Consciousness CLEARS the Mind

Stephen Grossberg[1]

Department of Cognitive and Neural Systems
Center for Adaptive Systems
and
Center of Excellence for Learning in Education, Science, and Technology
Boston University
677 Beacon Street, Boston, MA 02215, USA

Running title: Consciousness CLEARS the Mind

Key Words: Consciousness, learning, expectation, attention, resonance, synchrony, gamma oscillations, reset, cortical layers, thalamus, LAMINART

August, 2007

Technical Report CAS/CNS-TR-07-013

All correspondence should be addressed to
Professor Stephen Grossberg
Department of Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
Phone: 617-353-7858
Fax: 617-353-7755
Email:steve@bu.edu

Abstract

A full understanding of consciousness requires that we identify the brain processes from which conscious experiences emerge. What are these processes, and what is their utility in supporting successful adaptive behaviors? Adaptive Resonance Theory (ART) predicted a functional link between processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony (CLEARS), including the prediction that "all conscious states are resonant states." This connection clarifies how brain dynamics enable a behaving individual to autonomously adapt in real time to a rapidly changing world. The present article reviews theoretical considerations that predicted these functional links, how they work, and some of the rapidly growing body of behavioral and brain data that have provided support for these predictions. The article also summarizes ART models that predict functional roles for identified cells in laminar thalamocortical circuits, including the six layered neocortical circuits and their interactions with specific primary and higher-order specific thalamic nuclei and nonspecific nuclei. These predictions include explanations of how slow perceptual learning can occur without conscious awareness, and why oscillation frequencies in the lower layers of neocortex are sometimes slower beta oscillations, rather than the higher-frequency gamma oscillations that occur more frequently in superficial cortical layers. ART traces these properties to the existence of intracortical feedback loops, and to reset mechanisms whereby thalamocortical mismatches use circuits such as the one from specific thalamic nuclei to nonspecific thalamic nuclei and then to layer 4 of neocortical areas via layers 1-to-5-to-6-to-4.

Introduction
Adaptive Resonance Theory (ART) proposes that there is an intimate link between an animal's conscious awareness and its ability to learn quickly about a changing world throughout life. In particular, ART points to a critical role for "resonant" states in driving fast learning; hence the name *adaptive resonance*. These resonant states are bound together, using internal top-down feedback, into coherent representations of the world. In particular, ART proposes how learned bottom-up categories and learned top-down expectations interact to create these coherent representations. Learned top-down expectations can be activated in a data-driven manner by bottom-up processes from the external world, or by intentional top-down processes when they prime the brain to anticipate events that may or may not occur. In this way, ART clarifies one sense, but not the only one, in which the brain carries out predictive computation.

When such a learned top-down expectation is activated, matching occurs of the top-down expectations against bottom-up data. If the bottom-up and top-down patterns are not too different, such a matching process can lead to the focusing of attention upon the expected clusters of information, which are called *critical feature patterns*, at the same time that mismatched signals are suppressed. A resonant state emerges through sustained feedback between the attended bottom-up signal pattern and the active top-down expectation as they reach a consensus between what is expected and what is there in the outside world.

ART predicts that all conscious states in the brain are resonant states, and that these resonant states can trigger rapid learning of sensory and cognitive representations, without causing catastrophic forgetting. This prediction clarifies why it is easier to quickly learn about information to which one pays attention. ART hereby proposes that one reason why advanced animals are intentional and attentional beings is to enable rapid learning about a changing world throughout life.

Psychophysical and neurobiological data in support of ART have been reported in experiments on vision, visual object recognition, auditory streaming, variable-rate speech perception, somatosensory perception, and cognitive-emotional interactions, among others. Some of these data are summarized below. Others are reviewed in Carpenter and Grossberg (1991), Grossberg (1999b, 2003a, 2003b, 2003c), and Raizada and Grossberg (2003). In particular, ART mechanisms seem to be operative at all levels of the visual system, and it has been proposed how these mechanisms are realized by laminar circuits of visual cortex as they interact with specific and nonspecific thalamic nuclei (Grossberg, 2003b; Grossberg and Versace, 2007; Raizada and Grossberg, 2003; Versace and Grossberg, 2005, 2006). These laminar models of neocortex have been called LAMINART models because the laminar anatomy of neocortex embodies the types of attentional circuits that were predicted by ART (Grossberg, 1999a). Most recently, it has been proposed how a variation of these laminar neocortical circuits in the prefrontal cortex can carry out short-term storage of event sequences in working memory, learning of categories that selectively respond to these stored sequences, and variable-speed performance of the stored sequences under volitional control (Grossberg and Pearson, 2006; Pearson and Grossberg, 2005, 2006),. These examples from vision and cognition show how both spatial and temporal processes can be carried out by variations of the same neocortical design, and point the way towards a general theory of laminar neocortex that can explain aspects of all higher-order intelligent behavior.

What vs. Where: Why Procedural Memory is Not Conscious
Although ART-style learning and matching processes seem to be found in many sensory and cognitive processes, another type of learning and matching is found in spatial and motor processes. Spatial and motor processing in the brain's Where processing stream (Goodale and Milner, 1992). obey learning and matching laws that are often *complementary* (Grossberg, 2000b) to those used for sensory and cognitive processing in the What processing stream of the brain (Mishkin, Ungerleider,

and Macko, 1983; Ungerleider and Mishkin, 1982). Whereas sensory and cognitive representations use attentive matching to maintain their stability as we learn more about the world, spatial and motor representations are able to forget learned maps and gains that are no longer appropriate as our bodies develop and grow from infanthood to adulthood.

These memory differences can be traced to complementary differences in the corresponding matching and learning processes. ART-like sensory and cognitive learning occurs in an approximate *match* state, and matching is *excitatory*, which enables it to realize a type of excitatory priming. Spatial and motor learning often embodies Vector Associative Map (VAM) circuits (Gaudiano and Grossberg, 1991; Guenther, Bullock, Greve, and Grossberg, 1994) that occur in a *mismatch* state, and matching is realized by an *inhibitory* process. These complementary differences clarify why procedural memories are unconscious; namely, the inhibitory matching process that supports spatial and motor processes cannot lead to resonance.

A New Way to Compute: Digital and Binary, Feedforward and Feedback, Analog Coherence
The LAMINART models (e.g., Figure 1) are not merely anatomically more precise versions of previous ART ideas. They represent a breakthrough in computing that identifies new principles and processes that embody novel computational properties with revolutionary implications. LAMINART models embody a new type of hybrid between *feedforward* and *feedback* computing, and also between *digital* and *analog* computing (Grossberg, 2003b) for processing distributed data. These properties go beyond the types of Bayesian models that are so popular today. They underlie the fast but stable self-organization that is characteristic of cortical development and life-long learning.

The synthesis of feedback and feedback processing can be understood from the following example: When an unambiguous scene is processed, the LAMINART model can quickly group the scene in a fast feedforward sweep of activation that passes directly through layer 4 to 2/3 and then on to layers 4 to 2/3 in subsequent cortical areas (Figures 1c and e). This property clarifies how recognition can be so fast in response to unambiguous scenes; e.g., Thorpe et al. (1996). On the other hand, if there are multiple possible groupings in a scene, say in response to a complex textured scene, then competition among these possibilities due to inhibitory interactions in layers 4 and 2/3 (black cells and synapses in Figure 1) can cause all cell activities to become smaller. This happens because the competitive circuits in the model are *self-normalizing*; that is, they tend to conserve the total activity of the circuit. This self-normalizing property is related to the ability of the shunting on-center off-surround networks that realize the competitive circuits to process input contrasts over a large dynamic range without saturation (Douglas et al., 1995; Grossberg, 1973, 1980; Heeger, 1992).

In other words, these self-normalizing circuits carry out a type of real-time probability theory in which the amplitude of cell activity covaries with the certainty of the network's selection, or decision, about a grouping. Amplitude, in turn, is translated into processing speed and coherence of cell activities. Low activation slows down the feedforward processing in the circuit because it takes longer for cell activities to exceed output threshold and to activate subsequent cells above threshold. In the model, network uncertainty is resolved through feedback: Weakly active layer 2/3 grouping cells feed back signals to layers 6-then-4-then-2/3 to close a cortical feedback loop that contrast enhances and amplifies the winning grouping to a degree and at a rate that reflect the amount of statistical evidence for that grouping. As the winner is selected, and weaker groupings are suppressed, its cells become more active and synchronous, hence can again rapidly send the cortical decision to subsequent processing stages.
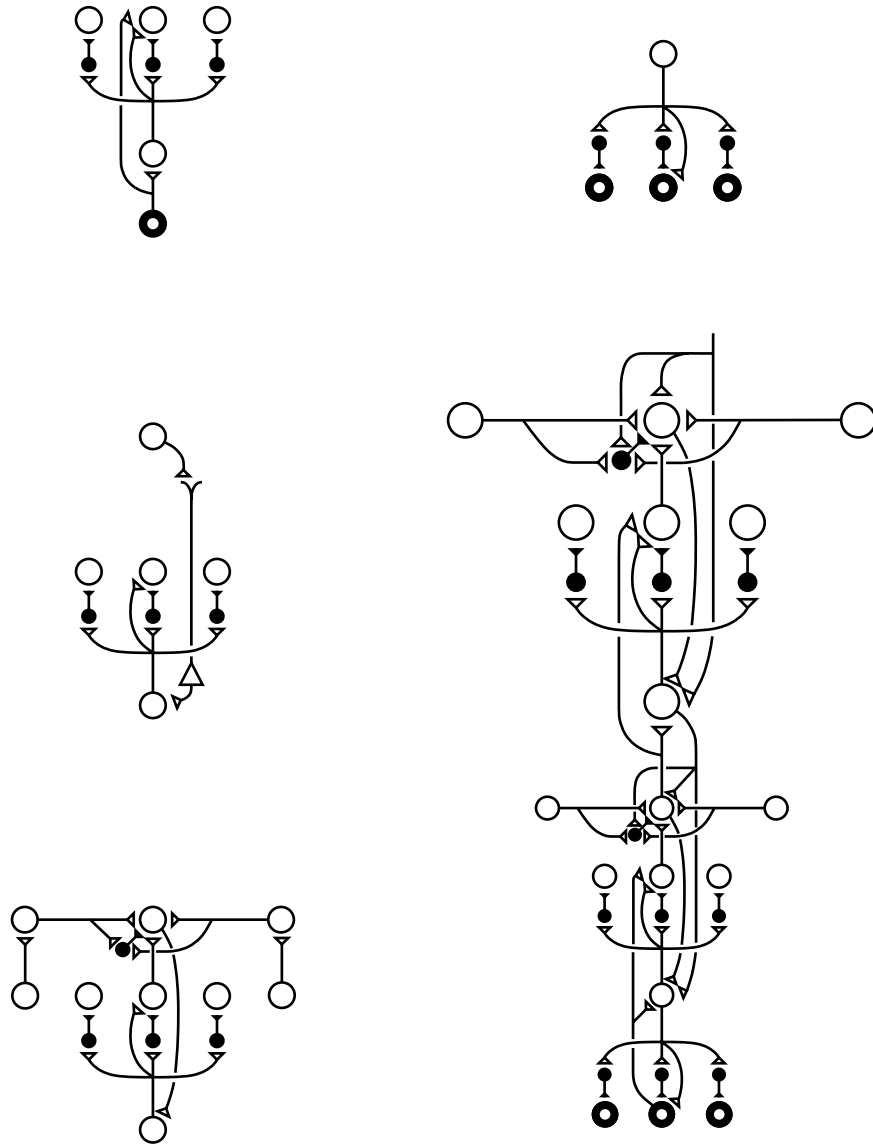
Figure 1. How known cortical connections join the layer 6 → 4 and layer 2/3 circuits to form an entire V1/V2 laminar model. Inhibitory interneurons are shown filled-in black. (a) The LGN provides bottom-up activation to layer 4 via two routes. First, it makes a strong connection directly into layer 4. Second, LGN axons send collaterals into layer 6, and thereby also activate layer 4 via the 6 → 4 on-center off-surround path. The combined effect of the bottom-up LGN pathways is to stimulate layer 4 via an on-center off-surround, which provides divisive contrast normalization (Grossberg, 1973, 1980; Heeger, 1992) of layer 4 cell responses. (b) Folded feedback carries attentional signals from higher cortex into layer 4 of V1, via the modulatory 6 → 4 path. Corticocortical feedback axons tend preferentially to originate in layer 6 of the higher area and to terminate in layer 1 of the lower cortex (Salin and Bullier, 1995, p.110), where they can excite the apical dendrites of layer 5 pyramidal cells whose axons send collaterals into layer 6. The triangle in the figure represents such a layer 5 pyramidal cell. Several other routes through which feedback can pass into V1 layer 6

exist (see Raizada and Grossberg (2001) for a review). Having arrived in layer 6, the feedback is then "folded" back up into the feedforward stream by passing through the 6 → 4 on-center off-surround path (Bullier et al., 1996). (c) Connecting the 6 → 4 on-center off-surround to the layer 2/3 grouping circuit: like-oriented layer 4 simple cells with opposite contrast polarities compete (not shown) before generating half-wave rectified outputs that converge onto layer 2/3 complex cells in the column above them. Just like attentional signals from higher cortex, as shown in (b), groupings that form within layer 2/3 also send activation into the folded feedback path, to enhance their own positions in layer 4 beneath them via the 6 → 4 on-center, and to suppress input to other groupings via the 6 → 4 off-surround. There exist direct layer 2/3 → 6 connections in macaque V1, as well as indirect routes via layer 5. (d) Top-down corticogeniculate feedback from V1 layer 6 to LGN also has an on-center off-surround anatomy, similar to the 6 → 4 path. The on-center feedback selectively enhances LGN cells that are consistent with the activation that they cause (Sillito et al., 1994), and the off-surround contributes to length-sensitive (endstopped) responses that facilitate grouping perpendicular to line ends. (e) The entire V1/V2 circuit: V2 repeats the laminar pattern of V1 circuitry, but at a larger spatial scale. In particular, the horizontal layer 2/3 connections have a longer range in V2, allowing above-threshold perceptual groupings between more widely spaced inducing stimuli to form (Amir, Harel, & Malach, 1993). V1 layer 2/3 projects up to V2 layers 6 and 4, just as LGN projects to layers 6 an 4 of V1. Higher cortical areas send feedback into V2 which ultimately reaches layer 6, just as V2 feedback acts on layer 6 of V1 (Sandell & Schiller, 1982). Feedback paths from higher cortical areas straight into V1 (not shown) can complement and enhance feedback from V2 into V1. Top-down attention can also modulate layer 2/3 pyramidal cells directly by activating both the pyramidal cells and inhibitory interneurons in that layer. The inhibition tends to balance the excitation, leading to a modulatory effect. These top-down attentional pathways tend to synapse in layer 1, as shown in Figure 2b. Their synapses on apical dendrites in layer 1 are not shown, for simplicity. (Reprinted with permission from Raizada and Grossberg (2001).)

In summary, the LAMINART circuit behaves like a real-time probabilistic decision circuit that operates as quickly as possible, given the evidence. It operates in a fast feedforward mode when there is little uncertainty, and automatically switches to a slower feedback mode when there is uncertainty. Feedback selects a winning decision which enables the circuit to speed up again, since activation amplitude, synchronization, and processing speed both increase with certainty.

The LAMINART model also embodies a novel kind of hybrid computing that simultaneously realizes the stability of digital computing and the sensitivity of analog computing. This is true because the feedback loop between layers 2/3-6-4-2/3 that selects or confirms a winning grouping (Figures 1c and e) has the property of *analog coherence* (Grossberg, 1999a; Grossberg, Mingolla, and Ross, 1997; Grossberg and Raizada, 2000); namely, this feedback loop can synchronously choose and store a winning grouping without losing analog sensitivity to amplitude differences in the input pattern. The coherence that is derived from synchronous storage in the feedback loop provides the stability of digital computing—the feedback loop exhibits hysteresis that can preserve the stored pattern against external perturbations— while preserving the sensitivity of analog computation.

All Conscious States are Resonant States, but Not the Converse: Biased Competition
Another property of note in a LAMINART circuit reflects the claim that the ability to rapidly learn throughout life without a loss of stability is related to consciousness: "all conscious states are

resonant states". However, the converse statement: "all resonant states are conscious states" is not predicted to be true. An example of such an exception will now be described.

LAMINART circuits can stabilize development and learning using the intracortical feedback loop between layers 2/3-6-4-2/3. This feedback loop supports an intracortical "resonance". This contrast-enhancing feedback loop selects winning groupings in the adult. It is also predicted to help stabilize development in the infant and learning throughout life, since cells that fire together wire together (Grossberg, 1999a). This *intra*cortical circuit can work even before *inter*cortical attentional feedback can develop (Figure 1e) . The LAMINART model clarified that both of these circuits can stabilize cortical development and learning, not only the top-down intercortical circuit that ART originally predicted. The intracortical feedback loops between different layers of the neocortex prevent an infinite regress from occurring, by stabilizing cortical development and learning before top-down intercortical feedback can develop and play its own role in stabilization.

Early versions of ART predicted that top-down attention can modulate and stabilize the learning process through a competitive matching process (Grossberg, 1976, 1980). Later modeling studies (e.g., Carpenter and Grossberg, 1987) refined this prediction to assert that this matching process is realized by a top-down, modulatory on-center, off-surround network. A great deal of perceptual and brain data have accumulated in support of this hypothesis; see Grossberg (2003b) and Raizada and Grossberg (2003) for reviews of these data, including the popular "biased competition" term for this process (Desimone, 1998).

The LAMINART model advanced this prediction by identifying intercortical and interlaminar circuits that can realize top-down, modulatory on-center, off-surround feedback (Figure b). This additional step also clarified how pre-attentive grouping and top-down attention share the same modulatory on-center, off-surround decision circuit from layer 6-to-4 with each other, and also with feedforward pathways that automatically activate cells in response to bottom-up inputs (Figures 1a - c). Because a "pre-attentive grouping is its own attentional prime," these intracortical feedback loops also solve another problem: ART predicted that, in order to prevent unstable development and learning, only bottom-up inputs can supraliminally activate brain sensory and cognitive cells that are capable of learning, since top-down attention is typically modulatory (except when volition enables top-down attention to generate visual imagery or thoughts; see Grossberg (2000a)). How, then, can illusory contours form without destabilizing brain circuits? Because a "pre-attentive grouping is its own attentional prime," it can use the layer 6-to-4 competitive decision circuit to select the correct grouping cells for learning, even without top-down attention.

This refinement of the ART prediction implies that, although top-down attention is needed for fast and stable learning of conscious experiences to occur, learning can also occur if a pre-attentive grouping competitively selects the correct cells with which to "resonate," and thereby synchronize, for a sufficiently long time using its intracortical 2/3-6-4-2/3 feedback circuit. Such learning may be slow and inaccessible to consciousness. Watanabe et al. (2003) have recently reported consistent data about slow perceptual learning without conscious awareness. More experiments need to be done to test if the predicted intracortical but interlaminar cortical mechanisms contribute to this sort of learning.

Three Types of Attention: Boundary, Surface, and Prototype Attention
The sharing by grouping and attention of the same decision circuit also enables the model to explain and simulate more data, including data about how attention can selectively activate an object by propagating along the object's boundary (Roelfsema et al. 1998); see Grossberg and Raizada (2000). Additional examples of the role of boundary attention have been described in simulations of the Necker cube (Grossberg and Swaminathan, 2004) and of bistable transparency (Grossberg and Yazdanbaksh, 2005).

Given the importance of attention in generating conscious experiences, it should be noted that at least three mechanistically-distinct types of attention have been distinguished by cortical modeling studies of visual perception and recognition: *Boundary attention*, whereby spatial attention can propagate along an object boundary to select the entire object for inspection; *surface attention*, whereby spatial attention can selectively fill in the surface shape of an object to form an "attentional shroud" (Tyler and Kontsevich, 1995); and *prototype attention* whereby critical feature patterns of a learned object category can be selectively enhanced. Boundary attention is summarized above. Surface attention helps to intelligently search a scene with eye movements and to learn view-invariant object categories (Fazl, Grossberg, and Mingolla, 2007). Prototype attention is the type of attention that is realized by ART top-down category learning circuits (Carpenter and Grossberg, 1987, 1991; Grossberg, 2003a). All three types of attention utilize one or another type of resonant feedback loops.

Distinguishing these three types of attention is difficult if only because they interact within and across the What and Where cortical processing streams. For example, boundary attention seems to be activated, at least in the experiments of Roelfsema et al. (1998), when spatial attentional maps in the Where cortical stream, notably parietal cortex, project to perceptual boundary representations in the What cortical stream, notably the pale stripes of cortical area V2. Surface attention can be activated when spatial attentional maps in the Where cortical stream, again from parietal cortex, project to perceptual surface representations, notably in the thin stripes of cortical area V2 and in V4, and conversely to form a surface-shroud resonance. Finally, prototype attention seems to act entirely within the What cortical stream from learned recognition categories in prefrontal cortex and inferotemporal cortex to perceptual representations in V2 and V4.

Characterizing these three mechanistically distinct types of attention is further complicated by the fact that feedback interactions occur between the boundary and surface representations in cortical areas V1 and V2, and are predicted to help separate figures from their background in depth, and to control saccadic eye movements (Fang and Grossberg, 2007; Fazl, Grossberg, and Mingolla, 2007; Grossberg, 1994, 1997). Modulation of either boundary or surface representations by spatial attention can therefore be expected to have effects on both types of representations due to these inter-stream feedback interactions. In addition, it is well-known that moving stimuli can activate both the What and Where streams and can automatically attract spatial attention in parietal cortex via Where stream pathways from V1 and MT to MST and parietal cortex,

Balancing Excitation and Inhibition: The Road to Synchrony
The dynamics of LAMINART circuits, whether in their pre-attentively or attentively activated modes, depend upon the existence of *balanced* excitatory and inhibitory signals in different cortical layers. In particular, a balance between excitation and inhibition is needed in the perceptual grouping circuit by bipole cells in layer 2/3 (Figures 1c and e). This balance helps to ensure that perceptual groupings can form inwardly between pairs or greater numbers of inducers, but not outwardly from a single inducer. Likewise, a balance between excitation and inhibition is required in the on-center of the circuit from layer 6-to-4 that can provide excitatory modulation of cell activities in layer 4, but not fire them fully (Figures 1a - c). As noted above, this latter circuit plays an important role in attention (Figure 2b) and in the pre-attentive selection of a correct perceptual grouping in response to a complicated scene (Figure 2c). Grossberg and Williamson (2001) proposed that such balanced circuits are needed for the cortex to develop and learn in a stable way, and simulated how such balanced connections could grow during cortical development. Indeed, if inhibition if inhibition develops to be too weak, then excitation can propagate uncontrollably, whereas if it is too strong, then cells cannot get sufficiently activated.

On the other hand, balanced excitatory and inhibitory connections have also been used to explain the observed variability in the number and temporal distribution of spikes emitted by cortical neurons (Shadlen and Newsome, 1998; van Vreeswijk and Sompolinsky, 1998). These spiking patterns are quite inefficient in firing cortical cells. Given the LAMINART model proposal that such variability may reflect mechanisms that are needed to ensure stable development and learning by cortical circuits—that is, "stability implies variability"— the cortex is faced with the difficult problem of how to overcome the inefficiency of variable spiking patterns in driving responses from cortical neurons. The LAMINART model shows how these balanced excitatory and inhibitory connections work together overcome the inefficiency of intermittent spiking by resynchronizing desynchronized signals that belong to the same object, and thereby ensuring that the cortex processes them efficiently. In other words, the very process that enables cortical cells to respond selectively to input patterns —namely, balanced excitation and inhibition—also ensures that cortical cells can fire vigorously and synchronously in response to those patterns that are selected by cortical bottom-up filtering, horizontal grouping, and top-down attention processes.

The remainder of this article summarizes properties of cortical circuits that enable them to realize the predicted CLEARS relationships, and illustrative data that support these predicted circuits.

The Stability-Plasticity Dilemma: Rapid Learning Throughout Life

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. How do we integrate them into unified experiences that cohere into a sense of self? One has only to see an exciting movie just once to marvel at this capacity, since we can then 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 remarkable degree, we can rapidly learn new facts without being forced to just as rapidly forget what we already know. As a result, we can confidently go out into the world without fearing that, in learning to recognize a new friend's face, we will suddenly forget our parents' faces. This is sometimes called the problem of *catastrophic forgetting*.

Many contemporary learning algorithms do experience catastrophic forgetting, particularly when they try to learn quickly in response to a changing world. Speaking technically, the brain solves a challenging problem that many current approaches to technology have not yet solved: It is a self-organizing system that is capable of rapid, yet stable, autonomous learning of huge amounts of data from a changing environment that can be filled with unexpected events. 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. ART has attempted to articulate some of these principles, and the neural mechanisms that realize them. The next sections summarize aspects of how this is proposed to occur.

How CLEARS Mechanisms Interact

Humans are *intentional* beings who learn expectations about the world and make predictions about what is about to happen. Humans are also *attentional* beings who focus processing resources upon a restricted amount of incoming information at any time. Why are we both intentional and attentional beings, and are these two types of processes related? The stability-plasticity dilemma and its solution using resonant states provides a unifying framework for understanding these issues.

To clarify the role of sensory or cognitive expectations, and of how a resonant state is activated, suppose you were asked to "find the yellow ball as quickly as possible, and you will win a $10,000 prize". Activating an expectation of a "yellow ball" enables its more rapid detection, and with a more energetic neural response. Sensory and cognitive top-down expectations hereby lead to *excitatory matching* with consistent bottom-up data. Mismatch between top-down expectations and bottom-up data can suppress the mismatched part of the bottom-up data, to focus attention upon the matched, or expected, part of the bottom-up data.

Excitatory matching and attentional focusing on bottom-up data using top-down expectations generates resonant brain states: When there is a good enough match between bottom-up and top-down signal patterns between two or more levels of processing, their positive feedback signals amplify and prolong their mutual activation, leading to a resonant state. Amplification and prolongation of activity triggers learning in the more slowly varying adaptive weights that control the signal flow along pathways from cell to cell. Resonance hereby provides a global context-sensitive indicator that the system is processing data worthy of learning, hence the name *Adaptive Resonance Theory*, or ART.

In summary, ART predicts a link between the mechanisms which enable us to learn quickly and stably about a changing world, and the mechanisms that enable us to learn expectations about such a world, test hypotheses about it, and focus attention upon information that we find interesting. ART clarifies this link by asserting that, in order to solve the stability-plasticity dilemma, only resonant states can drive rapid new learning.

It is just a step from here to propose that those experiences which can attract our attention and guide our future lives by being learned are also among the ones that are conscious. Support for this additional assertion derives from the many modeling studies whose simulations of behavioral and brain data using resonant states map onto properties of conscious experiences in those experiments.

The type of learning within the sensory and cognitive domain that ART mechanizes is *match learning*: Match learning occurs only if a good enough match occurs between bottom-up information and a learned top-down expectation that is read out by an active recognition category, or code. When such an approximate match occurs, previously learned knowledge can be refined. Match learning raises the concern about what happens if a match is not good enough? How does such a model escape perseveration on already learned representations?

If novel information cannot form a good enough match with the expectations that are read-out by previously learned recognition categories, then a memory search, or hypothesis testing, is triggered that leads to selection and learning of a new recognition category, rather than catastrophic forgetting of an old one. Figure 2 illustrates how this happens in an ART model; it will be discussed in greater detail below. In contrast, as noted above, learning within spatial and motor processes is proposed to be *mismatch learning* that continuously updates sensory-motor maps or the gains of sensory-motor commands. As a result, we can stably learn what is happening in a changing world, thereby solving the stability-plasticity dilemma, while adaptively updating our representations of where objects are and how to act upon them using bodies whose parameters change continuously through time.
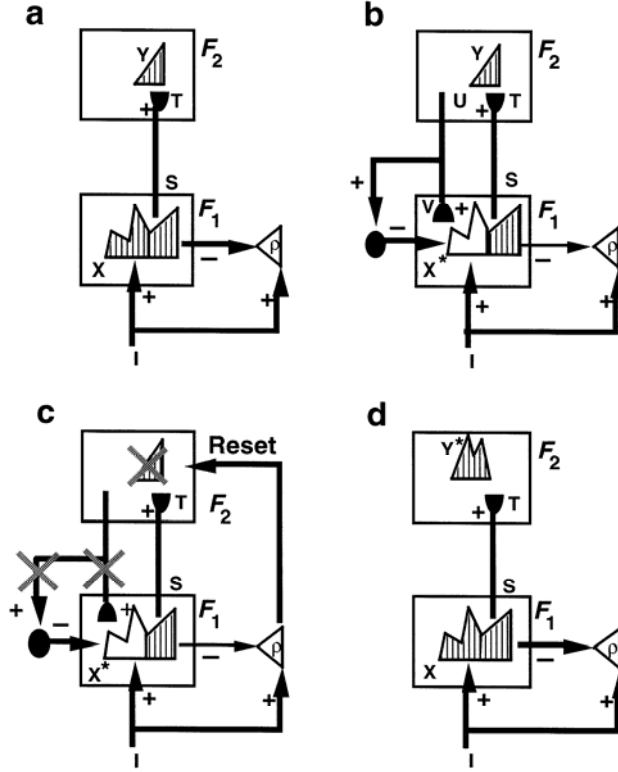
Figure 2. Search for a recognition code within an ART learning circuit: (a) The input pattern *I* is instated across the feature detectors at level F_1 as a short term memory (STM) activity pattern *X*. Input *I* also nonspecifically activates the orienting system with a gain that is called vigilance (ρ); that is, all the input pathways converge with gain ρ onto the orienting system and try to activate it. STM pattern *X* is represented by the hatched pattern across F_1. Pattern *X* both inhibits the orienting system and generates the output pattern *S*. Pattern *S* is multiplied by learned adaptive weights, also called long term memory (LTM) traces. These LTM-gated signals are added at F_2 cells, or 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 a *prototype* pattern *V* that encodes the learned expectation of the active F_2 nodes. Such a prototype represents the set of commonly shared features in all the input patterns capable of activating *Y*. If V mismatches I at F_1, then a new STM activity pattern *X** is selected *at F_1*. *X** is represented by the hatched pattern. It consists of the features of *I* that are confirmed by *V*. Mismatched features are inhibited. The 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 the orienting system. (c) If inhibition decreases sufficiently, the orienting system releases a nonspecific arousal wave to F_2; that is, a wave of activation that equally activates all F_2 nodes. This wave instantiates the intuition that "novel events are arousing". This arousal wave resets the STM pattern Y at F_2 by inhibiting Y. (d) After Y is inhibited, its top-down prototype signal is eliminated, and X can be reinstated at F_1. The prior reset event maintains inhibition of Y during the search cycle. As a result, X can activate a different STM pattern Y at F_2. If the top-down prototype due to this new Y pattern 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. Such a search cycle represents a type of nonstationary hypothesis testing. When search ends, an attentive resonance develops and learning of the attended data is initiated. [Adapted with permission from Carpenter and Grossberg (1993).]

It has been mathematically proved that match learning within an ART model leads to stable memories in response to arbitrary list of events to be learned (Carpenter and Grossberg, 1991). However, match learning also has a serious potential weakness: If you can only learn when there is a good enough match between bottom-up data and learned top-down expectations, then how do you ever learn anything that you do not already know? ART proposes that this problem is solved by the brain by using an interaction between complementary processes of *resonance* and *reset*, that are predicted to control properties of attention and memory search, respectively. These complementary processes help our brains to balance between the complementary demands of processing the familiar and the unfamiliar, the expected and the unexpected.

Organization of the brain into complementary processes is predicted to be a general principle of brain design that is not just found in ART (Grossberg, 2000b). A complementary process can individually compute some properties well, but cannot, by itself, process other complementary properties. In thinking intuitively about complementary properties, one can imagine puzzle pieces fitting together. Both pieces are needed to finish the puzzle. Complementary processes in the brain are much more dynamic than any such analogy, however: Pairs of complementary processes interact in such a way that their emergent properties overcome their complementary deficiencies to compute complete information about some aspect of the control of intelligent behavior.

The resonance process in the complementary pair of resonance and reset is predicted to take place in the What cortical stream, notably in the inferotemporal and prefrontal cortex. Here top-down expectations are matched against bottom-up inputs (Desimone, 1998; Miller, Li, and Desimone, 1991). When a top-down expectation achieves a good enough match with bottom-up data, this match process focuses attention upon those feature clusters in the bottom-up input that are expected. If the expectation is close enough to the input pattern, then a state of resonance develops as the attentional focus takes hold.

Figure 2 illustrates these ART ideas in a simple two-level example. Here, a bottom-up input pattern, or vector, I activates a pattern X of activity across the feature detectors of the first level F_1. For example, a visual scene may be represented by the features comprising its boundary and surface representations. This feature pattern represents the relative importance of different features in the inputs pattern I. In Figure 2a, the pattern peaks represent more activated feature detector cells, the troughs less activated feature detectors. This feature pattern sends signals S through an adaptive filter to the second level F_2 at which a compressed representation Y (also called a recognition category, or a symbol) is activated in response to the distributed input T. Input T is computed by multiplying the signal vector S by a matrix of adaptive weights that can be altered through learning. The representation Y is compressed by competitive interactions across F_2 that allow only a small subset of its most strongly activated cells to remain active in response to T. The pattern Y in the

figure indicates that a small number of category cells may be activated to different degrees. These category cells, in turn, send top-down signals U to F_1. The vector U is converted into the top-down expectation V by being multiplied by another matrix of adaptive weights. When V is received by F_1, a matching process takes place between the input vector I and V which selects that subset X* of F_1 features that were "expected" by the active F_2 category Y. The set of these selected features is the emerging "attentional focus".

Binding Distributed Feature Patterns and Symbols during Conscious Resonances
If the top-down expectation is close enough to the bottom-up input pattern, then the pattern X* of attended features reactivates the category Y which, in turn, reactivates X*. The network hereby locks into a resonant state through a positive feedback loop that dynamically links, or binds, the attended features across X* with their category, or symbol, Y.

Resonance itself embodies another type of complementary processing. Indeed, there seem to be complementary processes both within and between cortical processing streams (Grossberg, 2000). This particular complementary relation occurs between distributed feature patterns and the compressed categories, or symbols, that selectively code them:

Individual features at F_1 have no meaning on their own, just like the pixels in a picture are meaningless one-by-one. The category, or symbol, in F_2 is sensitive to the global patterning of these features, and can selectively fire in response to this pattern. But it cannot represent the "contents" of the experience, including their conscious qualia, due to the very fact that a category is a compressed, or "symbolic" representation. Practitioners of Artificial Intelligence have claimed that neural models can process distributed features, but not symbolic representations. This is not, of course, true in the brain, which is the most accomplished processor both of distributed features and of symbols that is known to humans. Nor is it true in ART.

The resonance between these two types of information converts the *pattern* of attended features into a coherent context-sensitive state that is linked to its category through feedback. It is this coherent state, which binds together distributed features and symbolic categories, that can enter consciousness. This resonant binding process joins spatially distributed features into either a stable equilibrium or a synchronous oscillation. The original ART article (Grossberg, 1976) predicted the existence of such synchronous oscillations, which were there described in terms of their mathematical properties as "order-preserving limit cycles". Since the first neurophysiological experiments reported such synchronous oscillations (Eckhorn *et al*., 1988; Gray and Singer, 1989; Gray *et al*., 1989), there have been a rapidly growing number of supportive experiments. Simulations of fast-synchronizing ART and perceptual grouping circuits were reported in Grossberg and Somers (1991) and Grossberg and Grunewald (1997), and in a laminar cortical model by Yazdanbakhsh and Grossberg (2004). The ability of neural circuits to synchronize quickly is a topic that is worthy of considerable discussion in its own right.

Resonance Links Intentional and Attentional Information Processing to Learning
In ART, the resonant state, rather than bottom-up activation, is predicted to drive the learning process. The resonant state persists long enough, and at a high enough activity level, to activate the slower learning processes in the adaptive weights that guide the flow of signals between bottom-up and top-down pathways between levels F_1 and F_2 in Figure 2. This viewpoint helps to explain how adaptive weights that were changed through previous learning can regulate the brain's present information processing, without learning about the signals that they are currently processing unless they can initiate a resonant state. Through resonance as a mediating event, one can understand from a deeper mechanistic view why humans are intentional beings who are continually predicting what may next occur, and why we tend to learn about the events to which we pay attention.

More recent laminar versions of ART, notably the Synchronous Matching ART (SMART) model (Grossberg and Versace, 2007; Versace and Grossberg, 2005, 2006), show how a match may lead to fast gamma oscillations that facilitate spike-timing dependent plasticity (STDP), whereas mismatch can lead to slower beta oscillations that greatly lower the probability that mismatched events can be learned by a STDP learning law. These new features will be summarized below after more basic concepts are reviewed.

Complementary Attentional and Orienting Systems Control Resonance vs. Reset
A sufficiently bad mismatch between an active top-down expectation and a bottom-up input, say because the input represents an unfamiliar type of experience, can drive a memory search. Such a mismatch within the attentional system is proposed to activate a complementary *orienting system*, which is sensitive to unexpected and unfamiliar events. ART suggests that this orienting system includes the hippocampal system, which has long been known to be involved in mismatch processing, including the processing of novel events (Deadwyler, West, and Lynch, 1979; Deadwyler, West, and Robinson, 1981; Otto and Eichenbaum, 1992; Sokolov, 1968, Vinogradova, 1975). More recent work on SMART also implicated the nonspecific thalamic nuclei; see below. Output signals from the orienting system rapidly reset the recognition category that has been reading out the poorly matching top-down expectation (Figures 2b and 2c). The cause of the mismatch is hereby removed, thereby freeing the system to activate a different recognition category (Figure 2d). The reset event hereby triggers memory search, or hypothesis testing, which automatically leads to the selection of a recognition category that can better match the input. Various data support the existence of this predicted hypothesis testing cycle. In particular, Banquet and Grossberg (1987) summarized evidence from an experiment on humans that was designed to test this ART prediction by measuring event-related potentials (ERP). This study showed that sequences of P120 – N200 – P300 ERPs have the properties of ART mismatch – arousal – reset that are predicted to occur during a hypothesis testing cycle. Many subsequent studies have provided additional evidence for predictive coding; e.g., Ahissar *et al.* (2002), Desimone *et al.* (1998), Engel *et al.* (2001), Gao *et al.* (1998), Hermann *et al.* (2004), Krupa *et al.* (1999), and Salin and Bullier (1995).

If no such recognition category exists, say because the bottom-up input represents a truly novel experience, then the search process automatically activates an as yet uncommitted population of cells, with which to learn about the novel information. In order for a top-down expectation to match a newly discovered recognition category, its top-down adaptive weights initially have large values, which are pruned by the learning of a particular expectation.

This learning process works well under both unsupervised and supervised conditions (Carpenter *et al.*, 1992). Unsupervised learning means that the system can learn how to categorize novel input patterns without any external feedback. Supervised learning uses predictive errors to let the system know whether it has categorized the information correctly. Supervision can force a search for new categories that may be culturally determined, and are not based on feature similarity alone. For example, separating the featurally similar letters E and F into separate recognition categories is culturally determined. Such error-based feedback enables variants of E and F to learn their own category and top-down expectation, or prototype. The complementary, but interacting, processes of attentive-learning and orienting-search together realize a type of error correction through hypothesis testing that can build an ever-growing, self-refining internal model of a changing world.

Controlling the Content of Conscious Experiences: Exemplars and Prototypes

What combinations of features or other information are bound together into conscious object or event representations? One view is that exemplars, or individual experiences, are learned, because humans can have very specific memories. For example, we can all recognize the particular faces of our friends. On the other hand, storing every remembered experience as exemplars can lead to a combinatorial explosion of memory, as well as to unmanageable problems of memory retrieval. A possible way out is suggested by the fact that humans can learn prototypes which represent general properties of the environment (Posner and Keele, 1968). For example, we can recognize that everyone has a face. But then how do we learn specific episodic memories? ART provides an answer to this question that overcomes problems faced by earlier models.

The first thing to realize is that ART prototypes are not merely averages of the exemplars that are classified by a category, as is typically assumed in classical prototype models. Rather, they are the actively selected *critical feature patterns* upon which the top-down expectations of the category focus attention. In addition, the generality of the information that is codes by these critical feature patterns is controlled by a gain control process, called *vigilance* control, which can be influenced by environmental feedback or internal volition (Carpenter and Grossberg, 1987). Low vigilance permits the learning of general categories with abstract prototypes. High vigilance forces a memory search to occur for a new category when even small mismatches exist between an exemplar and the category that it activates. As a result, in the limit of high vigilance, the category prototype may encode an individual exemplar.

Vigilance is computed within the orienting system of an ART model (Figures 2b-d). It is here that bottom-up excitation from all the active features in an input pattern I are compared with inhibition from all the active features in a distributed feature representation across F_1. If the ratio of the total activity across the active features in F_1 (that is, the "matched" features) to the total activity due to all the features in I is less than a *vigilance parameter* ρ (Figure 2b), then a reset wave is activated (Figure 2c), which can drive the search for another category with which to classify the exemplar. In other words, the vigilance parameter controls how bad a match can be before search for a new category is initiated. If the vigilance parameter is low, then many exemplars can all influence the learning of a shared prototype, by chipping away at the features which are not in common with all the exemplars. If the vigilance parameter is high, then even a small difference between a new exemplar and a known prototype (e.g., F vs. E) can drive the search for a new category with which to represent F.

One way to control vigilance is by a process of *match tracking*. Here a predictive error (e.g., D is predicted in response to F), the vigilance parameter increases until it is just higher than the ratio of active features in F_1 to total features in I. In other words, vigilance "tracks" the degree of match between input exemplar and matched prototype. This is the minimal level of vigilance that can trigger a reset wave and thus a memory search for a new category. Match tracking realizes a Minimax Learning Rule that conjointly *maximizes* category generality while it *minimizes* predictive error. In other words, match tracking uses the least memory resources that can prevent errors from being made.

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 views of faces and dogs. ART models hereby try 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 and Kemler, 1978; Ward, 1983).

If vigilance control is important for normal learning, one might expect breakdowns in vigilance control to contribute to certain mental disorders. It has been suggested that an abnormally low vigilance may contribute to medial temporal amnesia (Carpenter and Grossberg, 1993) and that an abnormally high vigilance may contribute to autism (Grossberg and Seidman, 2006). These proposals point to the utility of classifying certain mental disorders as "vigilance diseases".

A biologically relevant neural model must be able to explain and predict more behavioral and neural data than its competitors. One additional mark of maturity of such a model is that it "works" and can solve complicated real-world problems. Many benchmark studies of ART show that it useful in many large-scale engineering and technological applications. See http://profusion.bu.edu/techlab for some illustrative benchmark studies. In particular, vigilance control in the classification of complex data bases enables the number of ART categories that are learned to scale well with the complexity of the input data.

Memory Consolidation and the Emergence of Rules
As sequences of inputs are practiced over learning trials, the search process eventually converges upon stable categories. It has been mathematically proved (Carpenter and Grossberg, 1987) 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 some ART models can, at any stage of learning, 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. These examples show that the claims of some cognitive scientists and AI practioners that neural network models cannot learn rule-based behaviors are as incorrect as the claims that neural models cannot learn symbols.

From Spikes to Cognition: Gamma Oscillations and STDP in Laminar Thalamocortical Circuits
The Synchronous Matching Adaptive Resonance Theory (SMART) model advances ART in several ways (Grossberg and Versace, 2007; Versace and Grossberg, 2005, 2006); see Figure 3. SMART links attentive learning requirements to how laminar neocortical circuits interact with primary, higher-order (e.g., the pulvinar nucleus; Sherman and Guillery, 2001; Shipp, 2003), and nonspecific thalamic nuclei (van Der Werf *et al.*, 2002). Corticothalamocortical pathways work in parallel with corticocortical routes (Maunsell and van Essen, 1983; Salin and Bullier, 1995; Sherman and Guillery, 2002). Specific first-order thalamic nuclei (such as the Lateral Geniculate Nucleus, LGN) relay sensory information to the cerebral cortex, whereas specific second-order thalamic nuclei receive their main input from lower-order cortical areas, notably from layer 5, and relay this information to higher-order cortical areas (Sherman and Guillery, 2002). The model clarifies how a match between cortical and thalamic inputs at the level of specific first-order and higher-order

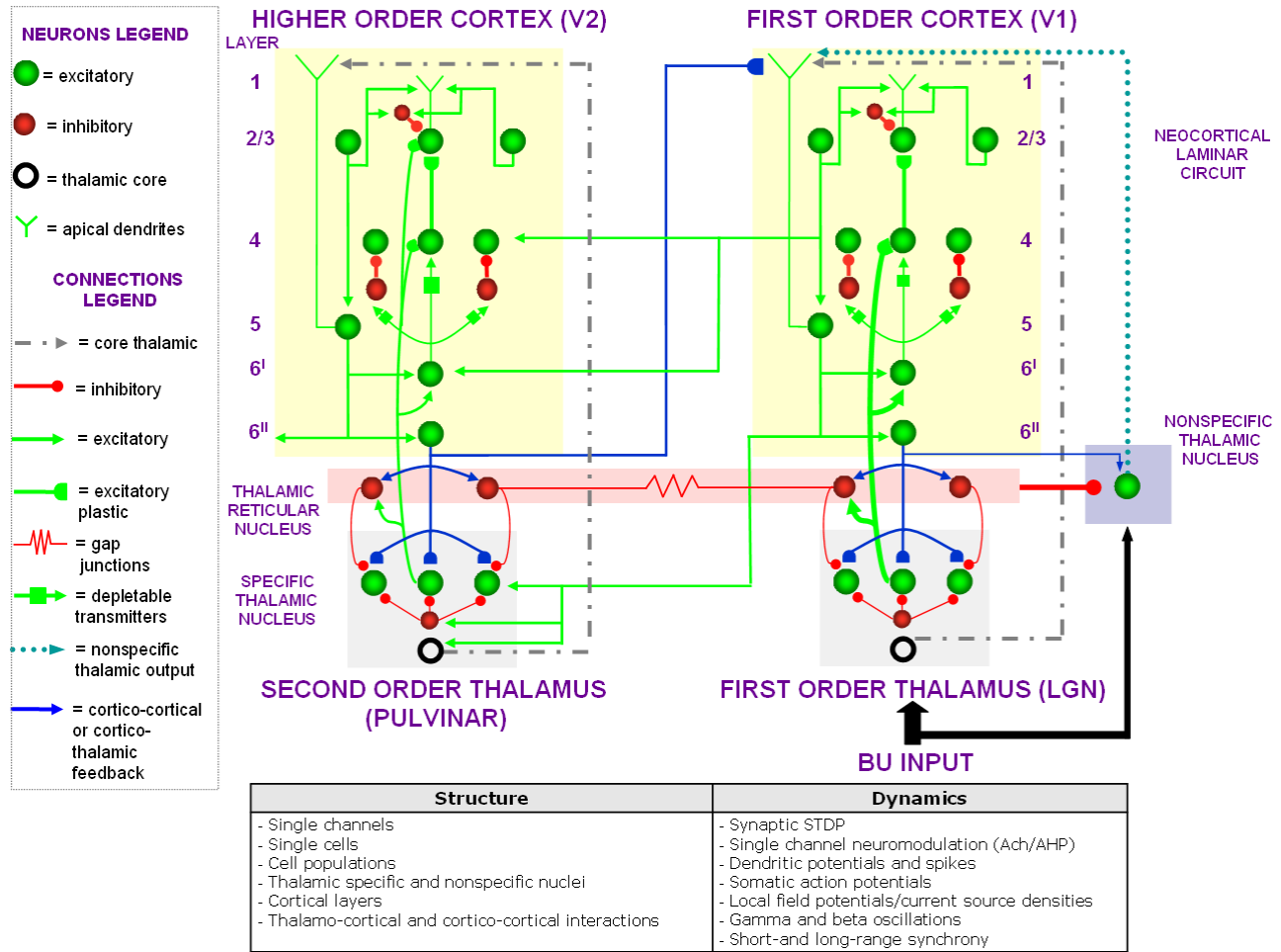thalamic nuclei might subserve fast stable learning of neural representations in the thalamocortical system.

Figure 3. SMART model overview. A first-order and higher-order cortical area are linked by corticocortical and corticothalamocortical connections. The thalamus is subdivided into specific first-order, second-order, nonspecific, and thalamic reticular nucleus (TRN). The thalamic matrix (one cell population shown as an open ring) provides priming to layer 1, where layer 5 pyramidal cell apical dendrites terminate. The specific thalamus relays sensory information (first-order thalamus) or lower-order cortical information (second-order thalamus) to the respective cortical areas via plastic connections. The nonspecific thalamic nucleus receives convergent BU input and inhibition from the TRN, and projects to layer 1 of the laminar cortical circuit, where it regulates reset and search in the cortical circuit (see text). Corticocortical feedback connections link layer 6^{II} of the higher cortical area to layer 1 of the lower cortical area, whereas thalamocortical feedback originates in layer 6^{II} and terminates in the specific thalamus after synapsing on the TRN. Layer 6^{II} corticothalamic feedback matches the BU input in the specific thalamus. V1 receives two parallel BU thalamocortical pathways. The LGN→V1 layer 4 pathway and the modulatory LGN→V1 layer 6^{I}→4 pathway provide divisive contrast normalization (Grossberg, 1980; Heeger, 1992) of layer 4 cell responses. The intracortical loop V1 layer 4→2/3→5→6^{I}→4 pathway (*folded feedback,* Raizada and Grossberg, 2003*)* enhances the activity of winning layer 2/3 cells at

their own positions via the $6^I \rightarrow 4$ on-center, and suppresses input to other layer 2/3 cells via the $6^I \rightarrow 4$ off-surround. V1 also activates the BU V1→V2 corticocortical pathways (V1 layer 2/3→V2 layers 6^I and 4) and the BU corticothalamocortical pathways (V1 layer 5 →PULV→V2 layers 6^I and 4), where the layer $6^I \rightarrow 4$ pathway provides divisive contrast normalization to V2 layer 4 cells analogously to V1. Corticocortical feedback from V2 layer $6^{II} \rightarrow$V1 layer $5 \rightarrow 6^I \rightarrow 4$ also uses the same modulatory $6^I \rightarrow 4$ pathway. TRN cells of the two thalamic sectors are linked via gap junctions, which provide synchronization of the two thalamocortical sectors when processing BU stimuli. [Reprinted with permission from Grossberg and Versace (2007).]

In particular, suppose that, at a specific thalamic nucleus, a sufficiently good match occurs between a bottom-up input pattern and a top-down expectation from layer 6 of its corresponding cortical area. Such a match can trigger fast synchronized gamma oscillations (γ, 20–70Hz), whose short period enables synchronized spikes to drive learning via a spike-timing-dependent plasticity (STDP; Levy and Steward, 1983; Markram *et al.*, 1997; Bi and Poo, 2001) learning rule. In particular, STDP is maximal when pre- and post-synaptic cells fire within 10-20ms of each other, and thus favors learning in match states whose synchronous fast oscillations fall within the temporal constraints of STDP (Traub *et al.*, 1998; Wespatat *et al.*, 2004). In contrast, mismatched cells undergo slower beta oscillations (β, 4–20Hz). whose spikes do not fall within the STDP learning window.

SMART hereby brings the new features of synchronized oscillation frequency and STDP into the discussion of how learning is selectively regulated. Aggregate and single-cell recordings from multiple thalamic and cortical levels of mammals have shown high- and low-frequency rhythmic synchronous activity correlated with cognitive, perceptual and behavioral tasks, and large-scale neuronal population models have been proposed to model oscillatory dynamics (Bazhenov *et al.*, 1998; Lumer *et al.*, 1997; Destexhe *et al.*, 1999; Siegel *et al.*, 2000). However, these models have not linked brain spikes, oscillations, STDP, and the brain states that subserve cognitive information processing.

SMART proposes that such a match or mismatch at a higher cortical level can occur as follows: Activation of layer 5 cells in a lower cortical area (e.g., V1) generates driving inputs to a higher-order specific thalamic area (e.g., pulvinar); see Rockland *et al.* (1999). Terminations arising from layer 5 are similar to retinogeniculate RL synapses, or driving, connections, often found in more proximal segments of the dendrites. This pattern of connectivity seems to be constant across species (Rouiller and Welcher, 2000).

A top-down expectation from layer 6^{II} of the corresponding cortical area (e.g., V2) is matched in this thalamic region against the layer 5 output pattern, similar to the way in which retinal inputs to the lateral geniculate nucleus are matched by top-down signals from layer 6^{II} of V1. If a sufficiently good match occurs, then synchronized gamma oscillations can be triggered in the pulvinar and V2, leading to learning of the critical features that are part of the matched pattern.

Thalamocortical Mismatch. Nonspecific Thalamic Nucleus, and Layer 1 Mediated Reset
If the match is not good enough, then the nonspecific thalamic nucleus gets activated by a mechanism that is similar to that summarized in Figure 2, but which is anatomically more precisely characterized in Figure 3. Nonspecific thalamic activation is nonspecifically broadcast as an arousal signal to many cortical areas via diffuse inputs across layer 1. We suggest that this nonspecific pathway is part of the orienting system that triggers reset in response to mismatch events. In particular, apical dendrites in layer 1 of layer 5 cells receive this arousal input. If some of these

layer 5 cells are active when the arousal burst occurs, their firing rate is enhanced in response to the arousal input. This enhancement of layer 5 cell firing triggers a selective reset of cortical and thalamic cells in the following way:

Layer 5 cells project to layer 4 via layer 6 (Figure 3). The signals from layer 6 to 4 are gated by habituative transmitters, also called depressing synapses. Activation patterns in these circuits just prior to the arousal burst bias the habituative network of cells feeding layer 4. The active circuits are presumably the ones that caused the predictive mismatch. When the arousal burst occurs, these previously active cells are disadvantaged relative to cells that were not active. A reset event can then occur that inhibits the previously active cells as it selects new cells with which to better code the novel input.

This model explanation is supported by quantitative simulations of data about single cell biophysics and neurophysiology, laminar neuroanatomy, aggregate cell recordings (current-source densities, local field potentials), and large-scale oscillations at beta and gamma frequencies, which the model functionally links to requirements about how to achieve fast stable attentive learning.

A Prediction About Why Oscillation Frequencies Are Slower in Lower Cortical Layers: Reset
As noted above, SMART predicts that thalamocortical mismatches may cause cortical reset via the deeper cortical layers 6 and 4. Model simulations also show that mismatches lead to slower beta oscillations. Putting these two properties together leads to the prediction that the deeper layers of neocortex may express beta oscillations more frequently than the superficial layers. Such a property has recently been experimentally reported (Buffalo *et al.*, 2004). It remains to test whether the observed experimental property is related to the SMART reset prediction.

Two issues may be noted in this regard. One concerns how the prediction may be tested: One possible test would be to carry out a series of experiments on the same animal in which the animal is exposed to environments with progressively more novel events. More novel events should cause more cortical resets. Do more cortical resets per unit time cause more beta oscillations in the lower cortical layers? The second issue notes that the differences between the oscillation frequencies in the deeper and more superficial cortical layers are averages over time. This is essential to realize because there exist interlaminar intracortical feedback loops that may be expected to synchronize all the cortical layers during a match event (Yazdanbakhsh and Grossberg, 2004). Indeed, these are the intracortical feedback loops whereby "a pre-attentive grouping is its own attentional prime", and thus enable neocortex to develop its circuits, without a loss of stability, even before intercortical attentional circuits can develop.

Discussion
This article summarizes how Adaptive Resonance Theory explains mechanistic relationships between the CLEARS properties of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony. ART proposes that these processes work together to solve the stability-plasticity dilemma, and thus to enable advanced animals, including humans, to learn quickly about a changing world throughout life without experiencing catastrophic forgetting. Conscious events are predicted to be a subset of resonant events. Resonance, in turn, controls learning to avoid catastrophic forgetting by involving learned-top-down expectations in a predictive matching process that focuses attention on the critical feature patterns that are worthy of learning, and also of consciousness. Recent developments of ART, such as the LAMINART model, and the spiking SMART model, show that predicted ART mechanisms can be given detailed anatomical and neurobiological interpretation within a laminar corticothalamic architecture that greatly expands the explanatory and predictive power of ART. An important research theme is to characterize which resonant events are not conscious events. A start in this direction is the observation that intracortical resonances may

stabilize learning without supporting conscious experiences, which seems to be what happens in the perceptual learning experiments of Watanabe et al. (2001).

References

Ahissar, M. & Hochstein, S. (2002). View from the Top: Herarchies and Reverse Hierarchies in the Visual System. Neuron, 36, 791-804.

Banquet, J.P. & Grossberg, S. (1987). Probing cognitive processes through the structure of event-related potentials during learning: An experimental and theoretical analysis. *Applied Optics, 26,* 4931-4946.

Bazhenov, M., Timofeev, I., Steriade, M. & Sejnowski, T.J. (1998). Computational models of thalamocortical augmenting responses. *The Journal of Neuroscience, 18*, 6444-6465.

Bi, G.Q. & Poo, M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience, 24*, 139-166.

Buffalo, E.A., Fries, P. & Desimone, R. (2004). Layer-specific attentional modulation in early visual areas. *Society for Neuroscience Abstract*, Project No. 717.6.

Carpenter, G.A. & Grossberg, S. (1987). 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. (1991). *Pattern recognition by self-organizing neural networks.* Cambridge, MA: MIT Press.

Carpenter, G.A. & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences, 16*, 131-137.

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.

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 1*, 101-107.

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

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.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences. 353*, 1245-1255.

Destexhe, A., Contreras, D. & Steriade, M. (1999). Cortically-induced coherence of a thalamic-generated oscillation. *Neuroscience. 92*, 427–443.

Douglas, R.J., Koch C., Mahowald M., Martin K.A.C. & Suarez H.H. (1995). Recurrent excitation in neocortical circuits. *Science, 269*, 981–985.

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.

Engel, A.K., Fries, P. & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Review Neuroscience, 2*, 704-716.

Fang, L. & Grossberg, S. (2007). From stereogram to surface: How the brain sees the world in depth. *Spatial Vision*, in press.

Fazl, A., Grossberg, S. & Mingolla, E. (2007) View-invariant object category learning, recognition, and search: How spatial and object attention are coordinated using surface-based attentional shrouds. Technical Report CAS/CNS-2007-011, Boston University.

Gao, E. & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences USA, 95*, 12663-12670.

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.

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. (1973) Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52*, 217-257.

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

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review. 87*, 1-51.

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. (1999a). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision, 12*, 163–187.

Grossberg, S. (1999b). The link between brain learning, attention, and consciousness. *Consciousness and Cognition, 8*, 1-44.

Grossberg, S. (2000a). How hallucinations may arise from brain mechanisms of learning, attention, and volition. *Invited article for the Journal of the International Neuropsychological Society, 6*, 579-588.

Grossberg, S. (2000b). The complementary brain: Unifying brain dynamics and modularity. T*rends in Cognitive Sciences, 4*, 233-246.

Grossberg, S. (2003a). Adaptive resonance theory. *The encyclopedia of cognitive science*. London : Macmillan Reference Ltd.

Grossberg, S. (2003b). How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behavioral and Cognitive Neuroscience Reviews, 2*, 47-76.

Grossberg, S. (2003c). Resonant neural dynamics of speech perception. *Journal of Phonetics, 31*, 423-445.

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

Grossberg, S., Mingolla, E. & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences, 20,* 106–111.

Grossberg, S. & Pearson, L.R. (2006). Laminar cortical dynamics of cognitive and motor working memory, sequence learning, and performance: Toward a unified theory of how the cerebral cortex works. Technical Report CAS/CNS TR-2006-002, Boston University. Submitted for publication.

Grossberg, S. & Raizada, R. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research. 40*, 1413–1432.

Grossberg, S. & Seidman, D. (2006). Neural dynamics of autistic behaviors: Cognitive, emotional, and timing substrates. *Psychological Review, 113*, 483-525.

Grossberg, S. & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks, 4*, 453–466.

Grossberg, S. & Swaminathan, G. (2004). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention and bistability. *Vision Research, 44*, 1147-1187.

Grossberg, S. & Versace, M. (2007). Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. Submitted for publication.

Grossberg, S. & Williamson, J.R. (2001). A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual grouping and learning. *Cerebral Cortex, 11*, 37-58.

Grossberg, S. & Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research, 45*, 1725-1743.

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.

Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience. 9*, 181-197.

Herrmann C.S., Munk, M.H.J. & Engel, A. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences, 8*, 347-355.

Krupa, D.J., Ghazanfar, A.A. & Nicolelis, M.A.L. (1999). Immediate thalamic sensory plasticity depends on corticothalamic feedback. *Proceedings of the National Academy of Sciences USA. 96*, 8200-8205.

Levy, W.B. & Steward, O. (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience, 8*, 791-797.

Lumer, E.D., Edelman, G.M. & Tononi, G. (1997). Neural dynamics in a model of the thalamocortical system I. Layers, loops and the emergence of fast synchronous rhythms. *Cerebral Cortex. 7*, 207-227.

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

Markram, H. Lubke, J., Frotscher, M. & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science, 275*, 213-215.

Maunsell, J.H.R., and Van Essen, D.C. (1983). Anatomical connections of the middle temporal visual area in the macaque monkey and their relationship to a hierarchy of cortical areas. The *Journal of Neuroscience*. 3, 2563–2586.

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

Otto, T. & Eichenbaum, H. (1992). Neuronal activity in the hippocampus during delayed non-match to sample performance in rats: Evidence for hippocampal processing in recognition memory. *Hippocampus, 2,* 323-334.

Pearson, L. R. & Grossberg, S. (2005). Neural dynamics of motor sequencing in lateral prefrontal cortex. *The Society of Neuroscience Abstract*. Washington, DC, November, 194.11.

Pearson, L. R. & Grossberg, S. (2006). Laminar cortical dynamics of cognitive and motor working memory. *The Society for Neuroscience Abstract.* Atlanta, GA, October, 610.7

Posner, M.I. & Keele, S.W. (1968). On the genesis of abstract ideas. J*ournal of Experimental Psychology, 77*, 353-363.

Raizada, R. & Grossberg, S. (2003) Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex, 13,*100-113.

Raizada, R.D.S. & Grossberg, S. (2003). Towards a Theory of the Laminar Architecture of Cerebral Cortex: Computational Clues from the Visual System. *Cerebral Cortex, 13*, 100-113.

Rockland, K.S., Andresen, J., Cowie, R.J. & Robinson, D.L. (1999). Single axon analysis of pulvinocortical connections to several visual areas in the macaque. *The Journal of comparative neurology, 406*, 221-50.

Roelfsema, P.R., Engel, A.K., König, P. & Singer, W. (1997). Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature, 385*, 157-161.

Roelfsema, P. R., Lamme, V. A. F. & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*, 376–381.

Rouiller, E.M. & Welker, E. (2000). A comparative analysis of the morphology of corticothalamic projections in mammals. *Brain Research Bulletin, 53*, 727-41.

Salin, P.A. & Bullier, J. (1995). Corticocortical connections in the visual system: structure and function. *Physiological reviews, 75*, 107-154.

Seitz, A., & Watanabe, T. (2003). Is subliminal learning really passive? *Nature, 422*, 36.

Shadlen, M. N. & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *The Journal of Neuroscience, 18,* 3870–3896.

Sherman, S.M. & Guillery, R. (2001). *Exploring the Thalamus*. San Diego: Academic Press.

Sherman, S.M. & Guillery, R.W. (2002). The role of the thalamus in the flow of information to the cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 357*, 1695-708.

Shipp, S. (2003). The functional logic of cortico-pulvinar connections. *Philosophical Transactions of the Royal Society of London, 358*, 1605-24.

Siegel, M., Körding, K.P. & König, P. (2000). Integrating top-down and bottom-up sensory processing by somato-dendritic interactions. *Journal of Computational Neuroscience, 8*, 161-173.

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.

Solokov, E.N. (1968). *Mechanisms of Memory*. Moscow University Press.

Thorpe, S., Fize, D. & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Traub, R.D., Spruston, N., Soltesz, I., Konnerth, A., Whittington, M.A. & Jefferys, G.R. (1998). Gamma-frequency oscillations: a neuronal population phenomenon, regulated by synaptic and intrinsic cellular processes, and inducing synaptic plasticity. *Progress in neurobiology, 55*, 563-75.

Tyler, C. W. & Kontsevich, L. L. (1995). Mechanisms of stereoscopic processing: stereo attention and surface perception in depth reconstruction. *Perception, 24*, 127-153.

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

van Der Werf, Y.D., Witter, M.P. & Groenewegen, H.J. (2002). The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain research Brain research reviews, 39*, 107-40.

van Vreeswijk, C. & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation, 10,* 1321–1371.

Versace, M. & Grossberg, S. (2005). Temporal binding and resonance in thalamocortical assemblies: learning and cognitive information processing in a spiking neuron model. *The Society for Neuroscience Abstract*, Washington, DC, November, 538.8

Versace, M. & Grossberg, S. (2006). From spikes to interareal synchrony: how attentive matching and resonance control learning and information processing by laminar thalamocortical circuits. *The Society for Neuroscience Abstract*, Atlanta, GA, 65.11.

Vinogradova, O.S. (1975). Functional organization of the limbic system in the process of registration of information: Facts and hypotheses. In R.L. Isaacson and K.H. Pribram (Eds.), *The hippocampus, Vol. 2* (p. 3-69). Plenum Press.

Ward, T.B. (1983). Response temp and separable-inte\-gral responding: Evidence for an integral-to-separable processing sequencing in visual perception. J*ournal of Experimental Psychology: Human Perception and Performance, 9*, 1029-1051.

Watanabe, T., Náñez, J., & Sasaki, Y. (2001). Perceptual learning without perception, *Nature*,413, 844-848.

Wespatat, V., Tennigkeit, F. & Singer, W. (2004). Phase sensitivity of synaptic modifications in oscillating cells of rat visual cortex. *The Journal of Neuroscience. 24*, 9067–9075.

Yazdanbakhsh, A. & Grossberg, S. (2004). Fast synchronization of perceptual grouping in laminar visual cortical circuits. *Neural Networks, 17*, 707-718.