similar type of resonant processing helps to explain cocktail party separation of distinct voices into auditory streams, as in the auditory continuity illusion of Fig. 2. This process goes on, however, at earlier stages of auditory processing than speech categorization. My colleagues Krishna Govindarajan, Lonce Wyse, and Michael Cohen and I have developed a model, called the ARTSTREAM model, of how distinguishable auditory streams are resonantly formed and separated (Grossberg, 1998b; Govindarajan, Grossberg, Wyse, & Cohen, 1995). Here the two main processing levels (Fig. 7) are a spectral stream level at which the frequencies of the sound spectrum are represented across a spatial map and a pitch stream level at which pitch nodes respond to the harmonics at the spectral stream level that comprise a given pitch. After the auditory signal is preprocessed, its spectral, or frequency, components are redundantly represented in multiple spectral streams; that is, the sound's preprocessed frequency components are represented in multiple spatial maps, each one of which can subserve the percept of a particular auditory stream.

Each of these spectral streams is filtered by bottom-up signals that activate its own pitch stream representation at the pitch stream level; that is, there are multiple pitch streams, one corresponding to every spectral stream. This multiple representation of a sound's spectral components and pitch interact to break up the entire sound stream

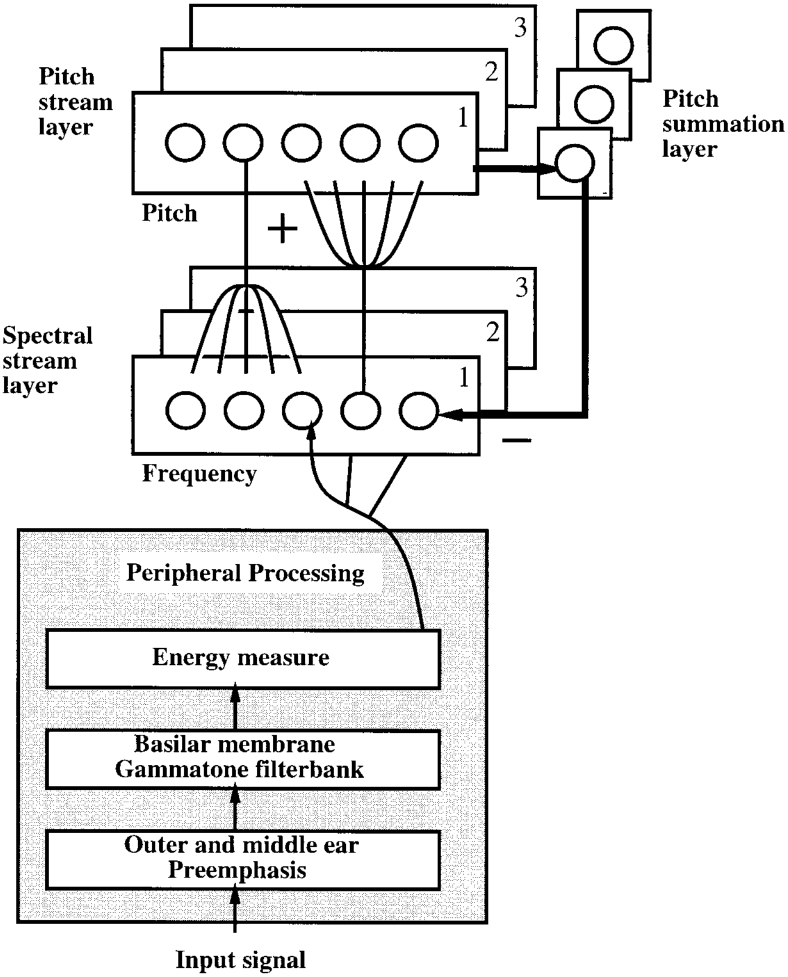


FIG. 7. Block diagram of the ARTSTREAM auditory streaming model. Note the nonspecific top-down inhibitory signals from the pitch level to the spectral level that realize ART matching within the network.

that is entering the system into distinct acoustic sources or voices. This happens as follows. A given sound spectrum is multiply represented at all the spectral streams and then redundantly activates all of the pitch nodes that are consistent with these sounds. These pitch representations compete to select a winner, which inhibits the representations of the same pitch across streams, while also sending top-down matching signals back to the spectral stream level. By the ART matching rule, the frequency components that are consistent with the winning pitch node are amplified, and all others are suppressed, thereby leading to a spectral-pitch resonance within the stream of the winning pitch node. In this way, the pitch layer coherently binds together the harmonically related frequency components that correspond to a prescribed auditory

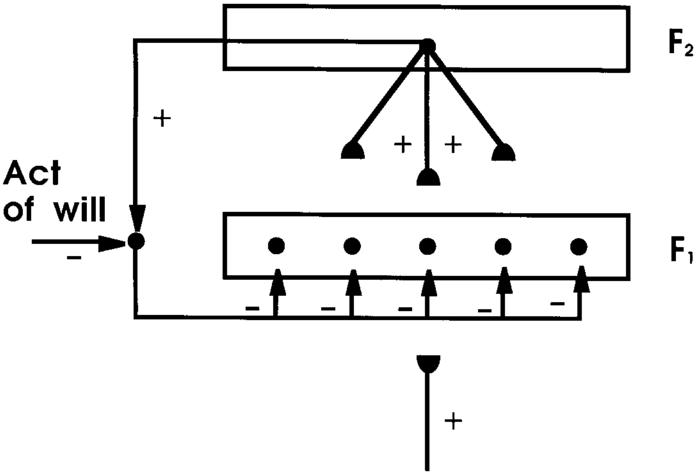


FIG. 8. One way to realize the ART matching rule using top-down activation of nonspecific inhibitory interneurons. Several mathematically possible alternative ways are suggested in the Appendix of G. A. Carpenter and S. Grossberg (1987a).

source. All the frequency components that are suppressed by ART matching in this stream are freed to activate and resonate with a different pitch in a different stream. The net result is multiple resonances, each selectively grouping together into pitches those frequencies that correspond to distinct auditory sources.

Using the ARTSTREAM model, we have simulated many of basic streaming percepts, including the auditory continuity illusion of Fig. 2. It occurs, I contend, because the spectral stream resonance takes a time to develop that is commensurate to the duration of the subsequent noise. Once the tone resonance develops, the second tone can quickly act to support and maintain it throughout the duration of the noise, much as [ba] fuses with [ib] during perception of [iba]. Of course, for this to make sense, one needs to accept the fact that the tone resonance does not start to get consciously heard until just about when the second tone occurs.

A CIRCUIT FOR ART MATCHING

Figure 7 incorporates one of the possible ways that Gail Carpenter and I proposed in the mid-1980s for how the ART matching rule can be realized (Carpenter & Grossberg, 1987a). This matching circuit is redrawn in Fig. 8 for clarity. It is perhaps the simplest such circuit, and I have found it in subsequent studies to be the one that is implicated by data time and time again.

In this circuit, bottom-up signals to the spectral stream level can excite their target nodes if top-down signals are not active. Top-down signals try to excite those spectral, or frequency component, nodes that are consistent with the pitch node that activates them. By themselves, top-down signals fail to activate spectral nodes because the pitch node also activates a pitch summation layer that nonspecifically inhibits all spectral nodes in its stream. The nonspecific top-down inhibition hereby prevents the specific top-down excitation from supraliminally activating any spectral nodes. On the other hand, when excitatory bottom-up and top-down signals occur together, then those spectral nodes that receive both types of signals can be fully activated. All other nodes in that stream are inhibited, including spectral nodes that were previously activated by bottom-up signals but received no subsequent top-down pitch support. Attention hereby selectively activates consistent nodes while nonselectively inhibiting all other nodes in a stream.

RESONANT DYNAMICS DURING BRIGHTNESS PERCEPTION

Having come this far, let us review how ART matching and resonance help to explain the enhanced brightness of the Ehrenstein disk in Fig. 1A. This apparently simple percept has attracted a great deal of attention from vision scientists because one could imagine many reasons why no brightness difference or the reverse brightness difference might have been seen instead. John Kennedy (Kennedy, 1979, 1988) has attempted to explain this percept by positing that ‘‘brightness buttons’’ occur at the ends of dark (low luminance) lines. The textbook mechanism for explaining these brightness buttons has, in turn, for decades been an appeal to the on-center, off-surround receptive fields of early visual processing. A cell that possesses such a receptive field is excited by inputs near the cell’s location (the on-center) but inhibited by inputs to more distant locations (the off-surround).

An analysis of how such cells respond to dark lines shows, however, that they cannot, by themselves, explain brightness buttons. I show below why neither on-center off-surround cells (called ON cells below) nor off-center on-surround cells (called OFF cells below) can explain this phenomenon. Such ON and OFF cells occur in the lateral geniculate nucleus (or LGN), which is a waystation from the photosensitive retina in the eye to the visual cortex. Thus the ON and OFF cells that occur in the LGN, and that are the source of cortical brightness percepts, cannot explain brightness buttons without further processing. Figure 9 shows that whatever contribution to area contrast is generated at the ends of thin lines by ON or OFF cells must be less in magnitude than that generated along their sides. As explained below, this should make the Ehrenstein disk appear darker, rather than brighter, than its surround.

To see why this is so, assume, as in Fig. 9B, that the thin line is black (low luminance) and surrounded by a white (high luminance) background. Since OFF cells respond best to low luminance in their receptive field center and high luminance in their surround, OFF cells whose centers lie inside the line will be activated. Furthermore, OFF cells near the line end (but still inside the line) will be more strongly activated than OFF cells in the middle of the line because the line end is more like a black disk surrounded by a white background than the line middle is (Fig. 9B). That is, an OFF cell whose center lies in the line end receives less inhibition from its surround than does a cell centered in the middle of the line because a larger area of the former cell’s surround lies in the white background.

A similar analysis can be applied to the ON cells. An ON cell is excited by high luminance in the center of its receptive field and low luminance in its surround. The ON cells that are active, then, are those centered outside the bar. An ON cell whose center is just outside the side of the line will respond more strongly than an ON cell centered just outside the end of the line (Fig. 9C).

Given that LGN ON and OFF cells, by themselves, cannot explain brightness but-

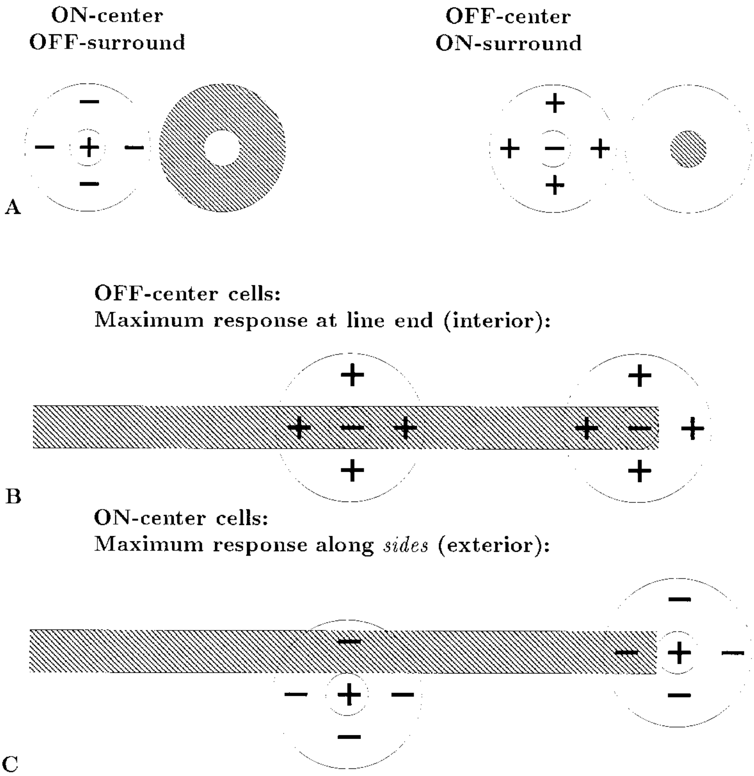


FIG. 9. Retinal center-surround cells and their optimal stimuli (A). The ON cell, on the left, responds best to a high-luminance disk surrounded by a low-luminance annulus. The OFF cell, on the right, responds best to a low-luminance disk surrounded by a high-luminance annulus (B). OFF cells respond to the inside of a black line. The OFF cell centered at the line end responds more strongly than the OFF cell centered in the middle because the surround region of the former cell is closer to optimal. In C ON cells respond to the white background just outside the black line. The amount of overlap of each ON cell's surround with the black line affects the strength of the cell's response. As seen in the ON cell's optimal stimulus (C), the more of the surround that is stimulated by a black region, the better the ON cell will respond. Thus, an ON cell centered just outside the side of the line will respond better than a cell centered just outside the end of the line because more of the off-surround is activated at the end of the line than along its side.

tons, an additional explanation needs to be found for how a brighter Ehrenstein disk could be generated. Clues were provided by John Kennedy, who analyzed a number of illusory contour stimuli. He argued that the effect of brightness buttons could often go unnoticed for isolated line segments, but could somehow be pooled and amplified in perceptual salience when several brightness buttons occurred in proximity or within a figurally complete region. In the mid-1980s, I worked with several colleagues to develop an analysis and interpretation of Kennedy's remarks by developing a neural model of visual boundary and surface representation (Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985a, b; Grossberg & Todorović, 1988).

In this model, the crucial mechanistic support for perceptually noticeable brightness buttons is a boundary segmentation that separates the region containing the buttons from other regions of a scene. Such a boundary segmentation may be generated by image edges, textures, or shading and may give rise to illusory contours such as the Ehrenstein circle. We suggested how brightness buttons could, at a later processing stage, activate a diffusion process that could ‘‘fill-in’’ a uniform level of brightness within the bounding illusory contour. The model successfully explained and predicted many facts about illusory contours and brightness percepts, among other phenomena, but it incorrectly predicted that the Ehrenstein disk should look darker than its surround. Given that so many brightness data had been correctly predicted by the model, including data collected after its publication, the question arose of how the model’s description was incomplete or incorrect. Such an analysis was recently carried out with Alan Gove and Ennio Mingolla (Gove, Grossberg, & Mingolla, 1995). We showed how the addition of a feedback loop from the visual cortex to the LGN helps to explain brightness buttons without disturbing the model’s previous explanations of other brightness phenomena.

The gist of this analysis can be summarized as follows. Brightness buttons are by definition an effect of an *oriented* structure such as a line or, more generally, a corner or sharp bend in a contour, on perceived brightness. Within the prior model, the computations leading to brightness perception were unoriented in the sense that they were initiated by ON and OFF cells with circularly symmetric receptive fields. How then could the effects of oriented filtering be used to modulate the inputs to the process that produces brightness buttons? Indeed, oriented filtering alone could not suffice. Interactions must exist among the oriented filters to determine the location of the ends of the lines at which the brightness buttons occur. A natural candidate for the latter interactions is the cortical *endstopping* process that has been known, since the Nobel-prize winning work of David Hubel and Thorstein Wiesel, to convert cortical complex cells into endstopped complex, or hypercomplex, cells (Hubel & Wiesel, 1977). These oriented cells are selectively activated at and near the ends of lines. Where should the results of this endstopped processing have their effect on brightness processing?

Having come this far, it is plausible to propose that the cortex influences LGN cells via top-down feedback, which it is well known to do. It is not plausible, however, that this massive feedback pathway exists just to make Ehrenstein disks appear bright. I had, however, earlier predicted that corticogeniculate feedback exists for a potentially important functional reason; namely, to enhance the activity of LGN cells that support the activity of presently active cortical cells and to suppress the activity of LGN cells that do not (Grossberg, 1976a,b, 1980). In addition, bottom-up retinal input, by itself, was hypothesized to supraliminally activate LGN cells, but top-down corticogeniculate feedback, by itself, was not. In other words, corticogeniculate feedback was predicted to realize an ART matching and resonance rule in order to control and stabilize learned changes in cortical LTM traces in response to the flood of visual experience.

Figure 10 summarizes how this type of corticogeniculate feedback can produce brightness buttons. Figure 11B summarizes a computer simulation of brightness buttons. The model’s boundary completion network generates the circular illusory contour of Fig. 11C. The brightness button activation pattern in Fig. 11B generates a

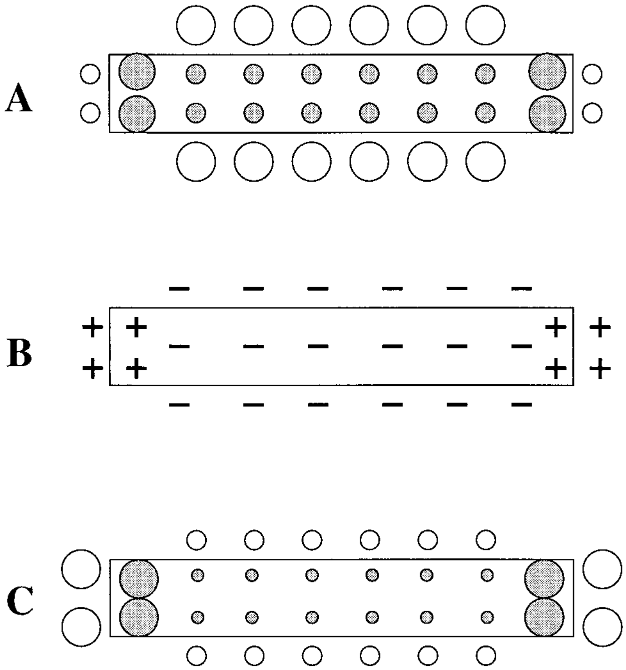


FIG. 10. Schematic diagram of brightness button formation in the model. In A the distribution of model LGN cell activities prior to receiving any feedback in response to a black bar is illustrated. Open circles code ON cell activity; filled circles code OFF cell activity. (B) The effect of feedback in bottom-up LGN activations. The plus (minus) signs designate the excitatory (inhibitory) top-down influence of an oriented endstopped cortical cell. (C) The LGN activity distribution after endstopped feedback, such as that in B, combines with the direct effect of ON and OFF cell processing, such as that in A. A brightness button is formed outside both ends of the line.

topographic input to a filling-in domain, wherein the inputs diffuse freely in all directions until they hit a barrier to filling-in that is imposed by the circular boundary signals in Fig. 11C. The result is an Ehrenstein disk with uniformly enhanced brightness relative to its surround in Fig. 11D.

Is there direct experimental evidence that corticogeniculate feedback can alter LGN cell properties as desired? Murphy and Sillito (1987) showed that cortical feedback causes significant length-tuning in cat LGN cells. As in cortical endstopping, the response to a line grows rapidly as a function of line length and then abruptly declines for longer lines. The response to long lines is hereby depressed. Redies *et al.* (1986) found that cat dorsal LGN cells and strongly endstopped cortical complex cells responded best at line ends. In other words, the response of the LGN cells to line ends was enhanced relative to the response to line sides.

Is there direct experimental evidence for the prediction that corticogeniculate feedback supports ART matching and resonance? In a remarkable 1994 *Nature* article, Sillito and his colleagues (Sillito *et al.*, 1994) published neurophysiological data that strikingly support this prediction. They wrote in particular that ‘‘cortically induced correlation of relay cell activity produces coherent firing in those groups of relay

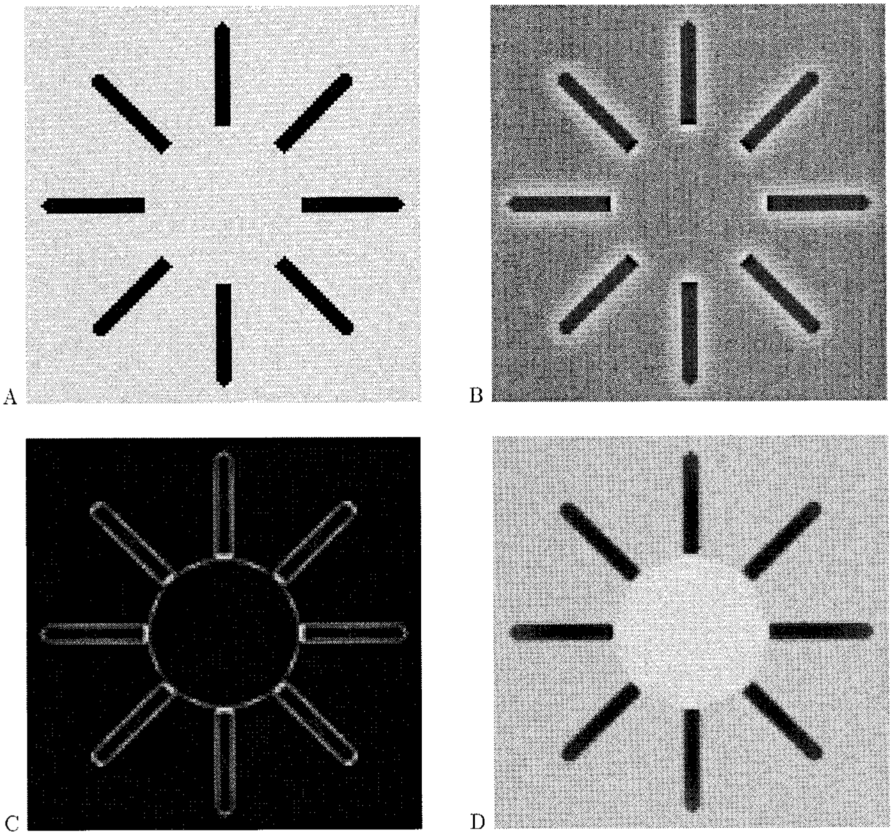


FIG. 11. (A) The Ehrenstein figure. (B) The LGN stage response. Both ON and OFF cell activities are coded as rectified deflections from a neutral gray. Note the brightness buttons at the line ends. (C) The equilibrium boundaries. (D) In the filled-in surface brightness, the central disk contains larger activities than the background, corresponding to the perception of increased brightness. (Reprinted with permission from Gove, Grossberg, & Mingolla, 1995.)

cells with receptive field alignments appropriate to signal the particular orientation of the moving contour to the cortex . . . this increases the gain of the input for feature-linked events detected by the cortex . . . the cortico-thalamic input is only strong enough to exert an effect on those dLGN cells that are additionally polarized by their retinal input . . . the feedback circuit searches for correlations that support the 'hypothesis' represented by a particular pattern of cortical activity.'' In short, Sillito verified all the properties of the ART matching rule.

HOW EARLY DOES ATTENTION ACT IN THE BRAIN?

If we take these results at face value, then it would appear that corticogeniculate feedback helps to ''focus attention'' upon expected patterns of LGN activity. However, it is typically argued that visual attention first acts at much higher levels of cortical organization, starting with the extrastriate visual cortex. Is there a contradiction here? The answer depends upon how you define attention. If attention refers only to processes that can be controlled voluntarily, then corticogeniculate feedback, being automatic, may not qualify. On the other hand, corticogeniculate feedback does appear to have the selective properties of an ‘‘automatic’’ attention process.

ATTENTION AT ALL STAGES OF SENSORY AND COGNITIVE NEOCORTEX?

It has, in fact, been suggested how similar automatic attentional processes are integrated within the laminar circuits of visual cortex, notably the circuits of cortical areas V1 and V2 that are used to generate perceptual groupings, such as the illusory contours in Fig. 1 (Grossberg, 1998a). In this proposal, the ART matching rule is realized as follows. Top-down attentional feedback from cortical area V2 to V1 is predicted to be mediated by signals from layer 6 of cortical area V2. These top-down signals attentionally prime layer 4 of cortical area V1 via an on-center off-surround network within V1 from layer 6 to layer 4. In this conception, layer 6 of V2 activates layer 6 of V1, possibly via a multisynaptic pathway, which in turn activates layer 4 of V1 via an on-center off-surround network from layer 6 to layer 4. This analysis predicts that the layer-6-to-layer-4 on-center off-surround circuit can *modulate* layer 4 cells, but cannot fully activate them because the top-down attentional prime, acting by itself, is subliminal. Such a modulatory effect is achieved by appropriately balancing the strength of the on-center and off-surround signals within the layer-6-to-layer-4 network.

Related modeling work has shown how such balanced on-center off-surround signals can lead to self-stabilizing development of the horizontal connections within layer 2/3 of V1 and V2 that subserve perceptual grouping (Grossberg & Williamson, 1997, 1998). It has also been shown how the top-down on-center off-surround circuit from area V1 to LGN can self-stabilize the development of disparity-sensitive complex cells in area V1 (Grunewald & Grossberg, 1998). Other modeling work has suggested how a similar top-down on-center off-surround automatical attentional circuit from cortical area MST to MT can be used to generate coherent representations of the direction and speed with which objects move (Chey, Grossberg, & Mingolla, 1997). Taken together, these studies show how the ART Matching Rule may be realized in known cortical circuits, and how it can self-stabilize development of these circuits as a precursor to its role in self-stabilizing later learning throughout life. Grossberg (1998a) has predicted that the same ART matching circuit exists within the laminar organization that is found universally in all sensory and cognitive neocortex, including the various examples of auditory processing that are reviewed above. This prediction does not, of course, deny that these circuits may be specialized in various ways to process the different types of information with which they are confronted.

Given that the cortical organization of top-down on-center off-surround attentional priming circuits seem to be ubiquitous in visual cortex, and by extension in other types of cortex, it is important to ask: What more does the brain need to add in order to generate a more flexible, task-dependent type of attention switching? This question

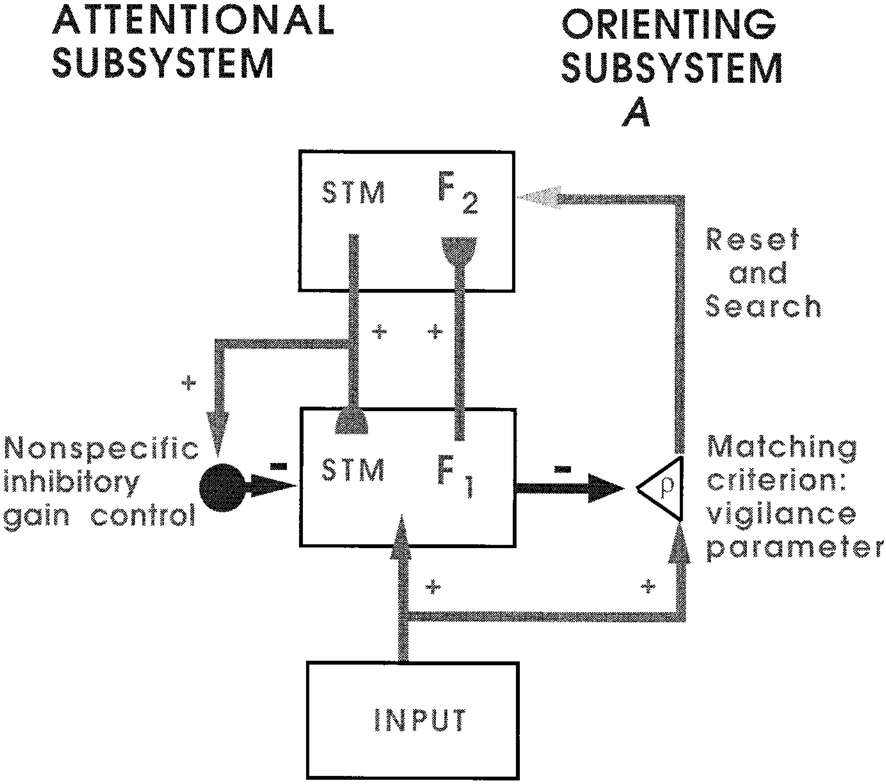


FIG. 12. An example of a model ART circuit in which attentional and orienting circuits interact. Level F_1 encodes a distributed representation of an event by a short-term memory (STM) activation pattern across a network of feature detectors. Level F_2 encodes the event using a compressed STM representation of the F_1 pattern. Learning of these recognition codes occurs at the long-term memory (LTM) traces within the bottom-up and top-down pathways between levels F_1 and F_2. The top-down pathways read-out learned expectations whose prototypes are matched against bottom-up input patterns at F_1. The size of mismatches in response to novel events are evaluated relative to the vigilance parameter ρ of the orienting subsystem A. A large enough mismatch resets the recognition code that is active in STM at F_2 and initiates a memory search for a more appropriate recognition code. Output from subsystem A can also trigger an orienting response.

leads us to our last example, that of visual object recognition, and how it breaks down during medial temporal amnesia. Various other models of object recognition, and their conceptual and explanatory weaknesses relative to ART, are reviewed in Grossberg and Merrill (1996).

SELF-ORGANIZING FEATURE MAPS FOR LEARNED OBJECT RECOGNITION

Let us begin with a two-level network that illustrates some of the main ideas in the simplest possible way. Level F_1 in Fig. 12 contains a network of nodes, or cell populations, each of which is activated by a particular combination of sensory features via inputs. Level F_2 contains a network of nodes that represent recognition codes, or categories, which are selectively activated by the activation patterns across F_1. Each F_1 node sends output signals to a subset of F_2 nodes. Each F_2 node thus receives inputs from many F_1 nodes. The thick bottom-up pathway from F_1 to F_2 in Fig. 12 represents in a concise way an array of diverging and converging pathways. Let learning take place at the synapses denoted by semicircular endings in the $F_1 \rightarrow F_2$ pathways. Pathways that end in arrowheads do not undergo learning. This bottom-up learning enables F_2 category nodes to become selectively tuned to particular combinations of activation patterns across F_1 feature detectors by changing their LTM traces.

Why is not bottom-up learning sufficient in a system that can autonomously solve the stability–plasticity dilemma? Why are learned top-down expectations also needed? To understand this, we consider a type of model that is often called a self-organizing feature map, competitive learning, or learned vector quantization. This type of model shows how to combine associative learning and lateral inhibition for purposes of learned categorization.

In such a model, as shown in Fig. 13A, an input pattern registers itself as a pattern of activity, or STM, across the feature detectors of level F_1. Each F_1 output signal is multiplied, or gated, by the adaptive weight, or LTM trace, in its respective pathway. All these LTM-gated inputs are added up at their target F_2 nodes. The LTM traces hereby *filter* the STM signal pattern and generate larger inputs to those F_2 nodes whose LTM patterns are most similar to the STM pattern. Lateral inhibitory, or competitive, interactions within F_2 contrast-enhance this input pattern. Whereas many F_2 nodes may receive inputs from F_1, lateral inhibition allows a much smaller set of F_2 nodes to store their activation in STM. These are the F_2 nodes, whose LTM patterns are most similar to the STM pattern. These inhibitory interactions also tend to conserve the total activity that is stored in STM (Grossberg, 1982), thereby realizing an interference-based capacity limitation in STM.

Only the F_2 nodes that win the competition and store their activity in STM can influence the learning process. STM activity opens a learning gate at the LTM traces that abut the winning nodes. These LTM traces can then approach, or track, the input signals in their pathways, a process called steepest descent. This learning law is thus often called gated steepest descent, or instar learning. This type of learning tunes the winning LTM patterns to become even more similar to the STM pattern and to thereby enable the STM pattern to more effectively activate the corresponding F_2 nodes. I introduced this learning law into neural network models in the 1960s (e.g., Grossberg, 1969) and into ART models in the 1970s (Grossberg, 1976a,b, 1978, 1980). Such an LTM trace can either increase (Hebbian) or decrease (anti-Hebbian) to track the signals in its pathway (Table 1). It has been used to model neurophysiological data about learning in the hippocampus (also called long-term potentiation and long-term depression) and about adaptive tuning of cortical feature detectors during early visual development (Artola & Singer, 1993; Levy, 1985; Levy & Desmond, 1985; Rauschecker & Singer, 1979; Singer 1983), thereby lending support to ART predictions that these systems would employ this type of learning.

Self-organizing feature map models were introduced and computationally characterized by Christoph von der Malsburg and me during the 1970s (Grossberg, 1972, 1976a, 1978; von der Malsburg, 1973; Willshaw & Malsburg, 1976). These models were subsequently applied and further developed by many authors, notably Teuvo Kohonen (Kohonen, 1984). They exhibit many useful properties, especially if not too many input patterns, or clusters of input patterns, perturb level F_1 relative to the number of categorizing nodes in level F_2. I proved that under these sparse environmental conditions category learning is stable in the sense that its LTM traces converge to fixed values as learning trials proceed. In addition, the LTM traces track the statistics of the environment, are self-normalizing, and oscillate a minimum number of times (Grossberg, 1976a). Also, the category selection rule, like a Bayesian classifier, tends to minimize error. I also proved, however, that under *arbitrary* environmental conditions, learning becomes unstable (Grossberg, 1976b). Such a model could forget your parents' faces when it learns a new face. Although a gradual switching off of plasticity can partially overcome this problem, such a mechanism cannot work in a learning system whose plasticity is maintained throughout adulthood.

This memory instability is due to basic properties of associative learning and lateral inhibition, which are two processes that occur ubiquitously in the brain. An analysis of this instability, together with data about human and animal categorization, conditioning, and attention, led me to introduce ART models to stabilize the memory of self-organizing feature maps in response to an arbitrary stream of input patterns.

HOW DOES ART STABILIZE LEARNING OF A SELF-ORGANIZING FEATURE MAP?

How does an ART model prevent such instabilities from developing? As noted above, in an ART model, learning does not occur when some winning F_2 activities are stored in STM. Instead, activation of F_2 nodes may be interpreted as ‘‘making a hypothesis’’ about an input at F_1. When F_2 is activated, it quickly generates an output pattern that is transmitted along the top-down adaptive pathways from F_2 to F_1. These top-down signals are multiplied in their respective pathways by LTM traces at the semicircular synaptic knobs of Fig. 13B. The LTM-gated signals from all the active F_2 nodes are added to generate the total top-down feedback pattern from F_2 to F_1. It is this pattern that plays the role of a learned expectation. Activation of this expectation may be interpreted as ‘‘testing the hypothesis,’’ or ‘‘reading out the prototype,’’ of the active F_2 category. As shown in Fig. 13B, ART networks are designed to match the ‘‘expected prototype’’ of the category against the bottom-up input pattern, or exemplar, to F_1. Nodes that are activated by this exemplar are suppressed if they do not correspond to large LTM traces in the top-down prototype pattern. The resultant F_1 pattern encodes the cluster of input features that the network deems relevant to the hypothesis based upon its past experience. This resultant activity pattern, called X* in Fig. 13B, encodes the pattern of features to which the network ‘‘pays attention.’’

If the expectation is close enough to the input exemplar, then a state of resonance develops as the attentional focus takes hold. The pattern X* of attended features reactivates the F_2 category Y which, in turn, reactivates X*. The network locks into a resonant state through a positive feedback loop that dynamically links, or binds,

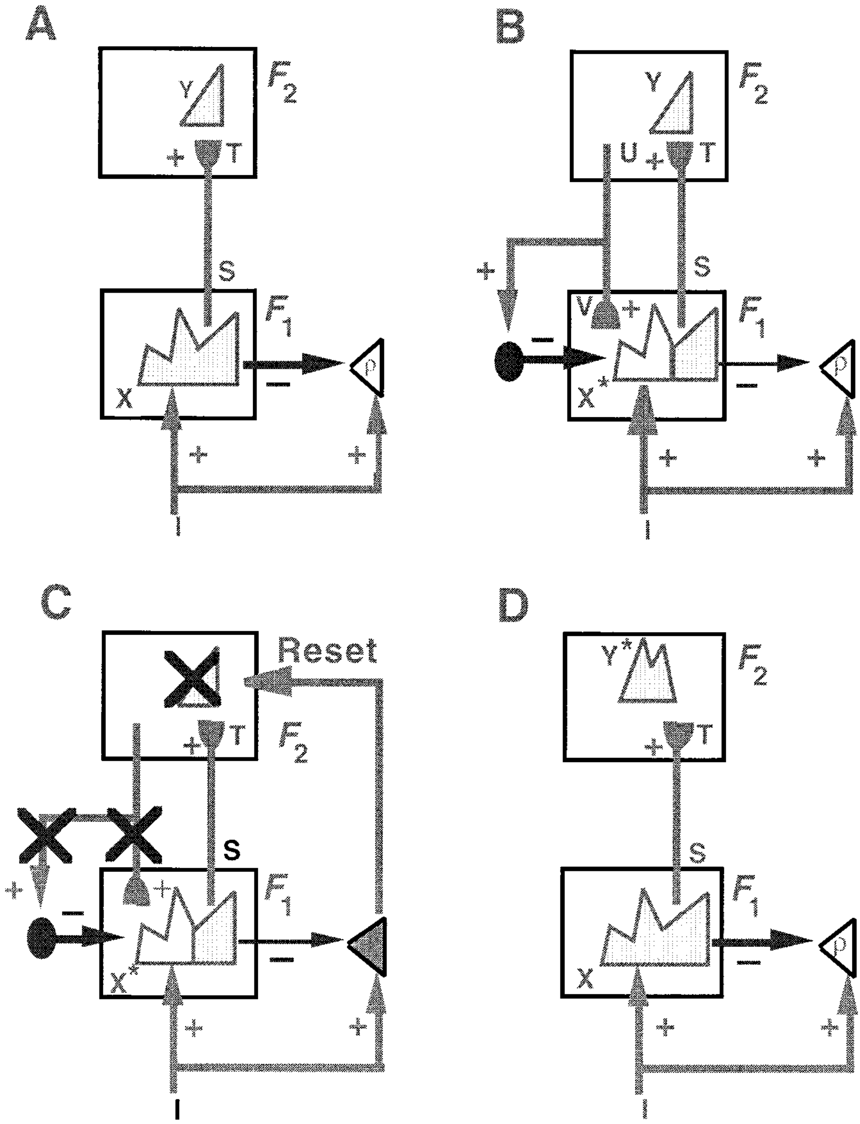
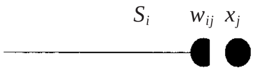


FIG. 13. ART search for a recognition code: (A) The input pattern I is instated across the feature detectors at level F_1 as a short-term memory (STM) activity pattern X. Input I also nonspecifically activates the orienting subsystem A; see Fig. 12. STM pattern X is represented by the hatched pattern across F_1. Pattern X both inhibits A and generates the output pattern S. Pattern S is multiplied by long-term memory (LTM) traces and added at F_2 nodes to form the input pattern T, which activates the STM pattern Y across the recognition categories coded at level F_2. (B) Pattern Y generates the top-down output pattern U, which is multiplied by top-down LTM traces and added at F_1 nodes to form the prototype pattern V that encodes the learned expectation of the active F_2 nodes. If V mismatches I at F_1, then a new STM activity pattern X* is generated at F_1. X* is represented by the hatched pattern. It includes the features of I that are confirmed by V. Inactivated nodes corresponding to unconfirmed features of X are unhatched. The reduction in total STM activity which occurs when X is transformed into X* causes a decrease in the total inhibition from F_1 to A. (C) If inhibition decreases sufficiently, A releases

TABLE 1
The Instar Learning, or Gated Steepest Descent Learning Rule, Embodies both Hebbian (LTP) and anti-Hebbian (LTD) Properties within a Single Process[a]



	Case 1	Case 2	Case 3	Case 4
State of S_i	+	−	+	−
State of x_j	+	+	−	−
State of w_{ij}	↑	↓	↔	↔

Note. Symbols: + = active; − = inactive; ↑ = increase; ↓ = decrease; ↔ = no change.

[a] Reprinted with permission from Grossberg and Merrill (1996).

X* with Y. The resonance binds spatially distributed features into either a stable equilibrium or a synchronous oscillation, much like the synchronous feature binding in visual cortex that has recently attracted so much interest after the experiments of Reinhard Eckhorn and Wolf Singer and their colleagues (Eckhorn *et al.,* 1988; Gray & Singer, 1989); also see Grossberg and Grunewald (1997).

In ART, the resonant state, rather than bottom-up activation, is predicted to drive the learning process. The resonant state persists long enough, at a high enough activity level, to activate the slower learning processes in the LTM traces. This helps to explain how the LTM traces can regulate the brain's fast information processing without necessarily learning about the signals that they process. Through resonance as a mediating event, the combination of top-down matching and attentional focusing helps to stabilize ART learning and memory in response to an arbitrary input environment. The stabilizing properties of top-down matching may be one reason for the ubiquitous occurrence of reciprocal bottom-up and top-down corticocortical and corticothalamic interactions in the brain.

HOW IS THE GENERALITY OF KNOWLEDGE CONTROLLED?

A key problem about consciousness concerns what combinations of features or other information are bound together into object or event representations. ART provides a new answer to this question that overcomes problems faced by earlier models. In particular, ART systems learn prototypes, rather than exemplars, because the at-

---

a nonspecific arousal wave to F_2, which resets the STM pattern Y at F_2. (D) After Y is inhibited, its top-down prototype signal is eliminated, and X can be reinstated at F_1. Enduring traces of the prior reset lead X to activate a different STM pattern Y* at F_2. If the top-down prototype due to Y* also mismatches I at F_1, then the search for an appropriate F_2 code continues until a more appropriate F_2 representation is selected. Then an attentive resonance develops and learning of the attended data is initiated. [Reprinted with permission from Grossberg and Merrill (1996).]

tended feature vector X*, rather than the input exemplar itself, is learned. Both the bottom-up LTM traces that tune the category nodes and the top-down LTM traces that filter the learned expectation learn to correlate activation of F_2 nodes with the set of all *attended* X* vectors that they have ever experienced. These attended STM vectors assign less STM activity to features in the input vector I that mismatch the learned top-down prototype V than to features that match V.

Given that ART systems learn prototypes, how can they also learn to recognize unique experiences, such as a particular view of a friend's face? The prototypes learned by ART systems accomplish this by realizing a qualitatively different concept of prototype than that offered by previous models. In particular, Gail Carpenter and I have shown with our students how ART prototypes form in a way that is designed to conjointly maximize category generalization while minimizing predictive error (Bradski & Grossberg, 1995; Carpenter & Grossberg, 1987a,b; Carpenter, Grossberg, & Reynolds, 1991; Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992). As a result, ART prototypes can automatically learn individual exemplars when environmental conditions require highly selective discriminations to be made. How the matching process achieves this is discussed below.

Before describing how this is achieved, let us note what happens if the mismatch between bottom-up and top-down information is too great for a resonance to develop. Then the F_2 category is quickly reset and a memory search for a better category is initiated. This combination of top-down matching, attention focusing, and memory search is what stabilizes ART learning and memory in an arbitrary input environment. The attentional focusing by top-down matching prevents inputs that represent irrelevant features at F_1 from eroding the memory of previously learned LTM prototypes. In addition, the memory search resets F_2 categories so quickly when their prototype V mismatches the input vector I that the more slowly varying LTM traces do not have an opportunity to correlate the attended F_1 activity vector X* with them. Conversely, the resonant event, when it does occur, maintains and amplifies the matched STM activities for long enough and at high enough amplitudes for learning to occur in the LTM traces.

Whether a resonance occurs depends upon the level of mismatch, or novelty, that the network is prepared to tolerate. Novelty is measured by how well a given exemplar matches the prototype that its presentation evokes. The criterion of an acceptable match is defined by an internally controlled parameter that Carpenter and I have called vigilance (Carpenter & Grossberg, 1987a). The vigilance parameter is computed in the orienting subsystem A; see Fig. 12. Vigilance weighs how similar an input exemplar I must be to a top-down prototype V in order for resonance to occur. Resonance occurs if $\rho|I| - |X^*| \leq 0$. This inequality says that the F_1 attentional focus X* inhibits A more than the input I excites it. If A remains quiet, then an $F_1 \leftrightarrow F_2$ resonance can develop.

Either a larger value of ρ or a smaller match ratio $|X^*|I|^{-1}$ makes it harder to satisfy the resonance inequality. When ρ grows so large or $|X^*|I|^{-1}$ is so small that $\rho|I| - |X^*| > 0$, then A generates an arousal burst, or novelty wave, that resets the STM pattern across F_2 and initiates a bout of hypothesis testing, or memory search. During search, the orienting subsystem interacts with the attentional subsystem (Figs. 13C and 13D) to rapidly reset mismatched categories and to select better F_2 representations with which to categorize novel events at F_1, without risking unselective forgetting of previous knowledge. Search may select a familiar category if its prototype is similar enough to the input to satisfy the resonance criterion. The prototype may then be refined by attentional focusing. If the input is too different from any previously learned prototype, then an uncommitted population of F_2 cells is selected and learning of a new category is initiated.

Because vigilance can vary across learning trials, recognition categories capable of encoding widely differing degrees of generalization or abstraction can be learned by a single ART system. Low vigilance leads to broad generalization and abstract prototypes. High vigilance leads to narrow generalization and to prototypes that represent fewer input exemplars, even a single exemplar. Thus a single ART system may be used, say, to learn abstract prototypes with which to recognize abstract categories of faces and dogs, as well as ‘‘exemplar prototypes’’ with which to recognize individual faces and dogs. A single system can learn both, as the need arises, by increasing vigilance just enough to activate A if a previous categorization leads to a predictive error. Thus the contents of a conscious percept can be modified by environmentally sensitive vigilance control.

Vigilance control hereby allows ART to overcome some fundamental difficulties that have been faced by classical exemplar and prototype theories of learning and recognition. Classical exemplar models face a serious combinatorial explosion, since they need to suppose that all experienced exemplars are somehow stored in memory and searched during performance. Classical prototype theories face the problem that they find it hard to explain how individual exemplars are learned, such as a particular view of a familiar face. Vigilance control enables ART to achieve the best of both types of model, by selecting the most general category that is consistent with environmental feedback. If that category is an exemplar, then a ‘‘very vigilant’’ ART model can learn it. If the category is at an intermediate level of generalization, then the ART model can learn it by having the vigilance value track the level of match between the current exemplar and the prototype that it activates. In every instance, the model tries to learn the most general category that is consistent with the data. This tendency can, for example, lead to the type of overgeneralization that is seen in young children until further learning leads to category refinement (Chapman *et al.*, 1986; Clark, 1973; Smith *et al.*, 1985; Smith & Kemler, 1978; Ward, 1983). Many benchmark studies of how ART uses vigilance control to classify complex data bases have shown that the number of ART categories that is learned scales well with the complexity of the input data; see Carpenter and Grossberg (1994) for a list of illustrative benchmark studies.

CORTICOHIPPOCAMPAL INTERACTIONS AND MEDIAL TEMPORAL AMNESIA

As sequences of inputs are practiced over learning trials, the search process eventually converges upon stable categories. Carpenter and I mathematically proved (Carpenter and Grossberg, 1987a) that familiar inputs directly access the category whose prototype provides the globally best match, while unfamiliar inputs engage the orienting subsystem to trigger memory searches for better categories until they become familiar. This process continues until the memory capacity, which can be chosen arbitrarily large, is fully utilized. The process whereby search is automatically disengaged is a form of memory consolidation that emerges from network interactions. Emergent consolidation does not preclude structural consolidation at individual cells, since the amplified and prolonged activities that subserve a resonance may be a trigger for learning-dependent cellular processes, such as protein synthesis and transmitter production. It has also been shown that the adaptive weights which are learned by an ART model at any stage of learning can be translated into IF-THEN rules (e.g., Carpenter *et al.*, 1992). Thus the ART model is a self-organizing rule-discovering production system as well as a neural network.

The attentional subsystem of ART has been used to model aspects of inferotemporal (IT) cortex, and the orienting subsystem models part of the hippocampal system. The interpretation of ART dynamics in terms of IT cortex led Miller, Li, and Desimone (1991) to successfully test the prediction that cells in monkey IT cortex are reset after each trial in a working memory task. To illustrate the implications of an ART interpretation of IT–hippocampal interactions, I review how a lesion of the ART model's orienting subsystem creates a formal memory disorder with symptoms much like the medial temporal amnesia that is caused in animals and human patients after hippocampal system lesions (Carpenter & Grossberg, 1993; Grossberg & Merrill, 1996). In particular, such a lesion *in vivo* causes unlimited anterograde amnesia; limited retrograde amnesia; failure of consolidation; tendency to learn the first event in a series; abnormal reactions to novelty, including perseverative reactions; normal priming; and normal information processing of familiar events (Cohen, 1984; Graf, Squire, & Mandler, 1984; Lynch, McGaugh, & Weinberger, 1984; Squire & Butters, 1984; Squire & Cohen, 1984; Warrington & Weiskrantz, 1974; Zola-Morgan & Squire, 1990).

Unlimited anterograde amnesia occurs because the network cannot carry out the memory search to learn a new recognition code. Limited retrograde amnesia occurs because familiar events can directly access correct recognition codes. Before events become familiar, memory consolidation occurs which utilizes the orienting subsystem (Figure 13C). This failure of consolidation does not necessarily prevent learning per se. Instead, learning influences the first recognition category activated by bottom-up processing, much as amnesics are particularly strongly wedded to the first response they learn. Perseverative reactions can occur because the orienting subsystem cannot reset sensory representations or top-down expectations that may be persistently mismatched by bottom-up cues. The inability to search memory prevents ART from discovering more appropriate stimulus combinations to attend. Normal priming occurs because it is mediated by the attentional subsystem.

Similar behavioral problems have been identified in hippocampectomized monkeys. Gaffan (1985) noted that fornix transection ‘‘impairs ability to change an established habit . . . in a different set of circumstances that is similar to the first and therefore liable to be confused with it.’’ In ART, a defective orienting subsystem prevents the memory search whereby different representations could be learned for similar events. Pribram (1986) called such a process a ‘‘competence for recombinant context-sensitive processing.’’ These ART mechanisms illustrate how, as Zola-Morgan and Squire (1990) have reported, memory consolidation and novelty detection may be mediated by the same neural structures. Why hippocampectomized rats have difficulty orienting to novel cues and why there is a progressive reduction in novelty-related hippocampal potentials as learning proceeds in normal rats is also clarified (Deadwyler, West, & Lunch, 1979; Deadwyler, West, & Robinson, 1981). In ART, the orienting system is automatically disengaged as events become familiar during the memory consolidation process. The ART model of normal and abnormal recognition learning and memory is compared with several other recent models of these phenomena in Grossberg and Merrill (1996).

At this point, it might also be useful to note that the processes of automatic and task-selective attention may not be independent *in vivo.* This is because higher-order attentional constraints that may be under task-selective control can in principle propagate downward through successive cortical levels via layer-6-to-layer-6 linkages. For example, recent modeling work has suggested how prestriate cortical areas may separate visual objects from one another and from their backgrounds during the process of figure–ground separation (Grossberg, 1994, 1997; Grossberg & McLoughlin, 1997). Such constraints may propagate top-down toward earlier cortical levels, possibly even area V1, to modulate the cells that get active there to be consistent with these figure–ground constraints. Still higher cortical processes, such as those involved in learned categorization, may also propagate their modulatory constraints to lower levels. How the strength of such top-down modulatory influences depends upon the source cortical area and on the number of synaptic steps to the target cortical area is a topic that has yet to be systematically studied.

HOW UNIVERSAL ARE ART PROCESSES IN THE BRAIN?

In all the examples discussed above—from early vision, visual object recognition, auditory streaming, and speech recognition—ART matching and resonance have played a central role in models that help to explain how the brain stabilizes its learned adaptations in response to changing environmental conditions. This type of matching can be achieved using a top-down nonspecific inhibitory gain control that downregulates all target cells except those that also receive top-down specific excitatory signals, as in Fig. 8. Are there yet other brain processes that utilize these mechanisms?

John Reynolds and colleagues in Bob Desimone's laboratory (Reynolds, Nicholas, Chelazzi, & Desimone, 1995) have reported neurophysiological data from cells in cortical areas V2 and V4 that are consistent with the ART attentional mechanism summarized in Fig. 8. Taken together with studies of the V1 → LGN attention circuit and of attentional control by frontal and inferotemporal cortex during visual object recognition, it may be concluded that ART-like top-down matching occurs throughout the brain's visual system.

With my colleagues Mario Aguilar, Dan Bullock, and Karen Roberts, a neural model has been developed to explain how the superior colliculus learns to use visual, auditory, somatosensory, and planned movement signals to control saccadic eye movements (Grossberg, Roberts, Aguilar, & Bullock, 1997). This model uses ART matching and resonance to help explain behavioral and neural data about multimodal eye movement control. The model clarifies how visual, auditory, and planned movement signals use learning to form a mutually consistent movement map and how attention gets focused on a movement target location after all these signals compete to determine where the eyes will move.

Recent experiments from Marcus Raichle’s lab at Washington University using positron emission tomography (PET) support the idea that ART top-down priming also occurs in human somatosensory cortex (Drevets, Burton, & Raichle, 1995). In their experiments, attending to an impending stimulus to the fingers caused inhibition of nearly cortical cells that code for the face, but not cells that code the fingers. Likewise, priming of the toes produced inhibition of nearby cells that code for the fingers and face, but not cells that code for the toes.

ART models have also been used to explain a great deal of data about cognitive–emotional interactions, notably about classical and instrumental conditioning (Grossberg, 1987b) and about human decision making under risk (Grossberg and Gutowski, 1987). In these examples, the resonances are between cognitive and emotional circuits and help to focus attention upon, and release actions toward, valued events in the world.

Thus all levels of vision, visual object recognition, auditory streaming, speech recognition, attentive selection of eye movement targets, somatosensory representation, and cognitive–emotional interactions may all incorporate variants of the circuit depicted in Fig. 8. These results suggest that a type of ‘‘automatic’’ attention operates even at early levels of brain processing, such as the lateral geniculate, but that higher processing levels benefit from an orienting subsystem that can be used to flexibly reset attention and to facilitate voluntary control of top-down expectations.

INTERNAL FANTASY, PLANNED MOVEMENT, AND VOLITIONAL GATING

Given this type of circuit, how could top-down priming be released from inhibition to enable us to voluntarily experience internal thinking and fantasies? This can be achieved through an ‘‘act of will’’ that activates inhibitory cells which inhibit the nonspecific inhibitory interneurons in the top-down on-center off-surround network of Fig. 8. This operation disinhibits the cells receiving the excitatory top-down signals in the on-center of the network. These cells are then free to generate supraliminal resonances. Such self-initiated resonances can, for example, be initiated by the readout of top-down expectations from higher-order planning nodes into temporally organized working memories, say in the prefrontal cortex (Fuster, 1996). It is, for example, well known that the basal ganglia can use such a disinhibitory action to gate the release of individual movements, sequences of movements, and even cognitive processes (Hikosaka, 1994; Middleton & Strick, 1994; Sakai *et al.,* 1998).

These examples also help to understand how top-down expectations can be used for the control of planned (viz., intentional) behavioral sequences. For example, once such planning nodes read-out their top-down expectations into working memory, the contents of working memory can be read-out and modified by on-line changes in ‘‘acts of will.’’ These volitional signals enable invariant representations of an intentional behavior to rapidly adapt themselves to changing environmental conditions. For example, Bullock, Grossberg, and Mannes (1993) have modeled how such a working memory can control the intentional performance of handwriting whose size and speed can be modified by acts of will, without a change of handwritten form.

Bullock, Grossberg, and Guenther (1994) have shown how a visual target that is stored in working memory can be reached with a novel tool that has never been used before. The latter study shows how a such a model can learn its own parameters through a type of Piagetian perform-and-test developmental cycle.

Thus we arrive at an emerging picture of how the adaptive brain works wherein the core issue of how a brain can learn quickly and stably about a changing world throughout life leads toward a mechanistic understanding of attention, intention, thinking, fantasy, and consciousness. The mediating events are adaptive resonances that effect a dynamic balance between the complementary demands of stability and plasticity and of expectation and novelty and which are a necessary condition for consciousness.

WHAT VS WHERE: WHY ARE PROCEDURAL MEMORIES UNCONSCIOUS?

Although the type of ART matching, learning, and resonance that have been reviewed above seem to occur in many sensory and cognitive processes, they are not the only types of matching and learning to occur in the brain. In fact, there seems to be a major difference between the types of learning that occur in sensory and cognitive processes versus those that occur in spatial and motor processes. In particular, sensory and cognitive processes are carried out in the What processing stream that passes through the inferotemporal cortex, whereas spatial and motor processes are carried out in the Where processing stream that passes through the parietal cortex. What processing includes object recognition and event prediction. Where processing includes spatial navigation and motor control. I suggest that the types of matching and learning that go on in the What and Where streams are different, indeed complementary, and that this difference is appropriate to their different roles. First, consider how we use a sensory expectation. Suppose, for example, that I ask you to ‘‘Look for the yellow ball, and if you find it within three hundred milliseconds, I will give you a million dollars.’’ If you believed me, you could activate a sensory expectation of ‘‘yellow balls’’ that would make you much more sensitive to yellow and round objects in your environment. As in ART matching, once you detected a yellow ball, you could then react to it much more quickly and with a much more energetic response than if you were not looking for it. In other words, sensory and cognitive expectations lead to a type of excitatory matching.

Now consider how we use a motor expectation. Such an expectation represents where we want to move (Bullock & Grossberg, 1988). For example, it could represent a desired position for the hand to pick up an object. Such a motor expectation is matched against where the hand is now. After the hand actually moves to the desired position, no further movement is required to satisfy the motor expectation. In this sense, motor expectations lead to a type of inhibitory matching. In summary, although the sensory and cognitive matching process is excitatory, the spatial and motor matching process is inhibitory. These are complementary properties. Models such as ART quantify how excitatory matching is accomplished. A different type of model, called a Vector Associative Map, or VAM, model, suggests how inhibitory matching is accomplished (Gaudiano & Grossberg, 1991; Grossberg, Guenther, Bullock, & Greve, 1993; Guenther, Bullock, Greve, & Grossberg, 1994).

As shown in the discussions of ART above, learning within the sensory and cognitive domain is often a type of *match* learning. It takes place only if there is a good enough match of top-down expectations with bottom-up data to risk altering previously stored knowledge within the system, or it can trigger learning of a new representation if a good enough match is not available. In contrast, learning within spatial and motor processes, such as VAM processes, is *mismatch* learning that is used to either learn new sensory-motor maps (e.g., Grossberg, Guenther, Bullock, & Greve, 1993) or to adjust the gains of sensory-motor commands (e.g., Fiala, Grossberg, & Bullock, 1996). These types of learning are also complementary.

Why are the types of learning that go into spatial and motor processes complementary to those that are used for sensory and cognitive processing? My answer is that ART-like learning allows the brain to solve the stability–plasticity dilemma. It enables us to continue learning more about the world in a stable fashion throughout life without forcing catastrophic forgetting of our previous memories. On the other hand, catastrophic forgetting is a good property when it takes place during spatial and motor learning. We have no need to remember all the spatial and motor maps that we used when we were infants or children. In fact, those maps would cause us a lot of trouble if they were used to control our adult limbs. We want our spatial and motor processes to continuously adapt to changes in our motor apparatus. These complementary types of learning allow our sensory and cognitive systems to stably learn about the world and to thereby be able to effectively control spatial and motor processes that continually update themselves to deal with changing conditions in our limbs.

Why, then, are procedural memories unconscious? The difference between cognitive memories and procedural, or motor, memories has gone by a number of different names, including the distinction between declarative memory and procedural memory, knowing that and knowing how, memory and habit, or memory with record and memory without record (Bruner, 1969; Miskin, 1982, 1993; Ryle, 1949; Squire & Cohen, 1984). The amnesic patient HM dramatically illustrated this distinction by learning and remembering motor skills like assembly of the Tower of Hanoi without being able to recall ever having done so (Bruner, 1969; Scoville & Milner, 1957; Squire & Cohen, 1984). We can now give a very short answer to the question of why procedural memories are unconscious: The matching that takes place during spatial and motor processing is often inhibitory matching. Such a matching process cannot support an excitatory resonance. Hence, it cannot support consciousness.

In this regard, Goodale and Milner (1992) have described a patient whose brain lesion has prevented accurate visual discrimination of object orientation, yet whose visually guided reaching behaviors toward objects are oriented and sized correctly. We have shown, in a series of articles, how head-centered and body-centered representations of an object’s spatial location and orientation may be learned and used to control reaches of the hand–arm system that can continuously adapt to changes in the sensory and motor apparatus that is used to plan and execute reaching behaviors (Bullock, Grossberg, & Guenther, 1993; Carpenter, Grossberg, & Lesher, 1998; Gaudiano & Grossberg, 1991; Grossberg, Guenther, Bullock, & Greve, 1993; Guenther, Bullock, Greve, & Grossberg, 1994). None of these model circuits has resonant loops; hence, they do not support consciousness.

When these models are combined into a more comprehensive system architecture for intelligent behavior, the sensory and cognitive match-based networks in the What processing stream through the inferotemporal cortex provide self-stabilizing representations with which to continually learn more about the world without undergoing catastrophic forgetting, while the Where/How processing stream's spatial and motor mismatch-based maps and gains can continually forget their old parameters in order to instate the new parameters that are needed to control our bodies in their present form. This larger architecture illustrates how circuits in the self-stabilizing match-based sensory and cognitive parts of the brain can resonate into consciousness, even while they are helping to direct the contextually appropriate activation of spatial and motor circuits that cannot.

SOME COMMENTS ABOUT AMODAL AND MODAL VISUAL PERCEPTS

There are many other aspects of perception and cognition, notably of vision and visual object recognition, which can be discussed in the light of recent modeling advances to shed light on consciousness. Here I make some summarizing remarks whose detailed analysis and justification can be found in the original articles. One issue of interest concerns the distinction between *modal* and *amodal* percepts. An amodal percept, such as the percept of a vertical boundary between the offset grating in Fig. 1C, is one which does not carry a visible perceptual sign. As noted above, it can be recognized without being seen; we are conscious of it even though it is perceptually invisible. A modal percept, such as a percept of brightness or color, does carry a visible perceptual sign. I believe that all theories of consciousness need to deal with how amodal percepts can occur because such percepts sharply distinguish between our consciously ‘‘knowing’’ that an event has occurred even though we do not consciously ‘‘perceive’’ it.

The FACADE theory of biological vision has provided an extensive analysis of some of the conditions that determine whether a percept will be modal or amodal (e.g., Francis, Grossberg, & Mingolla, 1994; Grossberg, 1994, 1997; Grossberg & McLoughlin, 1997; Grossberg & Mingolla, 1985b; Gove, Grossberg, & Mingolla, 1995). A key contribution of this theory is to suggest how visual scenes are processed in parallel by cortical boundary and surface systems, which are proposed to be realized by the interblob and blob processing streams from the LGN to cortical area V4. Boundaries include illusory contours (Fig. 1), as well as the boundaries that are formed in response to texture, shading, and stereo cues.

A key insight of this theory is that ‘‘all boundaries are invisible’’ (i.e., amodal) within the boundary processing stream, and that visibility is a property of the surface processing stream. Boundaries are invisible within the boundary processing stream because like-oriented signals from cortical simple cells that are sensitive to opposite contrast polarities are pooled at complex cells. Complex cells can hereby respond to contrasts that are either dark/light or light/dark, as can all subsequent stages of the boundary system. As a result of this pooling process, a boundary can be formed around an object whose relative contrasts with respect to its background may reverse along its perimeter. A secondary consequence is that a perceptual boundary, by pooling across opposite contrast polarities (as well as all opponent colors), cannot represent any visible property that depends upon knowing the direction of a brightness or color change.

Modal percepts are predicted to occur within the surface processing stream. Surface representations arise through interactions with the boundaries. First, the surface stream ‘‘discounts the illuminant,’’ or compensates for variable illumination (Helmholtz, 1962; Land, 1977, 1986). This discounting process eliminates brightness or color signals within homogeneously bright or colored regions of a scene, which could otherwise cause serious confusions between variable lighting conditions and the surface properties of objects in the world. At subsequent processing stages, the boundaries interact with the discounted surface signals. Here, the boundaries suppress surface signals that are not spatially coincident with them. Boundaries select surface signals that are spatially coincident with them and initiate a process of filling-in whereby these selected signals can diffuse within the controlling boundaries.

FACADE theory predicts that the boundaries which exercise this control occur subsequent to the cortical processing stage at which visual inputs from both eyes are binocularly fused. It was suggested in Grossberg (1987a) that, although the binocular matching process is initiated in cortical area V1, the stage at which the binocular boundaries are completed occurs no earlier than cortical area V2.

During binocular rivalry, the inputs to the two eyes are mismatched in such a way that image regions from only one eye at a time can be consciously perceived. FACADE theory suggests how boundary signals from the two eyes compete in a cyclical fashion through time, with the boundaries from one eye winning at any time in each position. Such competition has been traced to the mechanisms whereby a winning boundary is selected from among many possible boundary groupings, even when the inputs to both eyes represent the same scene. The cyclicity of the percept was traced to the habituative mechanisms whereby boundaries are rapidly reset in response to rapidly changing imagery in order to prevent them from persisting too long [see Francis, Grossberg, & Mingolla (1994) for an analysis of how long perceptual boundaries do persist]. Then the winning boundaries select those surface signals from the dominant eye which are spatially coincident with them while suppressing the spatially discordant surface signals from the losing eye. The first stage of such surface capture selects the surface properties from each eye separately. The selected surface representations are predicted to be amodal. These selected surface properties are then binocularly matched at a subsequent processing stage at which the modal, or visible, surface representation is predicted to form. This is also the processing stage at which visual figures are fully separated from one another and from their backgrounds.

Grossberg (1987a) predicted that this binocular modal, or visible, representation of the winning surface percept arises in cortical area V4, which resonates with inferotemporal cortex during consciousness. Logothetis *et al.* (1996) have reported consistent data on binocular rivalry from awake behaving monkeys. Schiller (1994) has reported data from awake behaving monkeys that is consistent with the prediction that figure–ground separation is completed in cortical area V4.

These results support the FACADE theory prediction that *amodal* percepts may form in cortical areas V2 or before and that *modal* representations of surfaces may first occur in cortical area V4. In further support of this hypothesis, Grossberg (1994) explained many data about 3D figure–ground separation in which, say, amodal representations of occluded object parts may be formed in cortical area V2 and used to recognize these occluded objects, even though they are not seen. Modal representations of both occluding objects and the unoccluded parts of the objects that they occlude may not be formed until cortical area V4. This is proposed despite the fact that all of these cortical processing stages may incorporate the ART matching rule within their laminar circuits and may resonate using both the intercortical and intracortical feedback pathways that activate the layer-6-to-layer-4 on-center off-surround networks, the former during attentional priming and the latter during the selection of winning perceptual groupings.

Grossberg (1997) proposed that the modally conscious surface representations in V4 may be used to recognize and to control reaching toward physically accessible objects, especially in infants, whereas the amodally conscious representations—both of boundaries and of surfaces—in V2 may be used to recognize partially occluded objects and to reach toward them via more indirect motor planning and control circuits. This proposal provides a functional reason for making some visual representations visible and others not visible; in particular, being able to distinguish between modal (e.g., occluding) and amodal (e.g., occluded) representations helps to prevent efforts to reach through an occluding object to the object that it is occluding. On the other hand, the proposal does not explain how the property of visibility is achieved by one type of representation but not the other, particularly since both types of representation may be assumed to be resonant. This fact does not contradict the hypothesis that all conscious states are resonant states. It does show, however, that further mechanisms are needed to explain why some of these resonant representations are modal whereas others are merely amodal.

The need for further mechanisms is well-illustrated by the following modeling prediction. It was predicted in Grossberg (1987a), and then used extensively to explain much more perceptual data in Grossberg (1994, 1997), that a network of double-opponent cells forms an important mechanism in the process whereby boundaries select only those surface brightness and color signals that are spatially coincident with them. Double-opponent cells are often cited as a key mechanism of color perception (e.g., Livingstone & Hubel, 1984). FACADE theory suggests that such networks are used to form both amodal and modal surface representations. In the amodal surface representations, double-opponent networks are predicted not to generate a percept of visible color. Some other factor must be sought, to whose discovery future research would be profitably directed.

REFERENCES

Abbott, L. F., Varela, K., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science,* **275,** 220–223.

Artola, A., & Singer, W. (1993). Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends in Neurosciences,* **16,** 480–487.

Bradski, G., & Grossberg, G. (1995). Fast learning VIEWNET architectures for recognizing 3-D objects from multiple 2-D views. *Neural Networks,* **8,** 1053–1080.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bruner, J. S. (1969). In G. A. Talland & N. C. Waugh (Eds.), *The pathology of memory.* New York: Academic Press.

Bullock, D., & Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review,* **95,** 49–90.

Bullock, D., Grossberg, S., & Guenther, F. H. (1993). A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. *Journal of Cognitive Neuroscience,* **5,** 408–435.

Bullock, D., Grossberg, S., & Mannes, C. (1993). A neural network model for cursive script production. *Biological Cybernetics,* **70,** 15–28.

Carpenter, G. A., & Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology,* **1,** 1–42. Reprinted in S. Grossberg (Ed.), *The adaptive brain* (1987, Vol. II, pp. 271–310). Amsterdam: Elsevier.

Carpenter, G. A., & Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing,* **37,** 54–115.

Carpenter, G. A., & Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics,* **26,** 4919–4930.

Carpenter, G. A., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences,* **116,** 131–137.

Carpenter, G. A., & Grossberg, S. (1994). Integrating symbolic and neural processing in a self-organizing architecture for pattern recognition and prediction. In V. Honavar & L. Uhr (Eds.), *Artificial intelligence and neural networks: Steps toward principled prediction* (pp. 387–421). San Diego: Academic Press.

Carpenter, G. A., Grossberg, S., & Lesher, G. W. (1998). The What-and-Where filter: A spatial mapping neural network for object recognition and image understanding. *Computer Vision and Image Understanding,* **69,** 1–22.

Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks,* **3,** 698–713.

Carpenter, G. A., Grossberg, S., & Reynolds, J. H. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks,* **4,** 565–588.

Chapman, K. L., Leonard, L. B., & Mervis, C. G. (1986). The effect of feedback on young children's inappropriate word usage. *Journal of Child Language,* **13,** 101–107.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America,* **14,** 2570–2594.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Morre (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics,* **36,** 428–456.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. Squire & N. Butters (Eds.), *The neuropsychology of memory* (pp. 83–103). New York: Guilford Press.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of a pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science,* **210,** 207–210.

Deadwyler, S. A., West, M. O., & Lynch, G. (1979). Activity of dentate granule cells during learning: Differentiation of perforant path inputs. *Brain Research,* **169,** 29–43.

Deadwyler, S. A., West, M. O., & Robinson, J. H. (1981). Entorhinal and septal inputs differentially control sensory-evoked responses in the rat dentate gyrus. *Science,* **211,** 1181–1183.

Drevets, W. C., Burton, H., & Raichle, M. E. (1995). Blood flow changes in human somatosensory cortex during anticipated stimulation. *Nature,* **373,** 249.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988).

Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics,* **60,** 121–130.

Fiala, J. C., Grossberg, S., & Bullock, D. (1996). Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. *Journal of Neuroscience,* **16,** 3760–3774.

Francis, G., & Grossberg, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception,* **35,** 543–567.

Francis, G., & Grossberg, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research,* **36,** 149–173.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research,* **34,** 1089–1104.

Fuster, J. M. (1996). Frontal lobe and the cognitive foundation of behavioral action. In A. R. Damasio, H. Damasio, & Y. Christen (Eds.), *The neurobiology of decision-making* (pp. 47–61). New York: Springer-Verlag.

Gaffan, D. (1985). Hippocampus: Memory, habit, and voluntary movement. *Philosophical Transactions of the Royal Society of London B,* **308,** 87–99.

Gaudiano, P., & Grossberg, S. (1991). Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks,* **4,** 147–183.

Goodale, M. A., & Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences,* **15,** 20–25.

Gove, A., S. Grossberg, & E. Mingolla. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience,* **12,** 1027–1052.

Govindarajan, K. K., Grossberg, S., Wyse, L. L., & Cohen, M. A. (1995). A neural network model of auditory scene analysis and source segregation. Technical Report: CAS/CNS-TR-94-039. Boston, MA: Boston University.

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **10,** 164–178, 183.

Gray, C. M., & Singer, W. 1989. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences,* **86,** 1698–1702.

Grossberg, S. (1964). *The theory of embedding fields with applications to psychology and neurophysiology.* New York: Rockefeller University.

Grossberg, S. (1969). On learning and energy–entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics,* **1,** 319–350.

Grossberg, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik,* **10,** 49–57.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding. I. Parallel development and coding of neural feature detectors. *Biological Cybernetics,* **23,** 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding. II. Feedback, expectation, olfaction, and illusions. *Biological Cybernetics,* **23,** 187–202.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5). New York: Academic Press.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review,* **1,** 1–51.

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control.* Norwell, MA: Kluwer.

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception. II. Binocular theory. *Perception and Psychophysics,* **41,** 117–158.

Grossberg, S. (1987b). *The adaptive brain* (Vol. I). Amsterdam: North-Holland.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks,* **1,** 17–61.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics,* **55,** 48–120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional figures. *Psychological Review,* **104,** 618–658.

Grossberg, S. (1998a). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. Technical report CAS/CNS TR-97-023. *Spatial Vision,* in press.

Grossberg, S. (1998b). Pitch based streaming in auditory perception. In N. Griffith & P. Todd (Eds.), *Musical networks: Parallel distributed perception and performance.* Cambridge, MA: MIT Press.

Grossberg, S., Boardman, I., & Cohen, M. A. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance,* **23,** 481–503.

Grossberg, S., & Grunewald, A. (1997). Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience,* **9,** 117–132.

Grossberg, S., Guenther, F., Bullock, D., & Greve, D. (1993). Neural representations for sensory-motor control, II: Learning a head-centered visuomotor representation of 3-D target position. *Neural Networks,* **6,** 43–67.

Grossberg, S., & Gutowski, W. (1987). Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review,* **94,** 300–318.

Grossberg, S., & Kuperstein, M. (1986). *Neural dynamics of adaptive sensory–motor control: Ballistic eye movements.* Amsterdam: North-Holland.

Grossberg, S., & McLoughlin, N. P. (1997). Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks,* **10,** 1583–1605.

Grossberg, S., & Merrill, J. W. L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience,* **3,** 257–277.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review,* **92,** 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics,* **38,** 141–171.

Grossberg, S., Roberts, K., Aguilar, M., & Bullock, D. (1997). A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience,* **17,** 9706–9725.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, **43,** 241–277.

Grossberg, S., & Williamson, J. R. (1997). Linking cortical development to visual perception. *Society for Neuroscience Abstract,* **23**(1), 568.

Grossberg, S., & Williamson, J. R. (1998). A neural model of how visual cortex develops a laminar architecture capable of adult perceptual grouping. Technical Report CAS/CNS TR-98-022. Boston, MA: Boston University.

Grunewald, A., & Grossberg, S. (1998). Self-Organization of binocular disparity tuning by reciprocal corticogeniculate interactions. *Journal of Cognitive Neuroscience,* **10,** 199–215.

Guenther, F. H., Bullock, D., Greve, D., & Grossberg, S. (1994). Neural representations for sensory-motor control. III. Learning a body-centered representation of 3-D target position. *Journal of Cognitive Neuroscience,* **6,** 341–358.

Helmholtz, H. L. F. von (1962). *Treatise on physiological optics* (J. P. C. Southall, translator). New York: Dover.

Hikosaka, O. (1994). Role of basal ganglia in control of innate movements, learned behavior and cognition—a hypothesis. In G. Percheron, J. S. McKenzie, & J. Feger (Eds.). The basal ganglia (Vol. IV, pp. 589–595). New York: Plenum.

Hubel, D. H., & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B,* **198,** 1–59.

Kennedy, J. M. (1979). Subjective contours, contrast, and assimilation. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation.* New York: Praeger.

Kennedy, J. M. (1988). Line endings and subjective controls. *Spatial Vision,* **3,** 151–158.

Kohonen, T. (1984). *Self-organization and associative memory.* New York: Springer–Verlag.

Land, E. (1977). The retinex theory of color vision. *Scientific American,* **237,** 108–128.

Land, E. (1986). Recent advances in retinex theory. *Vision Research,* **26,** 7–21.

Levy, W. B. (1985). Associative changes at the synapse: LTP in the hippocampus. In W. B. Levy, J. Anderson, & S. Lehmkuhle (Eds.), *Synaptic modification, neuron selectivity, and nervous system organization* (pp. 5–33). Hillsdale, NJ: Erlbaum.

Levy, W. B., & Desmond, N. L. (1985). The rules of elemental synaptic plasticity. In W. B. Levy, J. Anderson, & S. Lehmkuhle (Eds.), *Synaptic modification, neuron selectivity, and nervous system organization* (pp. 105–121). Hillsdale, NJ: Erlbaum.

Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience,* **4,** 309–356.

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature,* **360,** 621–624.

Lynch, G., McGaugh, J. L., & Weinberger, N. M. (Eds.) (1984). *Neurobiology of learning and memory.* New York: Guilford.

Michotte, A., Thines, G., & Crabbe, G. (1964). Les complements amodaux des structures perceptives. Louvain: Publications Universitaires de Louvain.

Middleton, F. A., & Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science,* **166,** 458–461.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science,* **254,** 1377–1379.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions Royal Society of London B,* **298,** 85–95.

Mishkin, M. (1993). Cerebral memory circuits. In T. A. Poggio & D. A. Glaser (Eds.), *Exploring brain functions: Models in neuroscience* (pp. 113–125). New York: Wiley.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences,* **6,** 414–417.

Murphy, P. C., & Sillito, A. M. (1987). Corticofugal feedback influences the generation of length tuning in the visual pathway. *Nature,* **329,** 727–729.

Parker, D. B. (1982). Learning-logic (Invention Report 581–64, File 1). Office of Technology Licensing. October, Stanford University.

Pribram, K. H. (1986). The hippocampal system and recombinant processing. In R. L. Isaacson & K. H. Pribram (Eds.), The hippocampus (Vol. 4, pp. 329–370). New York: Plenum.

Rauschecker, J. P., & Singer, W. (1979). Changes in the circuitry of the kitten's visual cortex are gated by postsynaptic activity. *Nature,* **280,** 58–60.

Redies, C., Crook, J. M., & Creutzfeldt, O. D. (1986). Neuronal responses to borders with and without luminance gradients in cat visual cortex and dLGN. *Experimental Brain Research,* **61,** 469–481.

Repp, B. H. (1980). A range-frequency effect on perception of silence in speech. *Haskins Laboratories Status Report on Speech Research,* **61,** 151–165.

Reynolds, J., Nicholas, J., Chelazzi, L., & Desimone, R. (1995). Spatial attention protects macaque V2 and V4 cells from the influence of non-attended stimuli. *Society for Neuroscience Abstracts,* **21**(III), 1759.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhard & J. L. McClelland (Eds.), *Parallel distributed processing.* Cambridge, MA: MIT Press.

Ryle, G. (1949). *The concept of mind.* Hutchinson Press.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., and Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. *Journal of Neuroscience,* **18,** 1827–1840.

Samuel, A. G. (1981a). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance,* **7,** 1124–1131.

Samuel, A. G. (1981b). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General,* **110,** 474–494.

Schiller, P. H. (1994). Area V4 of the primate visual cortex. *Current Directions in Psychological Science,* **3,** 89–92.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesion. *Journal of Neurology, Neurosurgery, and Psychiatry,* **20,** 11–21.

Sillito, A. M., Jones, H. E., Gerstein, G. L., and West, D. C. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature,* **369,** 479–482.

Singer, W. (1983). Neuronal activity as a shaping factor in the self-organization of neuron assemblies. In E. Basar, H. Flohr, H. Haken, & A. J. Mandell (Eds.), *Synergetics of the brain.* New York: Springer-Verlag.

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concept of size, weight, and density. *Cognition,* **21,** 177–237.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology,* **10,** 502–532.

Squire, L. R., & Butters, N. (Eds.) (1984). *Neuropsychology of memory.* New York: Guilford.

Squire, L. R., & Cohen, N. J. (1984). Human memory and amnesia. In G. Lunch, J. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 3–64). New York: Guilford.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems: Separation of appearance and location of objects. In D. L. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik,* **14,** 85–100.

Ward, T. B. (1983). Response tempo and separable-integral responding: Evidence for an integral-to-separable processing sequencing in visual perception. *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 1029–1051.

Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin,* **96,** 371–383.

Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. *Perception and Psychophysics,* **16,** 150–156.

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychology,* **12,** 419–428.

Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. Unpublished doctoral thesis, Harvard University, Cambridge, MA.

Willshaw, D. J., & Malsburg, C. von der (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London B,* **194,** 431–445.

Zola-Morgan, S. M., & Squire, L. R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science* **250,** 288–290.

Received July 6, 1998