How Does the Cerebral Cortex Work?
Development, Learning, Attention, and 3D Vision
by Laminar Circuits of Visual Cortex

Stephen Grossberg
Department of Cognitive and Neural Systems
and
Center for Adaptive Systems
Boston University
677 Beacon Street
Boston, MA 02215.
Phone: 617-353-7858
Fax: 617-353-7755
steve@bu.edu
http://www.cns.bu.edu/Profiles/Grossberg

Technical Report CAS/CNS TR-2003-005

January, 2003
Revised: March, 2003

Invited article for
Behavioral and Cognitive Neuroscience Reviews

Abstract

A key goal of behavioral and cognitive neuroscience is to link brain mechanisms to behavioral functions. The present article describes recent progress towards explaining how the visual cortex sees. Visual cortex, like many parts of perceptual and cognitive neocortex, is organized into six main layers of cells, as well as characteristic sub-lamina. Here it is proposed how these layered circuits help to realize processes of development, learning, perceptual grouping, attention, and 3D vision through a combination of bottom-up, horizontal, and top-down interactions. A key theme is that the mechanisms which enable development and learning to occur in a stable way imply properties of adult behavior. These results thus begin to unify three fields: infant cortical development, adult cortical neurophysiology and anatomy, and adult visual perception. The identified cortical mechanisms promise to generalize to explain how other perceptual and cognitive processes work.

Introduction

The advent of behavioral and cognitive neuroscience underscores the growing interest in mechanistically linking brain mechanisms to behavioral functions, or in explaining how a brain gives rise to a mind. Said in yet another way: How can the classical Mind/Body Problem be solved? Although there has been enormous experimental and theoretical progress on understanding brain *or* mind in the fields of neuroscience and psychology, respectively, establishing a mechanistic link between them has been very difficult, if only because these two levels of description often seem to be so different. Yet establishing such a linkage between brain and mind is crucial in any mature theory of how a brain or mind works. Without such a link, the mechanisms of the brain have no functional significance, and the functions of behavior have no mechanistic explanation.

In order to establish such a link with sufficient clarity for it to be scientifically predictive, rigorous models are needed. A rapidly growing number of models can now quantitatively simulate the neurophysiologically recorded dynamics of identified nerve cells in known anatomies *and* the behaviors that they control. Many predictions of these models have also been supported by subsequent experiments over the years. In this restricted sense, the Mind/Body Problem is at last starting to be understood. A particularly successful approach uses a theoretical method that has been systematically applied during the past thirty years (Grossberg, 1999c). In addition to leading to detailed models that quantitatively link brain and behavior, this method has led to the discovery of general computational principles and paradigms that represent a major shift away from earlier views of how the brain works.

Complementary Computing and Laminar Computing

Many scientists had earlier proposed that our brains possess independent modules, as in a digital computer. The brain's organization into distinct anatomical areas and processing streams supports the idea that brain processing is specialized, but that, in itself, does not imply that these streams contain independent modules. Independent modules should be able to fully compute

their particular processes on their own. Much behavioral data argue, however, against the existence of independent modules. For example, during visual perception, strong interactions are known to occur between perceptual qualities (Egusa, 1983; Faubert and von Grunau, 1995; Kanizsa, 1974; Pessoa, Beck, and Mingolla, 1996; Smallman and McKee, 1995). In particular, form and motion can interact, as can brightness and depth, among other combinations of qualities.

At least two new computational paradigms have gradually been identified from the cumulative experiences of modeling many aspects of brain and behavior over the past three decades: Complementary Computing and Laminar Computing (Grossberg, 1999a, 2000). *Complementary Computing* concerns the discovery that pairs of parallel cortical processing streams compute complementary properties in the brain. Each stream has complementary computational strengths and weaknesses, much as in physical principles like the Heisenberg Uncertainty Principle. Each cortical stream can also possess multiple processing stages. These stages realize a *hierarchical resolution of uncertainty*. "Uncertainty" here means that computing one set of properties at a given stage can suppress information about a complementary set of properties at that stage. The computational unit of brain processing that has behavioral significance is thus not a single processing stage, or any smaller entity such as the potential of a single cell or of a spike or burst of spikes. Instead, hierarchical interactions within a stream and parallel interactions between streams resolve their complementary deficiencies to compute complete information about a particular type of biological intelligence. These interactions have been used to clarify many of the data that do not support the hypothesis of independent modules. To model how the brain controls behavior, one thus needs to know how these complementary streams are organized with respect to one another.
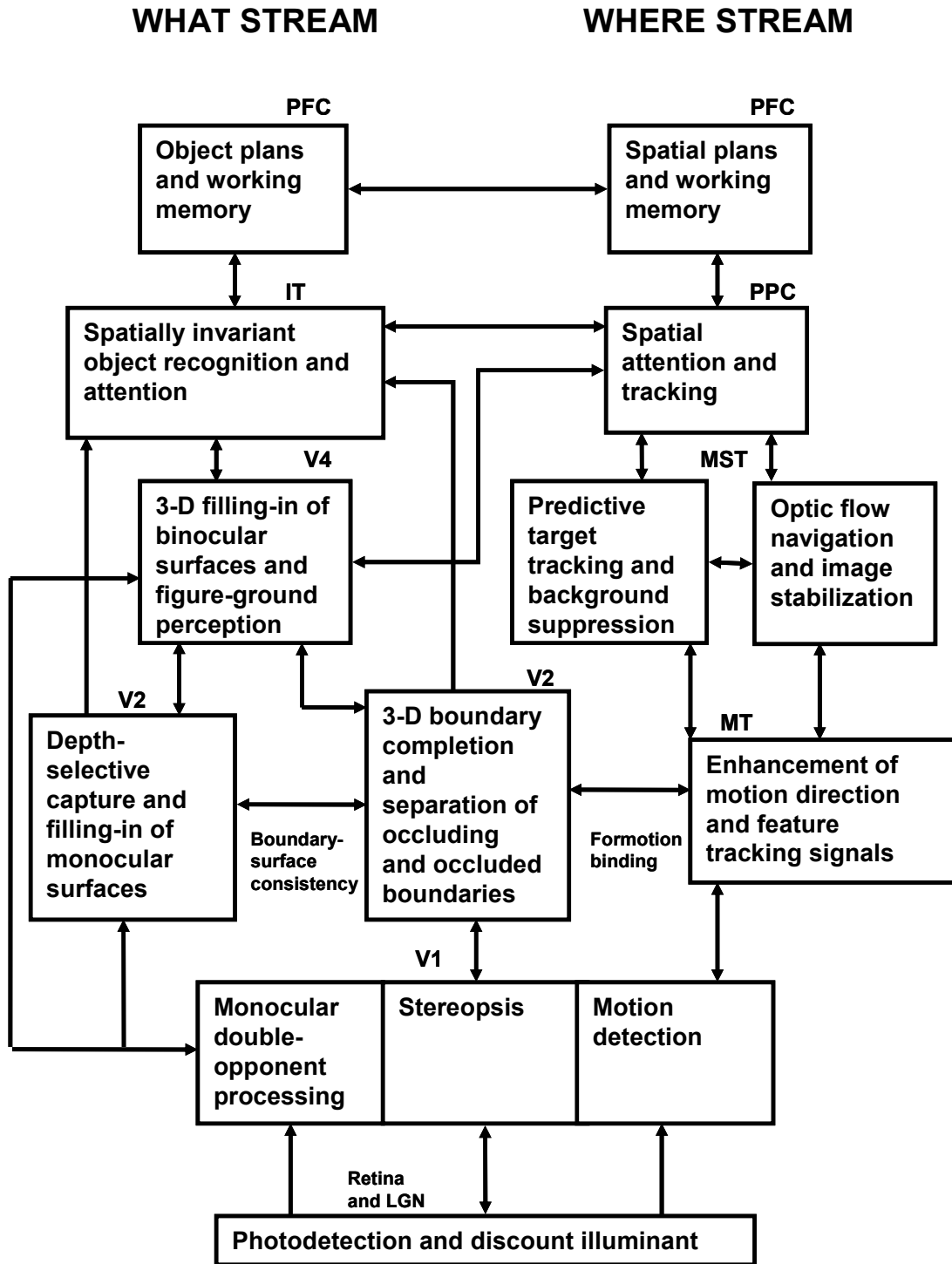
WHAT STREAM   **WHERE STREAM**

Figure 1. Some visual processes and their anatomical substates that are being modeled as part of a unified vision system. LGN = Lateral Geniculate Nucleus; V1 = striate visual cortex; V2, V4, MT, MST = prestriate visual cortex; IT = inferotemporal cortex; PPC = posterior parietal cortex; PFC = prefrontal cortex.

Understanding how the brain sees is one of the areas where experimental and modeling work have advanced the furthest, and this progress illustrates several different types of complementary interactions. Figure 1 provides a schematic macrocircuit of the types of processes that are being assembled into a unified theory of how the brain sees, including processes of vision, recognition, navigation, and cognition. In particular, key matching and learning processes within the What and Where cortical streams have been proposed to be complementary: The What stream, through cortical areas V1-V2-V4-IT-PFC, learns to recognize *what* objects and events occur. The Where stream, through cortical areas V1-MT-MST-PPC-PFC, spatially localizes *where* they are, and acts upon them. Complementary processes also occur within each stream: What stream boundary grouping via the (V1 interblob)-(V2 pale stripe)-V4 stages, and surface formation via the (V1 blob)-(V2 thin stripe)-V4 stages, have complementary properties. Where stream target tracking via MT-(MST ventral) and navigation via MT-(MST dorsal) have complementary properties. Such complementary processes are predicted to arise from symmetry-breaking operations during cortical development.

Laminar Computing concerns the fact that cerebral cortex, the seat of higher intelligence in all modalities, is organized into layered circuits (usually six main layers) which undergo characteristic bottom-up, top-down, and horizontal interactions (Brodmann, 1909; Martin, 1989). Differences in the thickness of these layers and the sizes and shapes of neurons led the German anatomist Korbinian Brodmann to identify more than fifty divisions, or areas, of neocortex. This classification has been invaluable as a basis for classifying distinct functions of different parts of neocortex. The functional utility of such a laminar organization in the control of behavior has, however, remained a mystery until recently. Understanding the functional uses of laminar computing should have an enormous payoff in understanding biological intelligence because, if one can understand how laminar circuits work in one part of the brain, then different intelligent capabilities may be expected to be understood as variations on a shared architecture theme.

Laminar Computing by Visual Cortex:

**Development, Learning, Grouping, and Attention**

Recent progress has clarified key properties of laminar computing through modeling aspects of how the laminar circuits of visual cortex are organized for seeing. This article will summarize some of this recent progress, as well as a number of general insights about cerebral cortex that flow from it, and directions for further research. Among the questions that will be treated are: How does the laminar organization of cortical circuits contribute to biological intelligence? What sorts of neural computations on the level of individual cells support visual perception as we know it? What is the link between processes of development in the infant and processes of perception and learning in the adult? What are the functional units of perception that determine perception in the adult, and how do developmental processes give rise to these units? What general design principles underlie the organization of neocortical circuits and systems?

A number of models have recently been proposed (Douglas *et al.,* 1995; Li, 1998; Stemmler *et al.,* 1995; Somers *et al.,* 1998; Yen and Finkel, 1998) to simulate aspects of visual cortical dynamics, but these models have not articulated why cortex has a laminar architecture. A different line of modeling work (Grossberg, 1999a; Grossberg and Howe, 2003; Grossberg, Mingolla, and Ross, 1997; Grossberg and Raizada, 2000; Grossberg and Williamson, 2001; Raizada and Grossberg, 2003) has suggested that the laminar organization of visual cortex accomplishes at least three things: (1) the developmental and learning processes whereby the cortex shapes its circuits to match environmental constraints in a *stable* way through time; (2) the binding process whereby cortex groups distributed data into coherent object representations that remain sensitive to analog properties of the environment; and (3) the attentional process whereby cortex selectively processes important events.

These results clarify that the visual cortex is not merely a bottom-up filtering device, as was proposed in the classical model of Hubel and Wiesel (1977). Instead, even early stages of visual cortex join together bottom-up filtering, horizontal grouping, and top-down attention. Perceptual grouping, the process that binds spatially distributed and incomplete information into 3D object

representations, starts at an early cortical stage; see Figure 2c. These grouping interactions are often cited as the basis of "non-classical" receptive fields that are sensitive to the context in which individual features are found (Bosking, Zhang, Schofield, and Fitzpatrick, 1997; Grosof, Shapley, and Hawken, 1993; Kapadia, Ito, Gilbert, and Westheimer, 1995; Knierim and van Essen, 1992; Peterhans and von der Heydt, 1989; Polat, Mizobe, Pettet, Kasamatsu, and Norcia, 1998; Sheth, Sharma, Rao, and Sur, 1996; von der Heydt, Peterhans, and Baumgartner, 1984; Sillito, Grieve, Jones, Cudeiro, and Davis, 1995). Likewise, even early visual processing is modulated by system goals via top-down expectations and attention (Motter, 1993; Roelfsema, Lamme, and Spekreijse, 1998; Sillito, Jones, Gerstein, and West, 1994; Somers, Dale, Seiffert, and Tootell, 1999; Watanabe, Sasaki, Nielsen, Takino, and Miyakawa, 1998). The laminar circuits provide an interface, called the *preattentive-attentive interface*, which exists between layers 6 and 4 (Figures 2b, 2c, and 2e) where data-driven bottom-up preattentive processing and task-directed top-down attentive processing are joined. Finally, mechanisms governing (1) in the infant lead to properties (2) and (3) in the adult, and properties (2) and (3) interact together intimately as a result. Thus, mechanisms that enable the cortex to develop and learn in a stable way *define* key properties of adult visual information processing, and there is no strict separation of attentive processes from pre-attentive processes such as perceptual grouping. A family of models that clarify these themes is called a LAMINART model (Figure 2) because it clarifies how mechanisms of Adaptive Resonance Theory, or ART, which have previously been predicted to occur in neocortex to help stabilize cortical development and learning (Grossberg, 1980, 1999c), are realized in identified laminar visual cortical circuits (Grossberg, 1999a).
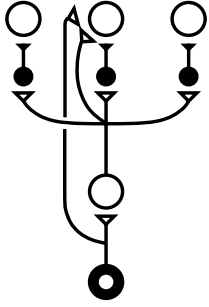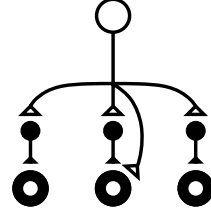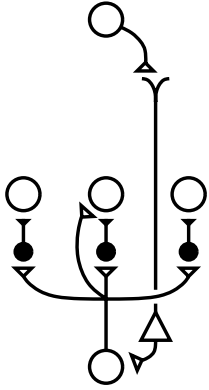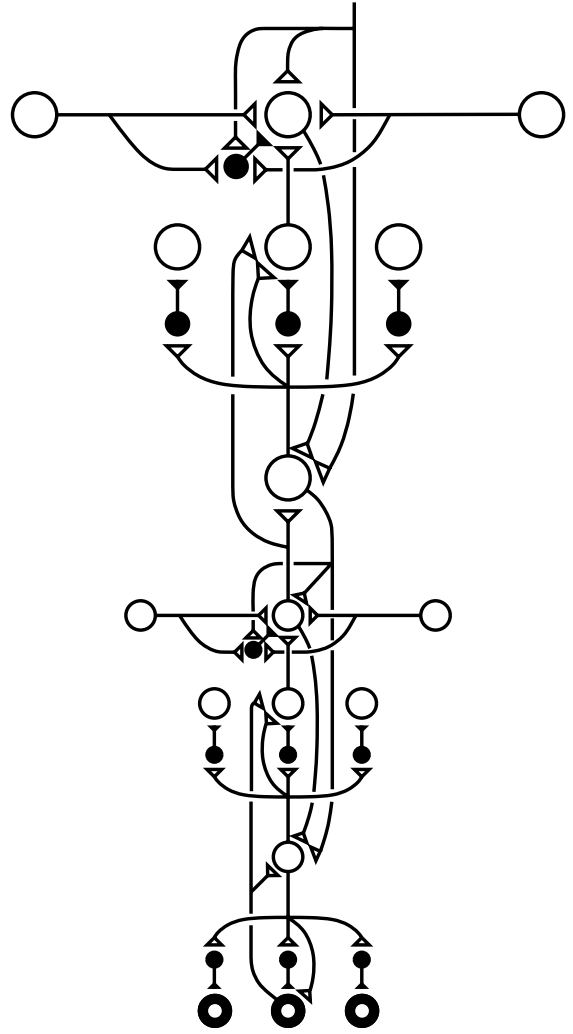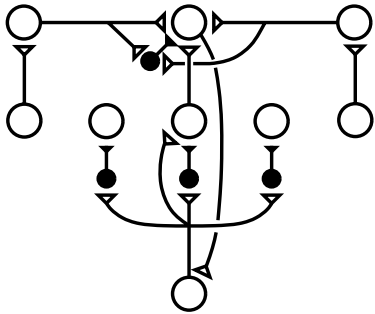
Figure 2. How known cortical connections join the layer 6 → 4 and layer 2/3 circuits to form an entire V1/V2 laminar model. Inhibitory interneurons are shown filled-in black. **(a)** The LGN provides bottom-up activation to layer 4 via two routes. First, it makes a strong connection directly into layer 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).)

Stable Development, Balanced Connections, Intermittent Spikes, and Synchrony

A number of other themes have also come to the fore through these discoveries. For one, the LAMINART model clarifies how excitatory and inhibitory connections in the cortex can develop in a stable way by achieving and maintaining a *balance* between excitation and inhibition (Grossberg and Williamson, 2001). It is believed that long-range excitatory horizontal connections between pyramidal cells in layer 2/3 of visual cortical areas play an important role in perceptual grouping (Hirsch and Gilbert, 1991; McGuire *et al*., 1991). The LAMINART model proposes how development enables the strength of long-range excitatory horizontal signals to become balanced against that of short-range disynaptic inhibitory signals which input to the same target pyramidal cells (Figure 1c). These balanced connections are proposed to realize properties of perceptual grouping in the adult. In a similar way, it is proposed that development enables the strength of excitatory connections from layer 6-to-4 to be balanced against those of inhibitory interneuronal connections (Wittmer, Dalva, and Katz, 1997); see Figures 2a and 2c. Thus the net excitatory effect of layer 6 on layer 4 is proposed to be modulatory. These approximately balanced excitatory and inhibitory connections exist within the on-center of an on-center off-surround network from layer 6-to-4. This network plays at least three functional roles that are intimately linked: maintaining a contrast-normalized response to bottom-up inputs at layer 4 (Figure 2a); forming perceptual groupings in layer 2/3 that maintain their sensitivity to analog properties of the world (Figure 2c); and biasing groupings via top-down attention from higher cortical areas (Figure 2b; also see Figure 2d).

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. Several model studies have shown how balanced excitation and inhibition can produce the highly variable interspike intervals that are found in cortical data (Shadlen and Newsome, 1998; van Vreeswijk and Sompolinsky, 1998). The LAMINART model proposes that such variability may reflect mechanisms that are needed to ensure stable development and learning by cortical circuits. Given that "stability implies variability," the cortex is faced with the difficult problem

that variable spikes are quite inefficient in driving responses from cortical neurons. On the other hand, when one analyses how these balanced excitatory and inhibitory connections work together to generate perceptual groupings and attentionally focused responses, it becomes clear that these particular circuits have the property of overcoming the inefficiency of intermittent spiking by preferentially responding to synchronized inputs; indeed they can rapidly resynchronize desynchronized signals (Grossberg and Grunewald, 1997; Grossberg and Somers, 1991). In fact, the article that introduced ART predicted a role for synchronous cortical processing, including synchronous oscillations, which were there called "order-preserving limit cycles", as part of the process of establishing resonant states (Grossberg, 1976). There is now a considerable amount of neurophysiological data pointing to the importance of synchronous processing in visual cortex, starting with the reports of Eckhorn, Bauer, Jordan, Brosch, Kruse, Munk, and Reitboeck (1988) and Gray and Singer (1989). The LAMINART model puts these data within a larger conceptual framework by predicting the existence of a functional link between properties of stable development, adult perceptual learning, perceptual grouping and attention, and synchronous cortical processing.

Attention, Competition, and Matching

When using a word as familiar as "attention," it is important to clarify what we mean and how it is proposed to work. LAMINART, and ART before it, predicted that an intimate link exists between processes of attention, competition, and bottom-up/top-down matching. LAMINART predicts, in particular, that top-town signals from higher cortical areas, such as area V2, can attentionally prime, or modulate, layer 4 cells in area V1 by activating the on-center off-surround network from layer 6-to-4 (Figures 2b and 2e). Because the excitatory and inhibitory signals in the on-center are balanced, attention can sensitize, or modulate, cells in the attentional on-center, without fully activating them, while also inhibiting cells in the off-surround.
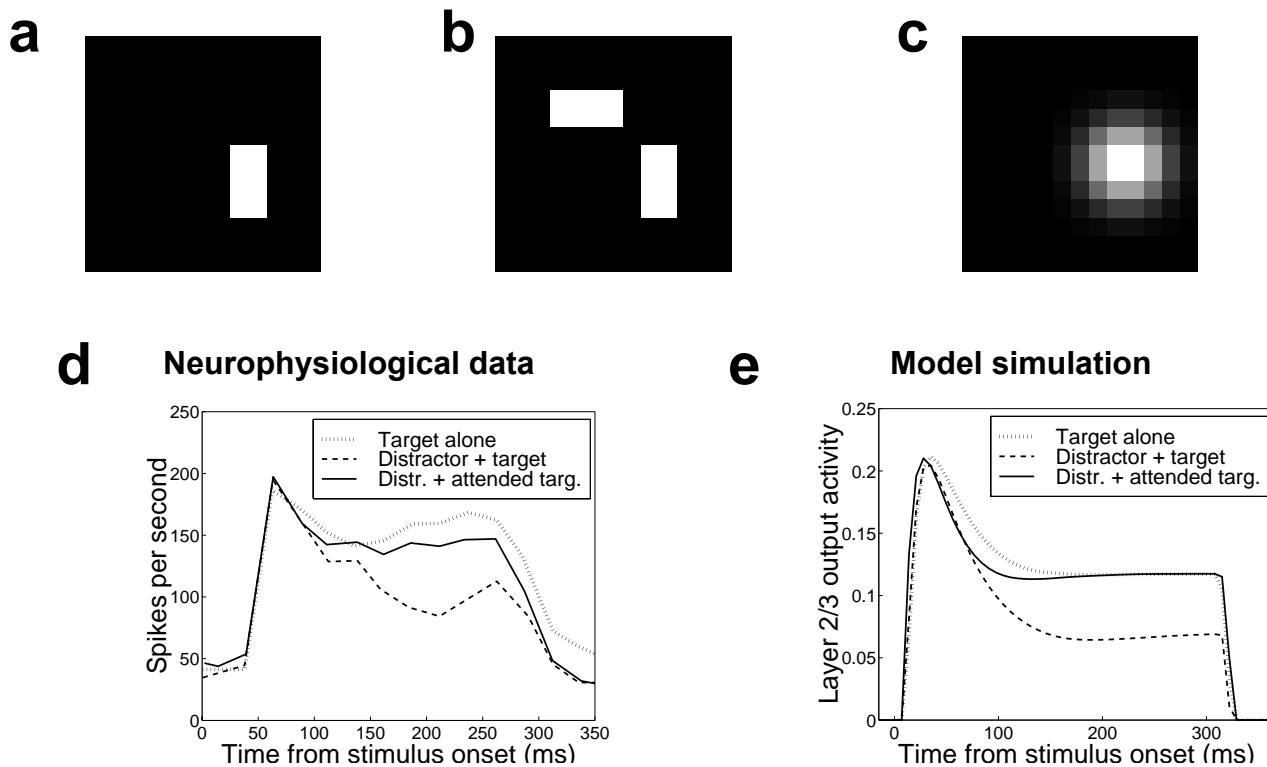
Figure 3. The effect of attention on competition between visual stimuli. A target stimulus, presented on its own (a), elicits strong neural activity at the recorded cell. When a second, distractor stimulus is presented nearby (b), it competes against the target, and activity is reduced. Directing spatial attention to the location of the target stimulus (c), protects the target from this competition, and restores neural activity to the levels elicited by the target on its own. The stimuli shown here, based on those used in the neurophysiological experiments of Reynolds et al. (1999), were presented to the model neural network. Spatial attention (c), was implemented as a Gaussian of activity fed back into layer 6. (d) Neurophysiological data from macaque V2 that illustrate the recorded activity patterns described above: strong responses to an isolated target (dotted line), weaker responses when a competing distractor is placed nearby (dashed line) and restored levels of activity when the target is attended (solid line). (Adapted with permission from Reynolds et al. (1999, Fig. 5).) (See also 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 Abstract*s, 21, 693.1). (e) Model simulation of the Reynolds et al. data. The time-courses illustrated show the activity of a vertically oriented cell stimulated by the target bar. If only the horizontal distractor bar were presented on its own, this cell would respond very weakly. If both target and distractor were presented, but with the horizontal distractor attended, the cell would respond, but more weakly than the illustrated case where the distractor and target are presented together, with neither attended. (Reprinted with permission from Grossberg and Raizada (2000).)

12

The importance of the conclusion that top-down attention is often expressed through a top-down modulatory on-center off-surround network cannot be overstated. Because of this organization, top-down attention can typically provide only excitatory modulation to cells in the on-center, while it can strongly inhibit cells in the off-surround. As Hupé *et al.* (1997, p. 1031) have noted: "feedback connections from area V2 modulate but do not create center-surround interactions in V1 neurons." When the top-down on-center matches bottom-up signals, it can amplify and synchronize them, while still suppressing mismatched signals in the off-surround. This prediction was first made as part of ART in the 1970's (Grossberg, 1976, 1978, 1980, 1999a, 1999c); see below for further discussion. The prediction that top-down attention has an on-center off-surround characteristic has since received a considerable amount of psychological and neurobiological empirical confirmation in the visual system (Bullier, Hupé, James, and Girard, 1996; Caputo and Guerra, 1998; Downing, 1988; Mounts, 2000; Reynolds, Chelazzi, and Desimone, 1999; Smith, Singh, and Greenlee, 2000; Somers, Dale, Seiffert, and Tootell, 1999; Sillito, Jones, Gerstein, and West, 1994; Steinman, Steinman, and Lehmkuhle, 1995; Vanduffell, Tootell, and Orban, 2000). Based on such data, this conclusion has recently been restated, albeit without a precise anatomical realization, in terms of the concept of "biased competition" (Desimone, 1998; Kastner and Ungerleider, 2001), in which attention biases the competitive influences within the network. Figure 3 summarizes data of Reynolds, Chelazzi, and Desimone (1999) and a simulation of these data from Grossberg and Raizada (2000) that illustrate the on-center off-surround character of attention in macaque V2.

Global Organization of Attention: Divided, Object/Spatial, Noun/Verb

It was noted above that bottom-up inputs arriving in the off-surround of an active top-down attentional prime may be suppressed. This property does not imply that inputs outside the off-surround will be suppressed. This is already clear in some conditions of the Reynolds, Chelazzi, and Desimone (1999) experiment that is summarized in Figure 3. Rather, this description shows how a particular attentional prime may act *if* it is deployed, and *if* bottom-up inputs fall within its domain of influence. Further analysis is required to understand the global allocation of attention throughout the brain in a task-sensitive manner.

In particular, many studies have shown that attention may be simultaneously divided among several targets (Pylyshyn and Storm, 1988; Yantis, 1992), and that object and spatial attention may both influence visual perception (Duncan, 1984; Posner, 1980). The distinction between object and spatial attention reflects the organization of visual cortex into parallel What and Where processing streams (Figure 1). Many cognitive neuroscience experiments have supported the hypotheses of Ungerleider and Mishkin (1982; see also Mishkin, Ungerleider, and Macko (1983)) and of Goodale and Milner (1992) that inferotemporal cortex and its cortical projections learn to categorize and recognize what objects are in the world, whereas the parietal cortex and its cortical projections learn to determine where they are and how to deal with them by locating them in space, tracking them through time, and directing actions towards them. This design into parallel streams thus separates sensory and cognitive processing from spatial and motor processing. The existence of parallel streams does not, however, imply that these streams do not interact. The very fact that the What stream strives to generate object representations that are independent of their spatial coordinates, whereas the Where stream generates representations of object location and tracking, indicates that the streams must interact to act upon recognized objects. Indeed, both object and spatial attention are needed to search for visual targets amid distractors. In this regard, Grossberg, Mingolla, and Ross (1994) quantitatively fit a a large human psychophysical database about visual search with a model, called the Spatial Object Search (SOS) model, that proposes how 3D boundary groupings and surface representations interact with object attention and spatial attention to find targets amid distractors. By clarifying how attention can be selectively allocated to desired targets, the SOS model also indicates how actions can be deployed towards them. This analysis needed to be sensitive to how object and spatial attention may respond to perceptual groupings and to surface properties, such as all occurrences of a color on a prescribed depth plane (Egeth, Virzi, and, Garbart, 1984; Nakayama and Silverman, 1986; Wolfe and Freedman-Hill, 1992).

Another aspect of the global control of attention concerns how it is allocated to objects and actions. In particular, visual form processing is elaborated within the What stream, whereas motion processing is elaborated within the Where stream (Maunsell and Van Essen, 1983). Form processing subserves the identification of objects in the world, whereas motion processing is

needed to identify many actions in the world. Indeed, early language development of a child often takes the form of simple noun and verb phrases that are learned in concert with visual examples of objects and their actions by a supervising parent. How attention across the What and Where streams is allocated to objects and actions, and how these events become a key substrate of spoken language, also requires a further analysis of global attentional control.

The present article focuses on the micro-architecture of attention but is consistent with, and illuminates, how attention may be globally organized by many brain regions acting together, as illustrated in Figure 1, particularly when keeping in mind how attention can leap between cortical areas via their layers 6, as illustrated in Figure 2e.

The Preattentive-Attentive Interface and Object-Based Attention

The manner in which top-down attention and pre-attentive perceptual grouping are interfaced within the cortical layers enables attention to focus on an entire object boundary, thereby not only influencing what objects are selectively attended, but also what groupings may be perceived. This is true, in particular, because the same layer 6-to-4 competition, or selection, circuit may be activated by pre-attentive grouping cells in layer 2/3, as well as by top-down attentional pathways (Figures 2b and 2c). Layer 4 cells can then, in turn, reactivate layer 2/3 cells (Figure 2c). This layer 6-to-4 circuit "folds" the feedback from top-down attention or a layer 2/3 grouping back into the feedforward flow of bottom-up inputs to layer 4. It is thus said to embody a "folded feedback" process. When ambiguous and complex scenes are being processed, intracortical folded feedback enables stronger groupings that are starting to form in layer 2/3 to inhibit weaker groupings, whereas intercortical folded feedback enables higher-order processing constraints to bias which groupings will be selected.

Figure 2e summarizes the hypothesis that top-down attentional signals to layer 1 may also directly modulate groupings via apical dendrites of both excitatory and inhibitory layer 2/3 cells in layer 1 (Lund and Wu, 1997; Rockland, and Virga, 1989). By activating both excitatory and

inhibitory cells in layer 2/3, the inhibitory cells may balance the excitatory cell activation, thereby creating a net modulatory response of grouping cells in layer 2/3.
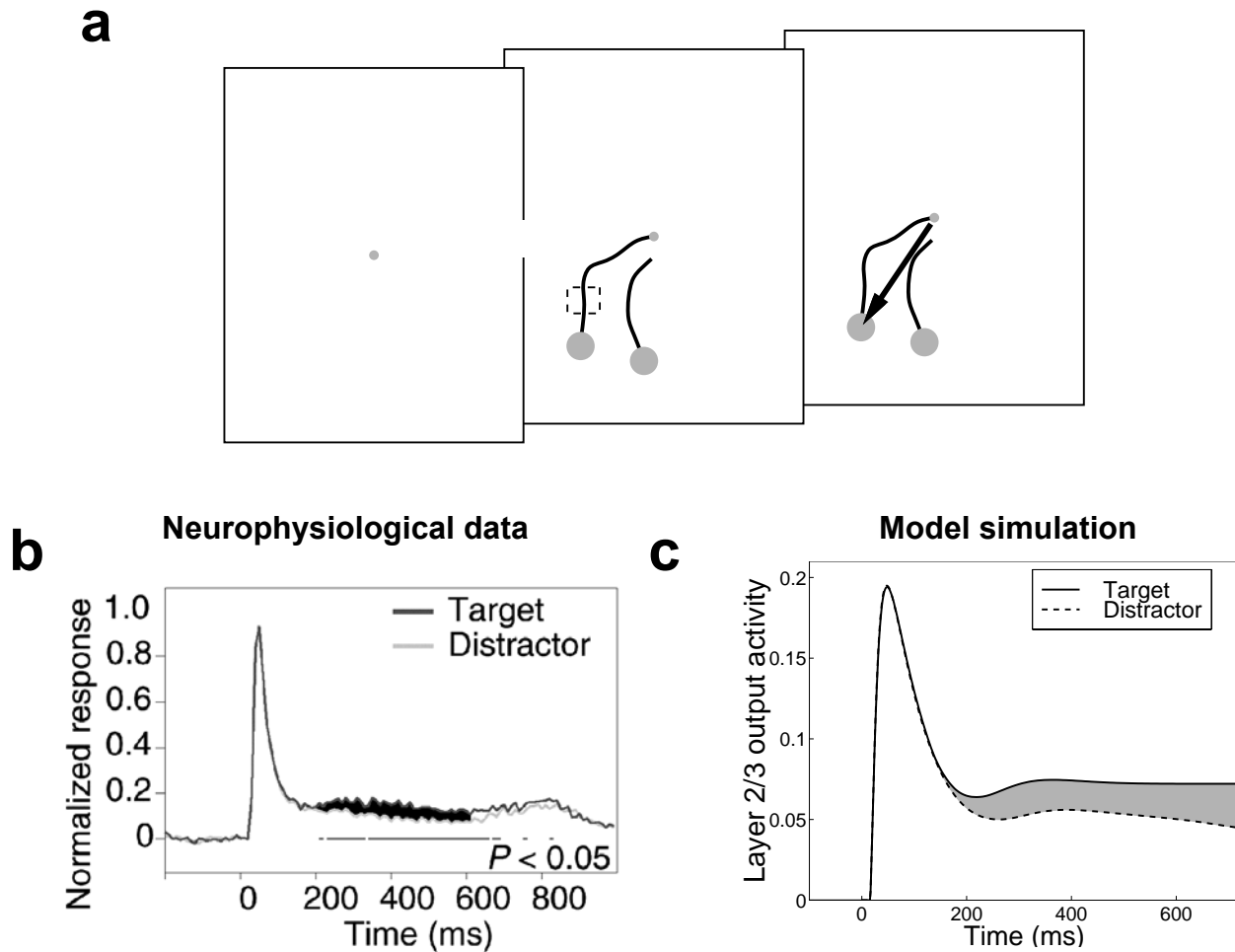

Figure 4. Spread of visual attention along an object boundary grouping, from an experiment by Roelfsema et al. (1998). (a) The experimental paradigm. Macaque monkeys performed a curve-tracing task, during which physiological recordings were made in V1. A fixation spot was presented for 300 ms, followed by a target curve and a distractor curve presented simultaneously. The target was connected at one end to the fixation point. While maintaining fixation, the monkeys had to trace the target curve, then, after 600 ms, make a saccade to its endpoint. (b) Neurophysiological data showing attentional enhancement of the firing of a neuron when its receptive field (RF) lay on the target curve, as opposed to the distractor. Note that the enhancement occurs about 200 ms after the initial burst of activity. Further studies have indicated that the enhancement starts later in distal curve segments, far from the fixation point, than it does in proximal segments, closer to fixation (Pieter Roelfsema, personal communication). This suggests that attentional signals propagate along the length of the target curve. (Figures (a) and (b) adapted with permission from Roelfsema et al. (1998).) (c) Model simulation of the Roelfsema et al. data. (Reprinted with permission from Grossberg and Raizada (2000).)

Because the cortex uses the same circuits to select groupings and to prime attention, attention can flow along perceptual groupings that define a discrete object (Roelfsema, Lamme, and Spekreijse, 1998). In particular, when attention causes an excitatory modulatory bias at some cells in layer 4, groupings that form in layer 2/3 can be enhanced by this modulation via their positive feedback loops from 2/3-to-6-to-4-to-2/3. The direct modulation of layer 2/3 by attention can also enhance these groupings. As a result, both infants and adults can focus their attention selectively upon whole objects. Figure 4 summarizes a LAMINART simulation of data from Roelfsema et al. (1998) of the spread of visual attention along an object boundary grouping. LAMINART has also been used to simulate the flow of attention along an illusory contour (Raizada and Grossberg, 2001), consistent with experimental data of Moore, Yantis, and Vaughan (1998). The ability of attention to selectively light up entire object representations has an obviously important survival value in adults. It is thus of particular interest that the intracortical and intercortical feedback circuits that control this property have been shown in modeling studies to play a key role in stabilizing infant development and adult perceptual learning within multiple cortical areas, including cortical areas V1 and V2.

Stable Development and Learning through Adaptive Resonance

Adaptive Resonance Theory, or ART (Engel, Fries, Singer, 2001; Grossberg, 1980, 1995, 1999b; Pollen, 1999) is a cognitive and neural theory which addresses a general problem that faces all adaptive brain processes; namely, the *stability-plasticity dilemma*: how can brain circuits be plastic enough to be rapidly fine-tuned by new experiences, and yet simultaneously stable enough that they do not get catastrophically overwritten by the new stimuli with which they are continually bombarded?

The solution that ART proposes to this problem is to allow neural representations to be modified only by those incoming stimuli with which they form a sufficiently close match. If the match is close enough, then learning occurs. Precisely because the match is sufficiently close, such learning fine-tunes the memories of existing representations. In this way, outliers cannot cause a

radical overwriting of an already learned representation. ART proposes how a learning individual can flexibly vary the criterion of how good a match is needed between bottom-up and top-down information in order for the presently active representation to be refined through learning. When coarse matches are allowed, the learned representations are capable of representing more general and abstract information. When only fine matches are allowed, the representations are more specific and concrete. If the active neural representation does not match with the incoming stimulus, then its neural activity is extinguished and hence unable to cause plastic changes. Suppression of an active representation enables a memory search to ensue whereby some other representation can become active instead through bottom-up signalling. This representation, in turn, reads out top-down signals that either gives rise to a match, thereby allowing learning, or a non-match, causing the search process to repeat until either a match is found or the incoming stimulus causes a totally new representation to be formed.

The connection with the LAMINART model of top-down attention is as follows: A key mechanism that implements the matching process is top-down attentional feedback directed to behaviorally relevant sensory stimuli. ART predicted that these top-down attentional matching signals are carried by a modulatory on-center off-surround network (e.g., Grossberg, 1980, 1982, 1999b), whose role is to select and enhance behaviorally relevant bottom-up sensory inputs (match), and suppress those that are irrelevant (non-match). Mutual excitation between the top-down feedback and the bottom-up signals that they match can amplify, synchronize, and maintain existing neural activity long enough for synaptic changes to occur. Thus, attentionally relevant stimuli are learned, while irrelevant stimuli are suppressed and hence prevented from destabilizing existing representations. See Grossberg (1999c) for a more detailed account.

The folded feedback layer 6-to-4 modulatory on-center off-surround attentional pathway in the LAMINART model (Figure 2b) can be thought of as an implementation of ART matching in cortical laminar circuitry. The claim that bottom-up sensory activity is enhanced when matched by top-down signals is in accord with an extensive neurophysiological literature showing the facilitatory effect of attentional feedback (Luck et al., 1997; Roelfsema et al., 1998; Sillito et al., 1994), but not with models in which matches with top-down feedback cause suppression (Mumford, 1992; Rao and Ballard, 1999). The ART proposal raises two key questions: First,

does top-down cortical feedback have the predicted modulatory on-center off-surround structure? Second, is there evidence that top-down feedback controls plasticity in the area to which it is directed?

Experimental evidence was summarized above in support of the modulatory on-center off-surround structure of top-down cortical attention in the visual system. ART predicts, in addition, that on-center off-surround attentional feedback should exist in all sensory and cognitive systems that are capable of stable on-line learning. Nobuo Suga and colleagues have shown that feedback from auditory cortex to the medial geniculate nucleus (MGN) and the inferior colliculus (IC) also has an on-center off-surround form (Zhang, Suga, and Yan, 1997) and Temereanca and Simons (2001) have produced evidence for a similar feedback architecture in the rodent barrel system.

The Link between Attention and Learning

A more stringent test of the ART claim is that top-down feedback should control plasticity. Psychophysically, the role of attention in controlling adult plasticity and perceptual learning was demonstrated by Ahissar and Hochstein (1993). Gao and Suga (1998) reported physiological evidence that acoustic stimuli caused plastic changes in the inferior colliculus (IC) of bats only when the IC received top-down feedback from auditory cortex. These authors also reported that plasticity is enhanced when the auditory stimuli were made behaviorally relevant, consistent with the ART proposal that top-down feedback allows attended, and thus relevant, stimuli to be learned, while suppressing unattended irrelevant ones. Evidence that cortical feedback also controls thalamic plasticity in the somatosensory system has been found by Krupa, Ghazanfar, and Nicolelis (1999) and by Parker and Dostrovsky (1999). These findings are reviewed by Kaas (1999).

Models of intracortical grouping-activated feedback and intercortical attention-activated feedback have shown that either type of feedback can rapidly synchronize the firing patterns of higher and lower cortical areas (Grossberg and Grunewald, 1997; Grossberg and Somers, 1991). ART puts this result into perspective by suggesting that resonance may lead to synchronization,

which may, in turn, trigger cortical learning by enhancing the probability that "cells that fire together wire together." An excellent recent discussion of top-down cortical feedback, synchrony, and their possible relations to the ART model is given by Engel, Fries, and Singer (2001).

Learning without Attention: The Pre-Attentive Grouping is its Own Attentional Prime

The hypothesis that attentional feedback exerts a controlling influence over plasticity in sensory cortex does not imply that unattended stimuli can never be learned. Indeed, it is clear that plasticity must be allowed to take place during early development, before top-down attention has even come into being. Grossberg (1999a) noted that, were this not possible, an infinite regress could be created, since a lower cortical level like V1 might then not be able to stably develop unless it received attentional feedback from V2, but V2 itself could not develop unless it had received reliable bottom-up signals from V1. How does the cortex avoid this infinite regress, so that during development, plastic changes in cortex may be driven by stimuli that occur with high statistical regularity in the environment (Grossberg and Williamson, 2001), without causing massive intability? How does this process continue to fine-tune sensory representations in adulthood even in cases where attention may not be explicitly allocated (Seitz and Watanabe, 2003; Watanabe, Nanez, and Sasaki, 2001)?

The subtlety of this issue is compounded by the fact that there is experimental and even mathematical support for the ART prediction that top-down attention plays a matching role which helps to control cortical plasticity, yet, in addition to the infinite regress problem mentioned above, it is also necessary to explain data about adult perception which, at the outset, seem to conflict with this prediction. In particular, how can pre-attentive groupings, such as illusory contours, form over positions that receive no bottom-up inputs? Illusory contours seem to contradict the ART hypothesis that feedback effects should merely be modulatory. How, then, can the cells that represent the illusory contour fire without bottom-up inputs? One cannot avoid this problem by claiming that illusory contour formation does not involve a process that includes learning, since many data show that the formation of horizontal connections can be influenced by

visual experience throughout life (see Grossberg and Williamson (2001) for a review). How, then, can we see illusory contours without destabilizing cortical development and learning?

As was described above, the ART matching rule has three aspects: first, incoming sensory signals that receive matching top-down excitatory feedback should be enhanced; second, non-matching inputs that do not receive excitatory feedback should be suppressed; and third, top-down feedback on its own should be only modulatory; that is, unable to produce above-threshold activity in the lower area in the absence of incoming bottom-up signals. The conceptual challenge is this: If ART matching is needed to stabilize cortical development and learning, and if ART matching requires that suprathreshold activation can occur only where there are bottom-up inputs, then does not the existence of illusory contours contradict the ART matching rule, since such groupings form over positions that receive no bottom-up inputs, and yet do not seem to destabilize cortical development or learning? If the brain had not solved this problem, anyone could roam through the streets of a city and destabilize pedestrians' visual systems simply by showing them images of Kanizsa squares! The absurdity of this possibility indicates how fundamental the issue at hand really is.

The LAMINART model proposes that the brain uses its laminar circuits to solve this problem in an ingenious, parsimonious, and simple way. Here is where the laminar cortical solution of the "preattentive-attentive interface" problem plays a key role: Both *inter*cortical attentional feedback and *intra*cortical grouping feedback share the same selection circuit from layer 6-to-4. In particular, when a horizontal grouping starts to form in layer 2/3, it also activates the intracortical feedback pathway from layer 2/3-to-6, which activates the modulatory on-center off-surround network from layer 6-to-4. This feedback pathway helps to select which cells will remain active to participate in a winning grouping. But this is the same network that ART requires attention to use when it stabilizes cortical development and learning through intercortical interactions. In other words, the layer 6-to-4 selection circuit, which in the adult helps to choose winning groupings, also helps to assure in the developing brain that the ART matching rule holds at every position along a grouping. Because the matching rule holds, only the correct combinations of cells can "fire together and wire together," and hence stability is achieved. *Intra*cortical feedback via layers 2/3-to-6-to-4-to-2/3 can realize this selection process

even before *inter*cortical attentional feedback can develop. This property is sometimes summarized with the phrase: "The pre-attentive grouping is its own attentional prime" (Grossberg, 1999a).

In summary, by joining together bottom-up (interlaminar) adaptive filtering, horizontal (intralaminar) grouping, top-down intracortical (but interlaminar) pre-attentive feedback, and top-down intercortical (and interlaminar) attentive feedback, the LAMINART model shows how some developmental and learning processes can occur without top-down attention, by using intracortical feedback processes that computationally realize the same stabilizing effects that top-down intercortical attentional processes were earlier predicted to realize. Because of this intimate link between intracortical and intercortical feedback processes, attention can modulate how pre-attentive grouping processes unfold, as in the case of the Roelfsema et al. (1998) data.

A Balancing Act between Excitatory and Inhibitory Circuits

It was noted above that, within key cortical circuits that realize these grouping and attentional processes, there seems to be a balance between excitatory and inhibitory interactions. The balance within layer 2/3 circuits is proposed to help achieve horizontal perceptual grouping. The balance between excitatory and inhibitory interactions within the on-center of the network from layer 6-to-4 helps to do several things, among them allow top-down attention to be modulatory. The schematic circuits in Figure 2 show only the above two types of balanced excitatory and inhibitory circuits. There are actually more types of interaction between excitation and inhibition in the cortex, but these are not displayed in Figure 2. One such interaction gives rise to monocular simple cell receptive fields in layer 4; see Liu, Gaska, Jacobson, and Pollen (1992), Palmer and Davis (1981), and Pollen and Ronner (1981) for relevant data and Olson and Grossberg (1998) for a modeling study. The next section proposes how another type of excitatory/inhibitory interaction within layer 3B gives rise to binocular simple cells that match monocular inputs from different eyes.
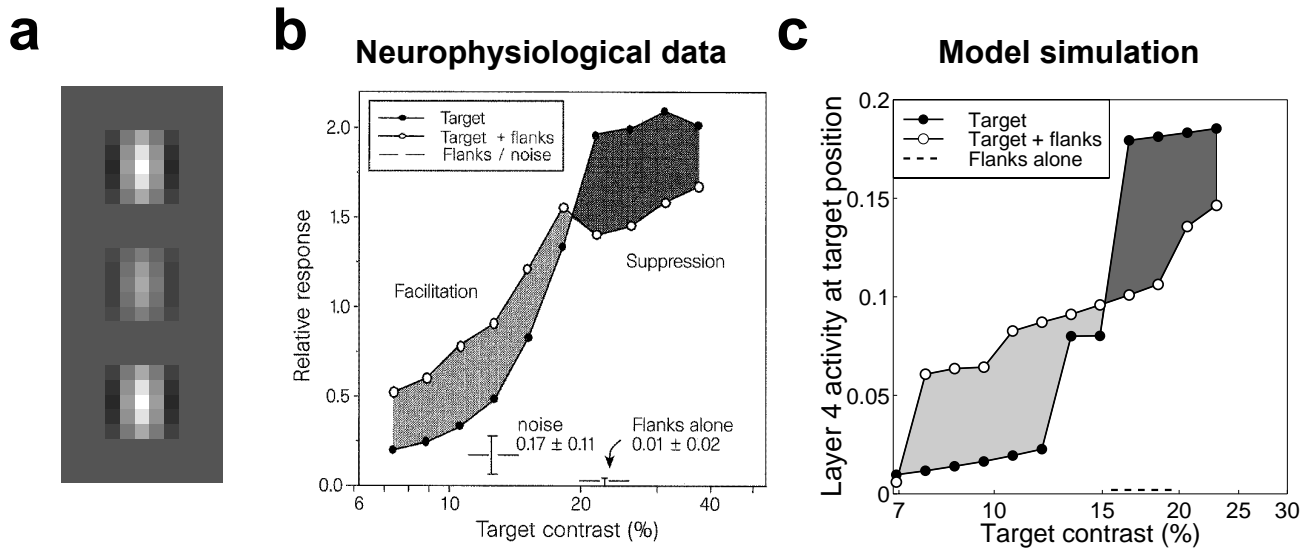
Figure 5. Contrast-dependent perceptual grouping in primary visual cortex. (a) Illustrative visual stimuli. A variable-contrast oriented Gabor patch stimulates the classical receptive field (CRF), with collinear flanking Gabors of fixed high-contrast outside of the CRF. The stimulus shown here, based on those used Polat et al. (1998), was presented to the model neural network. (b) Neural responses recorded from cat V1. The colinear flankers have a net facilitatory effect on weak targets which are close to the contrast-threshold of the cell, but they act to suppress responses to stronger, above-threshold targets. When the flankers are presented on their own, with no target present, the neural response stays at baseline levels. (Reproduced with permission from Polat et al. (1998).) (c) Model simulation of the Polat et al. data. (Reprinted with permission from Grossberg and Raizada (2000).)

Even interactions between the circuits within layer 2/3 and from layer 6-to-4, as in Figure 2, allows us to analyse and explain quite a bit of additional data, notably data in which the balance between excitatory and inhibitory interactions is altered due to different combinations of sensory inputs. A striking example of such a change in balance is illustrated in Figure 5, which summarizes data of Polat, Mizobe, Pettet, Kasamatsu, and Norcia (1998) on contrast-dependent perceptual grouping in primary visual cortex, and a simulation of these data from Grossberg and Raizada (2000). Roughly speaking, the excitatory effects that enable co-linear flankers to facilitate activation in response to a low-contrast target are mediated by layer 2/3 interactions, and the inhibitory effects that cause co-linear flankers to depress activation in response to a high-contrast target are mediated by the layer 6-to-4 off-surround. These two types of effects propagate throughout the network via layer 4-to-2/3 and layer 2/3-to-6 interactions, among

others. The fact that flanking stimuli alone yield no response could be due to either of two possibilities: Either the flanking stimuli occur outside the region where horizontal facilitation of the target V1 cell can reach, or the cell can only be modulated by horizontal interactions, but not fired by them, in the absence of bottom-up inputs.

In cortical area V2 of monkeys, it is known through the classical experiments of von der Heydt, Peterhans, and Baumgartner (1984) and subsequent experiments from this laboratory (e.g., Peterhans and von der Heydt (1989)) that approximately co-linear horizontal interactions from approximately co-oriented cells are capable of firing a target cell that does not receive bottom-up inputs. This capability is believed to be the basis for the brain's ability to form illusory contours in response to artificial stimuli such as the Kanizsa square, as well as in response to ecologically important scenic cues such as shading, texture, and depth cues. The von der Heydt et al. (1984) experiment confirmed a prediction of Grossberg and colleagues (Cohen and Grossberg, 1984; Grossberg, 1984; Grossberg and Mingolla, 1985a, 1985b) that perceptual grouping obeys a *bipole property*; namely, such a cell can fire if it gets approximately co-linear horizontal inputs from approximately co-oriented cells on both sides of its receptive field, even if it does not receive bottom-up input; or it can fire in response to bottom-up input alone, or bottom-up input plus any combination of horizontal signals. A receptive field structure for bipole cells was also predicted that was supported by later psychophysical experiments; e.g., Field, Hayes, and Hess (1993) and Kellman and Shipley (1991). The LAMINART model (Grossberg, 1999a; Grossberg, Mingolla, and Ross, 1997) extended this analysis by predicting how the bipole property may be realized by balanced excitatory and inhibitory interactions within layer 2/3, as summarized in Figure 6. Without these balanced inhibitory interactions, the growth of horizontal connections during development could proliferate uncontrollably if inhibition is too weak, or could be suppressed entirely if inhibition is too strong; see Grossberg and Williamson (2001) for model simulations.

It may be, however, that such a "firing" bipole property does not exist in cortical area V1. It is well-known from several experiments that co-linear horizontal interactions from co-oriented cells can enhance, or modulate, the activation of a V1 cell that also receives bottom-up inputs (e.g., Crook, Engelmann, and Lowel, 2002; Polat et al., 1998), but there seem to be few direct

experimental tests of whether cells can fire in response to horizontal interactions when they receive no bottom-up inputs. From a modeling viewpoint, changing from a firing bipole property to a modulatory bipole property might involve no more than a change in the relative strength of horizontal excitatory and disynaptic inhibitory interactions. It is nonetheless conceptually important to ascertain whether true perceptual grouping can occur in V1, or whether modulatory interactions exist that, for example, can increase the signal-to-noise ratio of cells which lie on perceptually salient contours without actually completing illusory contours across regions that receive no bottom-up inputs. It should also be noted that, when modulatory bipole interactions occur, then both the layer 2/3 inhibitory interneurons and the layer 6-to-4 inhibitory interneurons may both have a net inhibitory effect on target cells that receive no direct bottom-up inputs.

Data relevant to this issue have recently been collected by several labs. Ramsden, Hung, and Roe (2001) have done a study that compares V1 cortical responses to real contours and to illusory contours that are generated by offset gratings, using both optical imaging and single unit electrophysiology. They concluded that, in response to an offset grating illusory contour stimulus, cell responses are negatively signaled, or de-emphasized, in V1, whereas they are enhanced in V2, as previously reported. This result led these authors to propose that the differences in V1 and V2 responses to illusory and real contours might be a cue whereby the brain can tell the difference between a real and an illusory contour.

There is a simpler explanation of these data if the V1 cells under study can merely modulate, rather than fire vigorously, in response to purely horizontal signals. In particular, at the ends of lines in the offset grating, there will be strong inhibition that is due to a combination of layer 2/3 inhibitory interneurons and the inhibitory interneurons of the 6-to-4 off-surround, among other inhibitory circuitry. If there is no active boundary completion due to a firing bipole property within V1, then such inhibitory effects may not be offset by excitatory boundary completion signals, so that inhibition can dominate cell responses at positions within the gaps in the offset grating.
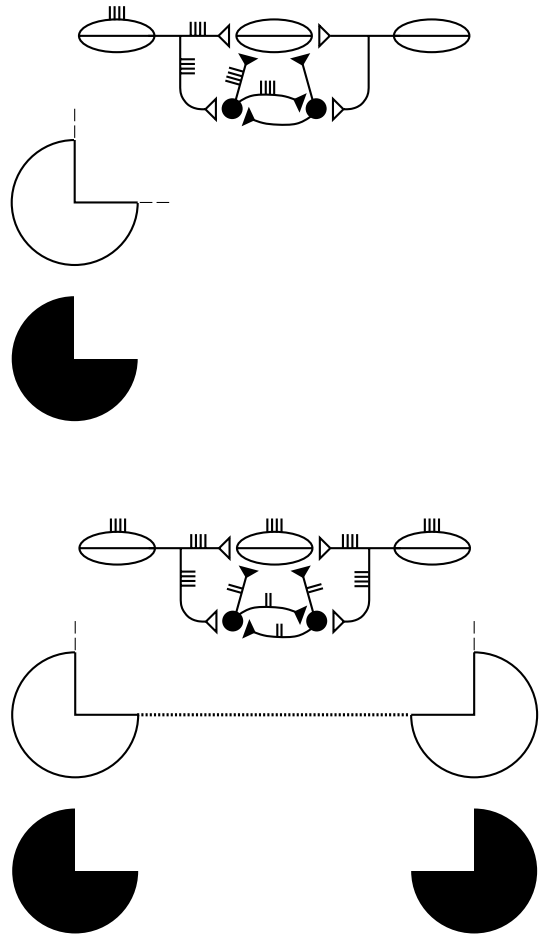
Figure 6. Schematic of the boundary grouping circuit in layer 2/3. Pyramidal cells with colinear, coaxial receptive fields (shown as ovals) excite each other via long-range horizontal axons (Bosking et al., 1997; Schmidt et al., 1997), which also give rise to short-range, disynaptic inhibition via pools of interneurons, shown filled-in black (McGuire et al., 1991). This balance of excitation and inhibition helps to implement the *bipole property*. (a) Illustration of how horizontal input coming in from just one side is insufficient to cause above-threshold excitation in a pyramidal cell (henceforth referred to as the target) whose receptive field does not itself receive any bottom-up input. The inducing stimulus (e.g. a Kanizsa 'pacman') excites the oriented receptive fields of layer 2/3 cells, which send out long-range horizontal excitation onto the target pyramidal. This excitation brings with it a commensurate amount of disynaptic inhibition. This balance of "one-against-one" prevents the target pyramidal cell from being excited above-threshold. The boundary representation of the solitary pacman inducer produces only weak, sub-threshold colinear extensions (thin dashed lines). (b) When two colinearly aligned inducer stimuli are present, one on each side of the target pyramidal cell receptive field, a boundary grouping can form. Long-range excitatory inputs fall onto the cell from both sides, and summate. However, these inputs fall onto a shared pool of inhibitory interneurons, which, as well as inhibiting the target pyramidal, also inhibit each other (Tamas, Somogyi, & Buhl, 1998), thus normalizing the total amount of inhibition emanating from the interneuron pool, without any individual interneuron saturating. The combination of summating excitation and normalizing inhibition together create a case of "two-against-one", and the target pyramidal is excited above-threshold. This process occurs along the whole boundary grouping, which thereby becomes represented by a line of suprathreshold-activated layer 2/3 cells (thick dotted line). Boundary strength scales in a graded analog manner with the strength of the inducing signals. (Reprinted with permission from Grossberg and Raizada (2000).)

Tucker and Katz (2003a, 2003b) have provided additional evidence that is consistent with this interpretation by studying tangential slice preparations of layer 2/3 cells of ferret primary visual cortex. Their electrophysiological and imaging methods probed cell responses when two stimulating electrodes up to several hundred microns apart were activated either simultaneously or with prescribed temporal phases. Such a set-up has the potential for directly testing for a firing or modulatory bipole property, or other form of horizontal interaction. Their studies showed that strong suppression occurs as a result of simultaneous activation. In order to interpret this result, it is important to know if the electrodes activated sites that converge on co-linear and co-oriented cells. If they did not, then the two sources could cause net inhibition even if a firing bipole property existed, since sufficiently non-colinear orientations are proposed to compete, not cooperate, using their disynaptic inhibitory interactions.

In fact, orientational competition was part of the earliest bipole model (Grossberg and Mingolla, 1985b). If at least some of their penetrations did represent such co-linear and co-oriented cells, then the absence of net suprathreshold activation favors a modulatory bipole property. In fact, the authors proposed a model to explain their data that strikingly resembles Figure 6, in which horizontal excitatory connections and connections from disynaptic inhibitory interneurons converge on target cells. In particular, they proposed that different excitatory connections synapse on multiple inhibitory interneurons, and that this summation causes the extra inhibition that was observed during simultaneous activation. This predicted property differs from that of a firing bipole model in its ability to allow inhibition to grow substantially when both co-linear input sources are active. If these data are interpreted at face value, then they suggest that one of the parameters in these V1 cells that may differ from a firing bipole interaction is the following: When a firing bipole property obtains, recurrent inhibitory interactions between the inhibitory interneurons are predicted to approximately normalize their total activity when more than one input source is active, while excitation summates at target cells, thereby creating net activation of the target cell. If these recurrent inhibitory interactions are weak, then a modulatory excitatory effect can be realized at target cells as a manifestation of the balance between excitation and inhibition that has been proposed to help stabilize the development of horizontal cortical connections. Due to the overlap in excitatory connections on inhibitory cells, inhibition can then easily increase more rapidly than excitation, as reported in the experiments.

27

What is a Visual Illusion?

Independent of whether this interpretation holds up when additional tests of a firing vs. modulatory bipole property are made in V1, available neural models of vision have provided an alternative explanation to the Ramsden et al. (2001) proposal of how the brain "knows" the difference between a real and an illusory contour. As noted in the discussion of Complementary Computing, visual boundaries and surfaces are predicted to obey complementary properties; see Figure 1. One of these complementary properties is that visual boundaries are not perceptually visible, in the sense that they are predicted not to generate percepts of visible lightness or color within the (V1 interblob)-(V2 pale stripe)-V4 boundary-processing cortical stream. Visual surfaces are proposed to generate visible lightness and color percepts within the (V1 blob)-(V2 thin stripe)-V4 cortical stream. A large body of perceptual and neural data are consistent with this prediction (see Grossberg (1994, 1997) for reviews), but it has not been directly tested. From the perspective of visual boundary and surface processing, a visual illusion is an unfamiliar combination of boundary and surface properties; e.g., in response to an offset grating, a boundary can be recognized without it being seen to separate regions of different visible surface lightness or color.

More generally, the proposal that the brain distinguishes real from illusory contours *within* the boundary system does not fit in well with data suggesting that no such distinction is needed. For example, perceptual boundaries play a key role in enabling the brain to fill-in surface representations after discounting of the illuminant takes place, and in bridging across retinal imperfections like the blind spot and retinal veins; see Grossberg (1994) for a review. Many of the boundaries that we believe to be real are, from a mechanistic viewpoint, really illusory contours that just happen to be correlated with surface lightness and color properties that "look real" based on past experience. If most illusory contours are regularly mistaken for real contours, then the need to distinguish illusory from real contours within the boundary system vanishes.

How Does the Visual Cortex See the World in Depth?

Our discussion so far has not considered how the brain sees the world in depth. The LAMINART model was recently combined with the FACADE model of 3D vision and figure-ground perception (Grossberg, 1994, 1997; Grossberg and McLoughlin, 1997; Kelly and Grossberg, 2000; McLoughlin and Grossberg, 1998) to propose how the laminar circuits of cortical areas V1, V2, and V4 are organized for purposes of stereopsis and 3D planar surface perception (Grossberg and Howe, 2003). This generalization, which is called the 3D LAMINART model, predicts how cellular and network mechanisms of 3D vision are linked to mechanisms of development, learning, grouping, and attention.

Despite some explanatory successes, many previous cortical models of 3D vision, for example the disparity energy model of Ohzawa, DeAngelis, and Freeman (1990), considered only stereopsis, by which image features from the two eyes are binocularly matched in cortical area V1. Stereopsis is important, but on its own is insufficient to explain the 3D surface percepts that form an integral part of our visual consciousness.

The LAMINART extension to 3D vision, shown in Figures 7 and 8, goes beyond these previous analyses in several ways. First, it provides a refined model of stereopsis in V1 which clarifies the role of cells in cortical layers 4, 3B, and 2/3A. In particular, the model revises how the disparity energy model achieves stereopsis, in a manner that is more consistent with recent data. Second, the model proposes how monocular and binocular information are combined and selected in V2 to form 3D boundary representations. Third, the model shows how these 3D boundaries give rise to visible 3D surface percepts in V4. Taken together, these model processes have been used to explain and simulate a much larger set of neurophysiological, anatomical, and psychophysical data about stereopsis and 3D surface perception than has previously been possible.
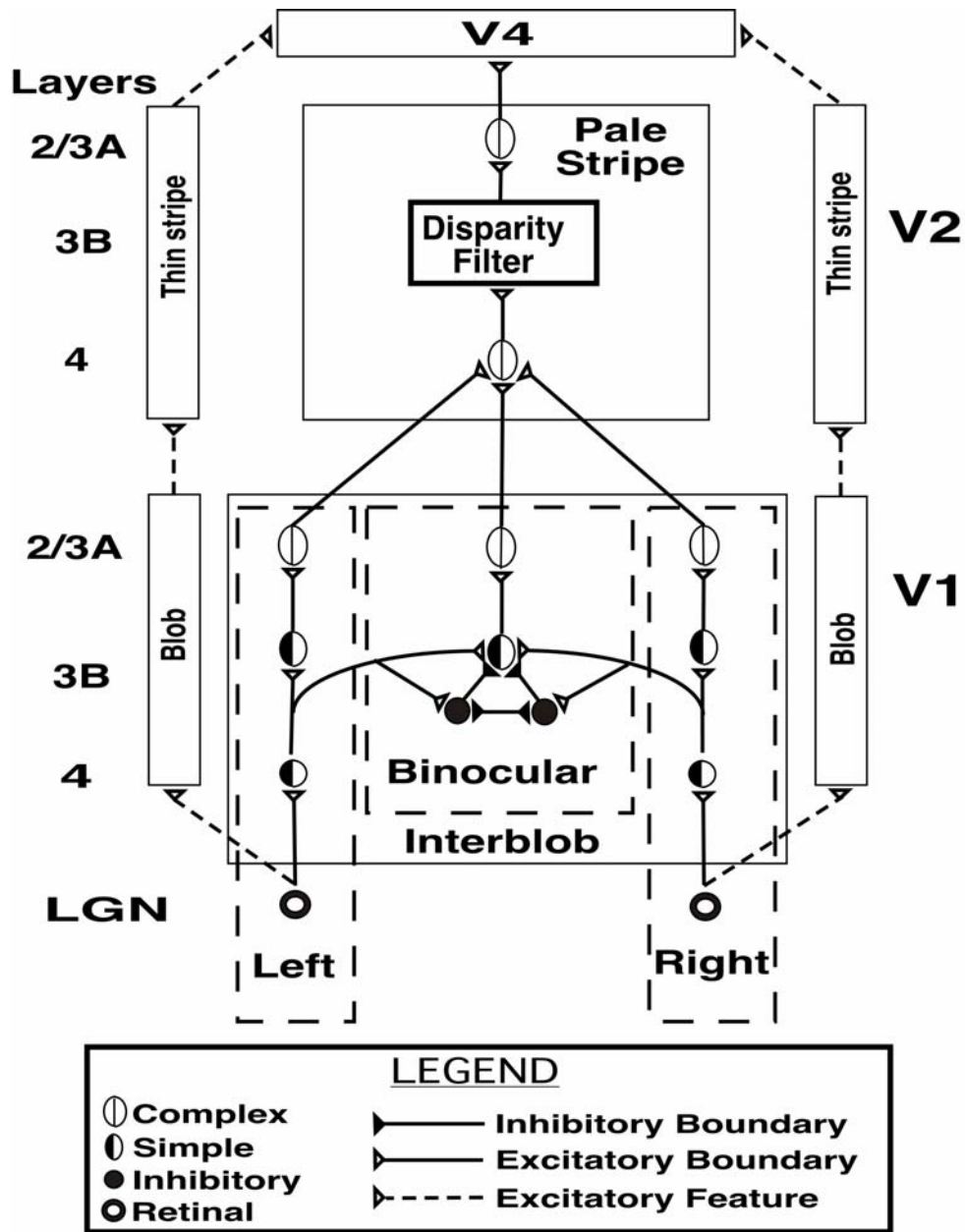
Figure 7. 3D Laminart model circuit diagram for simulating stereopsis and 3D planar surface percepts. See text for details. (Reprinted with permission from Grossberg and Howe (2003).)

The model achieves these goals by embodying five basic psychophysical constraints in its neural circuitry:

(1) *Reconciles contrast-specific binocular fusion with contrast-invariant boundary perception.* Only edges in the left and right retinal images that have the same contrast polarity—that is, their

luminance gradients have the same sign—can be binocularly fused to form a percept of depth (Howard and Rogers, 1995). Binocular fusion thus obeys the *same-sign hypothesis;* see Figure 9a. However, fused boundaries must also be able to form around objects whose contrast polarity with respect to the background can reverse along their perimeters (Grossberg, 1994); see the ellipse in Figure 9b. In other words, binocular boundaries need to be represented in a contrast-invariant way. How can the brain reconcile contrast-specific fusion with the need to form contrast-invariant object boundaries? The model proposes that both constraints are realized by interactions between cells in layers 4, 3B, and 2/3A of cortical area V1 interblobs; see Figure 7.
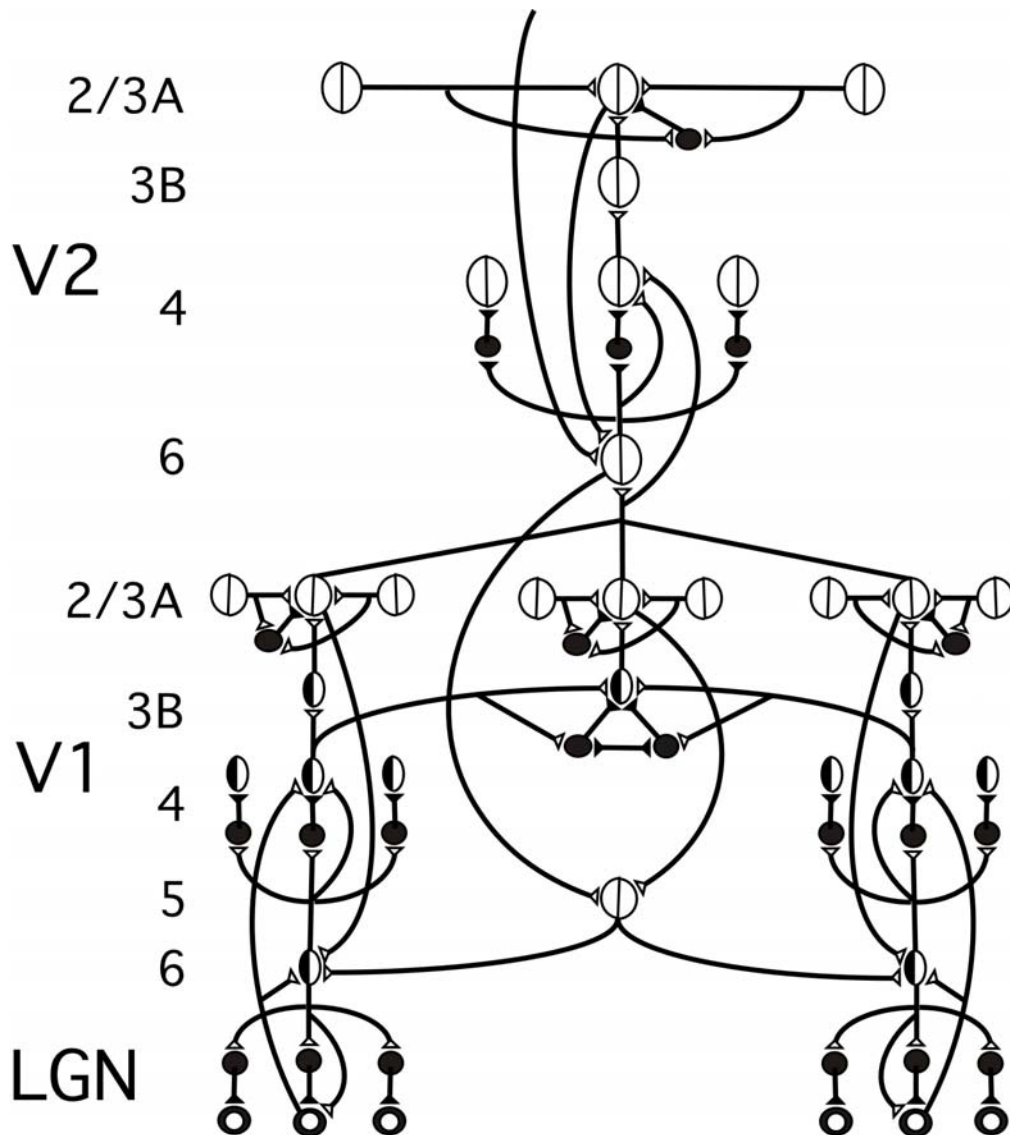

Figure 8. 3D LAMINART model, including 3D boundary completion and attention, as well as the binocular and monocular interactions summarized in Figure 7. See text for details. (Reprinted from Grossberg and Howe (2003).)

31

(2) *Implements the contrast constraint on binocular fusion.* The brain somehow figures out which of the many potential same-sign edges in the two retinal images should be binocularly fused, since veridical stereoscopic depth perception will occur only if the two edges belong to the same object. The brain hereby solves the *correspondence problem* (Julesz, 1971; Howard and Rogers, 1995). An early step in solving the correspondence problem is to binocularly fuse only edges with approximately the same magnitude of contrast (McKee et al., 1994). This constraint naturally arises when the brain fuses edges that derive from the same objects in the world. The model satisfies this constraint through interactions between excitatory and inhibitory cells in layer 3B of V1 (see Figure 7) that endow the binocular cells there with an *obligate property* (Poggio, 1991), whereby they respond preferentially to left and right eye inputs of approximately equal size.
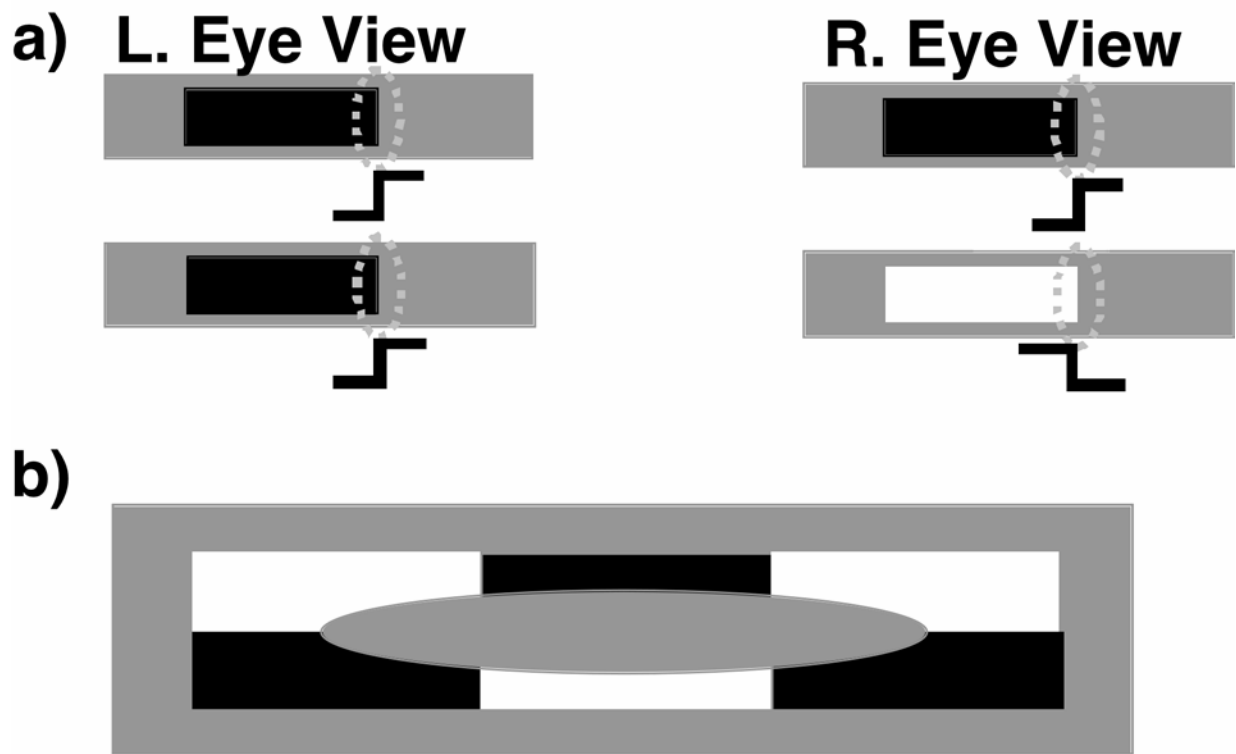
Figure 9. (a) The same-sign hypothesis: only edges that have the same contrast polarity can be stereoscopically fused to produce a percept of depth. (b) As it is traversed, the boundary of the ellipse changes its contrast polarity relative to the background, thereby illustrating the need for object boundaries to be represented in a contrast-invariant manner. See text for details. (Reprinted with permission from Grossberg and Howe (2003).)

(3) *Solves the correspondence problem.* Even if all binocular matches are of the same-sign and of similar contrast magnitude, there can still exist many false binocular matches between edges that do not correspond to the same objects. Some models have attempted to solve this problem by imposing a *unique-matching rule*, which states that any given feature in one retinal image is matched at most with one feature in the other retinal image (Marr and Poggio, 1976; Grimson, 1981; for a review see Howard and Rogers, 1995, pp.42-43). This rule is not satisfied, however, in situations like Panum's limiting case (Panum, 1858; Gillam et al., 1995; McKee et al., 1995) where a bar presented to one eye is simultaneously matched to two separate bars presented to the other eye. The present model does not enforce unique matches. Rather, the model encourages them by using a *disparity filter* that is proposed to occur in the pale stripes of cortical area V2, possibly in layer 3B; see Figures 7 and 10. This disparity filter uses two types of inhibitory interactions: line-of-sight inhibition and inhibition across depth but within cyclopean position, to encourage unique matches.
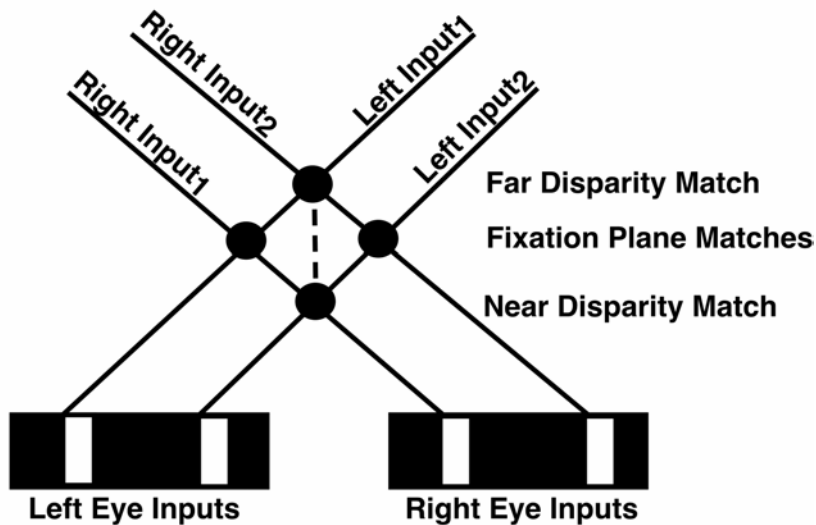

Figure 10. The V2 disparity filter. The V1 binocular boundary network matches an edge in one retinal image with every other edge in the other retinal image whose relative disparity is not too great, that has the same contrast polarity, and whose magnitude of contrast is not too different. In response to this image, the V1 boundary network creates four matches, with the two not in the fixation plane being false matches between edges that do not correspond to the same object. These false matches are suppressed by the disparity filter in V2, wherein each neuron is inhibited by every other neuron that shares either of its monocular inputs (i.e., shares a monocular line-of-sight represented by the solid lines) or is directly in front of or behind it (i.e., is connected to it by a dashed line). The solid lines that represent the monocular lines-of-sight also represent the positional shifts due to binocular matching: an edge in the left retinal image is shifted to the right for matches increasingly further away, whereas an edge in the right retinal image is shifted in the opposite direction. (Reprinted with permission from Grossberg and Howe (2003).)

(4) *Combines monocular and binocular information to form depth percepts*. Although Panum's limiting case may seem to be a laboratory curiosity, many naturally occurring situations also contain only one edge in one eye and two possible edges with which to match it in the other eye. For example, due to the lateral displacement of the eyes, an object's edge that is seen by one eye may be occluded in the other eye, as occurs during da Vinci stereopsis (Nakayama and Shimojo, 1990). Although it does not possess its own binocular information about binocular disparity, the monocularly viewed region has a definite depth conferred to it by the binocularly viewed parts of the scene. The brain can thus utilize monocular information to build up seamless 3D percepts of the world. In fact, in experiments involving Panum's limiting case, varying the relative contrast of the bars alters the perception of depth in a manner that reveals clear monocular-binocular interactions (Smallman and McKee, 1995). Dichoptic masking, where an object presented to one eye is obscured, or masked, by one presented to the other eye, illustrates a third way in which monocular and binocular information may interact (McKee et al., 1994).

Given that monocular information is important, the problem then arises about how to combine monocular and binocular boundaries. Because monocular boundaries do not have a definite depth associated with them, the brain needs to figure out to which depth they should be assigned. A proposed solution to this *monocular-binocular interface problem* was suggested in Grossberg (1994, 1997) in order to explain data about 3D figure-ground perception. This hypothesis is shown in the 3D LAMINART model to clarify many other data about 3D surface perception as well, In particular, the model assumes that the outputs of the monocular boundary cells are added to *all* depth planes in the pale stripes of cortical area V2 along their respective lines-of-sight, possibly in layer 4; see Figures 7 and 11a. The disparity filter, which helps to solve the correspondence problem, also solves the monocular-binocular interface problem by automatically eliminating most of the monocular boundaries that are not at the correct depths.
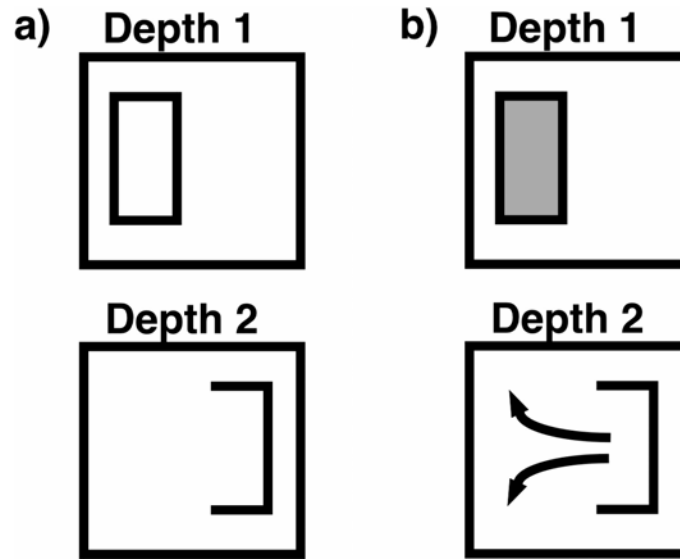
Figure 11. (a) A connected boundary is computed at Depth 1, but an open boundary is computed at Depth 2. (b) Filling-in of surface lightness is contained or not depending on the connectedness of the boundary. Monocular boundaries (i.e., two horizontal boundaries and the right vertical boundary) have been added to all depth planes whereas the binocular boundary (i.e., the left vertical boundary) is present only in the near depth plane, thereby creating a connected boundary, and thus containment of filling-in, only in the near depth plane. (Reprinted with permission from Grossberg and Howe (2003).)

(5) *Forms 3D surface percepts*. The above considerations summarize only how the brain may construct a 3D boundary representation of an object. As noted above, there is considerable evidence that boundary representations on their own do not give rise to visible percepts, which rather are a property of surface representations (Grossberg, 1994). Surface representations are proposed to derive from a filling-in process whereby lightness and color mark the depths at which the surfaces occur. Filling-in is needed to recover lightness and color estimates in regions where they have been suppressed by the process of discounting the illuminant (Grossberg & Todorovic, 1988). The existence of a filling-in process has been supported by psychophysical (Paradiso and Nakayama, 1991; Pessoa and Neumann, 1998; Pessoa, Thompson, and Noë, 1998; Rogers-Ramachandran and Ramachandran, 1998) and neurophysiological experiments (Lee, Mumford, Romero, and Lamme, 1998; Rossi, Rittenhouse, and Paradiso, 1996; Lamme, Rodriguez-Rodriguez, and Spekreijse, 1999; Rossi, Desimone, and Ungerleider, 1998). Boundaries are predicted to control the depths at which particular lightnesses and colors

can fill-in, a process called 3D *surface capture*. The 3D LAMINART model considers only the filling-in of achromatic lightness.

How does the brain ensure that lightness fills-in at only the correct depths? Grossberg (1994) proposed properties of this boundary-surface interaction that helped to explain many data about 3D figure-ground perception. In the 3D LAMINART model, one of these properties proved essential to explain 3D surface percepts that arise in stereopsis research. Namely, visible surfaces arise in cortical area V4 only if they are enclosed by *connected* boundaries. For example, a rectangular connected boundary may be composed of one vertical binocular boundary, one vertical monocularly viewed boundary, and two horizontal boundaries that code no disparity information, as in Figure 11. This connected boundary can support a visible surface percept at the depth corresponding to the binocular boundary if all other constraints are satisfied, because such a boundary can contain the filling-in process. However, if the vertical binocular boundary is missing, as it would be at a different depth plane, then the total boundary is not connected, and a visible percept will not be evident at that depth because filling-in can dissipate out of the boundary gap. This example illustrates how the monocular-binocular interface problem (item (4) above), and thus the correspondence problem (item (3) above), influence visible percepts of 3D surfaces.

Before the 3D LAMINART model was introduced, the FACADE model included a non-laminar model of stereopsis and 3D planar surface perception (Grossberg & McLoughlin, 1997; McLoughlin & Grossberg, 1998) that modified and generalized the disparity energy model of stereopsis (Ohzawa et al., 1990). This generalization incorporated rectification prior to binocular combination, absent from the original disparity energy model, which has recently received independent experimental support (Read et al., 2002; Cumming, 2002). It also proposed that positional shifts between left and right eye cortical inputs code disparities, rather than phase shifts, which has also received experimental support (Tsao & Livingstone, In Press). The FACADE model also incorporated a disparity filter to help solve the correspondence problem (Howard & Rogers, 1995), as well as mechanisms for filling-in 3D surface percepts from 3D boundary representations. In particular, the FACADE model explained the fact that stereoscopic fusion is *generally* impossible when the left and right eye stimuli differ too much in contrast

(Smallman & McKee, 1995). However, in the form developed by Grossberg and McLoughlin, the FACADE model could not explain why stereoscopic fusion is *always* possible in the special case where each eye sees only a single bar, regardless of the contrast difference of the two bars (McKee et al., 1994; Smallman & McKee, 1995).

The 3D LAMINART model was able to overcome this limitation by using properties of identified cells in laminar cortical circuits, and resimulated all the data previously simulated by McLoughlin & Grossberg (1998), in particular data on contrast variations of the correspondence problem and dichoptic masking. In addition, the new model can simulate still more psychophysical data than its non-laminar predecessors, including: the Venetian blind illusion, four different examples of da Vinci stereopsis (Nakayama & Shimojo, 1990; Gillam et al., 1999), stereopsis with opposite-contrast stimuli, the effect of interocular contrast differences on stereoacuities, and various lightness illusions. Simulations of some of these data are shown in Figures 12 and 13.

A Solution of the Correspondence Problem

Smallman and McKee (1995) initiated their extensive study of the correspondence problem by performing a control experiment in which each eye was presented with two bars, all four bars having the same high contrast. Subjects reported seeing two identical bars, both in the far disparity plane. Figure 12a shows the corresponding model simulation, which has the following explanation:
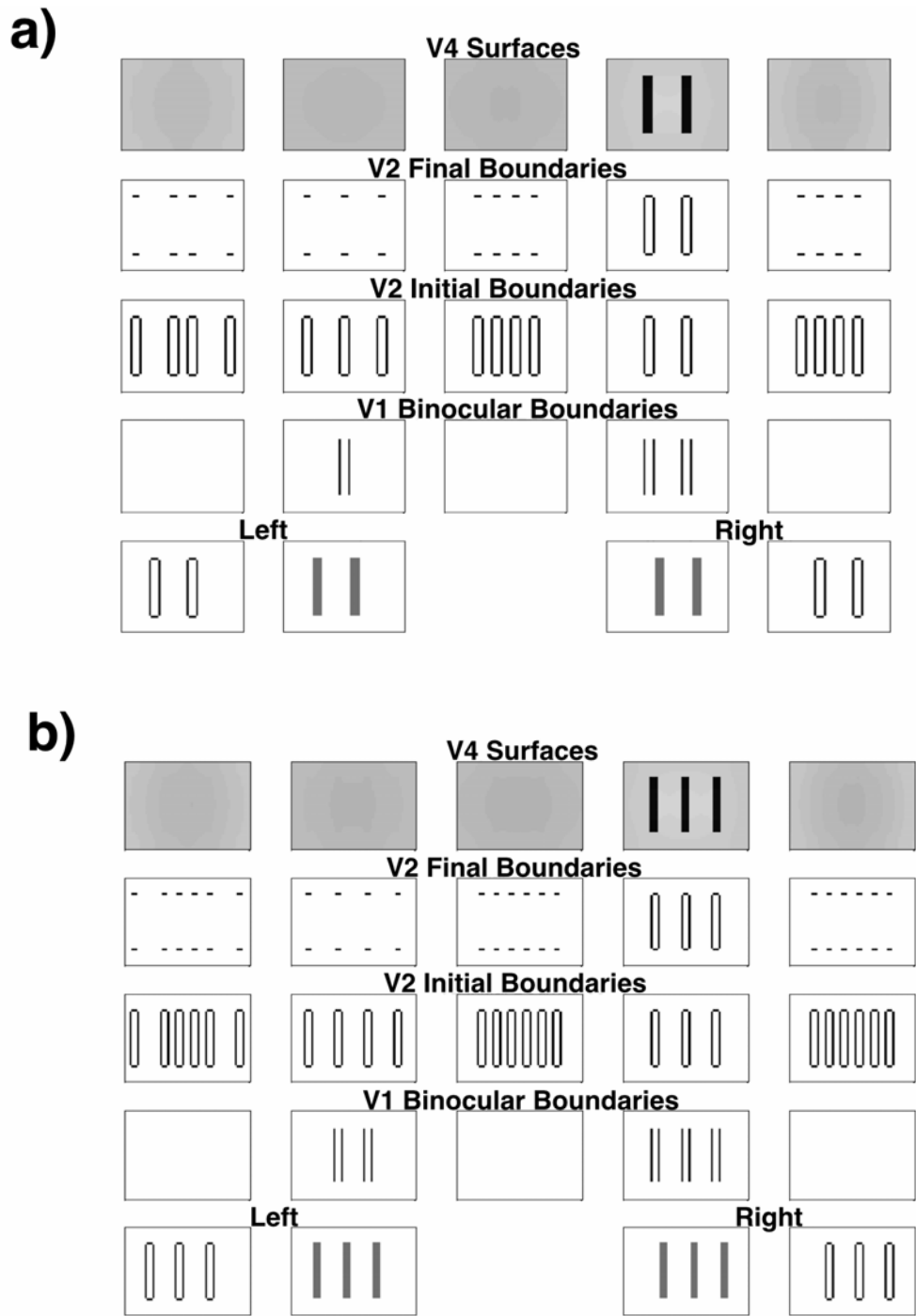
Figure 12. (a) Simulation of the control experiment Smallman and McKee (1995) used for subsequent studies of the correspondence problem. (b) Simulation of a more complicated version of the correspondence problem. See text for details. (Reprinted with permission from Grossberg and Howe (2003).)

Figure 12b summarizes the model simulation of a more complicated version of the correspondence problem that was experimentally studied by Smallman and McKee (1995). Once again the false matches are shown in the second plot of the second row and the correct matches in the fourth plot. Since there are more correct matches than false matches, the latter are again suppressed by the former via the line-of-sight inhibition of the V2 disparity filter. This simulation shows that the model can be applied to more general versions of the correspondence problem than that shown in Figure 12a. More generally, the model has also successfully simulated a still more complex version of the correspondence problem known as the Venetian blind illusion (Howard and Rogers, 1995, Figure 7.21), which consists of two gratings, a low frequency one that is presented to the left eye, and a high frequency one presented to the right eye. When fused, the frequencies of the gratings are such that every second bar of the left grating is in retinal correspondence with every third bar of the right grating. This stereogram produces a percept of short ramps, each containing three bars, sloping up from left to right interspaced with steep returns; viz., a Venetian blind. These simulations, along with those of various versions of da Vinci stereopsis, one of which is summarized below, clarify how the model can handle natural images by showing how it deals with different sorts of potentially confusing matches within the fusion range.

DaVinci Stereopsis

As noted above, DaVinci stereopsis images are of interest because, when viewing the 3D world, one eye sometimes can see parts of the world that the other eye cannot, due to 3D occlusions caused by the fact that each eye views the world from a slightly different position. In the experiments of Nakayama and Shimojo (1990), a thick bar was presented to both eyes and a thin bar only to the right eye, as shown in the first row of plots of Figure 13a. Subjects reported perceiving the thin bar behind the thick bar, at a depth that was consistent with the right edge of the thin bar of the right input being fused with the right edge of thick bar of the left input.The model explanation is as follows:
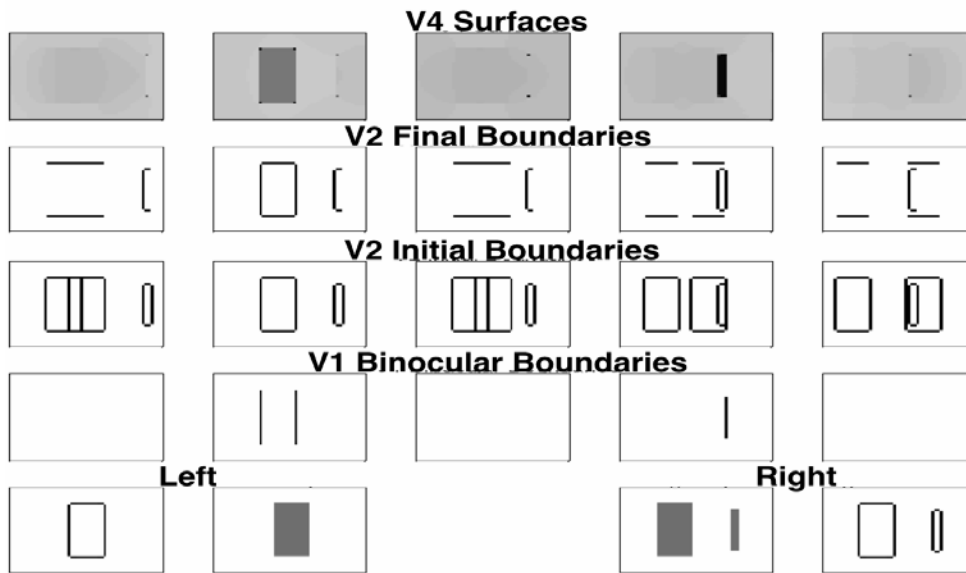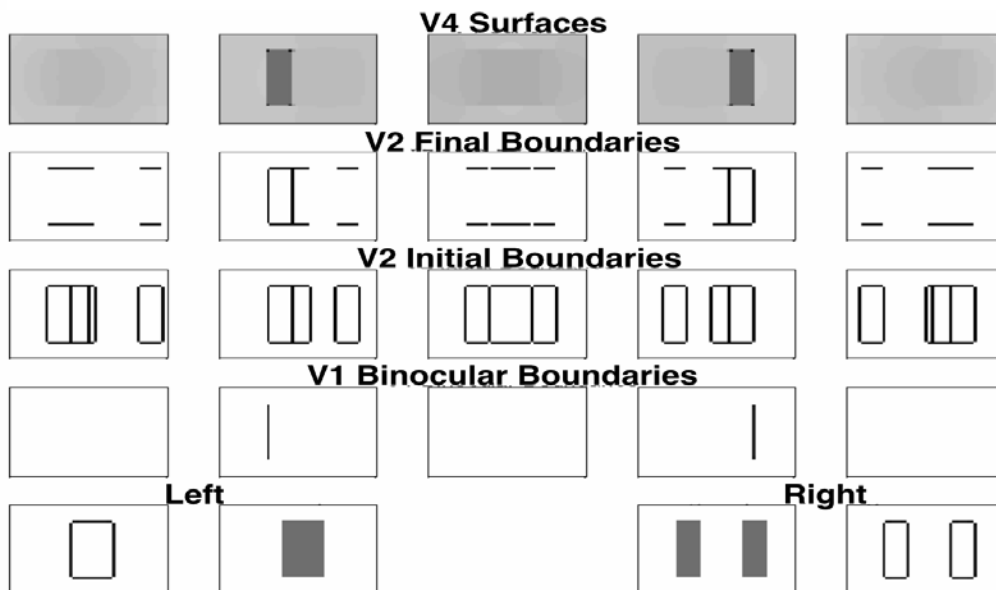
Figure 13. (a) Simulation of the depth percept invoked by the da Vinci stereopsis stimulus of Nakayama and Shimojo (1990). (b) Simulation of depth percept invoked by the da Vinci stereopsis stimulus of Gillam, Blackburn and Nakayama (1999). See text for details. (Reprinted from Grosssberg and Howe (2003).)

The vertical boundaries of the thick bar are fused binocularly in the near disparity plane in V1, as shown by the second plot of the second row. The right edge of the thin bar is matched with the right edge of the thick bar and is thus fused binocularly in the far disparity plane in V1, as shown by the fourth plot. The left edge of the thin bar is registered only monocularly because it cannot be matched with either of the edges of the left input. Monocular boundaries are added to all depth planes in the V2 disparity filter along their respective monocular lines-of-sight, as shown by plots in the third row. The vertical binocular boundaries are also added to the disparity filter, overlapping with the vertical boundaries of the thick bar representation in the second plot and with the rightmost vertical boundary in the fourth plot. These vertical boundaries, being stronger, eliminate all other vertical boundaries that share their lines-of-sight via the disparity filter's line-of-sight inhibition. However, they do not eliminate the vertical boundaries that originate from the left edge of the thin bar, because these do not share any of their lines-of-sight. The final V2 boundary representations are shown in the fourth row.

As usual, V4 fills in surfaces in those regions that are completely enclosed by a connected boundary. This produces a percept of a thick bar in a near disparity plane, represented by the second plot of the top row, and a thin bar in a far disparity plane, represented by the fourth plot. The very small squares seen in the top row are artifacts of the implementation of the filling-in process with a relatively small number of pixels and have no physiological significance. They disappear when the simulations are carried out at a sufficiently high resolution, which takes much more time. The model therefore correctly predicts that the thin bar will appear behind the thick bar at a depth that is consistent with the right edge of the thin bar being stereoscopically fused with the right edge of the thick bar, as in the data of Nakayama & Shimojo (1990).

The daVinci stereopsis stimulus of Gillam, Blackburn, and Nakayama (1999) illustrates a different point, that is simulated in Figure 13b. Here, the right eye sees two thin bars and the left eye a single thick bar. Subjects report seeing two thin bars, the left bar in the near disparity plane and other in the far disparity plane. Gillam et al. (1999) suggested that, because the right eye input contains a gap not present in the left eye input, this display demonstrates that stereopsis can be induced by monocular gaps. The display of Nakayama and Shimojo (1990) in Figure 13a

demonstrated the different point that depth perception could be determined by separation of a monocular bar from a binocular bar. The model explanation is as follows:

The left edge of the thick bar fuses with the left edge of the left thin bar to appear in a near disparity plane in V1, represented by the second plot of the second row. The right edge of the thick bar fuses with the right edge of the right thin bar to appear in a far disparity plane in V1, represented by the fourth plot of this row. In both these cases, the corresponding edges have the same contrast polarity. The two other vertical edges of the thin bars of the right input are registered only monocularly because they cannot be matched to either of the edges of the left input. As usual, the V1 monocular boundary representations are added to all depth planes in the V2 disparity filter along their respective lines-of-sight. As a result, two thin bar representations and one thick bar representation are seen in all disparity planes of the third row, with the slight complication that the thick bar representation overlaps with at least one of the two thin bar representations. The V1 binocular boundary representations are also added to the V2 disparity filter, overlapping with the leftmost vertical boundary in the second plot and the rightmost vertical boundary in the fourth plot. These vertical boundaries, being stronger, inhibit, via the recurrent line-of-sight inhibition of the disparity filter, all the other vertical boundaries that share any of their lines-of-sight. In particular, they do not inhibit those vertical boundary representations originating from the two monocularly viewed edges of the right input because these vertical boundaries do not share any of their lines-of-sight. The final V2 boundary representations are shown in the fourth row. V4 fills-in surfaces in those regions that are completely enclosed by boundaries, resulting in the percept of a thin near bar and a thin far bar, as reported by human subjects in the Gillam et al. (1999) experiments.

Nakayama and Shimojo (1990) have attempted to explain DaVinci stereopsis percepts in terms of an "ecological optics" hypothesis, which suggests that visual systems attempt to interpret unpaired image points in terms of occlusion. For example, in Figure 13a, both eyes see a thick bar but only the right eye a thin bar. According to the ecological optics hypothesis, the visual system interprets these stimuli by assuming that the thin bar is located behind the thick bar at the exact distance that would cause the thick bar to hide it from the left, but not from the right, eye. The 3D LAMINART model proposes, instead, that different organizational principles govern the

3D percepts that we see, including the five principles that have been used to explain data such as those summarized above. The 3D LAMINART model also simulates data that have no ecological optics interpretation with equal ease; see Grossberg and Howe (2003).

Supportive Neurophysiological and Anatomical Data

Given that the 3D LAMINART model can explain a great deal of challenging psychophysical data about stereopsis and 3D planar surface perception, it is important to ask if it forms a predictive link to known neurophysiological and anatomical data? In fact, all the relevant neurophysiological and anatomical data of that are known to us support the model. It should also be noted, however, that the model does not consider cortical areas V3, V3A and MT, even though there is evidence that these areas play a role in depth perception (e.g. Backus et al., 2001). These areas were not needed to simulate the model's targeted data. The function of area V3A appears to be particularly controversial with studies suggesting that it is variously concerned with relative disparity (Backus et al., 2001), saccades (Nakamura & Colby, 2000a, 2000b) and prehensile hand movements (Nakamura et al, 2001). Further complicating the situation is some evidence that the function of macaque V3A differs from that performed by human V3A (Tootell et al., 1997). The following sorts of data are functionally clarified by the model, as summarized in Figure 7:

V1 binocular boundaries. Consistent with the model, the LGN contains circularly symmetric on-center, off-surround receptive fields (Kandel, Schwartz, and Jessell, 2000, pp. 529). LGN lesion studies show that the parvocellular, but not the magnocellular, pathway is critical for fine stereopsis (Schiller, Logothetis, and Charles, 1990a, 1990b). Layer 4 of cortical area V1 is the major recipient of this parvocellular input in vivo (Callaway, 1998) and it is also the input layer of model V1. The model is also consistent with data showing that layer 4 outputs to layer 3B, but not to layer 2/3A, of V1 (Callaway, 1998), a large proportion of it is monocular (Hubel and Wiesel, 1968; Poggio, 1972), and many of its cells are simple (Hubel & Wiesel, 1968; Schiller, Finlay, and Volman, 1976).

The model assumes that polarity-specific binocular matching occurs in layer 3B. This is consistent with data showing that a significant proportion of layer 3B comprises simple cells (Dow, 1974), that layer 3 contains a significant number of binocular cells (Hubel and Wiesel, 1968; Poggio, 1972), and that projections to it can be independent of ocular dominance (Katz, Gilbert, and Wiesel, 1989).

The model suggests that binocular layer 2/3A cells pool responses from layer 3B cells of both contrast polarities so that they can represent the boundaries of objects whose contrast polarity, with respect to the background, changes as the boundary is transversed. In fact, layer 3B projects throughout layer 2/3A (Callaway, 1998), and layers 2 and 3 each contain significant numbers of binocular and complex cells (Poggio, 1972).

The model further predicts that there is a group of cells in layer 2/3A and 3B that respond only to binocular, and not to monocular, stimulation. Such "obligate cells" are known to exist in macaque V1 (Poggio and Fischer, 1977; Smith, Chino, Ni, and Cheng, 1997), with about 40% of tuned excitatory neurons being obligatory (Poggio and Talbot, 1981), including almost all "tuned zero" neurons (Poggio, 1991). Obligate cells do not appear to be as prevalent in cat (Anzai et al., 1995).

The model predicts that all these interactions occur in the V1 interblob regions, which is in keeping with observations that V1 interblobs are highly selective for orientation but relatively unselective for color (Merigan and Maunsell, 1993).

V1 monocular boundaries. The model proposes that the V1 monocular boundaries are formed by a process that is a simplified version of the process which forms V1 binocular boundaries. Much of the above data thus applies to the monocular boundaries network. Additional support for this network comes from observations that layer 3 (Hubel & Wiesel, 1968; Poggio, 1972) and layer 2 (Poggio, 1972) of V1 each comprise a large proportion of monocular cells.

V2 boundaries. The model is consistent with the older prediction of Grossberg and Mingolla (1985a, 1985b) that V2 boundaries are located in the V2 pale stripes. This hypothesis is consistent with observations that the V2 pale stripes receive the major projection from the V1 interblob regions, receive no significant projection from the V1 blob regions, are highly orientationally selective (Roe and Ts'o, 1997), and contain a complete map of visual space (Roe and Ts'o, 1995). The model is also supported by data showing that V2 is mainly binocular (Hubel and Livingstone, 1987; Roe and Ts'o, 1997), is mainly disparity-sensitive (Poggio and Fischer, 1977; von der Heydt, Zhou, and Friedman, 2000), contains many complex cells (Hubel and Livingstone, 1987), receives input into layer 4 (Rockland and Virga, 1990) and outputs to V4 (Xiao, Zych, and Felleman, 1999), which itself is highly selective for disparity (Merigan and Maunsell, 1993). In addition, the V2 pale stripes are disparity-selective (Peterhans, 1997).

The model predicts that one function of V2 is to suppress false matches by utilizing a disparity filter. This is consistent with observations that many cells exhibit false matches in V1 (Cumming and Parker, 2000), but not in V2 (Bakin et al., 2000).

Surfaces. Surfaces in the model are built up through interactions between the V1 blobs, the V2 thin stripes, and V4, consistent with the fact all these regions are linked by major projections (Livingstone and Hubel, 1984; Xiao et al., 1999), and that the V2 thin stripes are the least orientationally-selective area of V2 (Peterhans, 1997) and contain a complete map of visual space (Roe and Ts'o, 1995).

A Prediction: How Do Obligate Cells Work?

A key model prediction, that is worthy of experimental test, is that some binocular simple cells in layer 3B of V1 obey an obligate property whereby they can be activated only if they receive approximately equal inputs from both left and right eye monocular simple cells in layer 4. The obligate property at layer 3B binocular simple cells is predicted to be caused, as shown in Figure 7, by a balance between excitatory inputs from layer 4 monocular simple cells and inhibitory inputs from layer 3B inhibitory interneurons. The interneurons are themselves activated by layer 4 monocular simple cells and mutually inhibit each other, in addition to inhibiting the binocular simple cells. This balance in layer 3B between direct excitatory and indirect inhibitory

interneuronal inputs is reminiscent of the properties that were predicted for bipole cells in layer 2/3. It remains to be seen whether these apparent similarities are reflections of a deeper shared design.

A Synthesis of Stereo Vision, Attention, and Grouping

The data that Grossberg and Howe (2003) simulated using the 3D LAMINART model in Figure 7 did not include attentional manipulations or perceptual grouping by boundary completion. The LAMINART circuit in Figure 2 summarizes some of the key interactions that are proposed to govern perceptual grouping, and attention, without regard to its 3D representation. How, then, can perceptual grouping and attention be consistently joined to stereopsis and 3D planar surface perception processes to further develop the 3D LAMINART model? Figure 8 proposes how 3D boundaries can be completed and how attention can be selectively paid to objects in 3D. The following new features in Figure 8 are the basis for these properties:

First, layer 4 no longer directly activates layer 2/3, as in Figure 2c. Instead, layer 4 simple cells first activate layer 3B simple cells, which in turn activate layer 2/3A complex cells, as shown in Figure 7. The layer 2/3A cells can then interact via horizontal interactions, like those summarized in Figures 2c and 2e, to enhance cell activations due to approximately co-oriented and co-linear inputs. Second, binocular cells in layer 2/3A can represent different disparities, and thus different relative depths from an observer. Interactions between layer 2/3A cells that represent the same relative depth from the observer can be used to enhance boundaries between object contours that lie at that depth.

Because binocular fusion occurs in layer 3B, the binocular boundaries that are formed in layers 3B and 2/3A may be positionally displaced, or shifted, relative to their monocular input signals from layers 6 and 4. Figure 2c illustrates that these layer 2/3 boundaries feed signals back to layer 6 in order to select the winning groupings that are formed in layer 2/3, but issues about binocular shifts did not need to be considered in data explanations of the original LAMINART model. How can the positionally displaced binocular boundaries in layer 2/3A of Figure 8 contact the correct monocularly activated cells in layers 6 and 4, so that they can complete the

feedback loop between layers 2/3A-to-6-to-4-to-3B-to-2/3A that can select winning 3D groupings? Horizontal signals from the monocular layer 4 cells activate binocular obligate cells in layer 3B, which in turn activate layer 2/3A complex cells. This raises the question: How can such a layer 2/3A cell also use horizontal signals to activate its correct layer 6 monocular sources?

The 3D LAMINART model proposes that horizontal connections which are known to occur in layers 5 and 6 (Callaway & Wiser, 1996) accomplish this. Feedback signals from layer 2/3A propagate vertically to layer 5, whose cells activate horizontal axons in this layer that contact the appropriate layer 6 cells. These layer 5-to-6 horizontal contacts are assumed to be selectively formed during development. Grossberg and Williamson (2001) and Grossberg and Seitz (2003) have simulated how layer 2/3 connections and layer 6-to-4 connections may be formed during development. The selective layer 5-to-6 contacts are proposed to form according to similar laws. In summary, *inward* horizontal layer 4-to-3B and 2/3A-to-2/3A connections are proposed to form binocular cells and their groupings, while *outward* layer 5-to-6 connections are proposed to close the feedback loops that help to select the correct 3D groupings.

Once this role in 3D grouping for layer 6 horizontal connections is before us, the preattentive-attentive interface problem forces a proposal for how attention fits into the 3D circuit: namely, top-down attentional outputs from layer 6 of a higher cortical level like V2 activates the same layer 5 cells that contact monocular input sources in layer 6 via horizontal connections. Then the layer 6-to-4 modulatory on-center off-surround network controls attentional priming and matching, just like in Figure 2b. This proposal raises the question of how the top-down pathways from layer 6 of a higher cortical level know how to converge on the same layer 5 cells to which the layer 2/3 cells project at the lower cortical level? Since firing of the layer 2/3 cells activates the layer 5 cells as well as the layer 6 cells of the higher cortical level, as in Figures 7 and 8, this could occur due to associative learning.

3D Representations of Slanted and Curved Surfaces and their 2D Images

The model in Figure 8 can handle only the representation of 3D planar surfaces. However, most of the objects in the world are slanted or curved and span multiple depths with respect to an observer. Both binocular cues, such as disparity, and monocular cues, such as perspective, shading, and junctions, provide information about slant and tilt of an object. Slant is defined as deviation around the vertical axis and tilt is defined as deviation around the horizontal axis. These considerations raise the question: Can the 3D LAMINART model be consistently generalized to explain perceptual and neurobiological data about how we see slanted and curved surfaces in 3D? Moreover, can such a model also explain how we generate 3D representations of the projections of such surfaces into a 2D picture? Initial work in this direction by Grossberg and Swaminathan (2003) proposes how such an extension can be made.

Figure 14a clarifies one of the main difficulties that must be faced when trying to understand how this happens. Here, two different objects are made up of the same set of surfaces. Depending on how the individual surfaces are combined, two different 3D objects can be perceived. The same parallelogram can signal a near-to-far or a far-to-near slanted surface, depending upon the context. Contextual cues thus play a key role in disambiguating ambiguous local cues. In response to some 2D images, such as Necker cube images, the percept changes over time and depends on various factors such as attention and internal receptive field biases (Kawabata, 1986). The 3D LAMINART model clarifies how the different percepts in Figure 14a are generated, and also how Necker cube percepts can oscillate in a bistable fashion through time.

Information about binocular disparity can be used to determine the slant of an object. A slanted object is registered at multiple disparities and these representations need to be grouped across depth for it to be perceived as a single object. Information about tilt and curvature of an object can also be gleaned from disparity cues. Both neurophysiological and psychophysical studies have provided clues about the sorts of processes that provide the contextual disambiguation that is needed to unambiguously represent slanted and curved 3D surfaces.
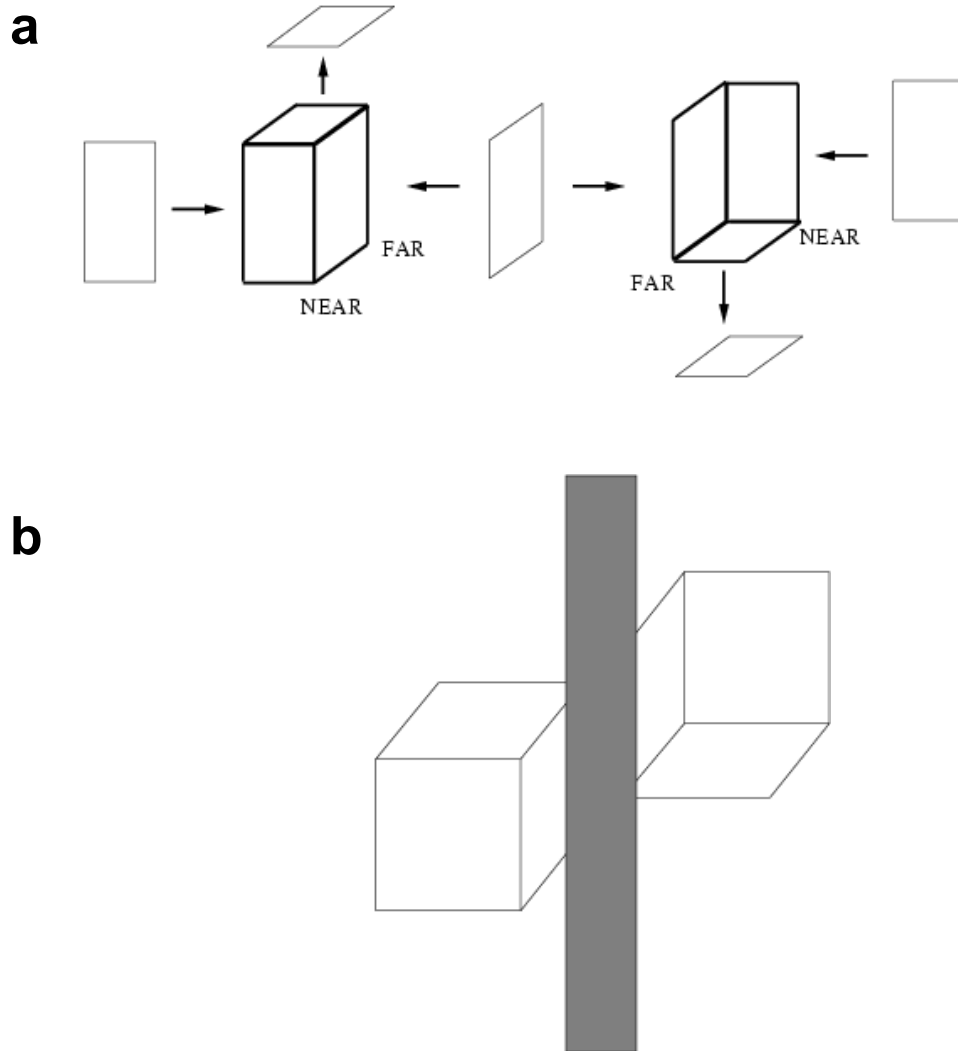
Figure 14. (a) The same 2D angles and shapes, when combined in different ways, can elicit percepts of different surface slants: The left bold figure has a positive slant (near to far) while the right bold figure has a negative slant (far to near). (b) Even though the sides of the cube are colinear in 2D, they are not colinear in their 3D interpretation (Tse, 1999). (Reprinted with permission from Tse (1999).)

Neurophysiological and Psychophysical Data about Disparity-Sensitive Cells

Neurophysiological studies have found cells in extrastriate cortex that are tuned to features important in 3D perception. In Macaque cortical area V2, cells are tuned to relative disparity (Thomas et al., 2002), disparity edges (von der Heydt et al., 2000), angles (Pasupathy and

Connor, 1999), border ownership (Zhou et al., 2000) and figure-ground relations (Bakin et al., 2000). There is evidence for cells tuned to slanted 3D boundaries in V4 (Hinkle and Connor, 2001). Curvature tuning is found in V4 (Pasupathy and Connor, 2001), inferotemporal cortex (Janssen et al., 2000), and parietal cortex (Taira et al., 2000).

Psychophysical studies have shown the importance of relative disparity, or disparity gradients, in human visual perception. Targets specified by a different stereoscopic slant than the distractors in an image or scene are detected pre-attentively (Holliday and Braddick, 1991), as are targets presented on a surface of different slant than that of the distractors (Nakayama and Silverman, 1986; He and Nakayama, 1995). Multi-element tracking results do not differ if the elements are on a planar or a slanted surface (Viswanathan and Mingolla, 1999). Ryan and Gillam (1993) provided evidence that three-dimensional aftereffects can result from disparity gradient adaptation by showing that the size of the aftereffect varied with the disparity gradient of the adapting lines. Lee (1999) showed that the size of aftereffects are also dependent on the difference in disparity between the adapting and test surfaces. Illusory surface experiments (Nakayama and Shimojo, 1992) illustrate that depth needs to be taken into account during grouping. Many studies show that, although 2D bipole grouping principles work well on the 2D projection of 3D images, in other cases, 2D grouping principles gives rise to a different result than the 3D percept. For example, in Figure 14b, even though two lines across the cubes are colinear in the 2D plane, they are not colinear in the 3D interpretation (Tse, 1999) and hence are not grouped.

3D Grouping using Context-Sensitive Angle and Disparity-Gradient Cells

The generalization of the 3D LAMINART model to the case of 3D slanted and curved surfaces proposes that object fragments at multiple depth planes can be grouped together by using context-sensitive interactions between angle cells and disparity-gradient cells that are sensitive to an object's slant and tilt. In particular, monocular cues in an image, notably combinations of angles, can bias the activation of some disparity-gradient cells more than others to form a 3D percept in response to 2D images, such as Necker cube images. These contextual interactions give the correct combinations of cells an advantage in the competitive processes that select the

final boundary groupings. Interactions between these cell types can also form illusory contours in response to curved 3D neon color displays. Figure 15a provides a schematic of the model, and Figure 15b shows some of the crucial interactions among several known cell types within the boundary system that are proposed to form 3D slanted and curved boundaries. Laminar interactions that are not crucial to this process are omitted in Figure 15, but they are included for completeness in Figure 16.
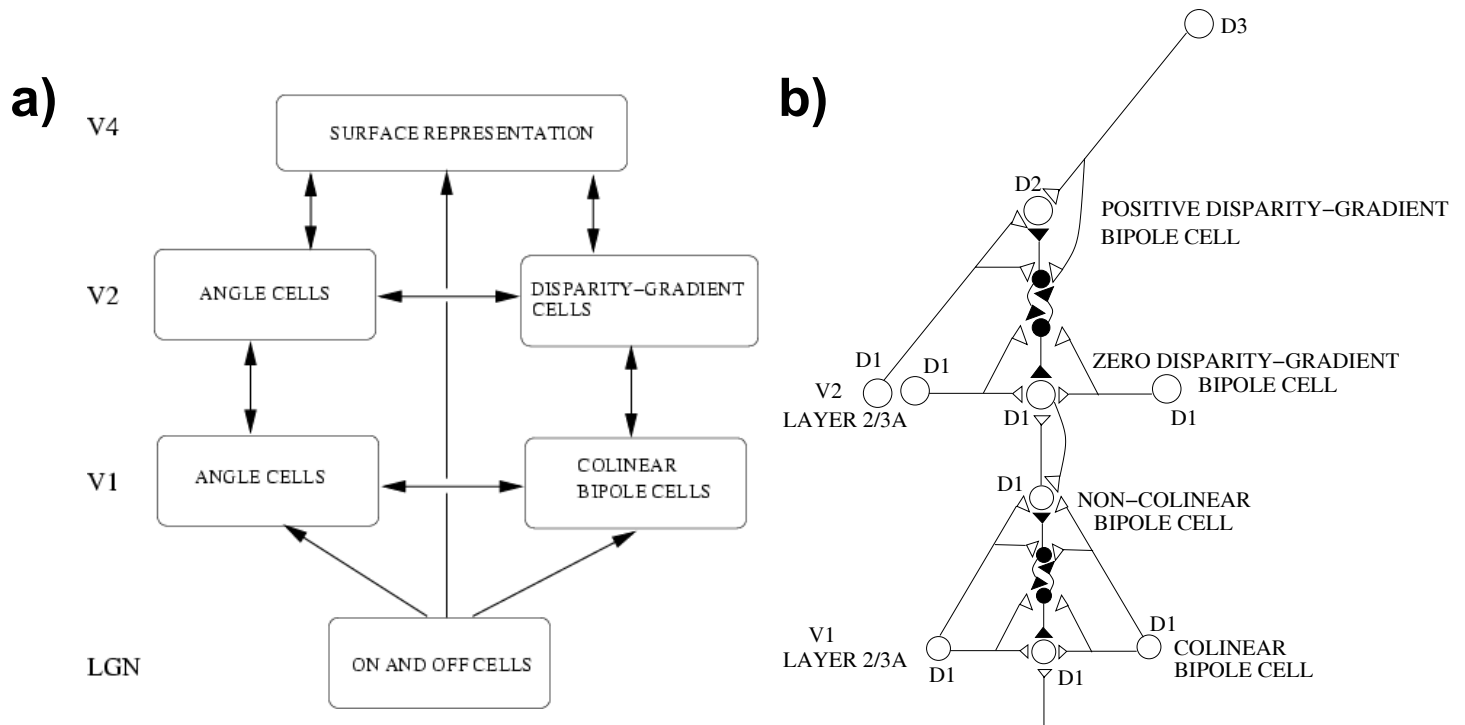

Figure 15. Block diagram of the 3D Laminart model of Grossberg and Swaminathan (2003) for representing slanted and curved 3D surfaces: The input image undergoes on-center, off-surround processing in the LGN. In layer 2/3A of V1, angle cells and colinear bipole cells get activated by angles and line segments in the images, respectively. Angle cells and colinear bipole cells interact with each other via long-range horizontal connections in layer 2/3A of V1. Colinear bipole cells activate disparity-gradient cells, while V1 angle cells activate V2 angle cells. V2 angle cells and disparity-gradient cells interact via long-range horizontal connections in layer 2/3A of V2 to disambiguate the types of ambiguity illustrated in Figure 14. Disparity-gradient cells group across position and disparity to form closed boundary segments, which are used as a barrier for filling-in of surfaces in V4 that receive filling-in signals from lower processing levels, such as the LGN. (b) Laminar circuit for 3D boundary grouping: V1 angle cells and colinear bipole cells are in layer 2/3A of V1. Layer 2/3A cells in V1 activate layer 2/3A cells in V2. Layer 2/3A of V2 contains V2 angle cells and disparity-gradient cells. Layer 2/3A of V2 feeds back to layer 2/3A of V1. Disparity-gradient cells group across disparity-gradient and disparities. D1, D2, and D3 represents various depths. Open (black) circles (triangles) represent excitatory (inhibitory) cells (connections).

A key insight of the 3D LAMINART model is that angle cells, disparity-gradient cells, and the more familiar within-depth colinear bipole cells can all develop from the same underlying long-range horizontal excitatory connections and disynaptic inhibitory interneurons, in response to image statistics. Thus these several cell types seem all to be variations of a single cortical design. Grossberg and Swaminathan (2003) have also proposed how these cell types may be organized within cortical maps.
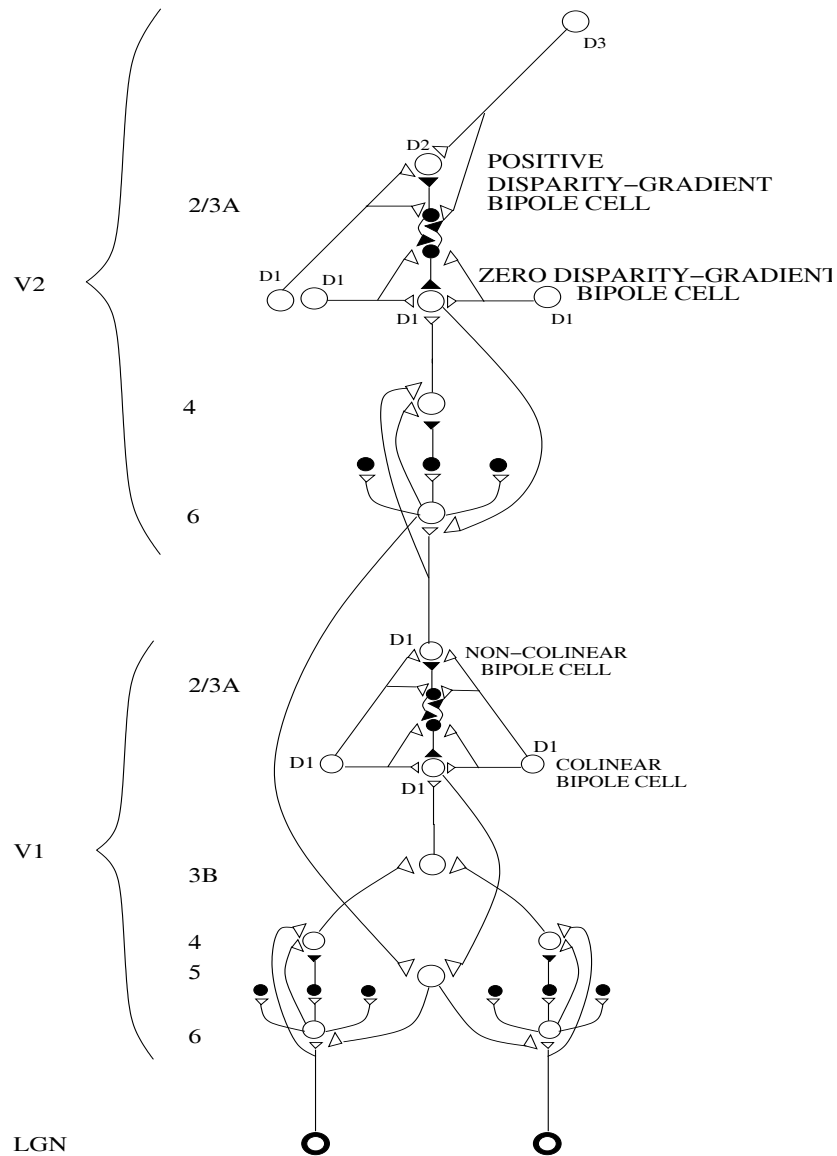

Figure 16. 3D LAMINART MODEL: The 3D LAMINART model in Figure 8 is extended to include the new types of cells in Figure 15 that are needed to form 3D representations of slanted and curved surfaces, and of 2D images. Open (black) circles (triangles) show excitatory (inhibitory) neurons (connections). See text for details.

52

Habituation, Development, Reset, and Bistability

Activity-dependent habituative mechanisms are needed for successful development of these cell types to occur. This sort of habituation has also proved to be essential in other studies of cortical development (Grunewald and Grossberg, 1998; Olson and Grossberg, 1998; Grossberg and Seitz, 2003); see Grossberg (2003) for a review. The habituative mechanisms prevent the developmental process from "getting stuck" with using the cells that initially win the competition for activation over and over again. Such perseveration would prevent multiple feature combinations from getting represented in a distributed fashion throughout the network. Habituative interactions help to solve this problem because habituation is activity-dependent: only those cells or connections habituate that are in active use. Thus, when habituation acts, it selectively weakens the competitive advantage of the initial winners, so that other cells can become activated to represent different input features.

Remarkably, these habituative mechanisms, which seem to play such an important role in cortical development, are also proposed to play several important roles in adult vision. One particularly striking property is that they can lead to multi-stable percepts when two or more 3D interpretations of a 2D image are approximately equally salient, as in Necker cube percepts. Grossberg and Swaminathan (2003) have used the same habituative and competitive mechanisms that they used to develop disparity-gradient receptive fields to also simulate how a 2D Necker cube image can generate two different 3D boundary and surface representations, which oscillate bistably from one to the other through time.

Surely the ecological value of habituative mechanisms in adult vision is not just to produce perceptual curiosities like bistable percepts of a Necker cube! In fact, these habituative mechanisms seem to play a crucially important role in adult vision; namely, to efficiently *reset* previously active visual representations when the scenes or images that induced them change or disappear. Without such an active reset process, visual representations could easily persist for a long time due to the hysteresis that could otherwise occur in circuits with as many feedback loops as those in Figures 2, 8, and 16. In many examples of this reset process, offset of a previously active input leads to an *antagonistic rebound* of activation in previously inactive cells,

and these newly activated cells help to inhibit the previously active cells, including grouping cells in layer 2/3. This reset process is not perfect, however, and there are large perceptual data bases concerning residual effects of previously active representations. In fact, such a reset process has elsewhere been used to explain psychophysical data about visual aftereffects (Francis and Grossberg, 1996a; Grunewald and Lankheet, 1996), visual persistence (Francis *et al*., 1994), and binocular rivalry (Arrington, 1993, 1995, 1996; Grossberg, 1987; Liang and Chow, 2002), among other data that are all proposed to be manifestations of the reset process. Ringach *et al*. (1999) have reported direct neurophysiological evidence for rebound phenomena using reverse correlation techniques to analyze orientational tuning in neurons of cortical area V1. Abbott et al. (1997) have provided direct experimental evidence in visual cortex of the habituative mechanisms that were predicted to cause the reset (Grossberg, 1968, 1969, 1980). Grossberg (1980, 1999b) also predicted that such reset processes play a role in driving the reset and memory search processes that help the adult brain to rapidly discover and learn new representations of the world, as part of Adaptive Resonance Theory. In summary, there is a predicted link, mediated by habituative transmitter mechanisms, between processes of cortical development in the infant, and processes of perceptual and cognitive reset, learning, and bistability in the adult. This link is worthy of a lot more experimental study than it has received to date.

How Does Surface Filling-in Span Multiple Depths?

As noted above, it has been proposed that the grouping of boundaries and the filling-in of surfaces are distinct, indeed complementary (Grossberg, 2000), processes. Whereas boundaries complete inwardly in an oriented fashion, surfaces fill-in outwardly in an unoriented fashion until a boundary is reached. The outward filling-in process needs to be controlled across multiple depth planes when it fills-in 3D curved surfaces. A potential problem is that a multiple-depth boundary may have gaps at some depths, but not others, which could allow spreading colors and brightnesses to spill out during filling-in. As noted in Figure 11, and illustrated in Figures 12 and 13, if there is a large gap in a boundary, then the filling-in within that boundary will not have an impact on conscious perception. I call this problem the *lightness dissipation problem*. A related problem is seen in 3D illusory displays that induce a percept of a 3D curved surface (Carman and

Welch, 1992; Liinasuo et al., 2000). Here the filling-in signal needs to spread in a controlled way across depths where there are no boundaries or filling-in inducers at all in the original images. How does the brain contain the filling-in process across a surface that spans multiple depths?
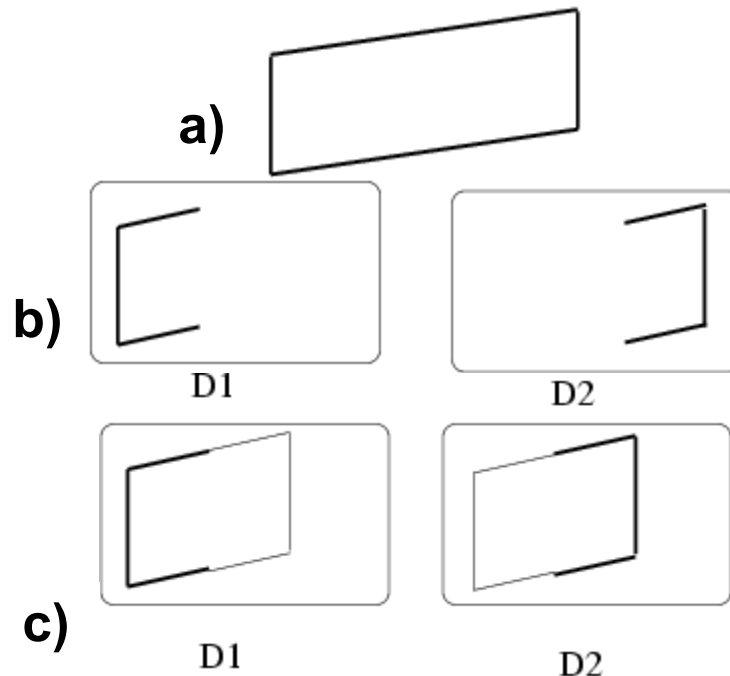

Figure 17. Filling-in of slanted surfaces. (a) The input is a slanted rectangle. (b) Multiple depth representation of the slanted rectangle. (c) Filling-in barriers: The boundary representation act as a strong filling-in barrier at the corresponding depth of the input representation, and as a weak barrier at nearby depths, thus creating closed boundary compartments within each depth. D1 (near) and D2 (far) represent different depths.

This problem is overcome in the 3D LAMINART model as follows: A boundary signal that acts as a strong barrier to filling-in at its preferred depth also weakly acts as a barrier to filling-in at other depths. For example, consider a slanted rectangle in depth, as in Figure 17a. Each boundary representation is activated at its preferred depths, as in Figure 17b, and this boundary representation has gaps at each depth. If no other boundaries existed, filling-in signals would flow out of the boundary gaps at each depth. The model proposes that the boundary at a particular depth is also represented, albeit weakly, at nearby other depths. This hypothesis has earlier been made to explain how a finite pool of depth-selective boundaries can control a continuous change in perceived depth (Grossberg, 1994, 1997). Here it is predicted to also contribute to percepts of slanted and curved surfaces in depth. In particular, the total boundary signal that acts as a barrier to filling-in at each depth is shown in Figure 17c. Now, a closed

boundary exists at each depth, and the filling-in signal is at least partially contained at each depth. Because of differences in boundary strength, however, the filled-in activity is not uniformly strong at each position. It is stronger wherever there is a strong boundary, since lightness and color can dissipate more through a weaker boundary than a stronger one. Grossberg and Swaminathan (2003) have simulated how a slanted surface representation can be generated by such differential filling-in across different depths.

<div align="center">

Future Research Directions

</div>

This review has considered some of the recent progress in modeling how the laminar circuits of visual cortex control processes of development, learning, attention, and 3D vision. These models make detailed predictions about how particular cell types and connections are linked to all of these processes. The models also identify previously unsuspected conceptual and mechanistic links between these processes that should provide inspiration for qualitatively new types of interdisciplinary experiments.

These results form part of a larger neural theory of 3D vision and figure-ground perception that is called FACADE theory, which derived its name from its goal of explaining how representations of Form-And-Color-And-DEpth are formed by visual cortex; see Figure 1. FACADE theory models many data that have not yet been analysed within a laminar framework, including data about 3D figure-ground perception (Grossberg, 1994, 1997; Kelly and Grossberg, 2000), 3D surface perception (Grossberg and Mingolla, 1987), texture perception (Grossberg and Pessoa, 1998), brightness and lightness perception (Grossberg and Kelly, 1999; Grossberg and Todorovic, 1988; Pessoa, Mingolla, and Neurmann, 1995), cortical synchronization (Grossberg and Grunewald, 1997; Grossberg and Somers, 1991), visual persistence (Francis, Grossberg, and Mingolla, 1994; Francis and Grossberg, 1996a, 1996b), and perceptual aftereffects (Francis and Grossberg, 1996a; Grossberg, Hwang, and Mingolla, 2002). Of immediate interest is the fact that grouping properties of 3D figure-ground perception seem to be significantly elaborated by cortical area V2 (Lamme, 1998). How such figure-ground constraints, many of which have already been modeled in a non-laminar context, will fit together with 3D LAMINART circuitry is of great current interest.

More generally, future research may be expected to expand the number of phenomena that can be explained using a unified laminar theory, with different specific processes being understood as variations of an underlying general theory of laminar cortical design. Due to the ubiquity of laminar neocortex, such progress may be expected to include a wide range of perceptual and cognitive processes, including modalities other than vision. For example, long-range horizontal connections are known to occur in many areas of neocortex, such as the auditory and language areas of the human temporal cortex (Schmidt, Schlote, Bratzke, Rauen, Singer, and Galuske, 1997). The proposed role of ART-based mechanisms in helping to ensure stable development and learning of LAMINART model circuits creates a context for such additional studies. In particular, neural models of visual object learning and recognition (Bradski and Grossberg, 1995; Carpenter, 1997; Carpenter and Ross, 1995; Grossberg, 1999c; Grossberg and Williamson, 1999), visual motion perception (Chey, Grossberg, and Mingolla, 1997; Grossberg, Mingolla, and Viswanathan, 2001), visual search (Grossberg, Mingolla, and Ross, 1994), auditory perception and streaming (Grossberg, 1999b), and speech perception and word recognition (Grossberg, Boardman, and Cohen, 1997; Grossberg and Myers, 2000; Grossberg and Stone, 1986), among other competences, have been developed in which ART-like cortical mechanisms play a key role in helping to explain the targeted data. It remains to be seen whether and how such ART-like mechanisms are specialized within the laminar circuits of other cortical areas to realize a variety of intelligent behaviors.

References

Abbott, L. G., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220–224.

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences USA*, *90*, 5718–5722.

Amir, Y., Harel, M., & Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in Macaque monkey visual cortex. *Journal of Comparative Neurology*, *334*, 19–46.

Anzai, A., Bearse, M. A., Freeman, R. D., & Cai, D. (1995). Contrast coding by cells in the cat's striate cortex: Monocular vs. binocular detection. *Visual Neuroscience*, *12,* 77–93.

Arrington, K. F. (1993). Binocular rivalry model using multiple habituating nonlinear reciprocal connections. *Neuroscience Abstracts, 19*, 1803.

Arrington, K. F. (1995). Neural model of rivalry between occlusion and disparity depth signals. *Neuroscience Abstracts, 21*, 125.

Arrington, K. F. (1996). Stochastic properties of segmentation-rivalry alternations. *Perception, 25, Supplement*, 62.

Backus, B. T., Fleet, D. J., Parker, A. J., & Heeger, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology, 86,* 2054–2068.

Bakin, J.S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three–dimensional surface configurations. *The Journal of Neuroscience*, *20*, 8188–8198.

Bosking, W.H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and

the arrangement of horizontal connections in the tree shrew striate cortex. *The Journal of Neuroscience*, *17*, 2112–2127.

Bradski, G., & Grossberg S. (1995). Fast-learning VIEWNET architectures for recognizing three-dimensional objects from multiple two-dimensional views. *Neural Networks, 8*, 1053–1080.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Leipzig: Barth.

Bullier, J., Hupé, J. M., James, A., & Girard, P. (1996). Functional interactions between areas V1 and V2 in the monkey. *Journal of Physiology (Paris), 90*, 217–220.

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, *21,* 47–74.

Callaway, E. M., & Wiser, A. K. (1996). Contributions of individual layer 2-5 spiny neurons to local circuits in macaque primary visual cortex. *Visual Neuroscience, 13,* 907–922.

Caputo, G., & Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research*, *38*, 669–689.

Carman, G. J. & Welch, L. (1992). Three-dimensional illusory contours and surfaces. *Nature*, *360*, 585–587.

Carpenter, G. A. (1997). Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. *Neural Networks, 8,* 1473–1494.

Carpenter, G. A., & Ross, W.D. (1995). ART-EMAP: a neural network architecture for object recognition by evidence accumulation. *IEEE Transactions on Neural Networks, 6*, 805–818.

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 A1,4,* 2570–2594.

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

Crook, J. M., Engelmann, R., & Löwel, S. (2002). Gaba-inactivation attenuates colinear facilitation in cat primary visual cortex. *Experimental Brain Research*, *143*, 295–302.

Cumming, B. G. (2002). Receptive field structure and disparity tuning in primate V1. *Vision Sciences Society (abstracts)*, 288.

Cumming, B. G., & Parker, A. J. (2000). Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. *The Journal of Neuroscience, 20*, 4758–4767.

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

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

Dow, B. M. (1974). Function classes of cells and their laminar distribution in monkey visual cortex. *Journal of Neurophysiology*, *37,* 927–946.

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psycholog*y: *Human Perception and Performanc*e, *14*, 188–202.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitbock, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics, 60*, 121–130.

Egeth, H., Virzi, R. A., and Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 32–39.

Egusa, H. (1983). Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field. *Perception, 12*, 167–175.

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

Faubert, J., & von Grunau, M. (1995) The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Research, 35*, 3119–3130.

Field, D.J., Hayes, A., & Hess, R.F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research, 33,* 173–193.

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.

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

Gillam, B., Blackburn, S., & Cook, M. (1995). Panum's limiting case: Double fusion, convergence error, or 'da Vinci stereopsis'. *Perception*, *24,* 333–346.

Gillam, B., Blackburn, S., & Nakayama, K. (1999). Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours. *Vision Research*, *39,* 493–502.

Goodale, M.A., & Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 10–25.

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

Grimson, W. E. (1981). A computer implementation of a theory of human stereo vision. *Philosophical Transactions of the Royal Society (B), 292,* 217–253.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal 'illusory' contours. *Natur*e, *365*, 550–552.

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, *60*, 758-765.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, *22*, 325–364.

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

Grossberg (1982), *Studies of mind and brain.* Dordrecht, The Netherlands: D. Reidel Publishing Company.

Grossberg, S. (1976). 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, Volume 5*. New York: Academic Press.

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

Grossberg, S. (1982). *Studies of mind and brain.* Amsterdam: Kluwer.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception. In E. Degreef & J. van Buggenhaut (Eds.), *Trends in mathematical psychology.* Amsterdam: North-Holland.

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

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

Grossberg, S. (1995). The attentive brain. *American Scientist, 83,* 438–449.

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 Visio*n, *12*, 163–187.

Grossberg, S. (1999b). Pitch-based streaming in auditory perception. In N. Griffith and P. Todd (Eds.), *Musical networks: Parallel distributed perception and performance.* Cambridge, MA: MIT Press, pp.117–140.

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

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, *4*, 233–246.

Grossberg, S. (2003). Linking visual cortical development to visual perception. In B. Hopkins & S. P. Johnson (Eds.). *Advances in infancy research: Neurobiology of infant vision.* Westport, CN: Praeger Publishers, in press.

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

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

Grossberg, S., & Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Research*, in press.

Grossberg, S., Hwang, S., & Mingolla, E. (2002). Thalamocortical dynamics of the McCollough effect: Boundary-surface alignment through perceptual learning. *Vision Research, 42,* 1259–1286.

Grossberg, S., & Kelly, F. J. (1999). Neural dynamics of binocular brightness perception. *Vision Research*, *39*, 3796–3816.

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

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., & Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics and Image processing*, *37*, 116–165.

Grossberg, S., Mingolla, E., & Ross, W. D. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review, 101*, 470–489.

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., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*, 2521–2553.

Grossberg, S., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, *4,* 735–767.

Grossberg, S., & Pessoa, L. (1998). Texture segregation, surface representation, and figure-ground separation. *Vision Research, 38,* 2657–2684.

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

Grossberg, S., & Seitz, A. (2003). Laminar development of receptive fields, maps, and columns in visual cortex: The coordinating role of the subplate. *Cerebral Cortex*, in press.

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

Grossberg, S. & Stone, G.O. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93,* 46–74.

Grossberg, S., & Swaminathan, G. (2003). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2d images: Development, attention, and bistability. Submitted for publication.

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. (1999). A self-organizing neural system for learning to recognize textured scenes. *Vision Research, 39,* 1385–1406.

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 groupings and learning. *Cerebral Cortex, 11,* 37–58.

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

Grunewald, A., & Lankheet, M.J. (1996). Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature*, *384*, 358–60.

He, Z.J., & Nakayama, K. (1995). Visual attention to surface in three-dimensional space. *Proceedings of the National Academy of Sciences USA, 21,* 11155–11159.

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

Hinkle, D., & Connor, C. E. (2001). Three-dimensional orientation tuning in macaque area V4. In *Society for Neuroscience Abstracts,* 286.7, San Diego, USA.

Hirsch, J. A., & Gilbert, C. D. (1991). Synpatic physiology of horizontal connections in the cat's visual cortex. *The Journal of Neuroscience, 11*, 1800–1809.

Holliday, I. E., & Braddick, O. J. (1991). Pre-attentive detection of a target defined by stereoscopic slant. *Perception*, *20*, 355–362.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis.* New York: Oxford University Press.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *The Journal of Neuroscience*, *7,* 3378–3415.

Hubel D. H., & Wiesel T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195,* 215–243.

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

Hupé, J. M., James A. C., Girard, D. C., & Bullier, J. (1997). Feedback connections from V2 modulate intrinsic connectivity within V1. *Society for Neuroscience Abstracts,* 406.15, 1031.

Janssen, P., Vogels, R., & Orban, G. A. (2000). Three-dimensional shape coding in inferior temporal cortex. *Neuron*, *27*, 385–397.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: The University of Chicago Press.

Kaas, J. H. (1999). Is most of neural plasticity in the thalamus cortical? *Proceedings of the National Academy of Sciences USA, 96*, 7622 –3.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science.* Chicago: University of Chicago Press (4th Ed.).

Kanizsa, G. (1974). Contours without gradients or cognitive contours. *Italian Journal of Psychology, 1*, 93–113.

Kapadia, M. K, Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*, 843–856.

Kastner, S. & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia, 39,* 1263–1276.

Katz, L. C., Gilbert, C. D., & Wiesel, T. N. (1989). Local circuits and ocular dominance columns in monkey striate cortex. *The Journal of Neuroscience, 9,* 1389–1399.

Kawabata, N. (1986). Attention and depth perception. *Perception*, *15*, 563–572.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception.

Cognitive Psychology, *23*, 141–221.

Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception and Psychophysics*, *62*, 1596–1618.

Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67,* 961–980.

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

Lamme, V. A. F. (1998). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience*, *15,* 1605–1615.

Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Corte*x, *9*, 406–413.

Lee, B. (1999). Aftereffects and the representation of stereoscopic surfaces. *Perception*, *28*, 1155–1169.

Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of primary visual cortex in higher level vision. *Vision Researc*h, *38*, 2429–2454.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computatio*n, *10*, 903–940.

Liang, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience,12*, 39–53.

Liinasuo, M., Kojo, I., Häkkinen, J., & Rovamo, J. (2000). Neon color spreading in three-dimensional illusory objects in humans. *Neuroscience Letters*, *281*, 119–122.

Liu, Z., Gaska, J. P., Jacobson, L. D., & Pollen, D. A. (1992). Interneuronal interaction between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research, 32*, 1193–1198.

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

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiolog*y, 77, 24–42.

Lund, J. S., & Wu, C. Q. (1997). Local circuit neurons of macaque monkey striate cortex: IV. Neurons of laminae 1-3A. *Journal of Comparative Neurolog*y, *384*, 109–126.

Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194,* 283–287.

Martin, J. H. (1989). *Neuroanatomy: Text and atlas*. Norwalk: Appleton and Lange.

Maunsell, J.H.R. & Van Essen, D.C. (1983). Response properties of single units in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus duration, speed and orientation. *Journal of Neurophysiology, 49*, 1127–1147.

McGuire, B. A., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *The Journal of Comparative Neurology, 305,* 370–392.

McKee, S. P., Bravo, M. J., Smallman, H. S., & Legge, G. E. (1995). The 'uniqueness

constraint' and binocular masking. *Perception*, *24,* 49–65.

McKee, S. P., Bravo, M. J., Taylor, D. G., & Legge, G. E. (1994). Stereo matching precedes dichoptic masking. *Vision Research*, *34,* 1047–1060.

McLoughlin, N. P., & Grossberg, S. (1998). Cortical computation of stereo disparity. *Vision Research*, *38*, 91–99.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways. *Annual Review of Neuroscience, 16,* 369–402.

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

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9,* 104–110.

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli. *Journal of Neurophysiolog*y, *70*, 909–919.

Mounts, J. R. W. (2000). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception and Psychophysics, 62,* 969–983.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of corticocortical loops. *Biological Cybernetics, 66,* 241–251.

Nakamura, H., Kuroda, T., Wakita, M., Kusunoki, M., Kato, A., Mikami, A., Sakata, H., & Itoh, K. (2001). From three-dimensional space vision to prehensile hand movements: The lateral intraparietal area links the area V3A and the anterior intraparietal area in macaque. *The Journal of Neuroscience*, *21*, 8174–8187.

Nakamura, K., & Colby, C. L. (2000a). Visual, saccade-related, and cognitive activation of single neurons in monkey extrastriate area V3A. *Journal of Neurophysiology, 84,* 677–692.

Nakamura, K., & Colby, C. L. (2000b). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences, 99,* 4026–4031.

Nakayama, K., & Shimojo, S. (1990). da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research*, *30,* 1811–1825.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*, 1357–1363.

Nakayama, K. & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264–265.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, *249,* 1037–1041.

Olson, S., & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, *11*, 189–208.

Palmer, L. A., & Davis, T. L. (1981). Receptive field structure in cat striate cortex. *Journal of Neurophysiology*, *46*, 260–276.

Panum, P. L. (1858). *Physiologische Untersuchungen ueber das Sehen mit zwei Augen.* Kiel: Schwerssche, Buchhandlung, translated by C. Hubscher (1940). Hanover, NH: Dartmouth Eye Institute.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Researc*h, *31*, 1221–1236.

Parker, J. L., & Dostrovsky, J. O. (1999). Cortical involvement in the induction, but not expression, of thalamic plasticity. *The Journal of Neuroscience, 19,* 8623 –9.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*, 2490–2502.

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, *86*, 2505–2519.

Pessoa, L., & Neumann, H. (1998). Why does the brain fill-in? *Trends in Cognitive Sciences*, *2*, 422–424.

Pessoa, L., Beck, J. & Mingolla, E. (1996). Perceived texture segregation in chromatic element-arrangement patterns: High intensity interference. *Vision Research, 36,* 1745–1760.

Pessoa, L., Mingolla, E., & Neumann, H. (1995). A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Research*, *35*, 2201–2223.

Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Beha*v*ioral and Brain Science*s, *21*, 723–802.

Peterhans, E. (1997). Functional organization of area V2 in the awake monkey. *Cerebral Cortex*, *12,* 335–357.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *The Journal of Neuroscience, 9,* 1749–1763.

Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized macaque monkey. *Investigative Ophthalmology*, *11*, 369–377.

Poggio, G. F. (1991). Physiological basis of stereoscopic vision. In D. Regan (Ed.), *Vision and visual dysfunction: Binocular vision* (pp. 224–238). Boston, MA: CRC.

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, *40*, 1392–1405.

Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology, 315*, 469–492.

Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, *391*, 580–584.

Pollen, D. A. (1999). On the neural correlates of visual perception. *Cerebral Cortex*, *9*, 4–19.

Pollen, D. A., & Ronner S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science, 212,* 1409–1411.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 2–25.

Pylyshyn, Z.W., & Storm, R.W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179–197.

Raizada, R.D.S., & Grossberg, S. (2001). Context-sensitive bindings by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition*, *8*, 341–466.

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.

Ramsden, B. M., Hung, C. P., & Roe, A. W. (2001). Real and illusory contour processing in Area V1 of the primate -- a cortical balancing act. *Cerebral Cortex*, *11*, 648–665.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects. *Nature Neuroscienc*e, *2*, 79–87.

Read, J. C. A, Cumming, B. C., & Parker, A. J. (2002). Simple cells can show non-linear binocular combination. *Vision Sciences Society (abstract)*, *287*.

Reynolds, J., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience*, *19*, 1736–1753.

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, 21.3*, 1759.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Properties of macaque V1 neurons studied with natural image sequences. *Investigative Ophthalmology and Vision Science 40*, abstract 989.

Rockland, K. S., & Virga, A. (1989). Terminal arbors of individual 'feedback' axons projecting from area V2 to V1 in the macaque monkey: A study using immunohistochemistry of anterogradely transported phaseolus vulgaris-leucoagglutinin. *Journal of Comparative Neurolog*y, *285* (1), 54–72.

Rockland, K. S., & Virga, A. (1990). Organization of individual cortical axons projecting from area V1 (area 17) to V2 (area 18) in the macaque monkey. *Visual Neuroscience*, *4*, 11–28.

Roe, A. W., & Ts'o, D. Y. (1995). Visual topography in primate V2: Multiple representation across functional stripes. *The Journal of Neuroscience*, *15,* 3689–3715.

Roe, A. W., & Ts'o, D. Y. (1997). The functional architecture of area V2 in the macaque monkey. *Cerebral Cortex*, *12,* 295–333.

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

Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Researc*h, *38*, 71–77.

Rossi, A. F., Desimone, R., & Ungerleider, L. G. (1998). Late onset responses to extra-receptive field stimulation in V1. *Society for Neuroscience Abstract*s, *28*, 789.2 (abstract).

Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, *273*, 1104–1107.

Ryan, C., & Gillam, B. (1993). A proximity-contingent stereoscopic depth aftereffect: evidence for adaptation to disparity gradients. *Perception*, *22*, 403–418.

Salin, P., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiological Reviews, 7*5, 107–154.

Sandell, J. H., & Schiller, P. H. (1982). Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *Journal of Neurophys*iolog y, *48*, 38–48.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976). Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology*, *39,* 1288–1319.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1990a). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, *5,* 321–346.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1990b). Functions of the colour-opponent and broad-band channels of the visual system. *Nature, 343,* 68–70.

Schmidt, K. E., Schlote, W., Bratzke, H., Rauen, T., Singer, W., & Galuske, R. A. W. (1997). Patterns of long range intrinsic connectivity in auditory and language areas of the human temporal cortex. *Society for Neuroscience Abstracts,* 415.13, 1058.

Seitz, A., & Watanabe, T. (2003). Is subliminal learning really passive? *Nature, 422*, 6927.

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.

Sheth, B. R., Sharma, J., Rao, S. C., & Sur, M. (1996). Orientation maps of subjective contours in visual cortex. *Scienc*e, *274,* 2110–2115.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492–496.

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

Smallman, H. S., & McKee, S. P. (1995). A contrast ratio constraint on stereo matching. *Proceedings of the Royal Society of London B*, *260,* 265–271.

Smith, A.T., Singh, K.D., & Greenlee, M.W. (2000). Attentional suppression of activity in the human visual cortex. *Neuroreport, 11,* 271–277.

Smith, E. L., Chino, Y., Ni, J., & Cheng, H. (1997). Binocular combination of contrast signals by striate cortical neurons in the monkey. *Journal of Neurophysiology, 78,* 366–382.

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Science*s USA, *96*, 1663–1668.

Somers, D. C., Todorov, E. V., Siapas, A. G., Toth, L. J., Kim, D., & Sur, M. (1998). A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cerebral Corte*x, *8*, 204–217.

Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a canter-surround organization. *Vision Research*, *35*, 1859–1869.

Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: A model bridging physiology and psycho-physics. *Scienc*e, *269*, 1877–1880.

Taira, M., Tsutshi, K. I., Jiang, M., Yara, K., & Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *Journal of Neurophysiology*, *83*, 3140–3146.

Tamas, G., Somogyi, P., & Buhl, E. H. (1998). Differentially interconnected networks of GABAergic interneurons in the visual cortex of the cat. *The Journal of Neuroscience, 18*, 4255–4270.

Temereanca, S., & Simons, D. J. (2001). Topographic specificity in the functional effects of corticofugal feedback in the whisker/barrel system. *Society for Neuroscience Abstracts,* 393.6.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, *5*, 472–478.

Tootell, R. B. H., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B, Sereno, M. I., & Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *The Journal of Neuroscience*, *17,* 7060–7078.

Tsao, D. Y., & Livingstone, M. S. (2003). Spatiotemporal maps of disparity-selective simple cells in macaque V1. *Neuron,* in press.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37–68.

Tucker, T.R. & Katz, L.C. (2003a). Spatiotemporal patterns of excitation and inhibition evoked by the horizontal network in layer 2/3 of ferret visual cortex. *Journal of Neurophysiology, 89,* 488–500.

Tucker, T.R. & Katz, L.C. (2003b). Recruitment of local inhibitory networks by horizontal connections in layer 2/3 of ferret visual cortex. *Journal of Neurophysiology, 89*, 501–512.

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.

Vanduffel, W., Tootell, R.B., & Orban, G.A. (2000). Attention-dependent suppression of metabolic activity in the early stages of the macaque visual system. *Cerebral Cortex, 10,* 109–126.

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

Viswanathan, L., & Mingolla, E. (1999). Dynamics of attention in depth: Evidence from multi-element tracking. Technical Report CAS/CNS-TR-99-010, Deptarment of Cognitive and Neural Systems, Boston University, Boston, MA.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science, 224,* 1260–1262.

von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, *40,* 1955–1967.

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

Watanabe, T., Sasaki, Y., Nielsen, M., Takino, R., & Miyakawa, S. (1998). Attention-regulated activity in human primary visual cortex. *Journal of Neurophysiology, 79,* 2218–2221.

Wittmer, L. L., Dalva, M. B., & Katz, L. C. (1997). Reciprocal interactions between layer 4 and layer 6 cells in ferret visual cortex. *Society Neuroscience Abstracts,* 651.5, 1668.

Wolfe, J. M., and Friedman-Hill, S. R. (1992). Part-whole relationships in visual search. *Investigative Opthalmology and Visual Science, 33,* 1355.

Xiao, Y., Zych, A., & Felleman, D. J. (1999). Segregation and convergence of functionally defined V2 thin stripe and interstripe compartment projections to area V4 of macaques. *Cerebral Cortex, 9,* 792–804.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Researc*h, *38*, 719–741.

Zhang, Y., Suga N., & Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature, 387*, 900 –903.

Zhou, H., Friedman, H. S., von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience*, *20,* 6594–6611.