Principles of Neural Science, 6e >

# Chapter 21: The Constructive Nature of Visual Processing

## Introduction

We are so familiar with seeing, that it takes a leap of imagination to realize that there are problems to be solved. But consider it. We are given tiny distorted upside-down images in the eyes and we see separate solid objects in surrounding space. From the patterns of stimulation on the retina we perceive the world of objects and this is nothing short of a miracle.

—Richard L. Gregory, *Eye and Brain,* 1966

MOST OF OUR IMPRESSIONS of the world and our memories of it are based on sight. Yet the mechanisms that underlie vision are not at all obvious. How do we perceive form and movement? How do we distinguish colors? Identifying objects in complex visual environments is an extraordinary computational achievement that artificial vision systems have yet to duplicate. Vision is used not only for object recognition but also for guiding our movements, and these separate functions are mediated by at least two parallel and interacting pathways.

The existence of parallel pathways in the visual system raises one of the central questions of cognition, the binding problem: How are different types of information carried by discrete pathways brought together into a coherent visual image?

## Visual Perception Is a Constructive Process

Vision is often incorrectly compared to the operation of a camera. A camera simply reproduces point-by-point the light intensities in one plane of the visual field. The visual system, in contrast, does something fundamentally different. It interprets the scene and parses it into distinct components, separating foreground from background. The visual system is less accurate than a camera at certain tasks, such as quantifying the absolute level of brightness or identifying spectral color. However, it excels at tasks such as recognizing a charging animal (or a speeding car) whether in bright sunlight or at dusk, in an open field or partly occluded by trees (or other cars). And it does so rapidly to let the viewer respond and, if necessary, escape.

A potentially unifying insight reconciling the visual system's remarkable ability to grasp the bigger picture with its inaccuracy regarding details of the input is that vision is a biological process that has evolved in step with our ecological needs. This insight helps explain why the visual system is so efficient at extracting useful information such as the identities of objects independent of lighting conditions, while giving less importance to aspects like the exact nature of the ambient light. Moreover, vision does so using previously learned rules about the structure of the world. Some of these rules appeared to have become wired into our neural circuits over the course of evolution. Others are more plastic and help the brain guess at the scene presented to the eyes based on the individual's past experience. This complex, purposeful processing happens at all levels of the visual system. It starts even at the retina, which is specialized to pick out object boundaries rather than creating a point-by-point representation of uniform surfaces.

This *constructive* nature of visual perception has only recently been fully appreciated. Earlier thinking about sensory perception was greatly influenced by the British empiricist philosophers, notably John Locke, David Hume, and George Berkeley, who thought of perception as an atomistic process in which simple sensory elements, such as color, shape, and brightness, were assembled in an additive way, component by component. The modern view that perception is an active and creative process that involves more than just the information provided to the retina has its roots in the philosophy of Immanuel Kant and was developed in detail in the early 20th century by the German psychologists Max Wertheimer, Kurt Koffka, and Wolfgang Köhler, who founded the school of Gestalt psychology.

The German term *Gestalt* means configuration or form. The central idea of the Gestalt psychologists is that what we see about a stimulus—the perceptual interpretation we make of any visual object—depends not just on the properties of the stimulus but also on its context, on other features in the visual field. The Gestalt psychologists argued that the visual system processes sensory information about the shape, color, distance, and movement of objects according to computational rules inherent in the system. The brain has a way of looking at the world, a set of expectations that derives in part from experience and in part from built-in neural wiring.

Max Wertheimer wrote: “There are entities where the behavior of the whole cannot be derived from its individual elements nor from the way these elements fit together; rather the opposite is true: the properties of any of the parts are determined by the intrinsic structural laws of the whole.” In the early part of the 20th century, the Gestalt psychologists worked out the laws of perception that determine how we group elements in the visual scene, including similarity, proximity, and good continuation.

We see a uniform six-by-six array of dots as either rows or columns because of the visual system’s tendency to impose a pattern. If the dots in each row are similar, we are more likely to see a pattern of alternating rows (Figure 21–1A). If the dots in each column are closer together than those in the rows, we are more disposed to see a pattern of columns (Figure 21–1B). The principle of good continuation is an important basis for linking line elements into unified shapes (Figure 21–1C). It is also seen in the phenomenon of contour saliency, whereby smooth contours tend to pop out from complex backgrounds (Figure 21–1D). The Gestalt features that we are disposed to pick out are also ones that characterize objects in natural scenes. Statistical studies of natural scenes show that object boundaries are likely to contain visual elements that lie in close proximity, are continuous across intersections, or form smooth contours. It is tempting to speculate that the formal features of objects in natural scenes created evolutionary pressure on our visual systems to develop neural circuits that have made us sensitive to those features.

Figure 21–1

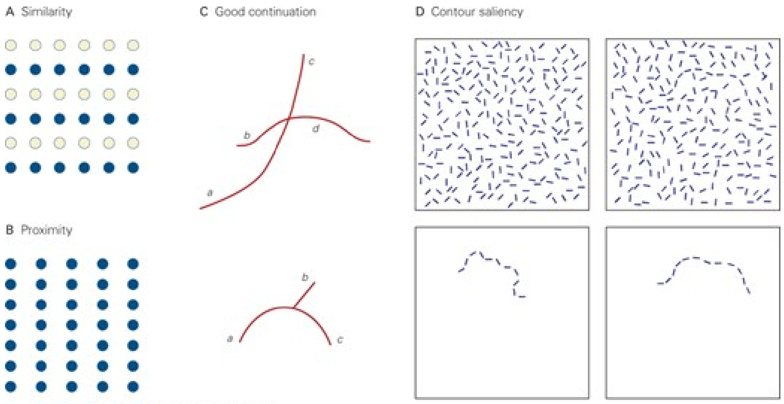
**Organizational rules of visual perception.** To link the elements of a visual scene into unified percepts, the visual system relies on organizational rules such as similarity, proximity, and good continuation.

**A.** Because the dots in alternating rows have the same color, an overall pattern of blue and white rows is perceived.

**B.** The dots in the columns are closer together than those in the rows, leading to the perception of columns.

**C.** Line segments are perceptually linked when they are collinear. In the top set of lines, one is more likely to see line segment **a** as belonging with **c** rather than **d**. In the bottom set, **a** and **c** are perceptually linked because they maintain the same curvature, whereas **a** and **b** appear to be discontinuous.

**D.** The principle of good continuation is also seen in contour saliency. On the right, a smooth contour of line elements pops out from the background, whereas the jagged contour on the left is lost in the background. (Adapted, with permission, from Field, Hayes, and Hess 1993. Copyright © 1993 Elsevier Ltd.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

Separating the figure and background in a visual scene is an important step in object recognition. At different moments, the same elements in the visual field can be organized into a recognizable figure or serve as part of the background for other figures (Figure 21–2). This process of segmentation relies not only on certain geometric principles, but also on cognitive influences such as attention and expectation. Thus, a priming stimulus or an internal representation of object shape can facilitate the association of visual elements into a unified percept (Figure 21–3). This internal representation can take many different forms reflecting the wide range of time scales and mechanisms of neural encoding. It could consist of transient reverberating spiking activity selective to a shape or a decision, lasting a fraction of a second, or the selective modulation of synaptic weights during a particular context of a task or an expected shape, or circuit changes that could comprise a long-term memory.

Figure 21–2

**Object recognition depends on segmentation of a scene into foreground and background.** Recognition of the white salamanders in this image depends on the brain "locating" the white salamanders in the foreground and the brown and black salamanders in the background. The image also illustrates the role of higher influences in segmentation: One can consciously select any of the three colors as the foreground. (Reproduced, with permission, from M.C. Escher's "Symmetry Drawing E56" © 2010 The M.C. Escher Company-Holland. All rights reserved. www.mcescher.com.)

Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

Figure 21–3

**Expectation and perceptual task play a critical role in what is seen.** It is difficult to separate the dark and white patches in this figure into foreground and background without additional information. This figure immediately becomes recognizable after viewing the priming image on page 501. In this example, higher-order representations of shape guide lower-order processes of surface segmentation. (Reproduced, with permission, from Porter 1954. Copyright 1954 by the Board of Trustees of the University of Illinois. Used with permission of the University of Illinois Press.)

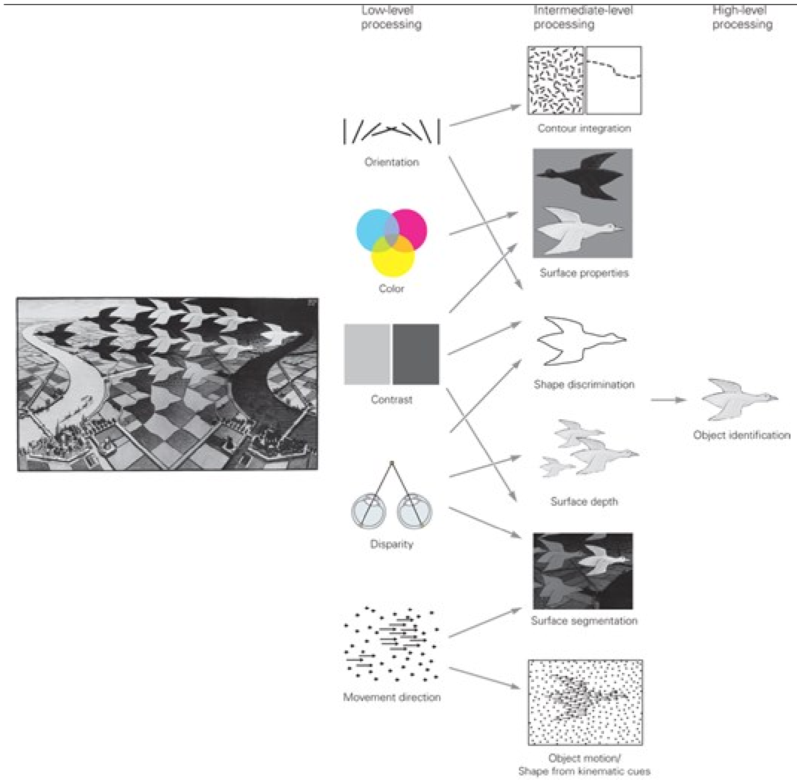
Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

The brain analyzes a visual scene at three levels: low, intermediate, and high (Figure 21–4). At the lowest level, which we consider in the next chapter (Chapter 22), visual attributes such as local contrast, orientation, color, and movement are discriminated. The intermediate level involves analysis of the layout of scenes and of surface properties, parsing the visual image into surfaces and global contours, and distinguishing foreground from background (Chapter 23). The highest level involves object recognition (Chapter 24). Once a scene has been parsed by the brain and objects recognized, the objects can be matched with memories of shapes and their associated meanings. Vision also has an important role in guiding body movement, particularly hand movement (Chapter 25).

Figure 21–4

**A visual scene is analyzed at three levels.** Simple attributes of the visual environment are analyzed (low-level processing), and these low-level features are then used to parse the visual scene (intermediate-level processing): Local visual features are assembled into surfaces, objects are segregated from background (surface segmentation), local orientation is integrated into global contours (contour integration), and surface shape is identified from shading and kinematic cues. Finally, surfaces and contours are used to identify the object (high-level processing). (M.C. Escher's "Day and Night". © 2020 The M.C. Escher Company—The Netherlands. All rights reserved. www.mcescher.com)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

In vision, as in other cognitive operations, various features—motion, depth, form, and color—occur together in a unified percept. This unity is achieved not by one hierarchical neural system but by multiple areas in the brain that are fed by parallel but interacting neural pathways. Because distributed processing is one of the main organizational principles in the neurobiology of vision, one must have a grasp of the anatomical pathways of the visual system to understand fully the physiological description of visual processing in later chapters.

In this chapter, we lay the foundation for understanding the neural circuitry and organizational principles of the visual pathways. These principles apply quite broadly and are relevant not only for the multiple areas of the brain concerned with vision but also for other types of sensory information processing by the brain.

## Visual Processing Is Mediated by the Geniculostriate Pathway

The brain's analysis of visual scenes begins in the two retinas, which transform visual input using a strategy of parallel processing (Chapter 22). This important neural computation strategy is utilized at all stages of the visual pathway as well as in other sensory areas. The pixel-like bits of visual input falling on individual photoreceptors—rods and cones—are analyzed by retinal circuits to extract some 20 local features, such as the local contrasts of dark versus light, red versus green, and blue versus yellow. These features are computed by different populations of specialized neural circuits forming independent processing modules that separately cover the visual field. Thus, each point in the visual field is processed in multiple channels that extract distinct aspects of the visual input simultaneously and in parallel. These parallel streams are then sent out along the axons of the retinal ganglion cells, the projection neurons of the retina, which form the optic nerves.

From the eye, the optic nerve extends to a midline crossing point, the optic chiasm. Beyond the chiasm, the fibers from each temporal hemiretina proceed to the ipsilateral hemisphere along the ipsilateral optic tract; fibers from the nasal hemiretinas cross to the contralateral hemisphere along the contralateral optic tract (Figure 21–5). Because the temporal hemiretina of one eye sees the same half of the visual field (hemifield) as the nasal hemiretina of the other, the partial decussation of fibers at the chiasm ensures that all the information about each hemifield is processed in the visual cortex of the contralateral hemisphere. The layout of the pathway also forms the basis for useful diagnostic information. As a consequence of the particular anatomy of this visual pathway, lesions at different points along the pathway lead to visual deficits with different geometric shapes (Figure 21–5) that can be distinguished reliably through clinical examination. The deficit could be entirely monocular; if present in both eyes, it could affect noncorresponding or corresponding parts of the visual field in the two eyes; it could be restricted to either the upper or the lower visual field or may extend into both, etc. Thus, the shape of the deficit could give valuable clues about type and location of the underlying nerve damage or occlusion (ranging from optic nerve degeneration, such as due to multiple sclerosis, to tumors, strokes, or physical trauma).

Figure 21–5

**Representation of the visual field along the visual pathway.** Each eye sees most of the visual field, with the exception of a portion of the peripheral visual field known as the monocular crescent. The axons of retinal neurons (ganglion cells) carry information from each visual hemifield along the optic nerve up to the optic chiasm, where fibers from the nasal hemiretina cross to the opposite hemisphere. Fibers from the temporal hemiretina stay on the same side, joining the fibers from the nasal hemiretina of the contralateral eye to form the optic tract. The optic tract carries information from the opposite visual hemifield originating in both eyes and projects into the lateral geniculate nucleus. Cells in this nucleus send their axons along the optic radiation to the primary visual cortex.

Lesions along the visual pathway produce specific visual field deficits, as shown on the *right*:

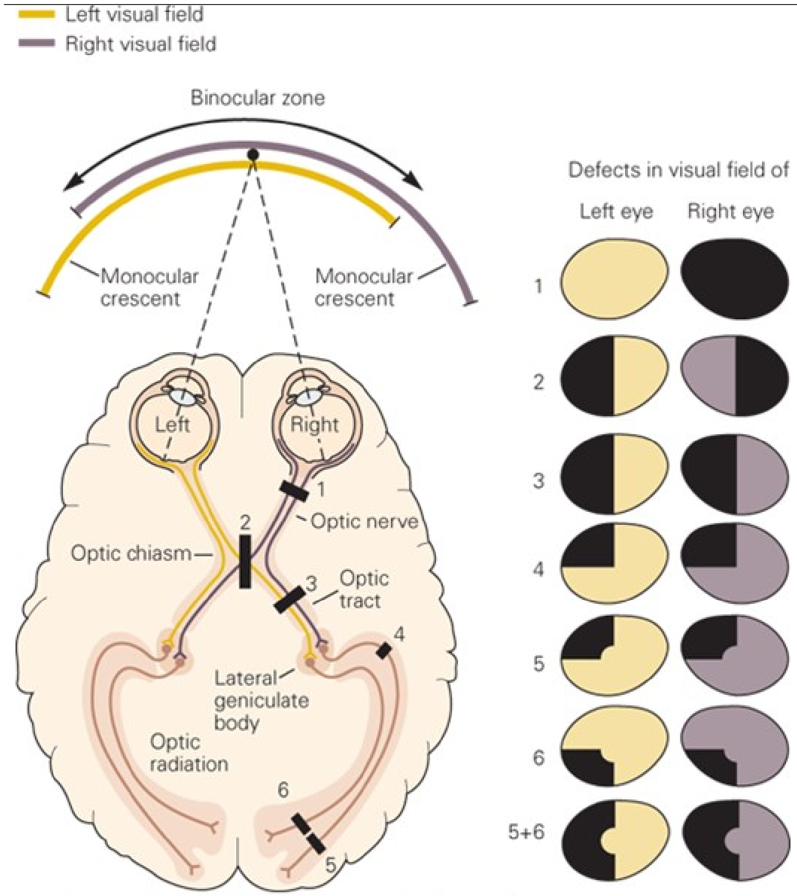
**1.** A lesion of an optic nerve causes a total loss of vision in one eye.

**2.** A lesion of the optic chiasm causes a loss of vision in the temporal half of each visual hemifield (bitemporal hemianopsia).

**3.** A lesion of the optic tract causes a loss of vision in the opposite half of the visual hemifield (contralateral hemianopsia).

**4.** A lesion of the optic radiation fibers that curve into the temporal lobe (Meyer's loop) causes loss of vision in the upper quadrant of the contralateral visual hemifield in both eyes (upper contralateral quadrantic anopsia).

**5,6.** Partial lesions of the visual cortex lead to deficits in portions of the contralateral visual hemifield. For example, a lesion in the upper bank of the calcarine sulcus (**5**) causes a partial deficit in the inferior quadrant, while a lesion in the lower bank (**6**) causes a partial deficit in the superior quadrant. The central area of the visual field tends to be unaffected by cortical lesions because of the extent of the representation of the fovea and the duplicate representation of the vertical meridian in the hemispheres.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Beyond the optic chiasm, the axons from nasal and temporal hemiretinas carrying input from one hemifield join in the optic tract, which extends to the lateral geniculate nucleus (LGN) of the thalamus. The LGN in primates consists of six primary layers: four parvocellular (Latin *Parvus*, small) and two magnocellular, each paired with a thin but dense intercalated or koniocellular (Greek *konio*, dust) layer (see Figure 21–14). The term "koniocellular" refers to the substantially smaller cell bodies in these layers relative to those of magnocellular or parvocellular layers. The parallel channels established in the retinas remain anatomically segregated through the LGN. Parvocellular layers get input from the midget retinal ganglion cells, which are the most numerous in the primate retina (~70%) and carry red-green opponent information (Chapter 22). Magnocellular layers get achromatic contrast information from the parasol ganglion cells (~10%). Koniocellular layers get input from the small and large bistratified ganglion cells, carrying blue-yellow information, that together make up the third most populous set of retinal projections to the LGN (~8%). Koniocellular layers also get inputs from a number of other numerically much smaller classes of retinal ganglion cells.

Each geniculate layer receives input from either the ipsilateral or the contralateral eye (see Figure 21–12) but is aligned so as to come from a matching region of the contralateral hemifield. Thus, they form a set of concordant maps stacked atop one another. The thalamic neurons then relay retinal information to the primary visual cortex. But the LGN is not simply a relay; the retinal information it receives can be strongly modulated by attention and arousal through inhibitory connections to this brain region and by feedback from the visual cortex.

The primary visual pathway is also called the geniculostriate pathway (Figure 21–6A) because it passes through the LGN on its way to the primary visual cortex (V1), also known as the striate cortex because of the myelin-rich stripe that runs through its middle layers. A second pathway extends from the retina to the pretectal area of the midbrain, where neurons mediate the pupillary reflexes that control the amount of light entering the eyes (Figure 21–6B). A third pathway from the retina runs to the superior colliculus and is important in controlling eye movements. This pathway continues to the pontine formation in the brain stem and then to the extraocular motor nuclei (Figure 21–6C).

Figure 21–6

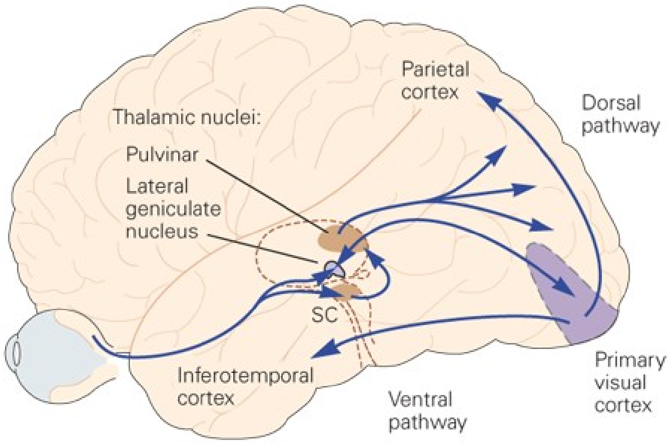
**Pathways for visual processing, pupillary reflex and accommodation, and control of eye position.**

**A.** *Visual processing.* The eye sends information first to thalamic nuclei, including the lateral geniculate nucleus and pulvinar, and from there to cortical areas. Cortical projections go forward from the primary visual cortex to areas in the parietal lobe (the dorsal pathway, which is concerned with visually guided movement) and areas in the temporal lobe (the ventral pathway, which is concerned with object recognition). The pulvinar also serves as a relay between cortical areas to supplement their direct connections. (Abbreviation: **SC**, superior colliculus).

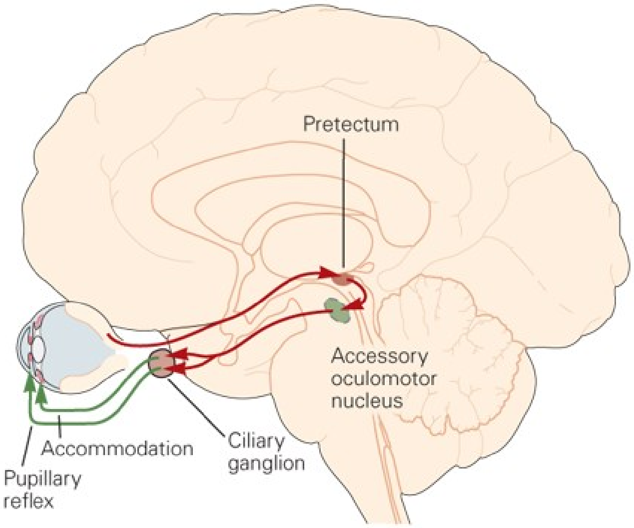
**B.** *Pupillary reflex and accommodation.* Light signals are relayed through the midbrain pretectum, to preganglionic parasympathetic neurons in the accessory oculomotor (Edinger-Westphal) nucleus, and out through the parasympathetic outflow of the oculomotor nerve to the ciliary ganglion. Postganglionic neurons innervate the smooth muscle of the pupillary sphincter, as well as the muscles controlling the lens.

**C.** *Eye movement.* Information from the retina is sent to the superior colliculus (**SC**) directly along the optic nerve and indirectly through the geniculostriate pathway to cortical areas (primary visual cortex, posterior parietal cortex, and frontal eye fields) that project back to the superior colliculus. The colliculus projects to the pons (**PPRF**), which then sends control signals to oculomotor nuclei, including the abducens nucleus, which controls lateral movement of the eyes. (Abbreviations: **FEF**, frontal eye field; **LGN**, lateral geniculate nucleus; **PPRF**, paramedian pontine reticular formation.)

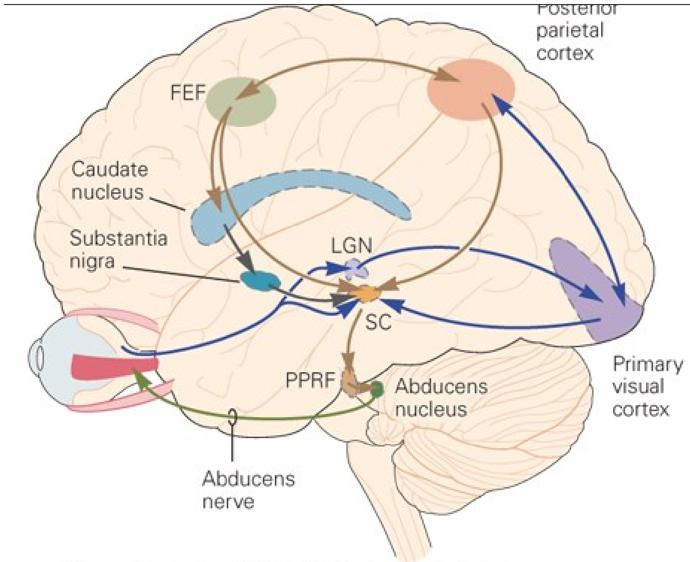
A Visual processing



B Pupillary reflex and accommodation



C Eye movement (horizontal)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Each LGN projects to the primary visual cortex through a pathway known as the optic radiation. These afferent fibers form a complete neural map of the contralateral visual field in the primary visual cortex. Beyond the striate cortex lie the extrastriate areas, a set of higher-order visual areas that are also organized as neural maps of the visual field. The preservation of the spatial arrangement of inputs from the retina is called retinotopy, and a neural map of the visual field is described as retinotopic or having a retinotopic frame of reference.

The primary visual cortex constitutes the first level of cortical processing of visual information. From there, information is transmitted over two major pathways. A ventral pathway into the temporal lobe carries information about what the stimulus is, and a dorsal pathway into the parietal lobe carries information about where the stimulus is, information that is critical for guiding movement.

A major fiber bundle called the corpus callosum connects the two hemispheres, transmitting information across the midline. The primary visual cortex in each hemisphere represents slightly more than half the visual field, with the two hemifield representations overlapping at the vertical meridian. One of the functions of the corpus callosum is to unify the perception of objects spanning the vertical meridian by linking the cortical areas that represent opposite hemifields.

## Form, Color, Motion, and Depth Are Processed in Discrete Areas of the Cerebral Cortex

In the late 19th and early 20th centuries, the cerebral cortex was differentiated into discrete regions by the anatomist Korbinian Brodmann and others using anatomical criteria. The criteria included the size, shape, and packing density of neurons in the cortical layers and the thickness and density of myelin. The functionally distinct cortical areas we have considered heretofore correspond only loosely to Brodmann's classification. The primary visual cortex (V1) is identical to Brodmann's area 17. In the extrastriate cortex, the secondary visual area (V2) corresponds to area 18. Beyond that, however, area 19 contains several functionally distinct areas that generally cannot be defined by anatomical criteria.

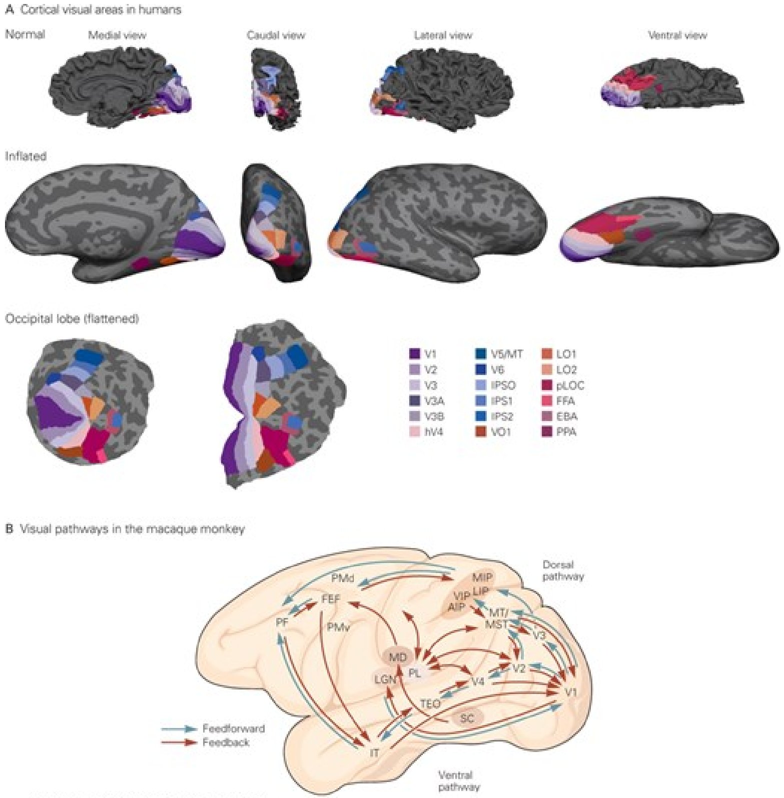
The number of functionally discrete areas of visual cortex varies between species. Macaque monkeys have more than 30 areas. Although not all visual areas in humans have yet been identified, the number is likely to be at least as great as in the macaque. If one includes oculomotor areas and prefrontal areas contributing to visual memory, almost half of the cerebral cortex is involved with vision. Functional magnetic resonance imaging (fMRI) has made it possible to establish homologies between the visual areas of the macaque and human brains (Figure 21–7). Based on pathway tracing studies in monkeys, we now appreciate that these areas are organized in functional streams (Figure 21–7B).

Figure 21–7

**Visual pathways in the cerebral cortex.**

**A.** Functional magnetic resonance imaging shows areas of the human cerebral cortex involved in visual processing. The **top row** shows areas on the gyri and sulci of a normal view of a brain; the **middle row** shows “inflated” views of the brain following a computational process that simulates inflating the brain like a balloon so as to stretch out the “wrinkles” of gyri and sulci into a smooth surface while minimizing local distortions. Light and dark gray regions identify gyri and sulci, respectively; the **bottom row** shows a two-dimensional representation of the occipital lobe (*left*) and a representation with less distortion by making a cut along the calcarine fissure. Different approaches are required for demarcating different functional areas. Retinotopic areas, by definition, contain continuous maps of visual space and are identified using stimuli such as rotating spirals or expanding circles that sweep through visual space. Maps in adjacent cortical areas run in opposite directions on the cortical surface and meet along boundaries of local mirror reversals. These mirror reversals can be used to identify area boundaries and thus demarcate each area. These retinotopic areas, including early visual areas V1, V2, and V3, and areas V3A, V3B, V6, hV4, VO1, LO1, LO2, and V5/MT, share boundaries in pairs; these boundaries converge (at the representation of the fovea) at the occipital pole. A different approach, identifying loci of attention, is used to map areas IPS1 and IPS2. Yet further sets of approaches or responsiveness to specific attributes or classes of objects (such as faces) are used for less strictly retinotopic areas. Functional specificity has been demonstrated for a number of visual areas: VO1 is implicated in color processing, the lateral occipital complex (**LO2**, **pLOC**) codes object shape, fusiform face area (**FFA**) codes faces, the parahippocampal place area (**PPA**) responds more strongly to places than to objects, the extrastriate body area (**EBA**) responds more strongly to body parts than objects, and V5/MT is involved in motion processing. Areas in the intraparietal sulcus (**IPS1** and **IPS2**) are involved in control of spatial attention and saccadic eye movements. (Images courtesy of V. Piech, reproduced with permission.)

**B.** In the macaque monkey, V1 is located on the surface of the occipital lobe and sends axons in two pathways. A dorsal pathway courses through a number of areas in the parietal lobe and into the frontal lobe and mediates attentional control and visually guided movements. A ventral pathway projects through V4 into areas of the inferior temporal cortex and mediates object recognition. In addition to feedforward pathways extending from primary visual cortex into the temporal, parietal, and frontal lobes (**blue arrows**), reciprocal or feedback pathways run in the opposite direction (**red arrows**). Feedforward and feedback can operate directly, between cortical areas, or indirectly, via the thalamus, in particular the pulvinar, which acts as a relay between cortical areas. The subcortical pathways involved include thalamic nuclei—the lateral geniculate nucleus (**LGN**), pulvinar nucleus (**PL**), and mediodorsal nucleus (**MD**)—and the superior colliculus (**SC**). (Abbreviations: **AIP**, anterior intraparietal area; **FEF**, frontal eye field; **IT**, inferior temporal cortex; **LIP**, lateral intraparietal area; **MIP**, medial intraparietal area; **MT**, middle temporal area; **PF**, prefrontal cortex; **PMd**, dorsal premotor cortex; **PMv**, ventral premotor cortex; **TEO**, posterior division of area IT; **V1**, primary visual cortex, Brodmann’s area 17; **V2**, secondary visual area, Brodmann’s area 18; **V3**, **V4**, third and fourth visual areas; **VIP**, ventral intraparietal area.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

The visual areas of cortex can be differentiated by the functional properties of their neurons. Studies of such functional properties have revealed that the visual areas are organized in two hierarchical pathways, a ventral pathway involved in object recognition and a dorsal pathway dedicated to the use of visual information for guiding movements. The ventral or object-recognition pathway extends from the primary visual cortex to the temporal lobe; it is described in detail in Chapter 24. The dorsal or movement-guidance pathway connects the primary visual cortex with the parietal lobe and then with the frontal lobes.

The pathways are interconnected so that information is shared. For example, movement information in the dorsal pathway can contribute to object recognition through kinematic cues. Information about movements in space derived from areas in the dorsal pathway is therefore important for the perception of object shape and is fed into the ventral pathway.

All connections between cortical areas are reciprocal—each area sends information back to the areas from which it receives input. These feedback connections provide information about cognitive functions, including spatial attention, stimulus expectation, and emotional content, to earlier levels of visual processing. The pulvinar in the thalamus serves as a relay between cortical areas (see Figure 21–7B).

The dorsal pathway courses through the parietal cortex, a region that uses visual information to direct the movement of the eyes and limbs, that is, for visuomotor integration. The lateral intraparietal area, named for its location in the intraparietal sulcus, is involved in representing points in space that are the targets of eye movements or reaching. Patients with lesions of parietal areas fail to attend to objects on one side of the body, a syndrome called *unilateral neglect* (see Figure 59–1 in Chapter 59).

The ventral pathway extends into the temporal lobe. The inferior temporal cortex stores information about the shapes and identities of objects; one portion represents faces, for damage to that region results in the inability to recognize faces (*prosopagnosia*).

The dorsal and ventral pathways each comprise a hierarchical series of areas that can be delineated by several criteria. First, at many relays, the array of inputs forms a map of the visual hemifield. The boundaries of these maps can be used to demarcate the boundaries of visual areas. This is particularly useful at early levels of the pathway where the receptive fields of neurons are small and visuotopic maps are precisely organized (see the next section for the definition of receptive field). At higher levels, however, the receptive fields become larger, the maps less precise, and visuotopic organization is therefore a less reliable basis to delineate the boundaries of an area.

Another means to differentiate one area from another, as shown by experiments in monkeys, depends upon the distinctive functional properties exhibited by the neurons in each area. The clearest example of this is an area in the dorsal pathway, the middle temporal area (MT or V5), which contains neurons with a strong selectivity for the direction of movement across their receptive fields. Consistent with the idea that the middle temporal area is involved in the analysis of motion, lesions of this area produce deficits in the ability to track moving objects.

A classical view of the organization of visual cortical areas is a hierarchical one, where the areas at the bottom of the hierarchy, such as V1 and V2, represent the visual primitives of orientation, direction of movement, depth, and color. In this view, the top of the ventral pathway's hierarchy would represent whole objects, with the areas in between representing intermediate level vision. This idea of "complexification" along the hierarchy suggests a mapping between the levels of visual perception and stages in the sequence of cortical areas. But more recent findings indicate a more complex story, where even the primary visual cortex plays a role in intermediate-level vision, and neurons in the higher areas may process information on components of objects. Moreover, as shown in Figure 21–7, one also has to take into account the fact that there is a powerful reverse flow of information, or feedback, from the "higher" to the "lower" cortical areas. As will be described in Chapter 23, this reverse direction of information contains higher order "top-down" cognitive influences including attention, object expectation, perceptual task, perceptual learning, and efference copy. Top-down influences may play a role in scene segmentation, object relationships, and perception of object details, as well as object recognition itself.

## The Receptive Fields of Neurons at Successive Relays in the Visual Pathway Provide Clues to How the Brain Analyzes Visual Form

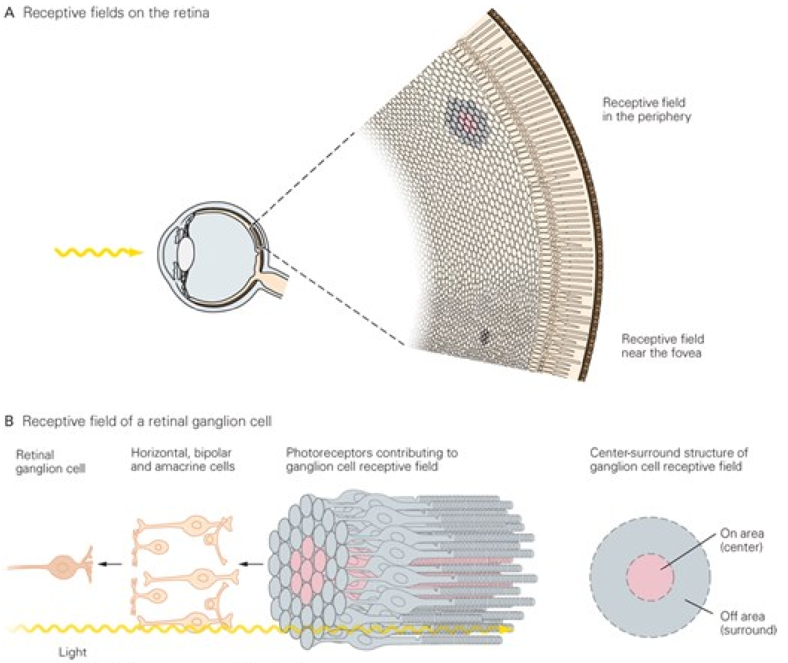
In 1906, Charles Sherrington coined the term *receptive field* in his analysis of the scratch withdrawal reflex: "The whole collection of points of skin surface from which the scratch-reflex can be elicited is termed the receptive field of that reflex." When it became possible to record from single neurons in the eye, H. Keffer Hartline applied the concept of the receptive field in his study of the retina of the horseshoe crab, *Limulus:* "The region of the retina which must be illuminated in order to obtain a response in any given fiber … is termed the receptive field of that fiber." In the visual system, a neuron's receptive field represents a small window on the visual field (Figure 21–8).

Figure 21–8

**Receptive fields of retinal ganglion cells in relation to photoreceptors**.

**A.** The number of photoreceptors contributing to the receptive field of a retinal ganglion cell varies depending on the location of the receptive field on the retina. A cell near the fovea receives input from fewer receptors covering a smaller area, whereas a cell farther from the fovea receives input from many more receptors covering a larger area (see Figure 21–10).

**B.** Light passes through nerve cell layers to reach the photoreceptors at the back of the retina. Signals from the photoreceptors are then transmitted by neurons in the outer and inner nuclear layers to a retinal ganglion cell.



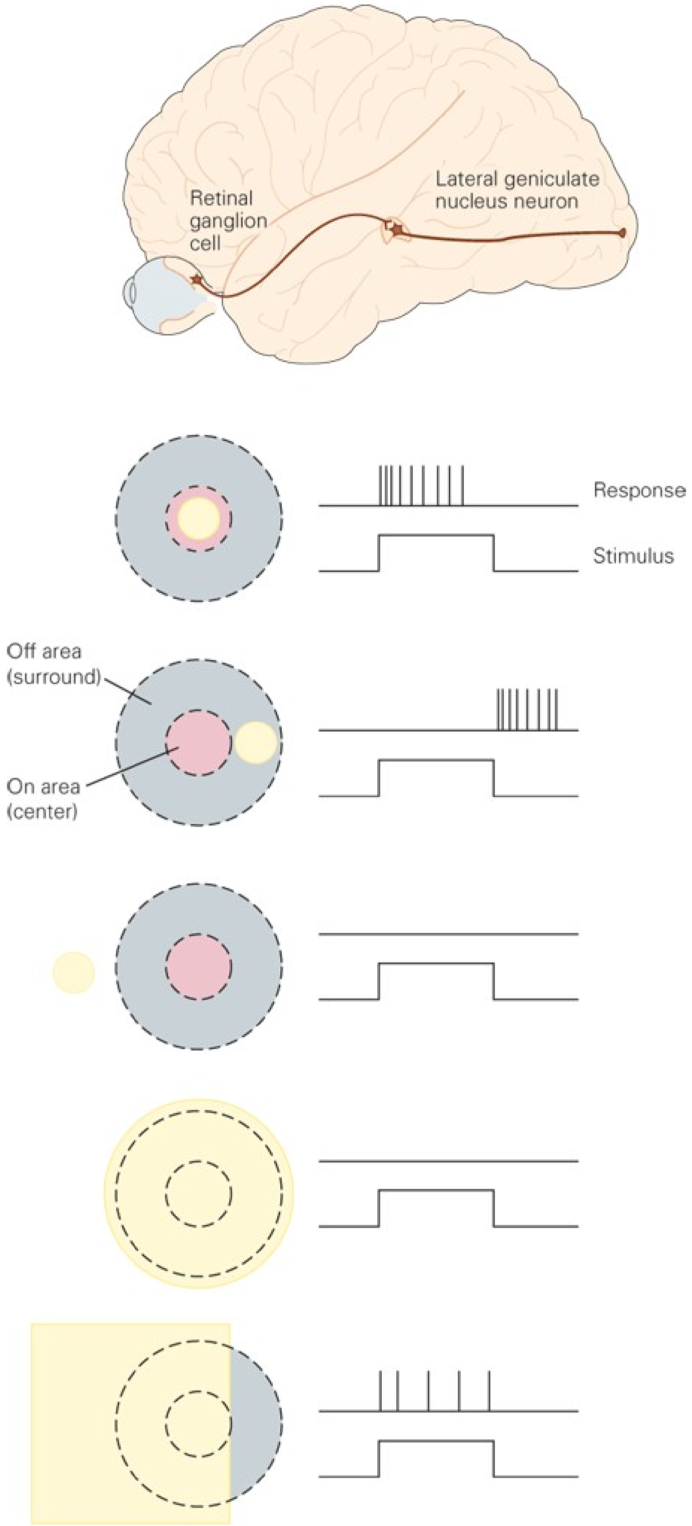
Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

But responses to only one spot of light yielded a limited understanding of a cell's receptive field. Using two small spots of light, both Hartline and Stephen Kuffler, who studied the mammalian retina, found an inhibitory surround or lateral inhibitory region in the receptive field. In 1953, Kuffler observed that "not only the areas from which responses can actually be set up by retinal illumination may be included in a definition of the receptive field but also all areas which show a functional connection, by an inhibitory or excitatory effect on a ganglion cell." Kuffler thus demonstrated that the receptive fields of retinal ganglion cells have functionally distinct subareas. These receptive fields have a center-surround organization and fall into one of two categories: *on-center* and *off-center*. Later work demonstrated that neurons in the LGN have similar receptive fields.

The on-center cells fire when a spot of light is turned on within a circular central region. Off-center cells fire when a spot of light in the center of their receptive field is turned off. The surrounding annular region has the opposite sign. For on-center cells, a light stimulus anywhere in the annulus surrounding the center produces a response when the light is turned off, a response termed *on-center, off-surround*. The center and surround areas are mutually inhibitory (Figure 21–9). When both center and surround are illuminated with diffuse light, there is little or no response. Conversely, a light–dark boundary across the receptive field produces a brisk response. Because these neurons are most sensitive to borders and contours—to differences in illumination as opposed to uniform surfaces—they encode information about contrast in the visual field.

Figure 21–9

**Receptive fields of neurons at early relays of visual pathways.** A circular symmetric receptive field with mutually antagonistic center and surround is characteristic of retinal ganglion cells and neurons in the lateral geniculate nucleus of the thalamus. The center can respond to the turning on or turning off of a spot of light (**yellow**) depending on whether the receptived field belongs to an "on-center" or "off-center" class, respectively. The surround has the opposite response. Outside the surround, there is no response to light, thus defining the receptive field boundary. The response is weak when light covers both the center and surround, so these neurons respond optimally to contrast (a light–dark boundary) in the visual field.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack,
Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

The size on the retina of a receptive field varies both according to the field's *eccentricity*—its position relative to the fovea, the central part of the retina where visual acuity is highest—and the position of neurons along the visual pathway. Receptive fields with the same eccentricity are relatively small at early levels in visual processing and become progressively larger at later levels. The size of the receptive field is expressed in terms of degrees of visual angle; the entire visual field covers nearly 180° (Figure 21–10A). In early relays of visual processing, the receptive fields near the fovea are the smallest. The receptive fields for retinal ganglion cells that monitor portions of the fovea subtend approximately 0.1°, whereas those in the visual periphery can be a couple of orders of magnitude larger.

Figure 21–10

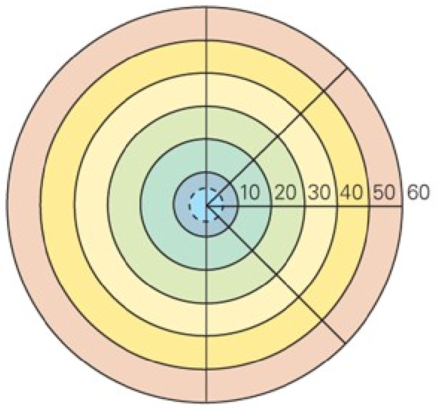
**Receptive field size, eccentricity, retinotopic organization, and magnification factor.** The color code refers to position in visual space or on the retina.

**A.** The distance of a receptive field from the fovea is referred to as the eccentricity of the receptive field.

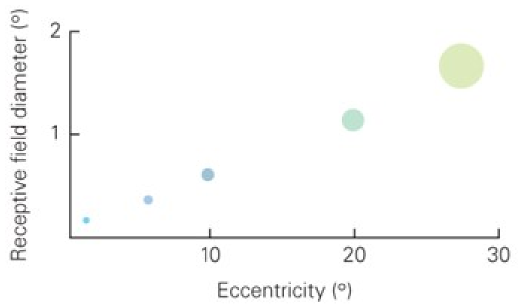
**B.** Receptive field size varies with distance from the fovea. The smallest fields lie in the center of gaze, the fovea, where the visual resolution is highest; fields become progressively larger with distance from the fovea.

**C.** The amount of cortical area dedicated to inputs from within each degree of visual space, known as the magnification factor, also varies with eccentricity. The central part of the visual field commands the largest area of cortex. For example, in area V1, more area is dedicated to the central 10° of visual space than to all the rest. The map of V1 shows the cortical sheet unfolded.

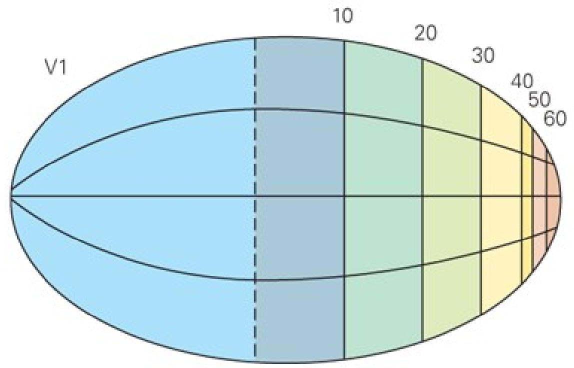
A Map of retinal eccentricity



B Receptive field size varies systematically with eccentricity



C Cortical magnification varies with eccentricity



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

The amount of cortex dedicated to a degree of visual space changes with eccentricity. More area of cortex is dedicated to the central part of the visual field, where the receptive fields are smallest and the visual system has the greatest spatial resolution (Figure 21–10C).

Receptive-field properties change from relay to relay along a visual pathway. By determining these properties, one can assay the function of each relay nucleus and how visual information is progressively analyzed by the brain. For example, the change in receptive-field structure that occurs between the LGN and cerebral cortex reveals an important mechanism in the brain's analysis of visual form. The key property of the form pathway is selectivity for the orientation of contours in the visual field. This is an emergent property of signal processing in primary visual cortex; it is not a property of the cortical input but is generated within the cortex itself.

Whereas retinal ganglion cells and neurons in the LGN have concentric center-surround receptive fields, those in the cortex, although equally sensitive to contrast, also analyze contours. David Hubel and Torsten Wiesel discovered this characteristic in 1958 while studying what visual stimuli provoked activity in neurons in the primary visual cortex. While showing an anesthetized animal slides containing a variety of images, they recorded extracellularly from individual neurons in the visual cortex. As they switched from one slide to another, they found a neuron that produced a brisk train of action potentials. The cell was responding not to the image on the slide but to the edge of the slide as it was moved into position.

## The Visual Cortex Is Organized Into Columns of Specialized Neurons

The dominant feature of the functional organization of the primary visual cortex is the visuotopic organization of its cells: the visual field is systematically represented across the surface of the cortex (Figure 21–11A).

Figure 21–11

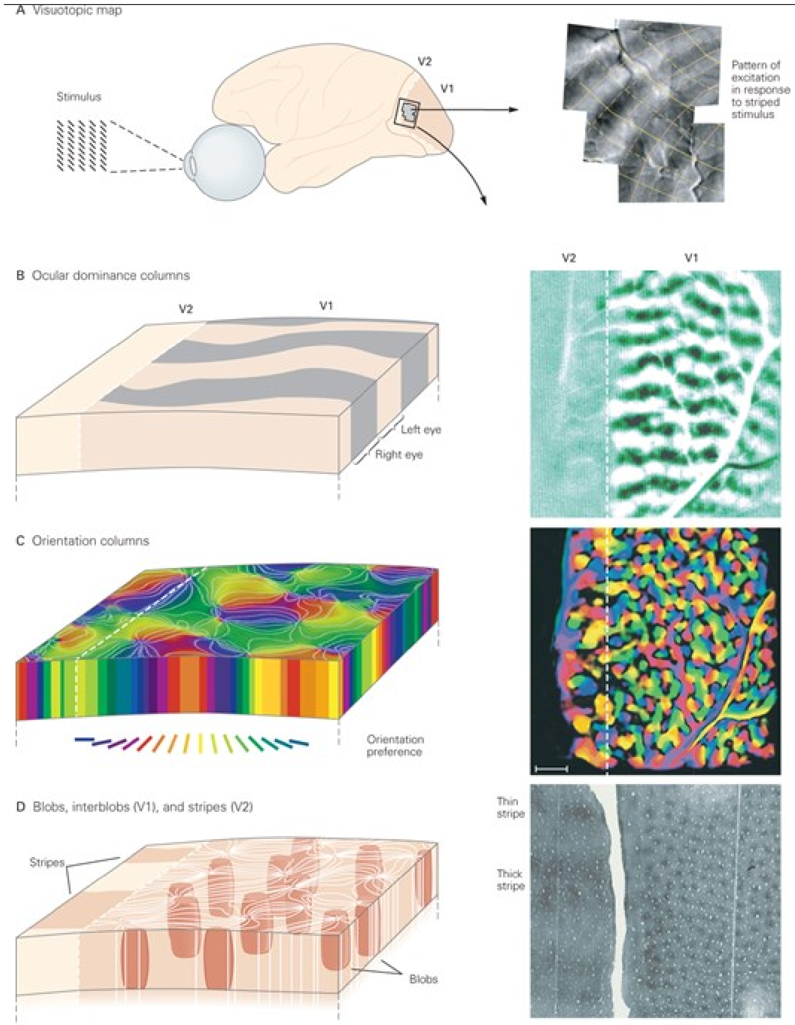
**Functional architecture of the primary visual cortex.** (Courtesy of M. Kinoshita and A. Das, reproduced with permission.)

**A.** The surface of the primary visual cortex is functionally organized as a map of the visual field. The elevations and azimuths of visual space are organized in a regular grid that is distorted because of variation in the magnification factor (see Figure 21–10). The grid is visible here in the dark stripes (visualized with intrinsic-signal optical imaging), which reflect the pattern of neurons that responded to a series of vertical candy stripes. Within this surface map, one finds repeated superimposed cycles of functionally specific columns of cells, as illustrated in **B**, **C**, and **D**.

**B.** The dark and light stripes represent the surface view of the left and right ocular dominance columns. These stripes intersect the border between areas V1 and V2, the representation of the vertical meridian, at right angles.

**C.** Some columns contain cells with similar selectivity for the orientation of stimuli. The different colors indicate the orientation preference of the columns. The orientation columns in surface view are best described as pinwheels surrounding singularities of sudden changes in orientation (the center of the pinwheel). The scale bar represents 1 mm. (Surface image of orientation columns on the left courtesy of G. Blasdel, reproduced with permission.)

**D.** Patterns of blobs in V1 and stripes in V2 represent other modules of functional organization. These patterns are visualized with cytochrome oxidase.



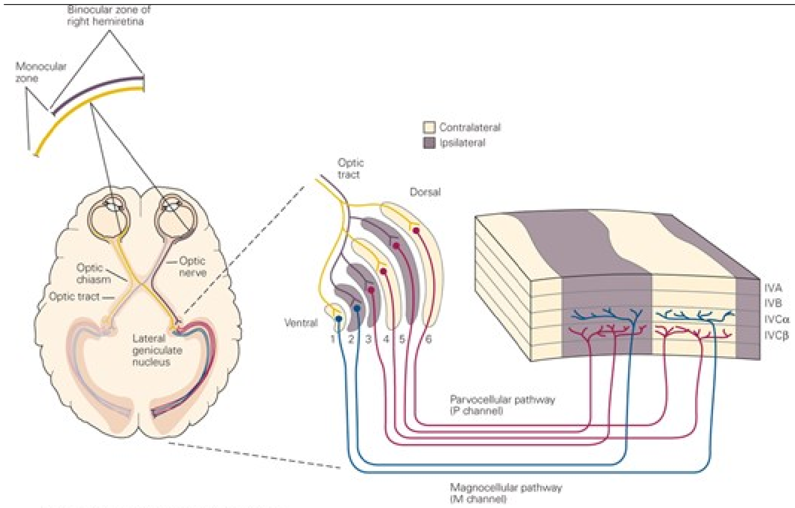
Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

In addition, cells in the primary visual cortex with similar functional properties are located close together in columns that extend from the surface of the cortex to the white matter. The columns are concerned with the functional properties that are analyzed in any given cortical area and thus reflect the functional role of that area in vision. The properties that are developed in the primary visual cortex include orientation specificity and the integration of inputs from the two eyes, which is measured as the relative strength of input from each eye, or ocular dominance.

Ocular-dominance columns reflect the segregation of thalamocortical inputs arriving from different layers of the LGN. Alternating layers of this nucleus receive input from retinal ganglion cells located in either the ipsilateral or contralateral retina (Figure 21–12). This segregation is maintained in the inputs from the LGN to the primary visual cortex, producing the alternating left-eye and right-eye ocular dominance bands (Figure 21–11B).

Figure 21–12

**Projections from the lateral geniculate nucleus to the visual cortex.** The lateral geniculate nucleus in each hemisphere receives input from the temporal retina of the ipsilateral eye and the nasal retina of the contralateral eye. The nucleus is a layered structure comprising four parvocellular layers (layers 3 to 6) and two magnocellular layers (layers 1 and 2). Each is paired with an intercalated koniocellular layer. (These layers are represented here by the gaps separating the primary layers. They are unlabeled to avoid clutter. See Figure 21–14.) The inputs from the two eyes terminate in different geniculate layers: The contralateral eye projects to layers 1, 4, and 6, whereas the ipsilateral eye sends input to layers 2, 3, and 5. Neurons from these geniculate layers then project to different layers of cortex. The parvocellular geniculate neurons project to layer IVCβ, the magnocellular ones project to layer IVCα, and the koniocellular ones project to “blobs” in the upper cortical layers (see Figures 21–14 and 21–15). In addition, the afferents from the ipsilateral and contralateral layers of the lateral geniculate nucleus are segregated into alternating ocular-dominance columns.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

Cells with similar orientation preferences are also grouped into columns. Across the cortical surface, there is a regular clockwise and counterclockwise cycling of orientation preference, with the full 180° cycle repeating every 750 µm (Figure 21–11C). Likewise, the left- and right-eye dominance columns alternate with a periodicity of 750 to 1,000 µm. One full cycle of orientation columns, or a full pair of left- and right-eye dominance columns, is called a *hypercolumn*. The orientation and ocular dominance columns at each point on the cortical surface are locally roughly orthogonal to each other. Thus, a cortical patch one hypercolumn in extent contains all possible combinations of orientation preference and left- and right-eye dominance.

Both types of columns were first mapped by recording the responses of neurons at closely spaced electrode penetrations in the cortex. The ocular-dominance columns were also identified by making lesions or tracer injections in individual layers of the LGN. More recently, a technique known as optical imaging has enabled researchers to visualize a surface representation of the orientation and ocular dominance columns in living animals. Developed for studies of cortical organization by Amiram Grinvald, this technique visualizes changes in surface reflectance associated with the metabolic requirements of active groups of neurons, known as intrinsic-signal optical imaging, or changes in fluorescence of voltage-sensitive dyes. Intrinsic-signal imaging depends on activity-associated changes in local blood flow and alterations in the oxidative state of hemoglobin and other intrinsic chromophores. These techniques are also now being complemented with imaging at cellular resolution using genetically encoded markers of neural activity.

An experimenter can visualize the distribution of cells with left or right ocular dominance, for example, by subtracting the image obtained while stimulating one eye from that acquired while stimulating the other. When viewed in a plane tangential to the cortical surface, the ocular dominance columns appear as alternating left- and right-eye stripes, each approximately 750 µm in width (Figure 21–11B).

The cycles of orientation columns form various structures, from parallel stripes to pinwheels. Sharp jumps in orientation preference occur at the pinwheel centers and “fractures” in the orientation map (Figure 21–11C).

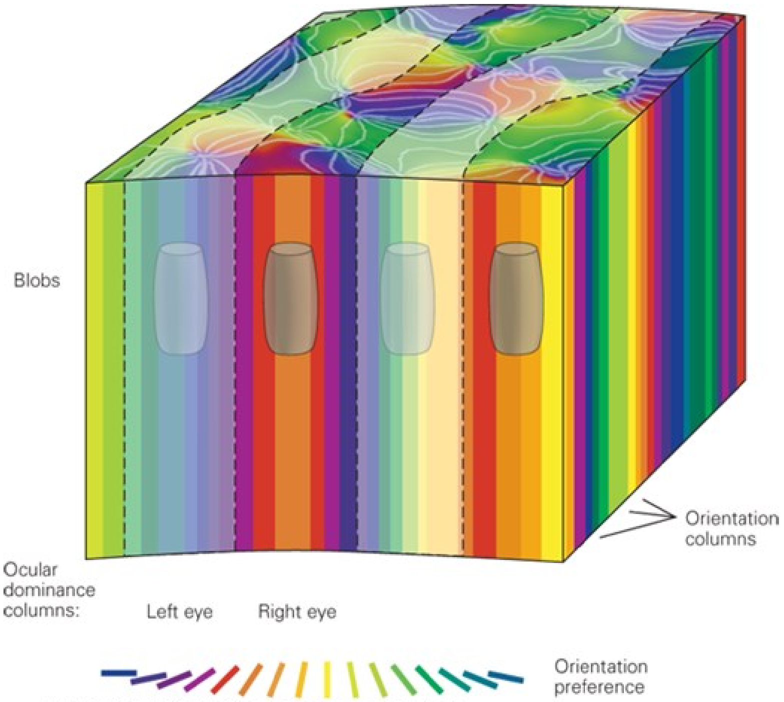
Embedded within the orientation and ocular-dominance columns are clusters of neurons that have poor orientation selectivity but strong color preferences. These units of specialization, located within the superficial layers, were revealed by a histochemical label for the enzyme cytochrome oxidase, which is distributed in a regular patchy pattern of blobs and interblobs. In the primary visual cortex, these blobs are a few hundred micrometers in diameter and 750 µm apart (Figure 21–11D). The blobs correspond to clusters of color-selective neurons. Because they are rich in cells with color selectivity and poor in cells with orientation selectivity, the blobs are specialized to provide information about surfaces rather than edges.

In area V2, thick and thin dark stripes separated by pale stripes are evident with cytochrome oxidase labeling (Figure 21–11D). The thick stripes contain neurons selective for direction of movement and for binocular disparity as well as cells that are responsive to illusory contours and global disparity cues. The thin stripes hold cells specialized for color. The pale stripes contain orientation-selective neurons.

For every visual attribute to be analyzed at each position in the visual field, there must be adequate tiling, or coverage, of neurons with different functional properties. As one moves in any direction across the cortical surface, the progression of the visuotopic location of receptive fields is gradual, whereas the cycling of columns occurs more rapidly. Any given position in the visual field can therefore be analyzed adequately in terms of the orientation of contours, the color and direction of movement of objects, and stereoscopic depth by a single computational module. The small segment of visual cortex that comprises such a module represents all possible values of all the columnar systems (Figure 21–13).

Figure 21–13

**A cortical computational module.** A chunk of cortical tissue roughly 1 mm square contains an orientation hypercolumn (a full cycle of orientation columns), one cycle of left- and right-eye ocular-dominance columns, and blobs and interblobs. This module would presumably contain all of the functional and anatomical cell types of primary visual cortex, which would be repeated hundreds of times to cover the visual field. (Adapted from Hubel 1988.)



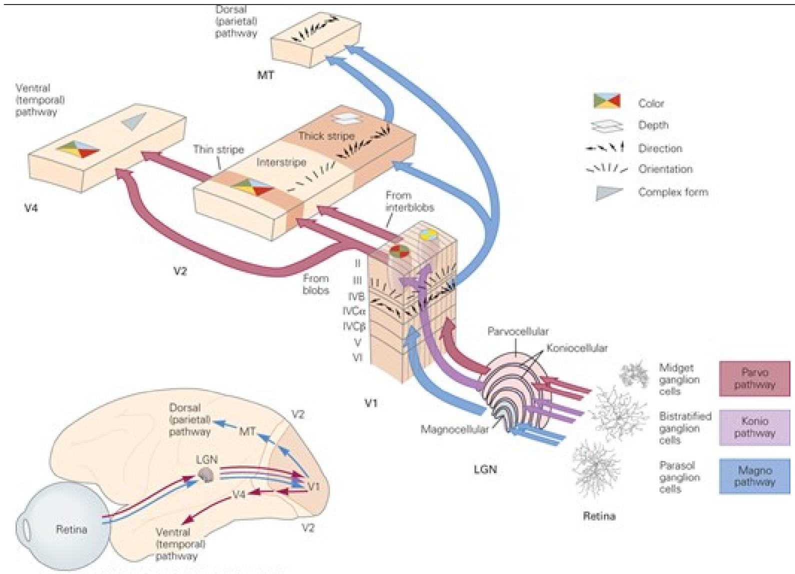
Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

The columnar systems serve as the substrate for two fundamental types of connectivity along the visual pathway. *Serial processing* occurs in the successive connections between cortical areas, connections that run from the back of the brain forward. At the same time, *parallel processing* occurs simultaneously in subsets of fibers that process different submodalities such as form, color, and movement, continuing the neural processing strategy started in the retina.

Many areas of visual cortex reflect this arrangement; for example, functionally specific cells in V1 communicate with cells of the same specificity in V2. These pathways are not absolutely segregated, however, for there is some mixing of information between different visual attributes (Figure 21–14).

Figure 21–14

**Parallel processing in visual pathways.** The ventral stream is primarily concerned with object identification, carrying information about form and color. The dorsal pathway is dedicated to visually guided movement, with cells selective for direction of movement. These pathways are not strictly segregated, however, and there is substantial interconnection between them even in the primary visual cortex. (Abbreviations: **LGN**, lateral geniculate nucleus; **MT**, middle temporal area.) (Retinal ganglion cell images courtesy of Dennis Dacey, reproduced with permission.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Columnar organization confers several advantages. It minimizes the distance required for neurons with similar functional properties to communicate with one another and allows them to share inputs from discrete pathways that convey information about particular sensory attributes. This efficient connectivity economizes on the use of brain volume and maximizes processing speed. The clustering of neurons into functional groups, as in the columns of the cortex, allows the brain to minimize the number of neurons required for analyzing different attributes. If all neurons were tuned for every attribute, the resultant combinatorial explosion would require a prohibitive number of neurons.

## Intrinsic Cortical Circuits Transform Neural Information

Each area of the visual cortex transforms information gathered by the eyes and processed at earlier synaptic relays into a signal that represents the visual scene. This transformation is accomplished by local circuits comprising both excitatory and inhibitory neurons.

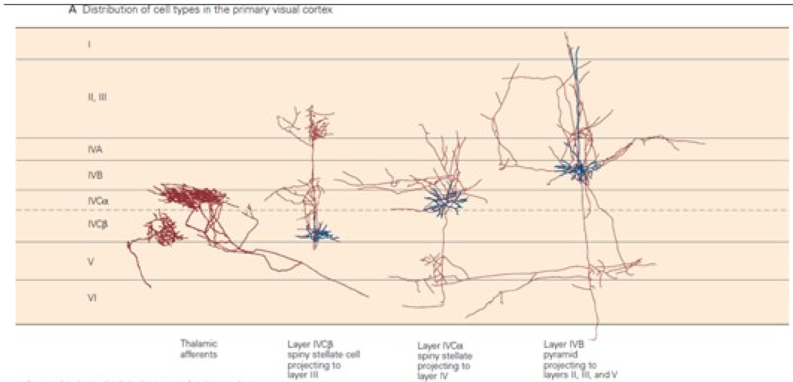
The principal input to the primary visual cortex comes from three parallel pathways that originate in the parvocellular, magnocellular, and the blue/yellow channels of koniocellular layers of the LGN (see Figure 21–12). Neurons in the parvocellular layers project to cortical layers IVCβ and 6, those in the magnocellular layers project to layer IVCα and 6, while the koniocellular neurons project to layer 1 and to the cytochrome oxidase blobs in layers 2 and 3. From there, a sequence of interlaminar connections, mediated by the excitatory spiny stellate neurons, processes visual information over a stereotyped set of connections (Figure 21–15).

Figure 21–15

**The intrinsic circuitry of the primary visual cortex.**

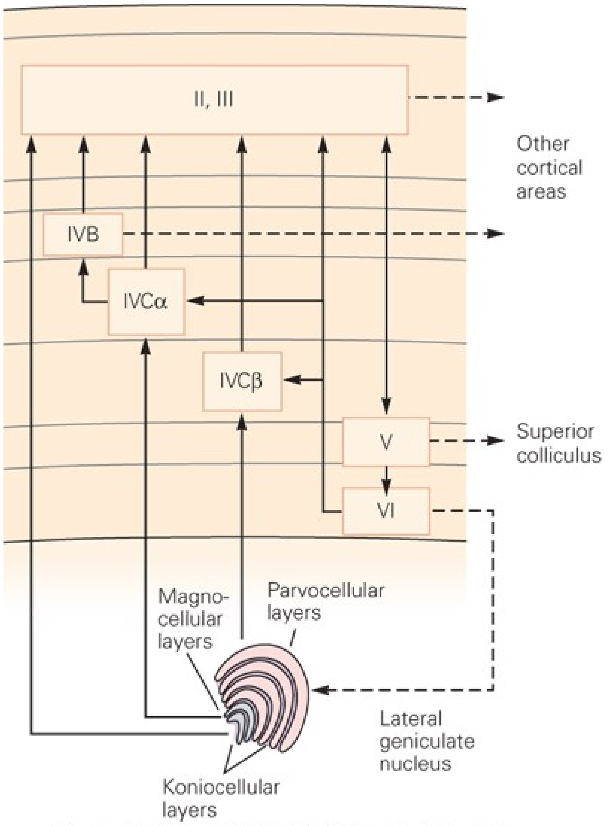
**A.** Examples of neurons in different cortical layers responsible for excitatory connections in cortical circuits. Layer IV is the principal layer of input from the lateral geniculate nucleus of the thalamus. Fibers from the parvocellular layer terminate in layer IVCβ, whereas the magnocellular fibers terminate in layer IVCα. The intrinsic cortical excitatory connections are mediated by spiny stellate and pyramidal cells. A variety of γ-aminobutyric acid (GABA)-ergic smooth stellate cells (not shown) are responsible for inhibitory connections. Dendritic arbors are colored **blue**, and axonal arbors are shown in **brown**. (Cortical neurons courtesy of E. Callaway, reproduced with permission. Thalamic afferents adapted, with permission, from Blasdel and Lund 1983. Copyright © 1983 Society for Neuroscience.)

**B.** Diagram of excitatory connections within the primary visual cortex. Output to other regions of cortex is sent from every layer of visual cortex.

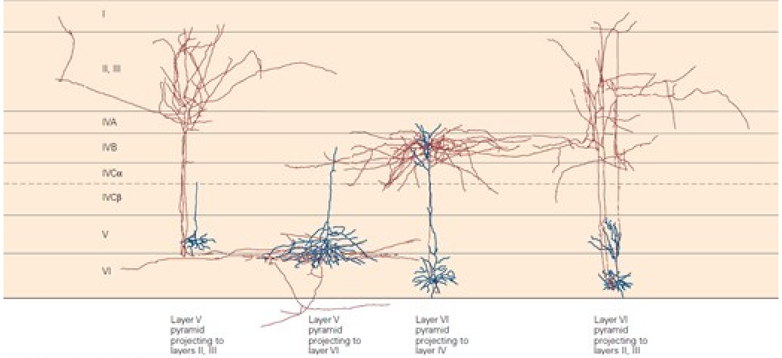


Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

B Simplified diagram of intrinsic circuitry



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

This characterization of parallel pathways is only an approximation, as there is considerable interaction between the pathways. This interaction is the means by which various visual features—color, form, depth, and movement—are linked, leading to a unified visual percept. One way this linkage, or binding, may be accomplished is through cells that are tuned to more than one attribute.

At each stage of cortical processing, pyramidal neurons extend output to other brain areas. Superficial-layer cells are responsible for connections to higher-order areas of cortex. Layer V pyramidal neurons project to the superior colliculus and pons in the brain stem. Layer VI cells are responsible for feedback projections, both to the thalamus and to lower-order cortical areas.

Neurons in different layers have distinctive receptive-field properties. Neurons in the superficial layer of V1 have small receptive fields, whereas neurons in deeper layers have large ones. The superficial-layer neurons are specialized for high-resolution pattern recognition. Neurons in the deeper layers, such as those in layer V that are selective for the direction of movement, are specialized for the tracking of objects in space.

Feedback projections are thought to provide a means whereby higher centers in a pathway can influence lower ones. The number of neurons projecting from the cortex to the LGN is 10-fold the number projecting from the LGN to the cortex. Although this feedback projection is obviously important, its function is largely unknown.

The activity of the excitatory pyramidal and spiny stellate neurons that mediate information flow into or out of cortical regions is also tightly controlled by local networks of inhibitory interneurons. The spike rates of excitatory neurons are constantly nonlinearly balanced by matched inhibition that maintains the stability of the neural response to an input. Inhibitory interneurons come in multiple classes distinguished by their morphology and their coexpression of distinct peptides such as parvalbumin, somatostatin, or vasoactive intestinal polypeptide (VIP). Some of these interneurons form cascading circuits where interneurons of one class target interneurons of another class, which then target excitatory neurons. This leads to multistep control mechanisms in the neural circuit whereby increasing activity in the first class of inhibitory interneurons reduces activity in the second class, disinhibiting and increasing responses in the excitatory targets at the end of the cascade. Such motifs of inhibitory control are likely to be common to multiple cortical sensory areas.

In addition to serial feedforward, feedback, and local recurrent connections, fibers that travel parallel to the cortical surface within each layer provide long-range horizontal connections (Figure 21–16). These connections and their role in the functional architecture of cortex were analyzed by Charles Gilbert and Torsten Wiesel, who used intracellular recordings and dye injection to correlate anatomical features with cortical function. Because the visual cortex is organized visuotopically, the horizontal connections allow target neurons to integrate information over a relatively large area of the visual field and are therefore important in assembling the components of a visual image into a unified percept.

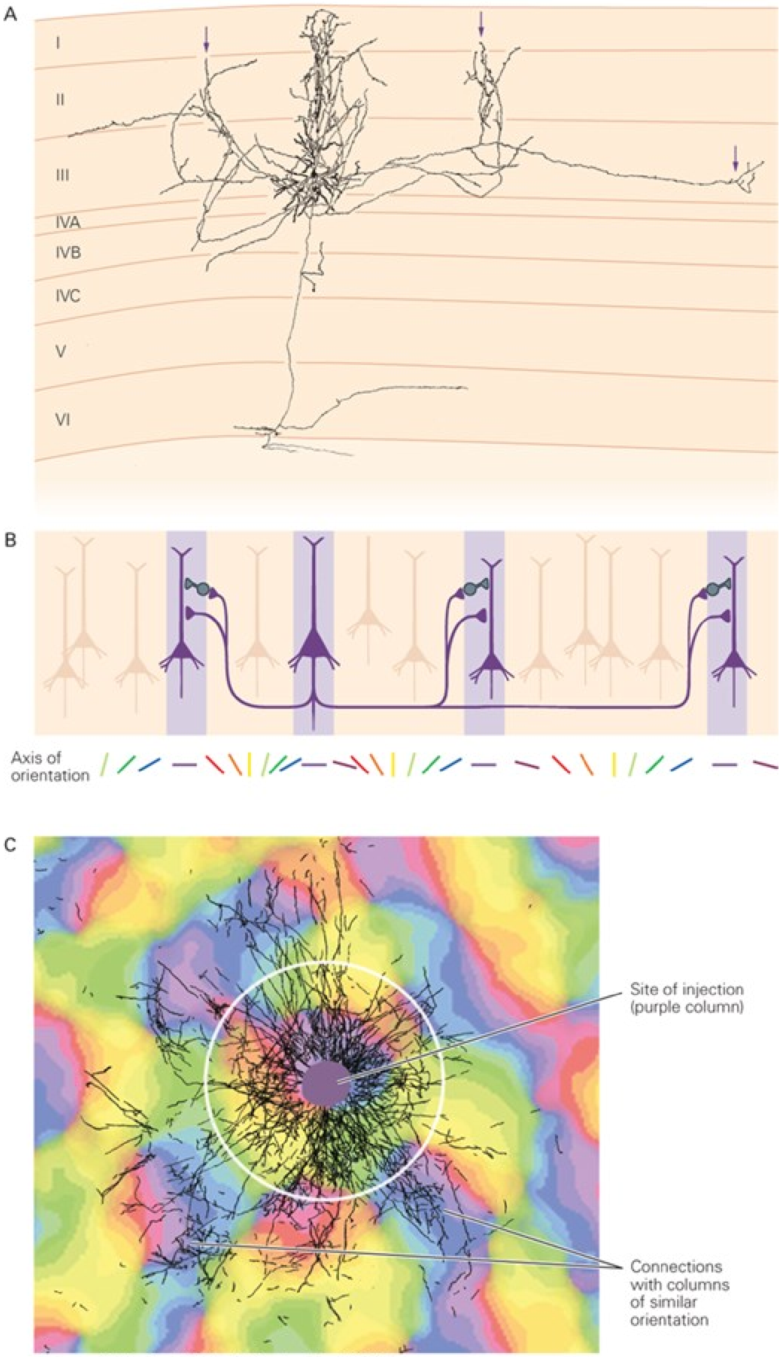
Figure 21–16

**Long-range horizontal connections in each layer of the visual cortex integrate information from different parts of the visual field.**

**A.** The axons of pyramidal cells extend for many millimeters parallel to the cortical surface. Axon collaterals form connections with other pyramidal cells as well as with inhibitory interneurons. This arrangement enables neurons to integrate information over large parts of the visual field. An important characteristic of these connections is their relationship to the functional columns. The axon collaterals are found in clusters (**arrows**) at distances greater than 0.5 mm from the cell body. (Reproduced, with permission, from Gilbert and Wiesel 1983. Copyright © 1983 Society for Neuroscience.)

**B.** Horizontal connections link columns of cells with similar orientation specificity.

**C.** The pattern of horizontal connections is visualized by injecting an adenoviral vector containing the gene encoding green fluorescent protein into one orientation column and superimposing the labeled image (**black**) on an optically imaged map of the orientation columns in the vicinity of the injection. (Diameter of white circle is 1 mm.) (Reproduced, with permission, from Stettler et al. 2002.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Integration can also be achieved by other means. The considerable convergence and divergence of connections at the synaptic relays of the afferent visual pathway imply that the receptive fields of neurons are larger and more complex at each successive relay and thus have an integrative function. Feedback connections may also support integration, both because of their divergence and because they originate from cells with larger receptive fields.

## Visual Information Is Represented by a Variety of Neural Codes

Individual neurons in a sensory pathway respond to a range of stimulus values. For example, a neuron in a color-detection pathway is not limited to responding to one wavelength but is instead tuned to a range of wavelengths. A neuron's response peaks at a particular value and tails off on either side of that value, forming a bell-shaped tuning curve with a particular bandwidth. Thus, a neuron with a peak response at 650 nm and a bandwidth of

100 nm might give identical responses at 600 nm and 700 nm.

To be able to determine the wavelength from neuronal signals, one needs at least two neurons representing filters centered at different wavelengths. Each neuron can be thought of as a *labeled line* in which activity signals a stimulus with a given value. When more than one such neuron fires, the convergent signals at the postsynaptic relay represent a stimulus with a wavelength that is the weighted average of the values represented by all the inputs.

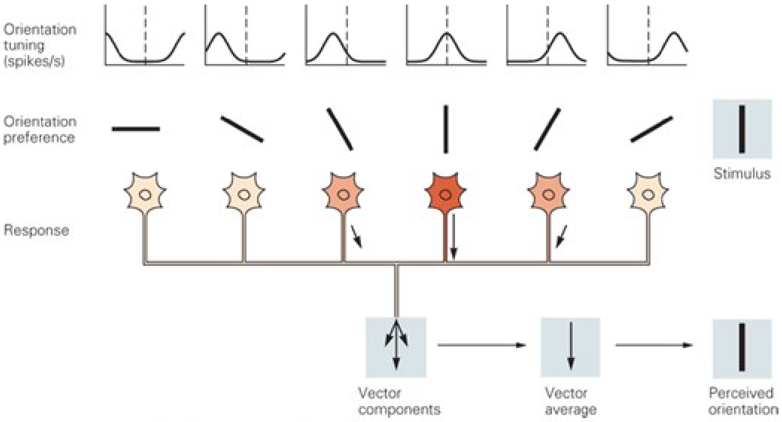
A single visual percept is the product of the activity of a number of neurons operating in a specific combinatorial and interactive fashion called a *population code.* Population coding has been modeled in various ways. The most prevalent model is called *vector averaging.*

We can illustrate population coding with a population of orientation-selective cells, each of which responds optimally to a line with a specific orientation. Each neuron responds not just to the preferred stimulus but rather to any line that falls within a range of orientations described by a Gaussian tuning curve with a particular bandwidth. A stimulus of a particular orientation most strongly activates cells with tuning curves centered at that orientation; cells with tuning curves centered away from but overlapping that orientation are excited less strongly.

Each cell's preferred orientation, the line label, is represented as a vector pointing in the direction of that orientation. Each cell's firing is a "vote" for the cell's line label, and the cell's firing rate represents the weighting of the vote. The cell's signal can thus be represented by a vector pointing in the direction of the cell's preferred orientation with a length proportional to the strength of the cell's response. For all the activated cells, one can calculate a vector sum with a direction that represents the value of the stimulus (Figure 21–17).

Figure 21–17

**Vector averaging is one model for population coding in neural circuits.** Vector averages describe the possible relationship between the responses in an ensemble of neurons, the tuning characteristics of individual neurons in the ensemble, and the resultant percept. Individual neurons respond optimally to a particular orientation of a stimulus in the visual field, but also respond at varying rates to a range of orientations. The stimulus orientation to which a neuron fires best can be thought of as a line label—when the cell fires briskly, its activity signifies the presence of a stimulus with that orientation. A number of neurons with different orientation preferences will respond to the same stimulus. Each neuron's response can be represented as a vector whose length indicates the strength of its response and whose direction represents its preferred orientation, or line label. (Adapted, with permission, from Kapadia, Westheimer, and Gilbert 2000.)



Another aspect of the population code is the variability of a neuron's response to the same stimulus. Repeated presentation of a stimulus to a neuron sensitive to that stimulus will elicit a range of responses. The most sensitive part of a neuron's tuning curve lies not at the peak but along the flanks, where the tuning curve is steepest. Here, small changes in the value of a stimulus produce the strongest change in response. Changes in stimulus value must, however, be sufficient to elicit a change in response that significantly exceeds the normal variability in the response of the neuron. One can compare that amount of change to the perceptual discrimination threshold. When many neurons contribute to the discrimination, the signal-to-noise ratio increases, a process known as probability summation, and the critical difference in stimulus value required for a significant change in neuronal response is less.

When the brain represents a piece of information, an important consideration is the number of neurons that participate in that representation. Although all information about a visual stimulus is present in the retina, the retinal representation is not sufficient for object recognition. At the other end of the visual pathway, some neurons in the temporal lobe are selective for complex objects, such as faces. Can an individual cell represent something as complex as a particular face? Such a hypothetical neuron has been dubbed a “grandmother cell” because it would represent exclusively a person’s grandmother, or a “pontifical cell” because it would represent the apex of a hierarchical cognitive pathway.

The nervous system does not, however, represent entire objects by the activity of single neurons. Instead, some cells represent parts of an object, and an ensemble of neurons represents an entire object. Each member of the ensemble may participate in different ensembles that are activated by different objects. This arrangement is known as a *distributed code.* Distributed codes can involve a few neurons or many. In any case, a distributed code requires complex connectivity between the cells representing a face and those representing the name and experiences associated with that person.

The foregoing discussion assumes that neurons signal information by their firing rate and their line labels. An alternative hypothesis is that the timing of action potentials itself carries information, analogous to Morse code. The code might be read from the synchronous firing of different sets of neurons over time. At one instant, one group of cells might fire together followed by the synchronous firing of another group. Over a single train of action potentials, a single cell could participate in many such ensembles. Whether sensory information is represented this way and whether the nervous system carries more information than that represented by firing rate alone are not known.

## Highlights

1. Vision is a constructive process fundamentally different from the mere recording of visual input as in a camera. Rather, the brain uses visual input to infer information about the world around it, including information about objects, such as their sizes, shapes, distances, and identities and how rapidly they are moving.

2. The tuning of neural circuits for visual features such as contrast, orientation, and motion often matches the distribution of the feature in the natural environment. This suggests an evolutionary, ethologically driven origin for the neural circuitry.

3. Visual circuitry, and thus vision, are modulated by individual visual experience.

4. Vision makes extensive use of parallel processing. The higher visual centers form two distinct pathways. The dorsal pathway, located in parietal cortex, is involved in motion perception, attention, and visually guided action. The ventral pathway, located in temporal cortex, processes form and objects. Further subdivisions of the ventral pathway are specialized, for example, for recognizing faces. These pathways, although distinct, communicate with each other; this is likely important for the perception of objects as coherent wholes.

5. Parallel processing starts at the retina. Distinct retinal circuits analyze each point of the visual input for different local features including local contrasts of achromatic bright versus dark, red versus green, and blue versus yellow. The information is sent out through distinct classes of retinal ganglion cells (magnocelluar, parvocelluar, and koniocellular, respectively, for the three features noted) whose axons form the optic nerves.

6. The optic nerves from the two eyes regroup at the optic chiasm such that all fibers from the left visual hemifield project to the right hemisphere of the brain, and vice versa. However, the parallel retinal channels remain anatomically segregated by eye and by visual feature, past a thalamic relay station, the lateral geniculate nucleus (LGN), up to primary visual cortex (V1).

7. The different channels enter V1 at different layers, although primarily they enter at the major input layers 4 and 6. The visual input is recombined to extract new sets of features. These include tuning for orientation, motion, and object depth (obtained by combining left- and right-eye inputs).

8. V1 neurons sharing basic properties such as spatial location or orientation preference form columns extending vertically from the pia to the white matter.

9. V1 neurons also form systematic horizontal maps of their response properties over cortex. The tuning for location forms a smooth “visuotopic” map of visual space, which changes gradually with distance, and is most finely resolved at the fovea, growing progressively coarser toward the periphery. Superimposed on the spatial map are locally smooth maps of orientation preference and left- versus right-eye preference, with interspersed columns that preferentially process color. These visual response features cycle over relatively short cortical distances, in effect

completing one full cycle over each partial shift of the spatial map. Thus, V1 circuits effectively analyze each visual location, in parallel, for the full set of V1 visual features.

10. Neural processing in V1 reflects its architecture, with local vertical processing along columns and lateral processing across columns. In addition, there is long-range processing that spans multiple columns.
11. The output of V1 feeds into progressively higher visual areas comprising more than 30 centers distributed along the dorsal and ventral pathways. The connectivity is reciprocal, with higher loci sending dense feedback targeting lower areas including the LGN.
12. A useful measure of visual processing is provided by changes in neuronal "receptive fields" along the visual pathway. The receptive field is the region of visual space from which the neuron receives input; it is further characterized by the neuron's optimal visual stimulus. Receptive fields grow larger and more complex at successive stages along the visual pathway. Their optimal stimuli also increase in complexity from simple pixel-like dots for photoreceptors, to oriented lines for V1, to faces in higher face-selective centers of the ventral pathway.
13. Looking forward, one of the most important unsolved questions is the interaction between feedforward visual processing through progressively "higher" neural computations and feedback mediated via the dense plexus of connections from higher to lower levels. Understanding this interaction may be the key to understanding how the brain effortlessly forms complex visual percepts.

**Charles D. Gilbert**

**Aniruddha Das**

## Selected Reading

Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106–154. [PubMed: 14449617]

Hubel DH, Wiesel TN. 1977. Functional architecture of macaque monkey visual cortex. Proc R Soc Lond B Biol Sci 198:1–59. [PubMed: 20635]

Hubener M, Shohan D, Grinvald A, Bonhoeffer T. 1997. Spatial relationships among three columnar systems in cat area 17. J Neurosci 17:9270–9284. [PubMed: 9364073]

Isaacson JS, Scanziani M. 2011. How inhibition shapes cortical activity. Neuron 72:231–243. [PubMed: 22017986]

Nassi JJ, Callaway EM. 2009. Parallel processing strategies of the primate visual system. Nat Rev Neurosci 10:360–372. [PubMed: 19352403]

Orban GA, Van Essen D, Vanduffel W. 2004. Comparative mapping of higher visual areas in monkeys and humans. Trends Cogn Sci 8:315–324. [PubMed: 15242691]

Stryker MP. 2014. A neural circuit that controls cortical state, plasticity, and the gain of sensory responses in mouse. Cold Spring Harb Symp Quant Biol 79:1–9. [PubMed: 25948638]

Tsao DY, Moeller S, Freiwald WA. 2008. Comparing face patch systems in macaques and humans. Proc Natl Acad Sci U S A 105:19514–19519. [PubMed: 19033466]

VanEssen DC, Anderson CH, Felleman DJ. 1992. Information processing in the primate visual system: an integrated systems perspective. Science 255:419–423. [PubMed: 1734518]

Wertheimer M. 1938. *Laws of Organization in Perceptual Forms*. London: Harcourt, Brace & Jovanovitch.

Wiesel TN, Hubel DH. 1966. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. J Neurophysiol 29:1115–1156. [PubMed: 4961644]

## References

Blasdel GG, Lund JS. 1983. Termination of afferent axons in macaque striate cortex. J Neurosci 3:1389–1413. [PubMed: 6864254]

Callaway EM. 1998. Local circuits in primary visual cortex of the macaque monkey. Annu Rev Neurosci 21:47–74. [PubMed: 9530491]

Field DJ, Hayes A, Hess RF. 1993. Contour integration by the human visual system: evidence for a local “association field.” Vision Res 33:173–193. [PubMed: 8447091]

Gilbert CD, Li W. 2012. Adult visual cortical plasticity. Neuron 75:250–264. [PubMed: 22841310]

Gilbert CD, Li W. 2013. Top-down influences on visual processing. Nat Rev Neurosci 14:350–363. [PubMed: 23595013]

Gilbert CD, Wiesel TN. 1983. Clustered intrinsic connections in cat visual cortex. J Neurosci 3:1116–1133. [PubMed: 6188819]

Hartline HK. 1941. The neural mechanisms of vision. Harvey Lect 37:39–68.

Hubel DH. 1988. *Eye, Brain and Vision*. New York: Scientific American Library.

Hubel DH, Wiesel TN. 1974. Uniformity of monkey striate cortex. A parallel relationship between field size, scatter and magnification factor. J Comp Neurol 158:295–306. [PubMed: 4436457]

Kapadia MK, Westheimer G, Gilbert CD. 2000. Spatial distribution of contextual interactions in primary visual cortex and in visual perception. J Neurophysiol 84:2048–2062. [PubMed: 11024097]

Kuffler SF. 1953. Discharge patterns and functional organization of mammalian retina. J Neurophysiol 16:37–68. [PubMed: 13035466]

Porter PB. 1954. Another puzzle-picture. Am J Psychol 67:550–551.

Stettler DD, Das A, Bennett J, Gilbert CD. 2002. Lateral connectivity and contextual interactions in macaque primary visual cortex. Neuron 36:739–750. [PubMed: 12441061]

Principles of Neural Science, 6e >

# Chapter 23: Intermediate-Level Visual Processing and Visual Primitives

## Introduction

WE HAVE SEEN IN Chapters 21 and 22 that the eye is not a mere camera, but instead contains sophisticated retinal circuitry that decomposes the retinal image into signals representing contrast and movement. These data are conveyed through the optic nerve to the primary visual cortex, which uses this information to analyze the shape of objects. It first identifies the boundaries of objects, represented by numerous short line segments, each with a specific orientation. The cortex then integrates this information into a representation of specific objects, a process referred to as *contour integration.*

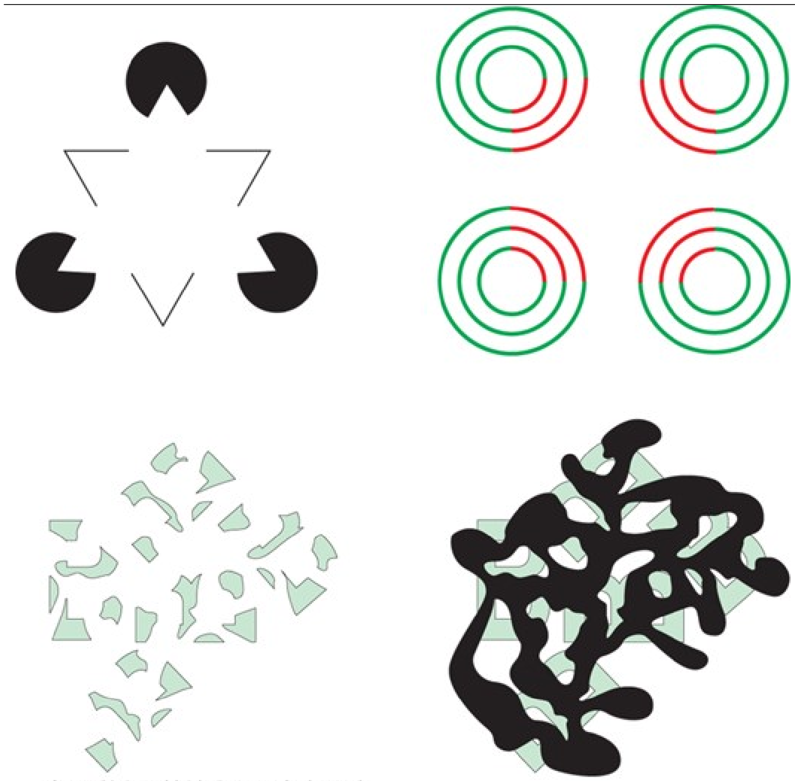
These two steps, local analysis of orientation and contour integration, exemplify two distinct stages of visual processing. Computation of local orientation is an example of low-level visual processing, which is concerned with identifying local elements of the light structure of the visual field. Contour integration is an example of intermediate-level visual processing, the first step in generating a representation of the unified visual field. At the earliest stages of analysis in the cerebral cortex, these two levels of processing are accomplished together.

A visual scene comprises many thousands of line segments and surfaces. Intermediate-level visual processing is concerned with determining which boundaries and surfaces belong to specific objects and which are part of the background (see Figure 21–4). It is also involved in distinguishing the brightness and color of a surface from the intensity and wavelength of light reflected from that surface. The physical characteristics of reflected light result as much from the intensity and color balance of the light that illuminates a surface as from the color of that surface. Determining the actual surface color of a single object requires comparison of the wavelengths of light reflected from multiple surfaces in a scene.

Intermediate-level visual processing thus involves assembling local elements of an image into a unified percept of objects and background. Although determining which elements belong together in a single object is a highly complex problem with an astronomical number of potential solutions, each relay in the visual circuity of the brain has built-in logic that allows assumptions to be made about the likely spatial relationships between elements. In certain cases, these inherent rules can lead to the illusion of contours and surfaces that do not actually exist in the visual field (Figure 23–1).

Figure 23–1

**Illusory contours and perceptual fill-in.** The visual system uses information about local orientation and contrast to construct the contours and surfaces of objects. This constructive process can lead to the perception of contours and surfaces that do not appear in the visual field, including those seen in illusory figures. *Top left:* In the Kanizsa triangle illusion, one perceives continuous boundaries extending between the apices of a white triangle, even though the only real contour elements are those formed by the Pac-Man–like figures and the acute angles. **Top right:** The inside and outside of the illusory pink square are the same white color as the page, but a continuous transparent pink surface within the square is perceived. *Bottom:* Occluding surfaces can also facilitate contour integration and surface segmentation. The irregular shapes on the left appear to be unrelated, but when they are partially occluded by black shapes (right), they are easily seen as fragments of the letter B.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

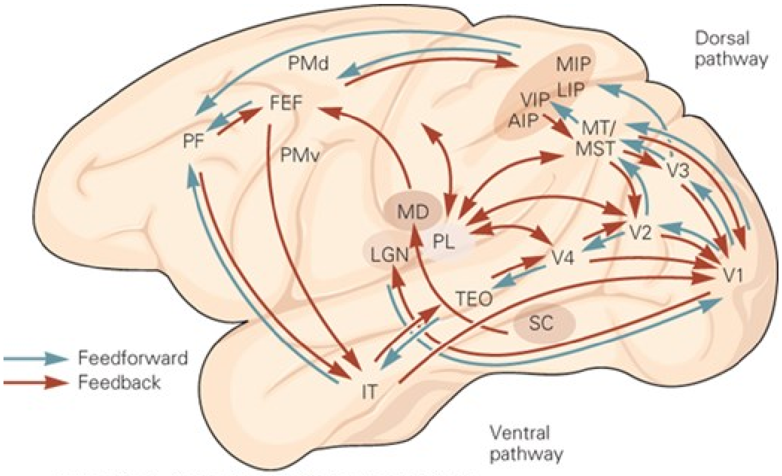
Three features of visual processing help overcome ambiguity in the signals from the retina. First, the way in which a visual feature is perceived depends on everything that surrounds it. The perception of a point, line, or surface, for example, depends on the relationship between that feature and what else is present in the scene. That is, the response of a neuron in the visual cortex is context-dependent: It depends as much on the presence of contours and surfaces outside the cell's receptive field as on the attributes within it. Second, the functional properties of neurons in the visual cortex can be altered by visual experience or perceptual learning. Finally, visual processing in the cortex is subject to the influence of cognitive functions, specifically attention, expectation, and "perceptual task" (the active engagement in visual discrimination or detection). The interaction between these three factors—the context or entire set of signals representing a scene, experience-dependent changes in cortical circuitry, and expectation—is vital to the visual system's analysis of complex scenes.

In this chapter, we examine how the brain's analysis of the local features in a visual scene, or *visual primitives,* proceeds in parallel with the analysis of more global features. Visual primitives include contrast, line orientation, brightness, color, movement, and depth. Each type of visual primitive is subject to the integrative action of intermediate-level processing. Lines with particular orientations are integrated into object contours, local contrast information into surface brightness and surface segmentation, wavelength selectivity into color constancy, and directional selectivity into object motion.

The analysis of visual primitives begins in the retina with the detection of brightness and color and continues in the primary visual cortex with the analysis of orientation, direction of movement, and stereoscopic depth. Properties related to intermediate-level visual processing are analyzed together with visual primitives in the visual cortex starting in the primary visual cortex (V1), which plays a role in contour integration and surface segmentation. Other areas of the visual cortex specialize in different aspects of this task: V2 analyzes properties related to object surfaces, V4 integrates information about color and object shape, and V5—the middle temporal area or MT—integrates motion signals across space (Figure 23–2).

Figure 23–2

**Cortical areas involved with intermediate-level visual processing.** Many cortical areas in the macaque monkey, including V1, V2, V3, V4, and middle temporal area (**MT**), are involved with integrating local cues to construct contours and surfaces and segregating foreground from background. The shaded areas extend into the frontal and temporal lobes because cognitive output from these areas, including attention, expectation, and behavioral task, contributes to the process of scene segmentation. (Abbreviations: **AIP**, anterior intraparietal cortex; **FEF**, frontal eye fields; **IT**, inferior temporal cortex; **LGN**, lateral geniculate nucleus; **LIP**, lateral intraparietal cortex; **MD**, medial dorsal nucleus of thalamus; **MIP**, medial intraparietal cortex; **MST**, medial superior temporal cortex; **MT**, middle temporal cortex; **PF**, prefrontal cortex; **PL**, pulvinar; **PMd**, dorsal premotor cortex; **PMv**, ventral premotor cortex; **SC**, superior colliculus; **TEO**, occipitotemporal cortex; **VIP**, ventral intraparietal cortex; **V1**, **V2**, **V3**, **V4**, primary, secondary, third, and fourth visual areas.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

## Internal Models of Object Geometry Help the Brain Analyze Shapes

A first step in determining an object's contour is identification of the orientation of local parts of the contour. This step commences in V1, which plays a critical role in both local and global analysis of form.

Neurons in the visual cortex respond selectively to specific local features of the visual field, including orientation, binocular disparity or depth, and direction of movement, as well as to properties already analyzed in the retina and lateral geniculate nucleus, such as contrast and color. Orientation selectivity, the first emergent property identified in the receptive fields of cortical neurons, was discovered by David Hubel and Torsten Wiesel in 1959.

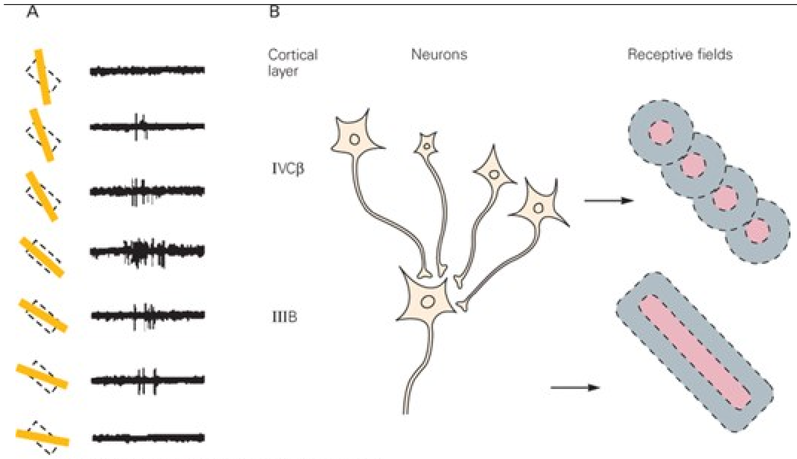
Neurons in both the retina (Chapter 22) and the lateral geniculate nucleus (Chapter 21) have circular receptive fields with a center-surround organization. They respond to the light–dark contrasts of edges or lines in the visual field but are not selective for the orientations of those edges (see Figure 21–9). In the visual cortex, however, neurons respond selectively to lines of particular orientations. Each neuron responds to a narrow range of orientations, approximately 40°, and different neurons respond optimally to distinct orientations. Hubel and Wiesel proposed that this orientation selectivity reflects the arrangement of the inputs from the lateral geniculate nucleus, and there is now a body of supportive evidence for the idea. Each V1 neuron receives input from several neighboring geniculate neurons whose center-surround receptive fields are aligned so as to represent a particular axis of orientation (Figure 23–3). Two principal types of orientation-selective neurons, simple and complex, have been identified.

Figure 23–3

**Orientation selectivity and mechanisms.**

**A.** A neuron in the primary visual cortex responds selectively to line segments that fit the orientation of its receptive field. This selectivity is the first step in the brain's analysis of an object's form. (Reproduced, with permission, from Hubel and Wiesel 1968. Copyright © 1968 The Physiological Society.)

**B.** The orientation of the receptive field is thought to result from the alignment of the circular center-surround receptive fields of several presynaptic cells in the lateral geniculate nucleus. In the monkey, individual neurons in layer IVCβ of V1 have unoriented receptive fields. However, when several neighboring IVCβ cells project to a neuron in layer IIIB they create a receptive field with a specific orientation for that postsynaptic cell.

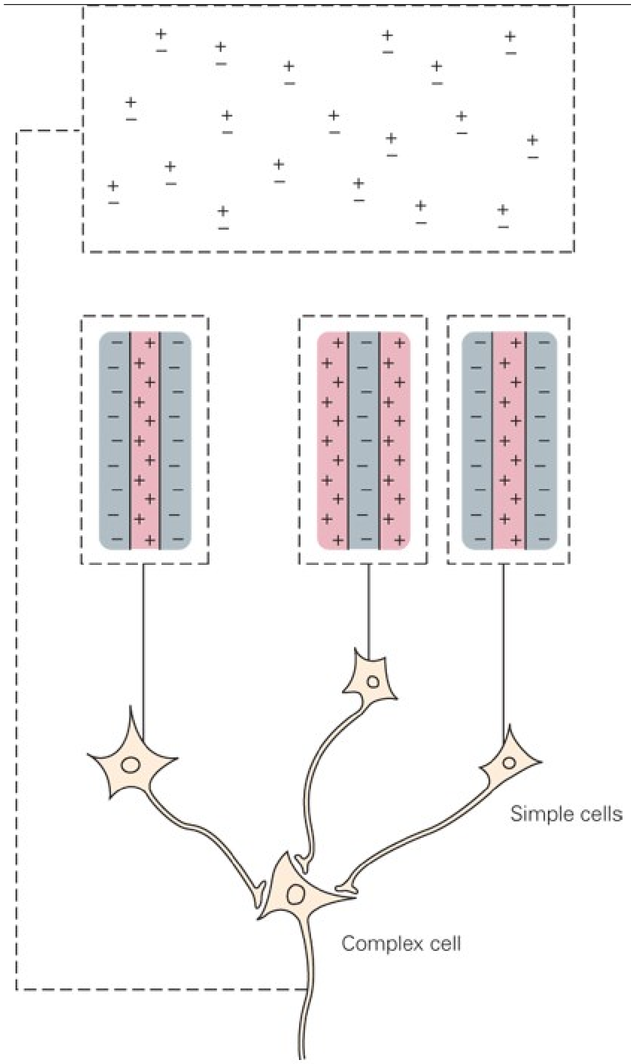


Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

*Simple cells* have receptive fields divided into ON and OFF subregions (Figure 23–4). When a visual stimulus such as a bar of light enters the receptive field's ON subregion, the neuron fires; the cell also responds when the bar leaves the OFF subregion. Simple cells have a characteristic response to a moving bar; they discharge briskly when a bar of light leaves an OFF region and enters an ON region. The responses of these cells are therefore highly selective for the position of a line or edge in space.

Figure 23–4

**Simple and complex cells in the visual cortex.** The receptive fields of simple cells are divided into subfields with opposite response properties. In an ON subfield (indicated by +), the onset of a light triggers a response in the neuron; in an OFF subfield (indicated by −), the extinction of a bar of light triggers a response. Complex cells have overlapping ON and OFF regions and respond continuously as a line or edge traverses the receptive field along an axis perpendicular to the receptive-field orientation.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

*Complex cells* are less selective for the position of object boundaries. They lack discrete ON and OFF subregions (Figure 23–4) and respond similarly to light and dark at all locations across their receptive fields. They fire continuously as a line or edge stimulus traverses their receptive fields. Hubel and Wiesel proposed that the complex cells are a second stage of the elaboration of receptive fields after simple receptive fields and are built by overlapping simple receptive fields.

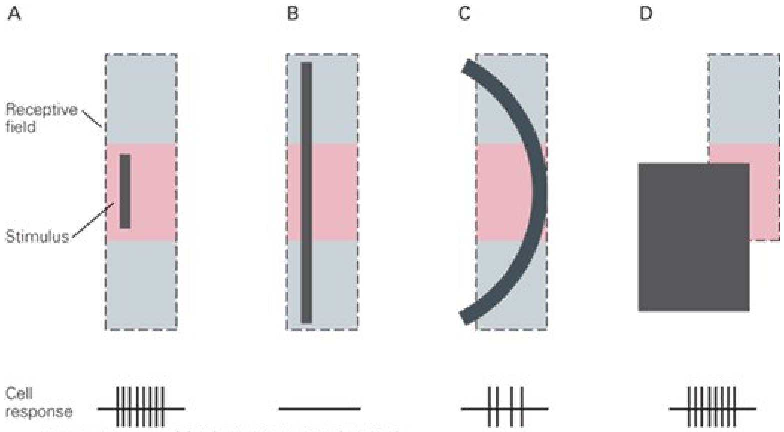
As one considers the range of receptive field properties that have been described in the early visual cortical areas, it is important to point out phylogenetic differences, with different species differing in the location in which these properties are first expressed and in the kinds of properties that are represented. In the cat, the target layer of the visual cortex for lateral geniculate neurons has oriented simple cells; it had been presumed that these cortical cells represent an obligatory first stage in the cortical processing of visual information, between the center-surround circularly symmetric receptive fields in the lateral geniculate nucleus and the receptive fields of complex cells in the superficial cortical layers. In primates, however, the geniculate target layers, 4Cα and β, have circularly symmetric, unoriented receptive fields. The postsynaptic target of the layer 4C cells, predominantly the superficial layers of the cortex, is populated with complex cells, therefore skipping a simple cell stage. In the mouse, orientation selectivity is seen in the lateral geniculate nucleus. The preceding comparison points out a few characteristics of the evolution of visual processing. One is the encephalization of function, where properties such as orientation are shifted to later stages of processing over stages of evolution. Another is the development of new pathways. It has been suggested that the magnocellular pathway in the monkey is equivalent to the entire geniculostriate pathway in the cat, whereas the parvocellular pathway, which mediates higher-resolution vision and color vision, is new to the primate.

Moving stimuli are often used to study the receptive fields of visual cortex neurons, not only to simulate the conditions under which an object moving in space is detected but also to simulate the conditions produced by eye movements. As we scan the visual environment, the boundaries of stationary objects move across the retina. In fact, visual perception requires eye movement. Visual cortex neurons do not respond to an image that is stabilized on the retina. These neurons require transient stimulation (moving or flashing stimuli) in order to be activated.

Some visual cortex neurons have receptive fields in which an excitatory center is flanked by inhibitory regions. Inhibitory regions along the axis of orientation, a property known as *end-inhibition*, restrict a neuron's responses to lines of a certain length (Figure 23–5). End-inhibited neurons respond well to a line that does not extend into the inhibitory flanks but lies entirely within the excitatory part of the receptive field. Because the inhibitory regions share the orientation preference of the central excitatory region, end-inhibited cells are selective for line curvature and also respond well to corners.

Figure 23–5

**End-inhibited receptive fields.** Some receptive fields have a central excitatory region flanked by inhibitory regions that have the same orientation selectivity. Thus, a short line segment or a long curved line will activate the neuron (**A** and **C**), but a long straight line will not (**B**). A neuron with a receptive field that displays only one inhibitory region in addition to the excitatory region can signal the presence of corners (**D**).



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

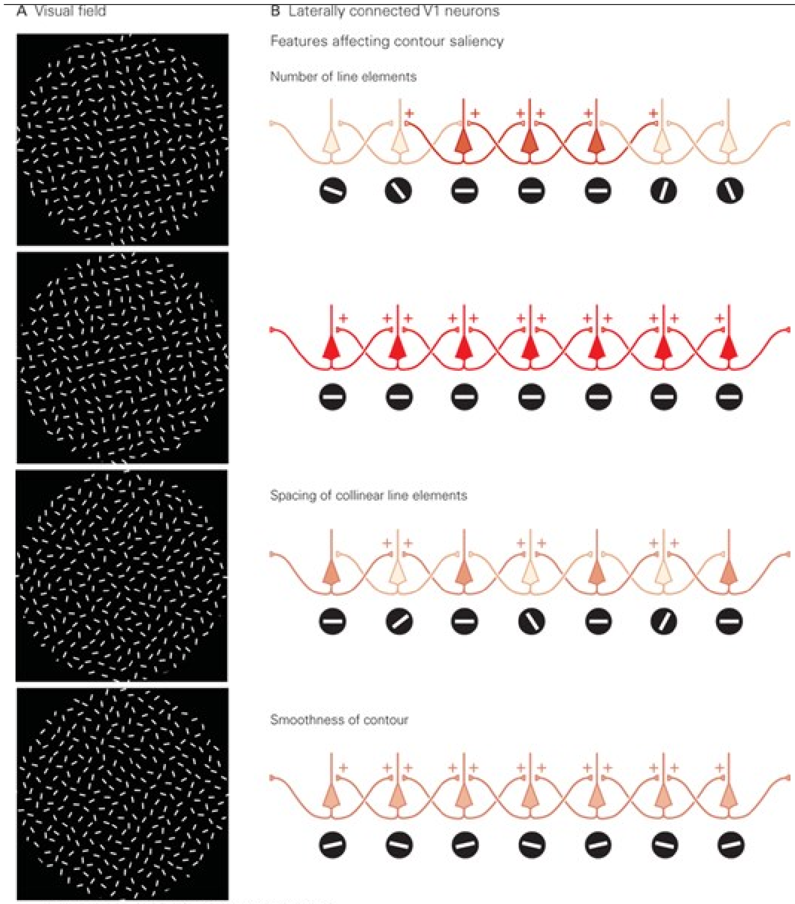
To define the shape of the object as a whole, the visual system must integrate the information on local orientation and curvature into object contours. The way in which the visual system integrates contours reflects the geometrical relationships present in the natural world (Figure 23–6). As originally pointed out by Gestalt psychologists early in the 20th century, contours that are immediately recognizable tend to follow the rule of good continuation (curved lines maintain a constant radius of curvature and straight lines stay straight). In a complex visual scene, such smooth contours tend to "pop out," whereas more jagged contours are difficult to detect.

Figure 23–6

**Contour integration reflects the perceptual rules of proximity and good continuation.** (Adapted, with permission, from Li and Gilbert 2002.)

**A.** A straight line composed of one or more contour elements with the same oblique orientation appears in the center of each of the four images here. In some images, the line pops out more or less immediately, without searching. Factors that contribute to contour saliency include the number of contour elements (compare the first and second frames), the spacing of the elements (third frame), and the smoothness of the contour (bottom frame). When the spacing between elements is too large or the orientation difference between them too great, one must search the image to find the contour.

**B.** These perceptual properties are reflected in the horizontal connections between columns of V1 neurons with similar orientation selectivity. As long as the visual elements are spaced sufficiently close together, excitation can propagate from cell to cell, thus facilitating the responses of V1 neurons. Each neuron in the network then augments the responses of neurons on either side, and the facilitated responses propagate across the network.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

The responses of a visual cortex neuron can be modulated by stimuli that themselves do not activate the cell and therefore lie outside the receptive field's core. This *contextual modulation* endows a neuron with selectivity for more complex stimuli than would be predicted by placing the components of a stimulus at different positions in and around the receptive field. The same visual features that facilitate the detection of an object in a complex scene (Figure 23–6A) also apply to contextual modulation. The properties of the features that confer perception of contours, even illusory ones, are reflected in the responses of neurons in the primary visual cortex, which are sensitive to the global characteristics of contours, even those that extend well outside their receptive fields.

Contextual influences over large regions of visual space are likely to be mediated by connections between multiple columns of neurons in the visual cortex that have similar orientation selectivity (Figure 23–6B). These connections are formed by pyramidal-cell axons that run parallel to the cortical surface (see Figure 21–16). The extent and orientation dependency of these horizontal connections provide the interactions that could mediate contour saliency (see Figure 21–14).

Central to the process of contour integration is the idea of the association field. The association field refers to the interactions across visual space required to perceptually link contour elements into global contours. It underlies the Gestalt principle of good continuation and the perceptual saliency of smooth contours embedded in complex scenes. Physiologically, it underlies the facilitation of neuronal responses by contour elements extending outside their "classical" receptive fields. Anatomically, it is mediated in part by the relationship between long-range horizontal connections and cortical functional architecture. Though it has been investigated most extensively in primary visual cortex, because of the ubiquity of horizontal connections across all areas of cortex, it is likely to be a strategy for associating bits of information that are mapped within every cortical area. The functional role of the association field in cortical areas outside of V1 depends on how information is mapped across the cortical surface and the relationship between these maps and the plexus of horizontal connections.

## Depth Perception Helps Segregate Objects From Background

Depth is another key feature in determining the perceived shape of an object. An important cue for the perception of depth is the difference between the two eyes' views of the world, which must be computed and reconciled by the brain. The integration of binocular input begins in the primary visual cortex, the first level at which individual neurons receive signals from both eyes. The balance of input from the two eyes, a property known as ocular dominance, varies among cells in V1.

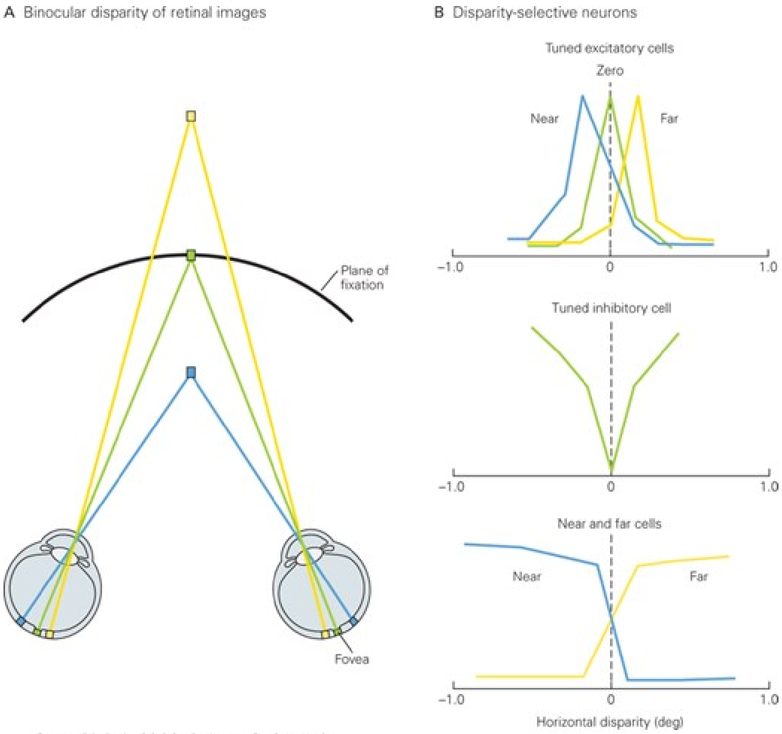
Binocular neurons in many visual cortical areas are also selective for depth, which is computed from the relative retinal positions of objects placed at different distances from the observer. An object that lies in the *plane of fixation* produces images at corresponding positions on the two retinas (Figure 23–7). The images of objects that lie in front of or behind the plane of fixation fall on slightly different locations in the two eyes, a property known as binocular disparity. Individual neurons can be selective for a narrow range of disparities and therefore positions in depth. Some are selective for objects lying on the plane of fixation (tuned excitatory or inhibitory cells), whereas others respond only when objects lie in front of the plane of fixation (near cells) or behind that plane (far cells).

Figure 23–7

### Stereopsis and binocular disparity.

**A.** Depth is computed from the positions at which images occur in the two eyes. The image of an object lying in the plane of fixation (**green**) falls on corresponding points on the two retinas. Images of objects lying in front of the plane of fixation (**blue**) or behind it (**yellow**) fall on noncorresponding locations on the two retinas, a phenomenon termed *binocular disparity*.

**B.** Neurons in many visual cortical areas are selective for particular ranges of disparity. Each plot shows the responses of a neuron to binocular stimuli with different disparities (abscissa). Some neurons are tuned to a narrow range of disparities and thus have particular disparity preferences (tuned excitatory or tuned inhibitory neurons), whereas others are tuned broadly for objects in front of the fixation plane (near cells) or beyond the plane (far cells). (Adapted, with permission, from Poggio 1995. Copyright © 1995 Oxford University Press.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

Depth plays an important role in the perception of object shape, in surface segmentation, and in establishing the three-dimensional properties of a scene. Objects that are placed near an observer can partially occlude those situated farther away. A surface passing behind an object is perceived as continuous even though its two-dimensional image on each retina represents two surfaces separated by the occluder. When the brain encounters a surface interrupted by gaps that have appropriate alignment and contrast, and lying in the near-depth plane, it fills in the gaps to create a continuous surface (Figure 23–8).

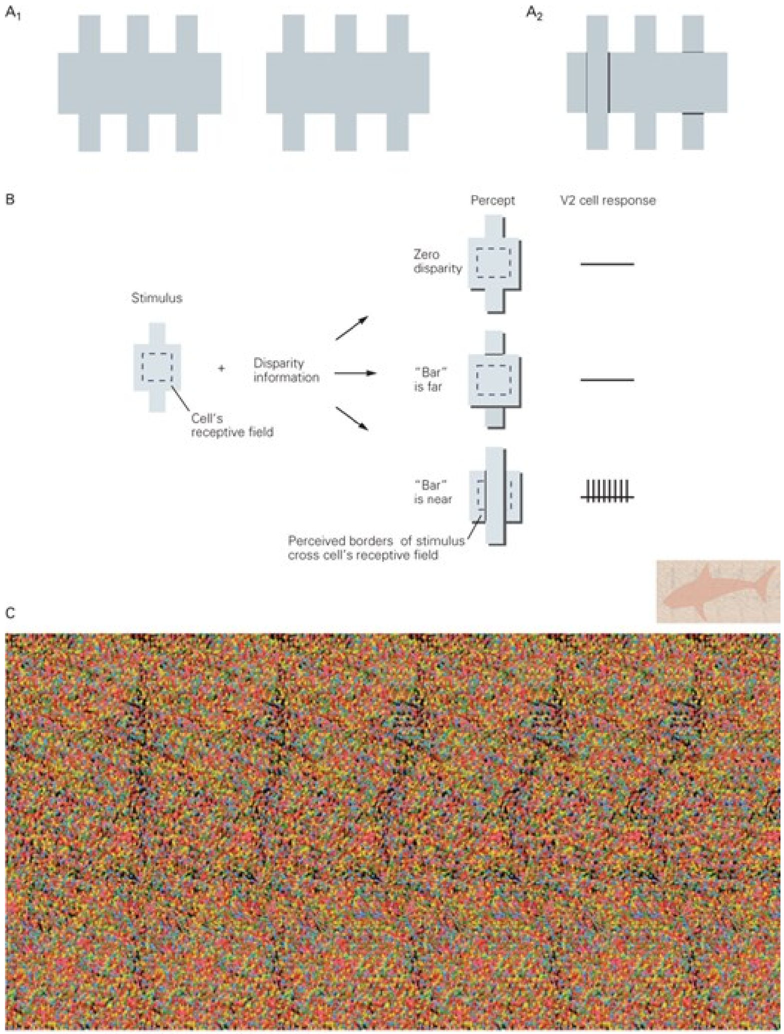
Figure 23–8

Global analysis of binocular disparity.

**A. 1.** Depth cues contribute to surface segmentation. If you view one of the images of three gray vertical bars crossing a gray horizontal rectangle, you see a uniform gray area within the rectangle. **2.** However, if you fuse the two rectangles with diverged eyes, the three vertical bars fall on the two retinas with near, zero, and far disparity. Seen this way, the bar at the left appears to hover in front of the rectangle with an illusory vertical edge crossing the rectangle, whereas the bar at the right appears to lie behind the edges of the horizontal rectangle.

**B.** A neuron in area V2 responds to illusory edges formed by binocular disparity cues. When the cell's receptive field is centered in the gray square, the cell does not respond to a vertical bar that has far disparity or the same disparity as the square. When the vertical bar has near disparity, the cell responds as the illusory vertical edge crosses its receptive field. (Reproduced, with permission, from Bakin, Nakayama, and Gilbert 2000. Copyright © 2000 Society for Neuroscience.)

**C.** A random-dot stereogram is seen as a random array of colored dots until you diverge or converge your eyes to bring the adjacent dark vertical stripes into register, producing a three-dimensional image of a shark that emerges from the background. This effect stems from systematic disparity for selected sets of dots. (© Fred Hsu/ Wikimedia Commons/CC-BY-SA-3.0.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Although the depth of a single object can be established easily, determining the depths of multiple objects within a scene is a much more complex problem that requires linking the retinal images of all objects in the two eyes. The disparity calculation is therefore a global one: The calculation in one part of the visual image influences the calculation for other parts. When the assignment of depth is unambiguous in one part of an image, that information is applied to other parts of the image where there is insufficient information to determine depth, a phenomenon known as disparity capture.

Random-dot stereograms provide a dramatic demonstration of the global scope of disparity analysis. The visual information presented to each eye appears to be incoherent, but when the stereogram is viewed binocularly, the disparity between the random array of dots in the two images allows an embedded shape to become visible (Figure 23–8C). The calculation underlying this percept is not simple, but requires determining which features shown to the left eye correspond to features seen by the right eye and propagating local disparity information across the image.

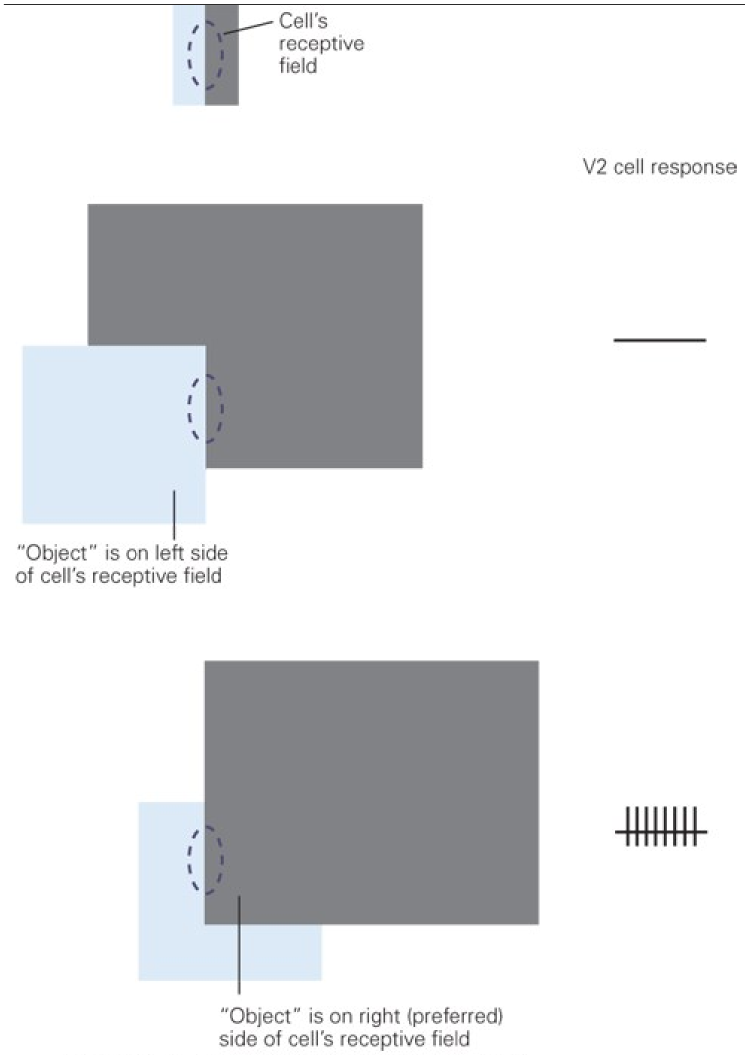
Neurons in area V2 display sensitivity to global disparity cues. Distant depth cues can be used to link contour elements that belong to an object, and to separate them from the object background (Figure 23-B).

In addition to binocular disparity, the visual system also uses many monocular cues to discriminate depth. Depth determination through monocular cues, such as size, perspective, occlusion, brightness, and movement, is not difficult. Another cue that originates outside the visual system is vergence, the angle between the optical axes of the two eyes for objects at varying distances. Yet another binocular cue, known as DaVinci stereopsis, is the presence of features visible to one eye but occluded in the other eye's view.

Neurons in areas V1 and V2 also signal foreground–background relationships. A cell with its receptive field in the center of a pattern within a larger surface may respond even when the boundary of that surface is distant from the receptive field. This response helps differentiate the object from its background. In making sense of an image, the brain must identify which edge belongs to which object and differentiate the edge of each object from the background. Some cells in area V2 have the property of "border ownership," firing only when a figure but not the background is to one side of the edge, even when the local edge information is identical in both instances (Figure 23–9).

Figure 23–9

**Border ownership.** Cells in area V2 are sensitive to the boundaries of whole objects. Even though the local contrast is the same for the two rectangles within a cell's receptive field, the cell responds only when the boundary is part of the full rectangle that lies on the preferred side of the receptive field. (Adapted, with permission, from Zhou, Friedman, and von der Heydt 2000. Copyright © 2000 Society for Neuroscience.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

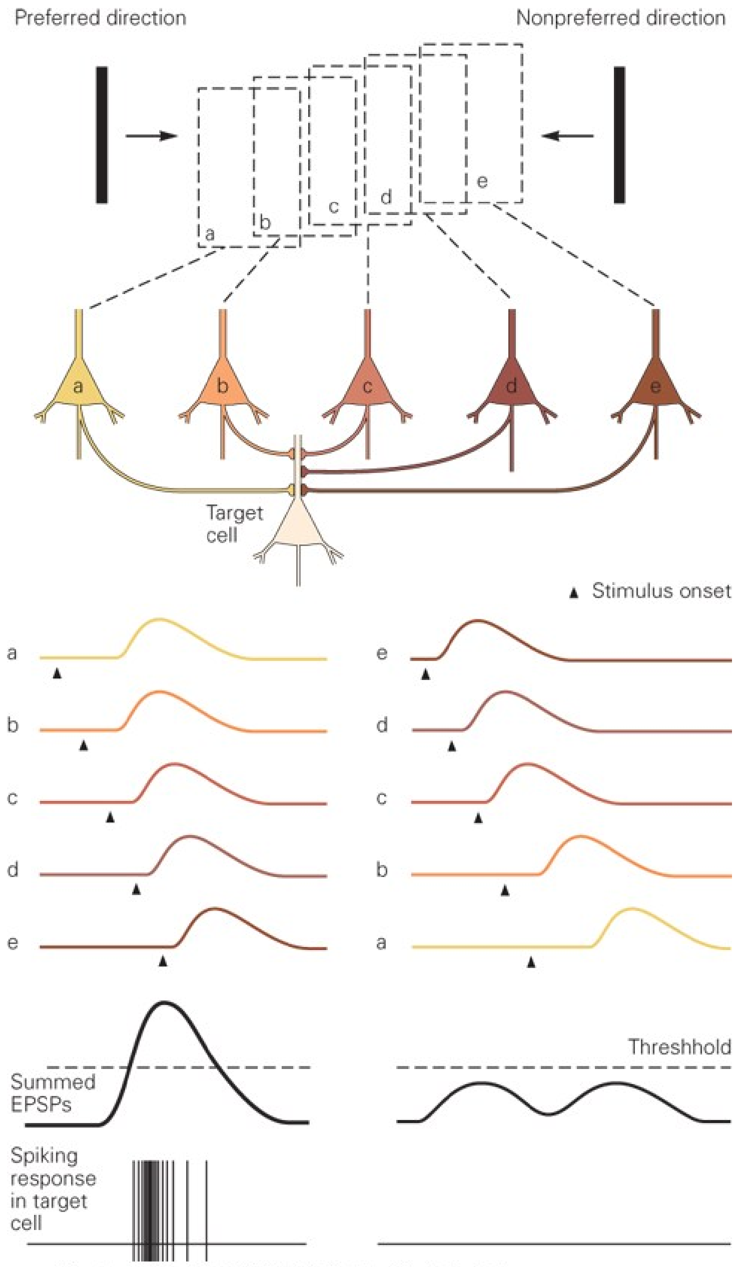
## Local Movement Cues Define Object Trajectory and Shape

The primary visual cortex determines the direction of movement of objects. Directional selectivity in neurons likely involves sequential activation of regions on different sides of the receptive field.

If an object moving at an appropriate velocity first encounters a region of a neuron's receptive field with long response latencies and then passes into regions with progressively shorter latencies, signals from throughout the receptive field will arrive at the cell simultaneously and the neuron will fire vigorously. If the object moves in the opposite direction, signals from the different regions will not summate and the cell may never reach the threshold for firing (Figure 23–10).

Figure 23–10

**Directional selectivity of movement.** A neuron's selectivity for direction of movement depends on the response latencies of presynaptic neurons relative to the onset of a stimulus. The response latencies of presynaptic neurons *a* and *b* are somewhat longer than those of neurons *d* and *e*. When a stimulus moves from left to right, neurons *a* and then *b* are activated first, but because their response latencies are longer, their inputs arrive at the target neuron superimposed with the inputs from neurons *d* and *e*, and the summated inputs cause the neuron to fire. In contrast, stimuli moving leftward produce signals that arrive in the target neuron at different times and therefore do not reach the cell's threshold for firing. (Abbreviation: **EPSP**, excitatory postsynaptic potential.) (Adapted, with permission, from Priebe and Ferster 2008. Copyright © 2008 Elsevier.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Early in the visual pathways, analysis of the movement of an object is limited by the size of the receptive fields of the sensory neurons. Even in the initial cortical areas V1 and V2, the receptive fields of neurons are small and might encompass only a fraction of an object. Eventually, however, information about the direction and speed of movement of discrete aspects of an object must be integrated into a computation of the movement of a whole object. This problem is more difficult than one might expect.

If one observes a complex shape moving through a small aperture, the part of the object's boundary within the aperture appears to move in a direction perpendicular to the boundary's orientation (Figure 23–11A). One cannot detect a line's true direction of movement if the line's ends are not visible. The image of a line appears the same if it is moving slowly along an axis perpendicular to its orientation or more quickly along an oblique axis. This is the quandary presented by the receptive field of a V1 neuron. The visual system's solution is to assume that the movement of a contour is perpendicular to its orientation. Thus, an object is first presented to the visual system in countless small pieces with boundaries of different orientations, all of which appear to be moving in different directions and at different velocities (Figure 23–11A).

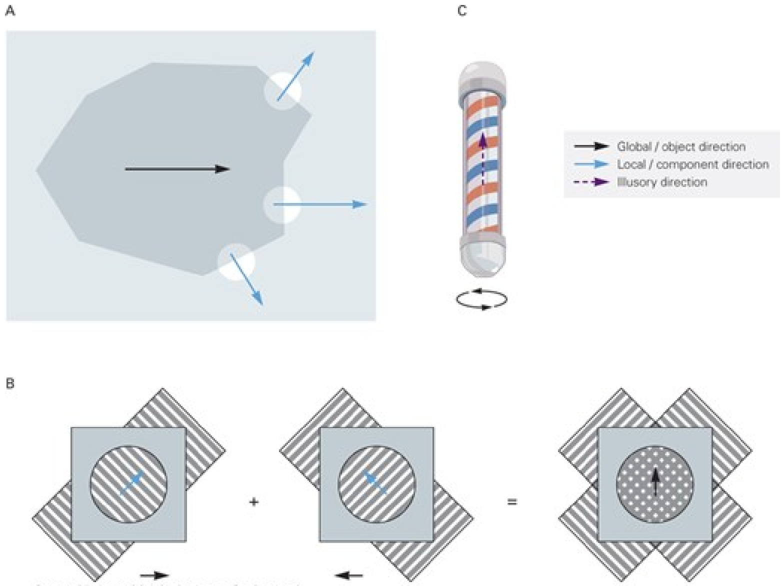
Figure 23–11

**The aperture problem and barber-pole illusion.**

**A.** Although an object moves in one direction, each component edge when viewed through a small aperture appears to move in a direction perpendicular to its orientation. The visual system must integrate such local motion signals into a unified percept of a moving object.

**B.** Gratings are used to test whether a neuron is sensitive to local or global motion signals. When the gratings are superimposed and moved independently in different directions, one does not see the two gratings sliding past each other but rather a plaid pattern moving in a single, intermediate direction. Neurons in the middle temporal area of monkeys are responsive to such global motion rather than to local motion.

**C.** Motion perception is influenced by scene segmentation cues, as seen in the barber-pole illusion. Even though the pole rotates around its axis, one perceives the stripes as moving vertically, due to the global vertical rectangle surround of the barber pole enclosure.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

Determining the direction of motion of an object requires resolving multiple cues. This can be demonstrated readily by placing one grating on top of another and moving the two in different directions. The resulting checkerboard pattern appears to move in an intermediate direction between the trajectories of the individual gratings (Figure 23–11B). This percept depends on the relative contrast of the gratings and the area of grating overlap. With large relative contrasts, the gratings appear to slide across each other, moving in their individual directions rather than together in a common direction.

An important determinant of perceived direction is scene segmentation, the separation of moving elements into foreground and background. In a scene with moving objects, segmentation is not based on local cues of direction; instead, perception of direction depends on scene segmentation. The barber-pole illusion provides another example of the predominance of global relationships over the perception of simple attributes. The rotating stripes are perceived as moving vertically along the long axis of the pole (Figure 23–11C). The perception of motion in the visual field uses a complex algorithm that integrates the bottom-up analysis of local motion signals with top-down scene segmentation.

Integration of local motion signals in monkeys has been observed in the middle temporal area (area MT or V5), an area specializing in motion. The neurons in this area are selective for a particular direction of movement of an overall pattern, rather than individual components of the pattern. This dependency on the overall pattern is also seen in the correspondence of their responses with the perceived direction in the barber-pole effect.

## Context Determines the Perception of Visual Stimuli

### Brightness and Color Perception Depend on Context

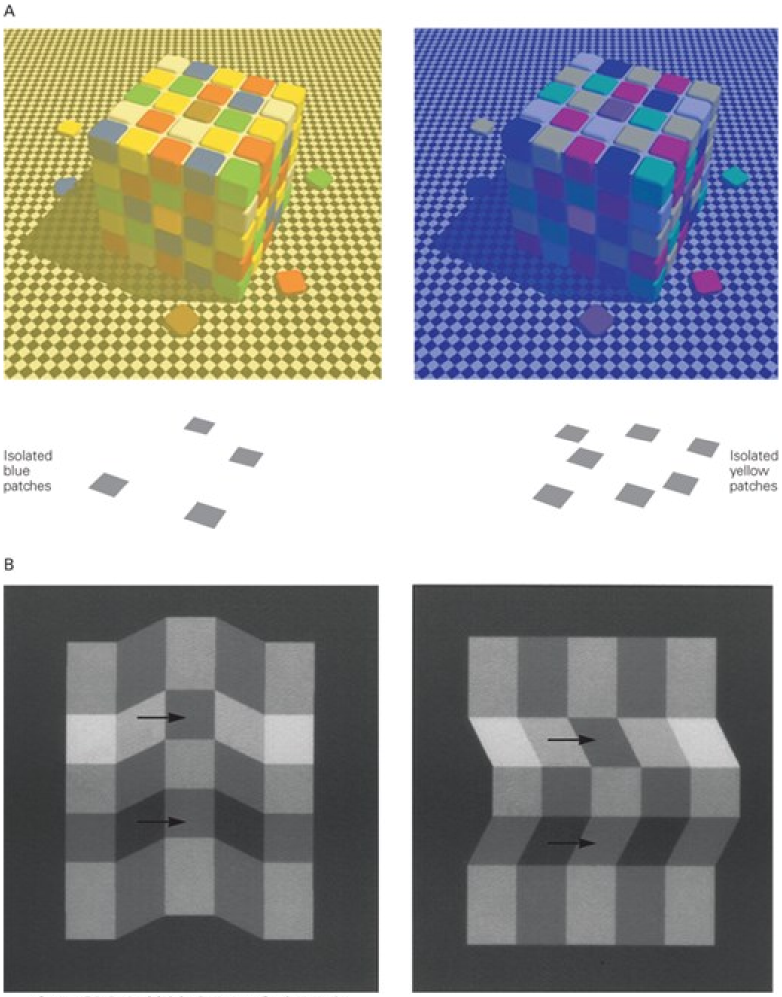
The visual system measures the surface characteristics of objects by comparing the light arriving from different parts of the visual field. As a result, the perception of brightness and color is highly dependent on context. In fact, perceived brightness and color can be quite different from what is expected from the physical properties of an object. At the same time, perceptual constancies make objects appear similar even when the brightness and wavelength distribution of the light that illuminates them changes from natural to artificial light, from sunlight to shadow, or from dawn to midday (Figure 23–12A).

Figure 23–12

**Color and brightness perception depend on contextual cues.**

**A.** Perceived surface colors remain relatively stable under different illumination conditions and the consequent changes in wavelength of the light reflected from the surface. The **yellow squares** on the left and right cubes appear similar despite the fact that the wavelengths of light coming from the two sets of surfaces are very different. In fact, if the **blue squares** on the top of the left cube and the **yellow squares** on the top of the right cube are isolated from their contextual squares, their colors appear identical. (Reproduced, with permission, from www.lottolab.org.)

**B.** Brightness perception is also influenced by three-dimensional shape. The four gray squares indicated by **arrows** all have the same luminance. The apparent brightnesses are similar in the left illustration but different in the right illustration. This is because the visual system has an inherent expectation that illumination comes from above (the position of the sun relative to us), and thus the perception that the surface below the fold in the illustration on the right is brighter than the surface of the same luminance that lies above. (Reproduced, with permission, from Adelson 1993. Copyright © 1993 AAAS.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

As we move about or as the ambient illumination changes, the retinal image of an object—its size, shape, and brightness—also changes. Yet under most conditions, we do not perceive the object itself to be changing. As we move from a brightly lit garden into a dimly lit room, the intensity of light reaching the retina may vary a thousandfold. Both in the room's dim illumination and in the sun's glare, we nevertheless see a white shirt as white and a red tie as red. Likewise, as a friend walks toward you, she is seen as coming closer; you do not perceive her to be growing larger even though the image on your retina does expand. Our ability to perceive an object's size and color as constant illustrates again a fundamental principle of the visual system: It does not record images passively, like a camera, but instead uses transient and variable stimulation of the retina to construct representations of a stable, three-dimensional world.

Another example of contextual influence is color induction, whereby the appearance of a color in one region shifts toward that in an adjoining region. Shape also plays an important role in the perception of surface brightness. Because the visual system assumes that illumination comes from above, gray patches on a folded surface appear very different when they lie on the top or bottom of the surface, even when they are in fact the same shade of gray (Figure 23–12B).

The responses of some neurons in the visual cortex correlate with perceived brightness. Most visual neurons respond to surface boundaries; the center-surround structure of the receptive fields of retinal ganglion cells and geniculate neurons is suited to capturing boundaries. Most such cells do not respond to the interior parts of surfaces, for uniform interiors produce no contrast gradients across receptive fields. However, a small percentage of neurons do respond to the interiors of surfaces, signaling local brightness, texture, or color, and the responses of these neurons are influenced by context. The cell's response changes as the brightness of surfaces *outside* a cell's receptive field change, even when the brightness of the surface within the receptive field remains fixed.

Because most neurons respond to surface boundaries and not to areas of uniform brightness, the visual system calculates the brightness of surfaces from information about contrast at the edges of surfaces. The brain's analysis of surface qualities from boundary information is known as perceptual fill-in. If one fixates the boundary between a dark disk and a surrounding bright area for a few seconds, the disk will "fill in" with the same brightness as the surrounding area. This occurs because the cells that respond to edges fire only when the eye or stimulus moves. They gradually cease to respond to a stabilized image and no longer signal the presence of the boundary. Neurons with receptive fields within the disk gradually begin to respond in a fashion similar to those with receptive fields in the surrounding area, demonstrating short-term plasticity in their receptive-field properties.

An object's color always appears more or less the same despite the fact that under different conditions of illumination the wavelength distribution of light reflected from the object varies widely. To identify an object, we must know the properties of its surface rather than those of the reflected light, which are constantly changing. Computation of an object's color is therefore more complex than analyzing the spectrum of reflected light. To determine a surface's color, the wavelength distribution of the incident light must be determined. In the absence of that information, surface color can be estimated by determining the balance of wavelengths coming from different surfaces in a scene. Some neurons in V4 respond similarly to different illumination wavelengths if the perceived color remains constant. By being responsive to the light across an extensive surface, these neurons are selective for surface color rather than wavelength.

## Receptive-Field Properties Depend on Context

The distinction between local and global effects—between stimuli that occur within a receptive field and those beyond—poses the problem of how the receptive field itself is defined. Because the original characterization of the receptive fields of visual cortex neurons did not take into account contextual influences, some investigators now distinguish between "classical" and "nonclassical" receptive fields.

However, even the earliest description of the sensory receptive field allowed for the possibility of influences from portions of the sensory surface outside the narrowly defined receptive field. In 1953, Steven Kuffler, in his pioneering observations on the receptive-field properties of retinal ganglion cells, noted that "not only the areas from which responses can actually be set up by retinal illumination may be included in a definition of the receptive field but also all areas which show a functional connection, by an inhibitory or excitatory effect on a ganglion cell. This may well involve areas which are somewhat remote from a ganglion cell and by themselves do not set up discharges."

A more useful distinction contrasts the response of a neuron to a simple stimulus, such as a short line segment, with its response to a stimulus with multiple components. Even in the primary visual cortex, neurons are highly nonlinear; their response to a complex stimulus cannot be predicted from their responses to a simple stimulus placed in different positions around the visual field. Their responses to local features are instead dependent on the global context within which the features are embedded. Contextual influences are pervasive in intermediate-level visual processing, including contour integration, scene segmentation, and the determination of object shape, object motion and surface properties.

## Cortical Connections, Functional Architecture, and Perception Are Intimately Related

Intermediate-level visual processing requires sharing of information from throughout the visual field. The relationship of interconnections within the primary visual cortex to the functional architecture of this area suggests that this circuitry mediates contour integration.

Cortical circuits include a plexus of long-range horizontal connections formed by the axons of pyramidal neurons running parallel to the cortical surface. Horizontal connections exist in every area of the cerebral cortex, but their function varies from one area to the next depending on the functional architecture of each area. In the visual cortex, these connections mediate interactions between orientation columns of similar specificity, thus integrating information over a large area of visual cortex that represents a great expanse of the visual field (see Figure 21–16).

The fact that these horizontal connections link neurons similar in function but representing distant locations in the visual field suggests that these connections have a role in contour integration. Contour integration and the related property of contour saliency reflect the Gestalt principle of good continuation. Both are mediated by the horizontal connections in V1 (see Figure 23–6).

A final feature of cortical connectivity important for visuospatial integration is feedback projections from higher-order cortical areas. Feedback connections are as extensive as the feedforward connections that originate in the thalamus or at earlier stages of cortical processing. Little is known about the function of these feedback projections. They likely play a role in mediating the top-down influences of attention, expectation, and perceptual task, all of which are known to affect early stages in cortical processing.

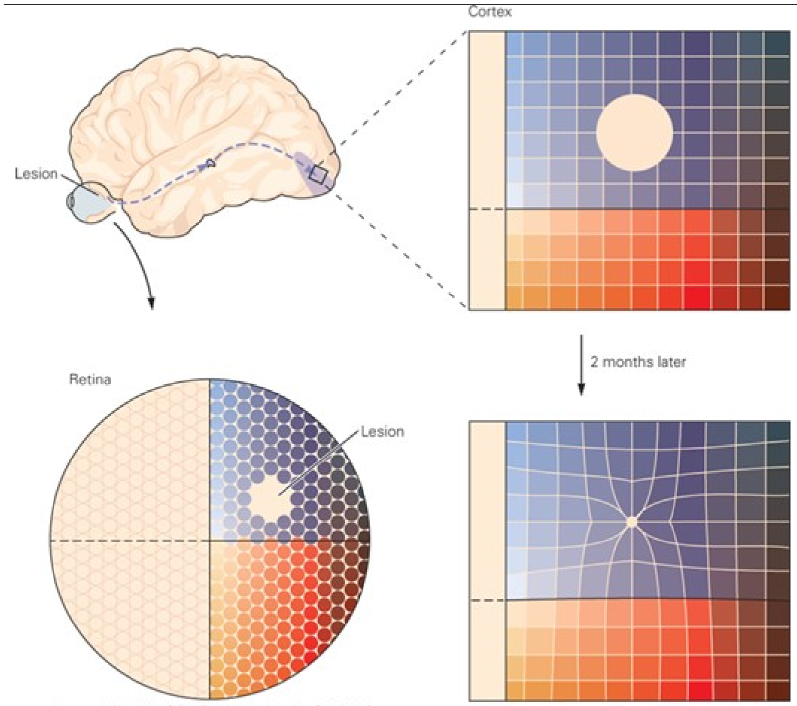
### Perceptual Learning Requires Plasticity in Cortical Connections

The synaptic connections in ocular-dominance columns are adaptable to experience only during a critical period in development (Chapter 49). This suggests that the functional properties of visual cortex neurons are fixed in adulthood. Nevertheless, many properties of cortical neurons remain mutable throughout life. For example, changes in the visual cortex can occur following retinal lesions.

When focal lesions occur in corresponding positions on the two retinas, the corresponding part of the cortical map, referred to as the lesion projection zone, is initially deprived of visual input. Over a period of several months, however, the receptive fields of cells within this region shift from the lesioned part of the retina to the functioning area surrounding the lesion. As a result, the cortical representation of the lesioned part of the retina shrinks while that of the surrounding region expands (Figure 23–13).

Figure 23–13

**Adult cortical plasticity.** When corresponding positions in both eyes are lesioned, the cortical area receiving input from the lesioned areas—the lesion projection zone—is initially silenced. The receptive fields of neurons in the lesion projection zone eventually shift from the area of the lesion to the surrounding, intact retina. This occurs because neurons surrounding the lesion projection zone sprout collaterals that form synaptic connections with neurons inside the zone. As a result, the cortical representation of the lesioned part of the retina shrinks while that of the surrounding retina expands.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

The plasticity of cortical maps and connections did not evolve as a response to lesions but as a neural mechanism for improving our perceptual skills. Many of the attributes analyzed by the visual cortex, including stereoscopic acuity, direction of movement, and orientation, become sharper with practice. Hermann von Helmholtz stated in 1866 that “the judgment of the senses may be modified by experience and by training derived under various circumstances, and may be adapted to the new conditions. Thus, persons may learn in some measure to utilize details of the sensation which otherwise would escape notice and not contribute to obtaining any idea of the object.” This perceptual learning is a variety of implicit learning that does not involve conscious processes (Chapter 52).

Perceptual learning involves repeating a discrimination task many times and does not require error feedback to improve performance. Improvement manifests itself, for example, as a decrease in the threshold for discriminating small differences in the attributes of a target stimulus or in the ability to detect a target in a complex environment. Several areas of visual cortex, including the primary visual cortex, participate in perceptual learning.

An important aspect of perceptual learning is its specificity: Training on one task does not transfer to other tasks. For example, in a three-line bisection task, the subject must determine whether the centermost of three parallel lines is closer to the line on the left or the one on the right. The amount of offset from the central position required for accurate responses decreases substantially after repeated practice.

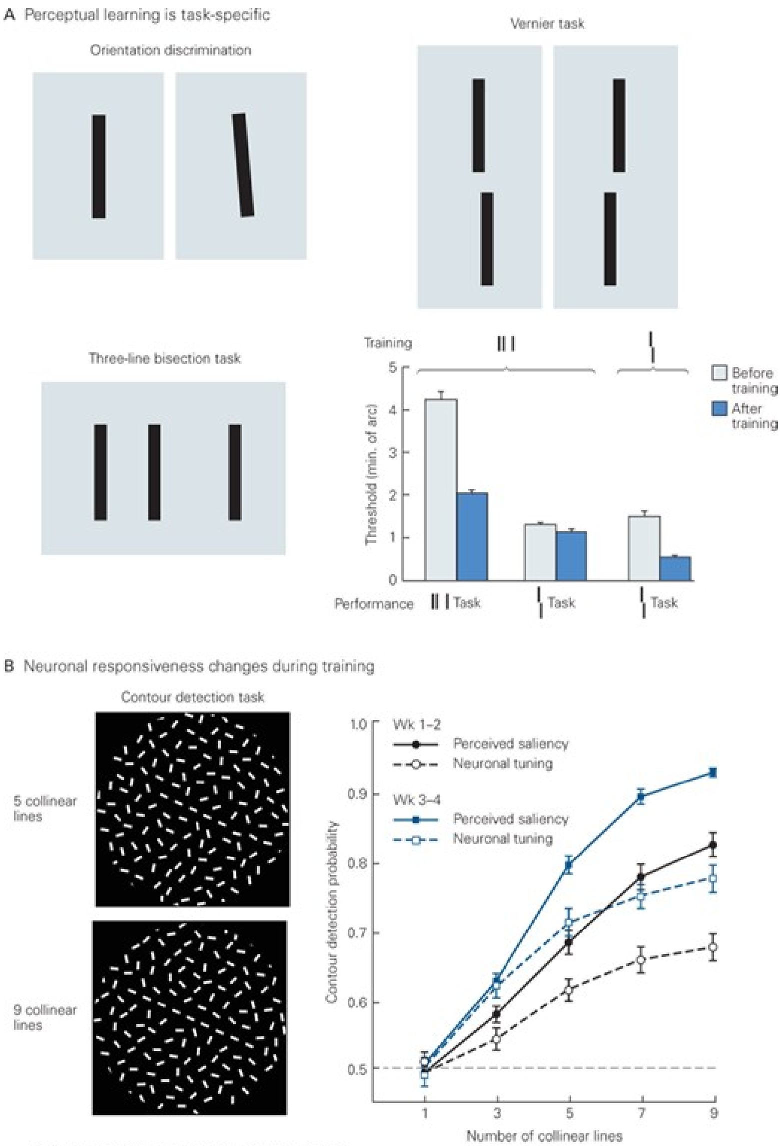
Learning of this task is specific to the location in the visual field and to the orientation of the lines. This specificity suggests that early stages of visual processing are responsible, for in the early stages, receptive fields are smallest, visuotopic maps are most precise, and orientation tuning is sharpest. The learning is also specific for the stimulus configuration. Training on three-line bisection does not transfer to a vernier discrimination task in which the context is a line that is collinear with the target line (Figure 23–14A).

Figure 23–14

**Perceptual learning.** Perceptual learning is a form of implicit learning. With practice, one can learn to discriminate smaller differences in orientation, position, depth, and direction of movement of objects.

**A.** The improvement is seen as a reduction in the amount of change required to reliably detect a tilted line or one positioned to the left or right of a nearly collinear line (vernier task). Perceptual learning is highly specific, so that training on a three-line bisection task leads to substantial improvement in that task (*left pair of bars* in the bar graph) without affecting performance on the vernier discrimination task (*central pair of bars*). However, training specifically on vernier discrimination does enhance performance on that task (*right pair of bars*).

**B.** Subjects can detect collinear line segments embedded in a random background more easily as the number of collinear segments is increased. The responses of neurons in V1 grow correspondingly stronger with the increase in the number of line segments. After practice, a line with fewer segments stands out more easily, and with this improvement, the responses in V1 also increase. (Reproduced, with permission, from Crist, Li, and Gilbert 2001; Li, Piech, and Gilbert 2008.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

The response properties of neurons in the primary visual cortex change during the course of perceptual learning in a way that tracks the perceptual improvement. An example of this is seen in contour saliency. With practice, subjects can more easily detect contours embedded in complex backgrounds. Detection improves with contour length, as do the responses of neurons in V1. With practice, subjects improve their ability to detect shorter contours and V1 neurons become correspondingly more sensitive to shorter contours (Figure 23–14B).

## Visual Search Relies on the Cortical Representation of Visual Attributes and Shapes

The detectability of features such as color, orientation, and shape is related to the process of visual search. In a complex image, certain objects stand out or "pop out" because the visual system processes simultaneously, in parallel pathways, the features of the target and the surrounding distractors (Figure 23–15). When the features of a target are complex, the target can be identified only through careful inspection of an entire image or scene.

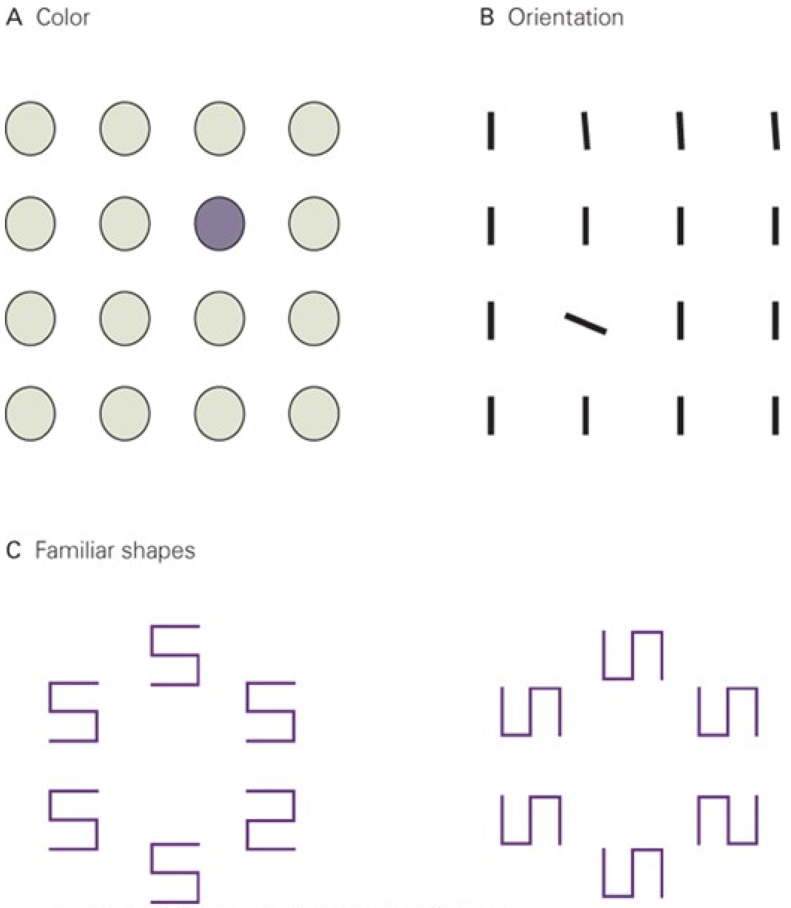
Figure 23–15

**One object in a complex image stands out under certain conditions.**

**A.** A differently colored object pops out.

**B.** A differently oriented line also pops out.

**C.** More complex shapes can pop out when they are very familiar, such as the numeral 2 embedded in a field of 5s. Rotating the image by 90° renders the elements of the figure less recognizable, making it more difficult to find the one figure that differs from the rest. (Reproduced, with permission, from Wang, Cavanagh, and Green 1994. Copyright © 1994 Springer Nature.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

The pop-out phenomenon can be influenced by training. A stimulus that initially cannot be found without effortful searching will pop out after training. The neuronal correlate of such a dramatic change is not certain. Parallel processing of the features of an object and its background is possible because feature information is encoded in retinotopically mapped areas at multiple locations in the visual cortex. Pop-out probably occurs early in the visual cortex. The pop-out of complex shapes such as numerals supports the idea that early in visual processing neurons can represent, and be selective for, shapes more complex than line segments with a particular orientation.

## Cognitive Processes Influence Visual Perception

Scene segmentation—the parsing of a scene into different objects—involves a combination of bottom-up processes that follow the Gestalt rule of good continuation and top-down processes that create object expectation.

One strong top-down influence is spatial attention, which can change focus without any movement of an observer's eyes. Spatial attention can be object-oriented in that the focus of attention is distributed over the area occupied by the attended object, allowing the visual cortex to analyze the shape and attributes of objects one at a time.

Attentional mechanisms can solve the superposition problem. Before we can recognize an object in a scene that includes many objects, we must determine which features correspond to which objects. Our sense that we identify all objects in the visual field simultaneously is illusory. Instead, we serially process objects in rapid succession by shifting attention from one to the next. The results of each analysis build up the perception of a complex environment populated with many distinct objects. A dramatic demonstration of the importance of attention in object recognition is *change blindness.* If a subject rapidly shifts between two slightly different views of the same scene, he will not be able to detect the absence of an important component of the scene in one view without considerable scrutiny (see Figure 25–8).

Another top-down influence is perceptual task. At early stages in visual processing, the properties of the same neuron vary with the type of visual discrimination being performed. Object identification involves a process of hypothesis testing in which information arriving from the retina is compared with internal representations of objects. This process is reflected in studies that have shown that early stages in processing, such as the primary visual cortex, are activated when scenes are imagined without visual input.

## Highlights

1. Vision requires segregating objects from their backgrounds, a process involving contour integration and surface segmentation.
2. This process is simplified by relying on the statistical properties of natural forms. As recognized by the Gestalt psychologists early in the 20th century, we naturally link scene components based on grouping rules of similarity, proximity, and contour smoothness (referred to as "good continuation").
3. Neurons in visual cortical areas have properties consonant with Gestalt grouping rules. They perform a local and global analysis of scene properties in parallel. The local properties are the visual primitives, which include orientation selectivity, direction selectivity, contrast sensitivity, disparity selectivity, and color selectivity. The corresponding global properties include contour integration, object movement, border ownership, disparity capture, and color constancy.
4. Perception of visual features is dependent on context; similarly, neuronal responses are context dependent. The principle underlying these interactions is the association field, a pattern of interactions between bits of information that are mapped across each cortical area. The association field mediates contour integration in visual cortex but is likely to be a general feature of processing throughout the cerebral cortex. The anatomical substrate for the association field includes a network of long-range horizontal connections formed by the axons of cortical pyramidal cells, which extend for long distances parallel to the cortical surface.
5. Different visual cortical areas contribute to the various global properties, and interactions between areas, including top-down influences, are required for their development. Though there has been considerable emphasis on selectivity for increasing stimulus complexity as one ascends a hierarchy of cortical areas through feedforward connections extending from the primary visual cortex to areas in the temporal (ventral pathway) and parietal (dorsal pathway) cortex, feedback connections are of equal importance.
6. Future studies will elucidate the relative contributions of intrinsic, feedforward, and feedback cortical connections, and the interactions between them, in cortical processing. Evidence is emerging that rather than having fixed functions, neurons are adaptive processors, taking on different functional roles under different behavioral contexts. Neurons may mediate this functional diversity by input selection, expressing task-relevant inputs and suppressing task-irrelevant inputs. When operating abnormally, these functional and connectivity dynamics may account for perceptual and behavioral phenomena associated with disorders such as autism and schizophrenia.

**Charles D. Gilbert**

## Selected Reading

Albright TD, Stoner GR. 2002. Contextual influences on visual processing. Annu Rev Neurosci 25:339–379. [PubMed: 12052913]

Gilbert CD, Sigman M. 2007. Brain states: top-down influences in sensory processing. Neuron 54:677–696. [PubMed: 17553419]

Gilbert CD, Sigman M, Crist R. 2001. The neural basis of perceptual learning. Neuron 31:681–697. [PubMed: 11567610]

Li W, Piech V, Gilbert CD. 2004. Perceptual learning and top-down influences in primary visual cortex. Nat Neurosci 7:651–657. [PubMed: 15156149]

Li W, Piech V, Gilbert CD. 2006. Contour saliency in primary visual cortex. Neuron 50:951–962. [PubMed: 16772175]

Priebe NJ, Ferster D. 2008. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. Neuron 57:482–497. [PubMed: 18304479]

## References

Adelson EH. 1993. Perceptual organization and the judgment of brightness. Science 262:2042–2044. [PubMed: 8266102]

Bakin JS, Nakayama K, Gilbert CD. 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. J Neurosci 20:8188–8198. [PubMed: 11050142]

Crist RE, Li W, Gilbert CD. 2001. Learning to see: experience and attention in primary visual cortex. Nat Neurosci 4:519–525. [PubMed: 11319561]

Cumming BG, DeAngelis GC. 2001. The physiology of stereopsis. Annu Rev Neurosci 24:203–238. [PubMed: 11283310]

Ferster D, Miller KD. 2000. Neural mechanisms of orientation selectivity in the visual cortex. Annu Rev Neurosci 23:441–471. [PubMed: 10845071]

He ZJ, Nakayama K. 1994. Apparent motion determined by surface layout not by disparity or three-dimensional distance. Nature 367:173–175. [PubMed: 8114913]

Hubel DH, Wiesel TN. 1968. Receptive fields and functional architecture of monkey striate cortex. J Physiol 195:215–243. [PubMed: 4966457]

Li W, Gilbert CD. 2002. Global contour saliency and local colinear interations. J Neurophysiol 88:2846–2856. [PubMed: 12424317]

Li W, Piech V, Gilbert CD. 2008. Learning to link visual contours. Neuron 57:442–451. [PubMed: 18255036]

Movshon JA, Adelson EH, Gizzi MS, Newsome WT. 1985. The analysis of moving visual patterns. In: C Chagas, R Gattass, CG Gross (eds). *Study Group on Pattern Recognition Mechanisms,* pp. 67–86. Vatican City: Pontifica Academia Scientiarum.

Nakayama K. 1996. Binocular visual surface perception. Proc Natl Acad Sci U S A 93:634–639. [PubMed: 8570607]

Nakayama K, Joseph JS. 2000. Attention, pattern recognition and popout in visual search. In: R Parasuraman (ed). *The Attentive Brain.* Cambridge, MA: MIT Press.

Poggio GE. 1995. Mechanisms of stereopsis in monkey visual cortex. Cereb Cortex 5:193–204. [PubMed: 7613075]

Purves D, Lotto RB, Nundy S. 2002. Why we see what we do. Am Sci 90:236–243.

Wang Q, Cavanagh P, Green M. 1994. Familiarity and pop-out in visual search. Percept Psychophys 56:495–500. [PubMed: 7991347]

Zhou H, Friedman HS, von der Heydt R. 2000. Coding of border ownership in monkey visual cortex. J Neurosci 20:6594–6611. [PubMed: 10964965]

Principles of Neural Science, 6e >

# Chapter 24: High-Level Visual Processing: From Vision to Cognition

## Introduction

AS WE HAVE SEEN, LOW-LEVEL visual processing is responsible for detecting various types of contrasts in the patterns of light projected onto the retina. Intermediate-level processing is concerned with the identification of so-called visual primitives, such as contours and fields of motion, and the segregation of surfaces. High-level visual processing integrates information from a variety of sources and is the final stage in the visual pathway leading to visual perception.

High-level visual processing is concerned with identifying behaviorally meaningful features of the environment and thus depends on descending signals that convey information from short-term working memory, long-term memory, and executive areas of cerebral cortex.

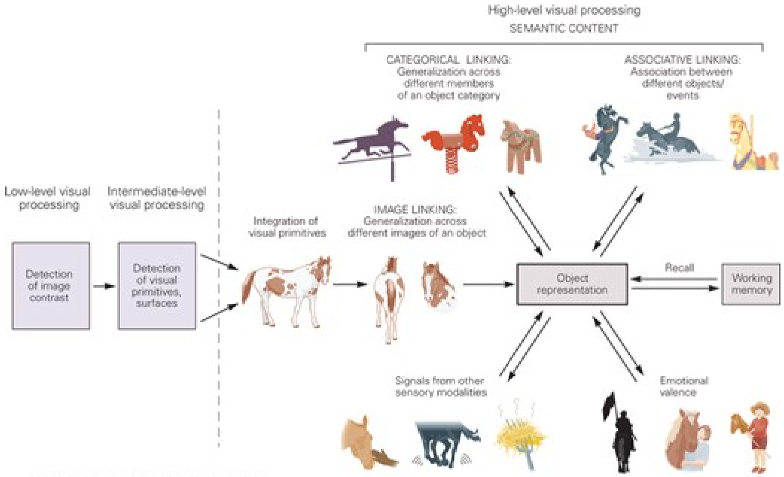
## High-Level Visual Processing Is Concerned With Object Recognition

Our visual experience of the world is fundamentally object-centered. We can recognize the same object even when the patterns of light it casts onto the retina vary greatly with viewing conditions, such as lighting, angle, position, and distance. And this is the case even for visually complex objects, those that include a large number of conjoined visual features.

Moreover, objects are not mere visual entities, but are commonly associated with specific experiences, other remembered objects, and sensations—such as the hum of the coffee grinder or the aroma of a lover's perfume—and a variety of emotions. It is the behavioral significance of objects that guides our action based on visual information. In short, object recognition establishes a nexus between vision and cognition (Figure 24–1).

Figure 24–1

**Representation of entire objects is central to high-level visual processing.** Representation of entire objects involves integration of visual features extracted at earlier stages in the visual pathways. This integration is a generalization of the numerous retinal images generated by the same object and of different members of an object category. The representation also incorporates information from other sensory modalities, attaches emotional valence, and associates the object with the memory of other objects or events. Object representations can be stored in working memory and recalled in association with other memories.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

## The Inferior Temporal Cortex Is the Primary Center for Object Recognition

Primate studies implicate neocortical regions of the temporal lobe, principally the inferior temporal cortex, in object perception. Because the hierarchy of synaptic relays in the cortical visual system extends from the primary visual cortex to the temporal lobe, the temporal lobe is a site of convergence of many types of visual information.

Neuropsychological studies have found that damage to the inferior temporal cortex can produce specific failures of object recognition. Neurophysiological and functional imaging studies have, in turn, yielded remarkable insights into the ways in which the activity of inferior temporal neurons represents objects, how these representations relate to perceptual and cognitive events, and how they are modified by experience.

Visual signals originating in the retina are processed in the lateral geniculate nucleus of the thalamus before reaching the primary visual cortex (V1). Ascending visual pathways from V1 follow two main parallel and hierarchically organized streams: the ventral and dorsal streams (Chapter 21). The ventral stream extends ventrally and anteriorly from V1 through V2, via V4, into inferior temporal cortex, which, in macaque monkeys, comprises the lower bank of the superior temporal sulcus and the ventrolateral convexity of the temporal lobe (Figure 24–2). Neurons at each synaptic relay in this ventral stream receive convergent input from the preceding stage. At the top of the hierarchy, inferior temporal neurons are in a position to integrate a large and diverse quantity of visual information over a vast region of visual space.

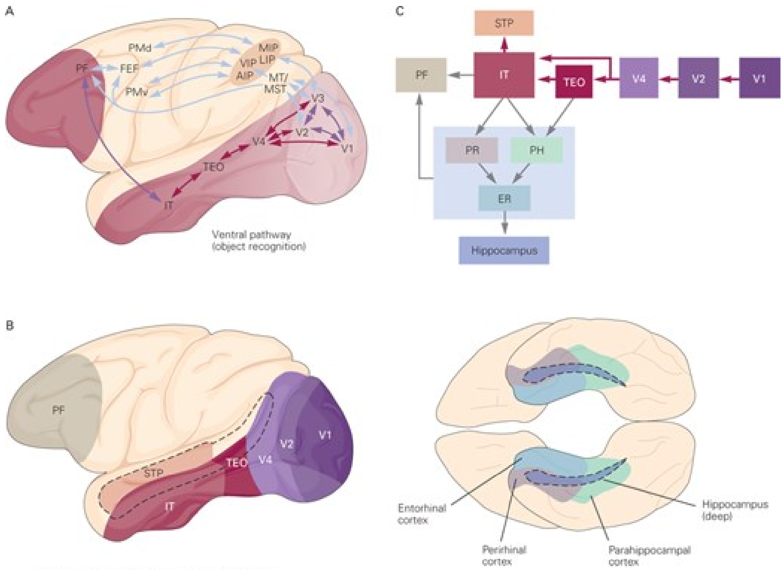
Figure 24–2

**Cortical pathway for object recognition.**

**A.** A lateral view of the macaque brain shows the major pathways involved in visual processing, including the pathway for object recognition (**red**). (Abbreviations: **AIP**, anterior intraparietal cortex; **FEF**, frontal eye fields; **IT**, inferior temporal cortex; **LIP**, lateral intraparietal cortex; **MIP**, medial intraparietal cortex; **MST**, medial superior temporal cortex; **MT**, middle temporal cortex; **PF**, prefrontal cortex; **PMd**, dorsal premotor cortex; **PMv**, ventral premotor cortex; **TEO**, temporo-occipital cortex; **VIP,** ventral intraparietal cortex.)

**B.** Lateral and ventral views of the macaque monkey brain show the cortical areas involved in object recognition. (Abbreviations: **IT**, inferior temporal cortex; **PF**, prefrontal cortex; **STP**, superior temporal polysensory area; **TEO**, temporo-occipital cortex.)

**C.** The inferior temporal cortex (**IT**) is the end stage of the ventral stream (**red arrows**) and is reciprocally connected with neighboring areas of the medial temporal lobe and prefrontal cortex (**gray arrows**). This chart illustrates the main connections and predominant direction of information flow. (Abbreviations: **ER**, entorhinal cortex; **PF**, prefrontal cortex; **PH**, parahippocampal cortex; **PR**, perirhinal cortex; **STP**, superior temporal polysensory area; **TEO**, temporo-occipital cortex.)



The inferior temporal cortex is a large brain region. The patterns of anatomical connections to and from this region indicate that it comprises at least two main functional subdivisions—the posterior area temporo-occipital cortex and the anterior area temporal cortex—and functional evidence suggests further subdivisions into multiple functionally specialized areas. As we shall see, the distinction between anterior and posterior parts of the inferior temporal cortex is supported by both neuropsychological and neurophysiological evidence.

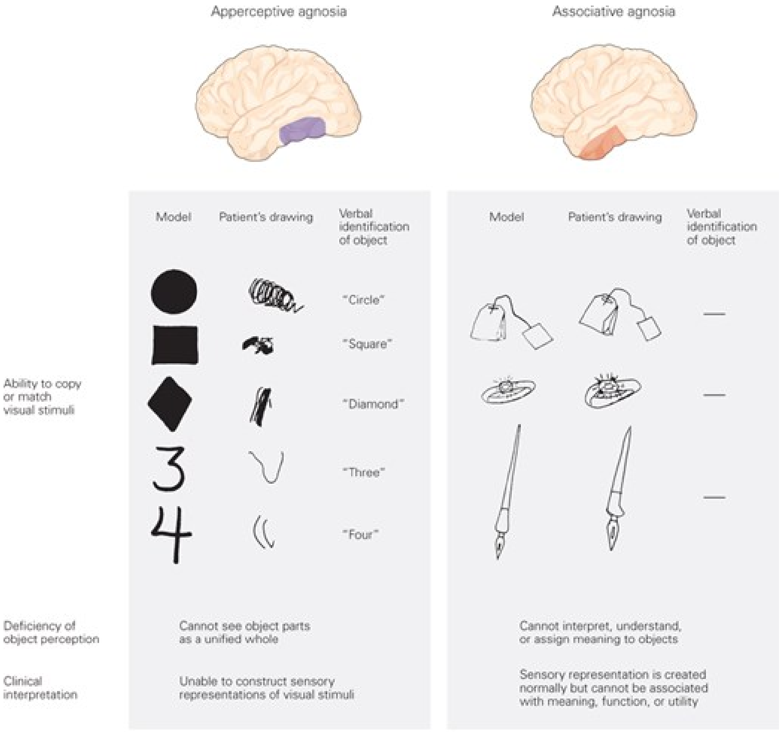
## Clinical Evidence Identifies the Inferior Temporal Cortex as Essential for Object Recognition

The first clear insight into the neural pathways mediating object recognition was obtained in the late 19th century when the American neurologist Sanger Brown and the British physiologist Edward Albert Schäfer found that experimental lesions of the temporal lobe in primates abolished the ability to recognize objects. Unlike the deficits that accompany lesions of occipital cortical areas, temporal lobe lesions do not impair sensitivity to basic visual attributes, such as color, motion, and distance. Because of the unusual type of visual loss, the impairment was originally called psychic blindness, but this term was later replaced by *visual agnosia* ("without visual knowledge"), a term coined by Sigmund Freud.

In humans, there are two basic categories of visual agnosia, apperceptive and associative, the description of which led to a two-stage model of object recognition in the visual system. With apperceptive agnosia, the ability to match or copy complex visual shapes or objects is impaired (Figure 24–3). This impairment results from disruption of the first stage of object recognition: integration of visual features into sensory representations of entire objects. With associative agnosia, the ability to match or copy complex objects remains intact, but the ability to identify objects is impaired. This impairment results from disruption of the second stage of object recognition: association of the sensory representation of an object with knowledge of the object's meaning or function.

Figure 24–3

**Neurons in the temporal lobe of humans are involved in object recognition.** Damage to the inferior temporal cortex impairs the ability to recognize visual objects, a condition known as visual agnosia. There are two major categories of visual agnosia: Apperceptive agnosia results from damage to the posterior region, and associative agnosia results from damage of the anterior region. (Reproduced, with permission, from Farah 1990. © 1990 Massachusetts Institute of Technology.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

Consistent with this functional hierarchy, apperceptive agnosia is most common following damage to the posterior inferior temporal cortex, whereas associative agnosia, a higher-order perceptual deficit, is more common following damage to the anterior inferior temporal cortex. Neurons in the anterior subdivision exhibit a variety of memory-related properties not seen in the posterior area.

More focal lesions within temporal cortex can lead to specific deficits. Damage to a small region of the human temporal lobe results in an inability to recognize faces, a form of associative agnosia known as *prosopagnosia*. Patients with prosopagnosia can identify a face as a face, recognize its parts, and even detect specific emotions expressed by the face, but they are unable to identify a particular face as belonging to a specific person.

Prosopagnosia is an example of a *category-specific agnosia*, in which patients with temporal lobe damage fail to recognize particular items belonging to a specific semantic category. Category-specific agnosias for living things, fruits, vegetables, tools, or animals have also been reported. Owing to the pronounced behavioral significance of faces and the normal ability of people to recognize an extraordinarily large number of faces, prosopagnosia may simply be the most commonly diagnosed variety of category-specific agnosia.

## Neurons in the Inferior Temporal Cortex Encode Complex Visual Stimuli and Are Organized in Functionally Specialized Columns

The coding of visual information in the temporal lobe has been studied extensively using electrophysiological techniques, beginning with the work of Charles Gross and colleagues in the 1970s. Neurons in this region have distinctive response properties. They are relatively insensitive to simple stimulus features such as orientation and color. Instead, the vast majority possess large, centrally located receptive fields and encode complex stimulus features. These selectivities often appear somewhat arbitrary. An individual neuron might, for example, respond strongly to a crescent-shaped pattern of a particular color and texture. Cells with such unique selectivities likely provide inputs to higher-order neurons that respond to specific meaningful objects.

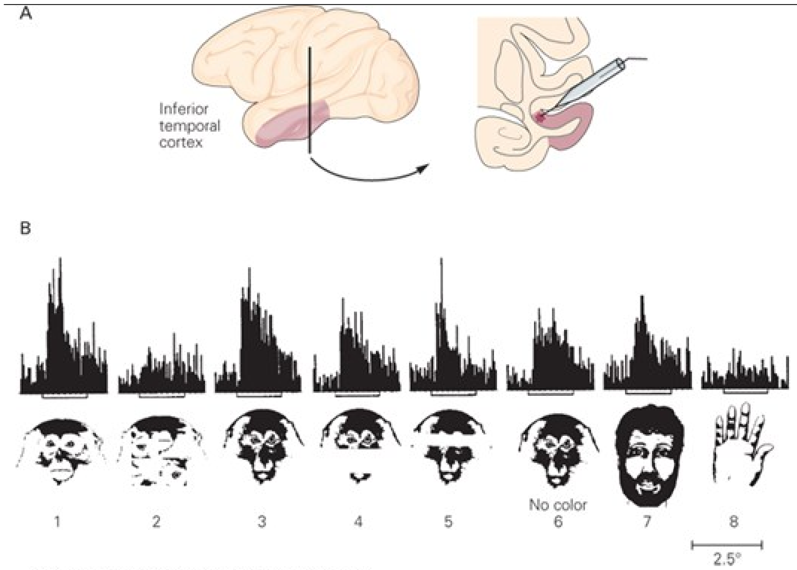
In fact, within the inferior temporal cortex, several small subpopulations of neurons are activated by highly meaningful objects, such as faces and hands (Figure 24–4), as Charles Gross discovered. For cells that respond to the sight of a hand, individual fingers are particularly critical. Among cells that respond to faces, the most effective stimulus for some cells is a frontal view of the face, whereas for others it is a side view. Although some neurons respond preferentially to faces in general, others respond only to specific facial expressions. It seems likely that such cells contribute directly to face recognition.

Figure 24–4

**Neurons in the inferior temporal cortex of the monkey are involved in face recognition.** (Reproduced, with permission, from Desimone et al. 1984. Copyright © 1984 Society for Neuroscience.)

**A.** The location of the inferior temporal cortex of the monkey is shown in a lateral view and coronal section. The colored area is the location of the recorded neurons.

**B.** Peristimulus time histograms illustrate the frequency of action potentials in a single neuron in response to different images (shown below the histograms). This neuron responded selectively to faces. Masking of critical features, such as the mouth or eyes (**4**, **5**), led to a substantial but not complete reduction in response. Scrambling the parts of the face (**2**) nearly eliminated the response.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

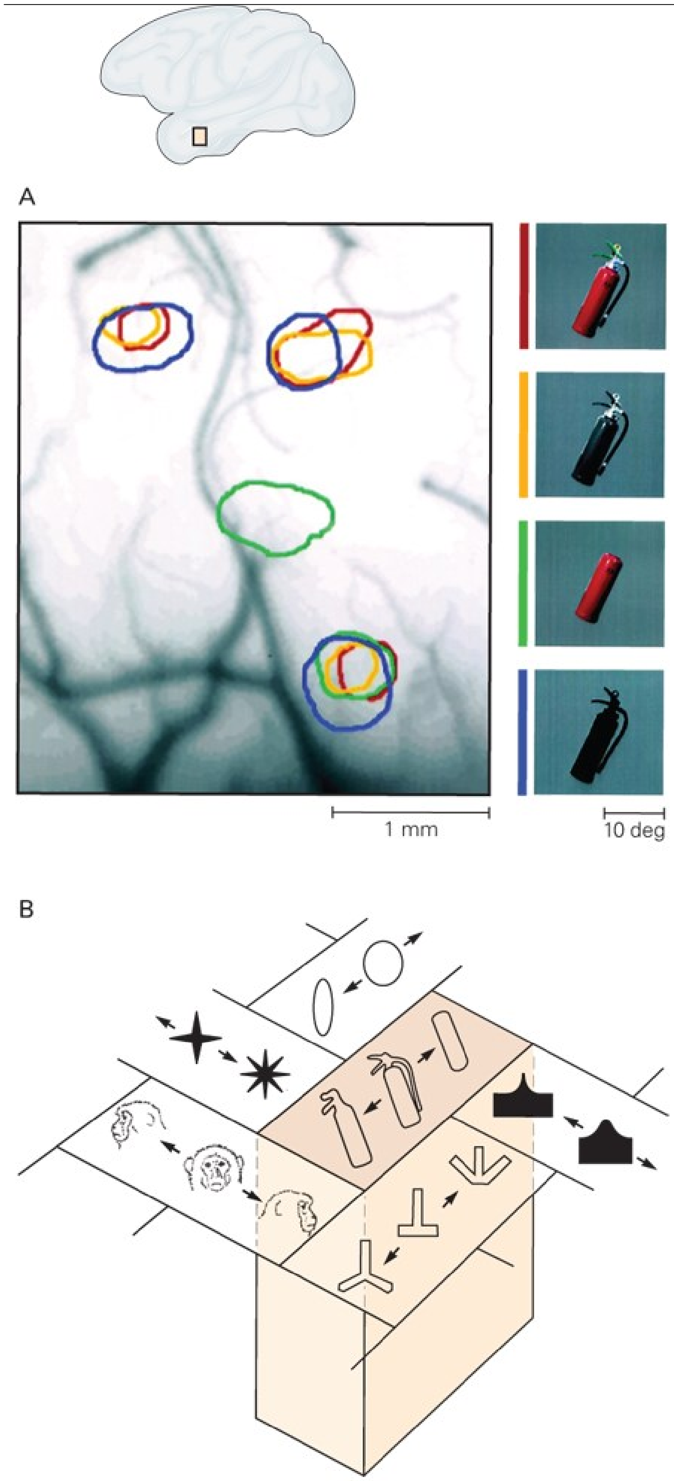
In initial relays in the cortical visual system, neurons that respond to the same stimulus features, such as orientation or direction of motion, but from different parts of the visual field are organized in columns. Cells within the inferior temporal cortex are similarly organized. Columns of neurons representing the same or similar stimulus properties commonly extend throughout the cortical thickness and over a range of approximately 400 µm. The columns are arranged such that different stimuli that possess some similar features are represented in partially overlapping columns (Figure 24–5). Thus, one stimulus can activate multiple columns. Horizontal connections can span many millimeters and may facilitate the formation of distributed networks for encoding objects.

Figure 24–5

**Neurons in the anterior portion of the inferior temporal cortex that respond to complex visual stimuli are organized into columns.** (Reproduced, with permission, from Tanaka 2003. Copyright © 2003 Oxford University Press.)

**A.** Optical images of the surface of the anterior inferior temporal cortex illustrate regions selectively activated by the objects shown at the right.

**B.** Neurons of the inferior temporal cortex are organized in functionally specialized columns that extend from the surface of the cortex. According to this model, each column includes neurons that respond to a specific visually complex object. Columns of neurons that represent variations of an object, such as different faces or different fire extinguishers, constitute a hypercolumn.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

## The Primate Brain Contains Dedicated Systems for Face Processing

Prosopagnosia often occurs in the absence of any other form of agnosia. Such a highly specific perceptual deficit could be explained by focal lesions of face-selective neurons located in exclusive clusters. This idea was strengthened by the discovery of face-selective regions in the human brain by Nancy Kanwisher and colleagues using functional magnetic resonance imaging (fMRI) and by Gregory McCarthy and colleagues using direct electrophysiological recordings from the surface of the human brain. Kanwisher and colleagues found that during the presentation of pictures of faces and other objects one area in the human temporal lobe, the fusiform face area, responded significantly more during the presentation of faces compared to other objects.

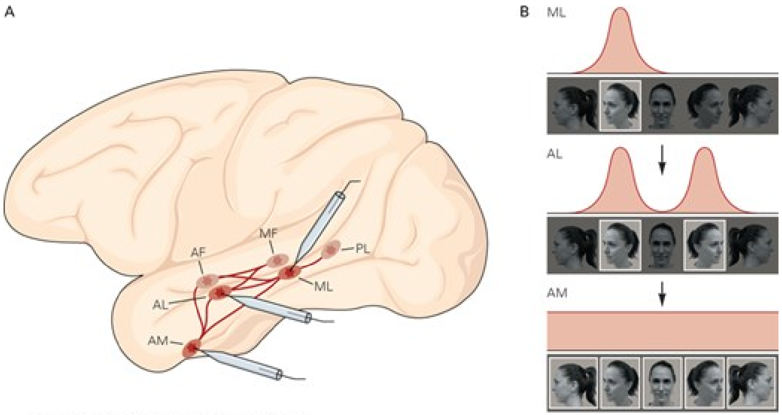
Subsequently, several more face-selective areas were found, primarily in temporal but also in prefrontal cortex. Early studies of these areas provided circumstantial evidence for clustering of face-selective neurons. In later studies, Doris Tsao, Winrich Freiwald, and colleagues directly demonstrated such clustering and showed that face processing might be performed by a dedicated face-processing network spanning from the posterior part of inferior temporal cortex to prefrontal cortex. Using fMRI, they found six areas in temporal cortex and three in prefrontal cortex of the macaque monkey that responded more selectively to faces than to other objects. These areas, called face patches, are found at highly consistent locations across individuals and thus are named based on their location. Each face patch is a few millimeters in diameter and thus differs organizationally from the inferior temporal columns. Intracellular recordings from the face patches revealed that the vast majority of cells respond selectively more to faces than to other objects. Thus, millions of face cells are clustered into a fixed number of small areas. These areas are directly connected to each other, thus forming a face-processing network. Within this network, each node appears to be functionally specialized. From posterior to anterior locations within the temporal lobe, the initial face patches respond to particular views of the face, and then face patches become gradually more selective to identity and less selective for angle of view. Furthermore, dorsal face areas within the temporal lobe exhibit a selectivity for natural facial motion, which ventral areas lack. Thus, a highly specialized network, located primarily in temporal cortex, processes the multiple dimensions of information conveyed by a face (Figure 24–6).

Figure 24–6

**The temporal lobe contains a network of face-selective areas.**

**A.** Functional magnetic resonance imaging of macaque monkeys watching pictures of faces and other objects identified six face-selective areas in the temporal lobe, inside and around the superior temporal sulcus. These areas occur at the same locations across subjects and have been given names based on their anatomical location (**PL**, posterior lateral; **ML**, medial lateral; **MF**, medial fundus of the superior temporal sulcus; **AL**, anterior lateral; **AF**, anterior fundus; **AM**, anterior medial). These areas are interconnected to form a face-processing network.

**B.** Single-neuron recordings from areas ML, AL, and AM show tuning to head orientation. ML cells are tuned to specific head orientations, many AL cells are tuned to multiple orientations that are mirror-symmetric versions of each other, and AM cells are broadly and more weakly tuned to head orientation. These three representations in interconnected areas can be thought of as transformations of each other (**arrows**).



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

## The Inferior Temporal Cortex Is Part of a Network of Cortical Areas Involved in Object Recognition

Object recognition is intimately intertwined with visual categorization, visual memory, and emotion, and the outputs of the inferior temporal cortex contribute to these functions (see Figure 24–2). Among the principal projections are those to the perirhinal and parahippocampal cortices, which lie medially adjacent to the ventral surface of the inferior temporal cortex (Figure 24–2C). These regions project, in turn, to the entorhinal cortex and the hippocampal formation, both of which are involved in long-term memory storage and retrieval. A second major projection from the inferior temporal cortex is to the prefrontal cortex, an important site for high-level visual processing. As we shall see, prefrontal neurons play important roles in object categorization, visual working memory, and memory recall.

The inferior temporal cortex also provides input—directly and indirectly via the perirhinal cortex—to the amygdala, which is believed to apply emotional valence to sensory stimuli and to engage the cognitive and visceral components of emotion (Chapter 42). Finally, the inferior temporal cortex is a major source of input to multimodal sensory areas of cortex such as the superior temporal polysensory area (Figure 24–2B), which lies dorsally adjacent to the inferior temporal cortex.

## Object Recognition Relies on Perceptual Constancy

The ability to recognize objects as the same under different viewing conditions, despite the sometimes markedly different retinal images, is one of the most functionally important requirements of visual experience. The invariant attributes of an object—for example, spatial and chromatic relationships between image features or characteristic features such as the stripes of a zebra—are cues to the identity and meaning of the objects.

For object recognition to take place, these invariant attributes must be represented independently of other image properties. The visual system does this with proficiency, and its behavioral manifestation is termed *perceptual constancy.* Perceptual constancy has many forms ranging from invariance across simple transformations of an object, such as changes of size or position, to more difficult ones, such as rotation in depth or changes in lighting, and even to the sameness of objects within a category: All zebras look alike.

One of the best examples is *size constancy.* An object placed at different distances from an observer is perceived as having the same size, even though the object produces images of different absolute size on the retina. Size constancy has been recognized for centuries, but only in the past several decades has it been possible to identify the neural mechanisms responsible. An early study found that lesions of the inferior temporal cortex lead to failures of size constancy in monkeys, suggesting that neurons in this area play a critical role in size constancy. Indeed, one of the most striking properties of individual inferior temporal neurons is the invariance of their shape selectivity even to very big changes in stimulus size (Figure 24–7A).

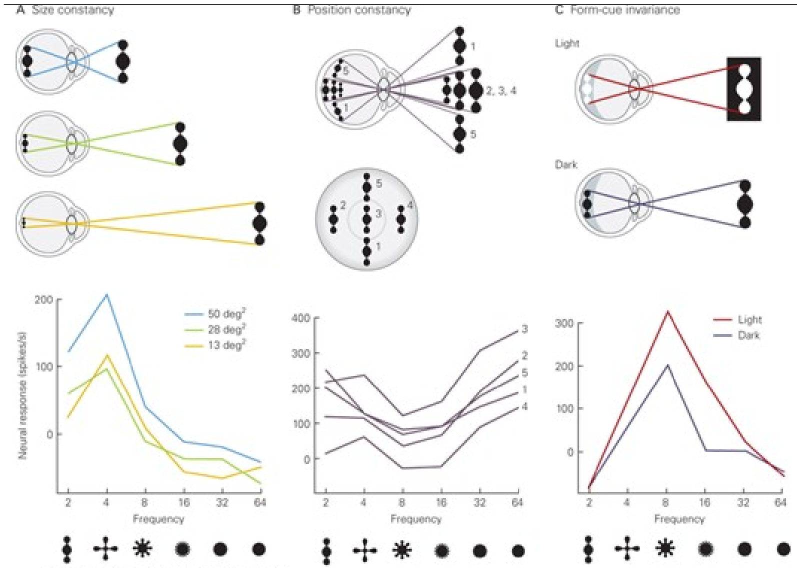
Figure 24–7

**Perceptual constancy is reflected in the behavior of neurons in the inferior temporal cortex.** The responses of many inferior temporal neurons are selective for stimuli with a particular frequency (number) of lobes but invariant to object size, position, and reflectance. (Reproduced, with permission, from Schwartz et al. 1983.)

**A.** *Size constancy.* An object is perceived to be the same even when the retinal image size decreases with the distance of the object in the visual field. The response of the vast majority of inferior temporal neurons to substantial changes in retinal image size is invariant, as illustrated here by the record of a single cell.

**B.** *Position constancy.* An object is perceived to be the same despite changes in position in the retinal image. Almost all inferior temporal neurons respond similarly to the same stimulus in different positions in the visual field, as illustrated here by the record of a single neuron.

**C.** *Form-cue invariance.* An object is perceived to be the same despite changes in reflectance. Most inferior temporal neurons respond similarly to the two images illustrated, as shown in the record of an individual neuron.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e Copyright © McGraw-Hill Education. All rights reserved.

Another type of perceptual constancy is *position constancy*, in which objects are recognized as the same regardless of their location in the visual field. The pattern of selective response of many inferior temporal neurons does not vary when an object changes position within their large receptive fields (Figure 24–7B). *Form-cue invariance* refers to the constancy of a form when the cues that define the form change. The silhouette of Abraham Lincoln's head, for example, is readily recognizable whether it is black on white, white on black, or red on green. The responses of many inferior temporal neurons do not change with changes in contrast polarity (Figure 24–7C), color, or texture.

*Viewpoint invariance* refers to the perceptual constancy of three-dimensional objects observed from different angles. Because most objects we see are three-dimensional and opaque, when looked at from different viewpoints, some parts become invisible, while others are revealed, and all others change in appearance. Yet despite the limitless range of retinal images that might be cast by a familiar object, an observer can readily recognize an object independently of the angle at which it is viewed. There are notable exceptions to this rule, which generally occur when an object is viewed from an angle that yields an uncharacteristic retinal image, such as a bucket viewed from directly above.

Thus, object recognition mechanisms must infer the identity of objects from apparent complex shapes. Many neurons in inferior temporal cortex do not exhibit viewpoint invariance. In fact, many are systematically tuned to viewing angle. Yet at more anterior locations, neurons are not only more size and position invariant, but they also exhibit greater invariance to viewpoint. The face-processing system is a case in point. Neurons in posterior face patches are tuned to viewing angle, while neurons in anterior face patches exhibit great robustness to changes in viewpoint. Thus, population responses in posterior face areas contain more information about head orientation than those in anterior areas, while the anterior face patches provide more information about face identity across head orientations compared to posterior face areas. The degree of viewpoint invariance achieved in anterior inferior temporal cortex, by individual neurons and populations of neurons, might be sufficient to account for perceptual viewpoint invariance. But this has not been directly shown yet. Alternatively, viewpoint invariance may be achieved at a higher stage of cortical processing, such as the prefrontal cortex.

Studies of the conditions under which viewpoint invariance fails may lead to insights into the neural mechanisms of the behavior. One such condition is presentation of mirror images. Although mirror images are not identical, they are frequently perceived as such, a confusion reflecting a false-positive identification by the system for viewpoint invariance. Carl Olson and colleagues examined the responses of neurons in a particular region of the inferior temporal cortex to stimuli that were mirror images. Consistent with the perceptual confusion, many inferior temporal neurons responded similarly to both images. Similarly, in one face area between the posterior and anterior ones described earlier, profile-selective cells respond similarly to the left and right profile of a face. These results reinforce the conclusion that activity in the inferior temporal cortex reflects perceptual invariance, albeit incorrectly in this case, rather than the actual features of a stimulus.

## Categorical Perception of Objects Simplifies Behavior

All forms of perceptual constancy are the product of the visual system's attempts to generalize across different retinal images generated by a single object. A still more general type of constancy is the perception of individual objects as belonging to the same semantic category. The apples in a basket or the many appearances of the letter *A* in different fonts, for example, are physically distinct but are effortlessly perceived as *categorically* identical.

Categorical perception is classically defined as the ability to distinguish objects of different categories better than objects of the same category. For example, it is more difficult to discriminate between two red lights that differ in wavelength by 10 nm than to discriminate between red and orange lights with the same wavelength difference.

Categorical perception simplifies behavior. For example, it usually does not matter whether an apple is completely spherical or slightly mottled on the left side or whether the seat we are offered is a Windsor or a Chippendale side chair. Similarly, reading ability requires that one be able to recognize the alphabet in a broad variety of type styles. Like the simpler forms of perceptual constancy, categorical perception relies on the brain's ability to extract invariant features of objects seen.

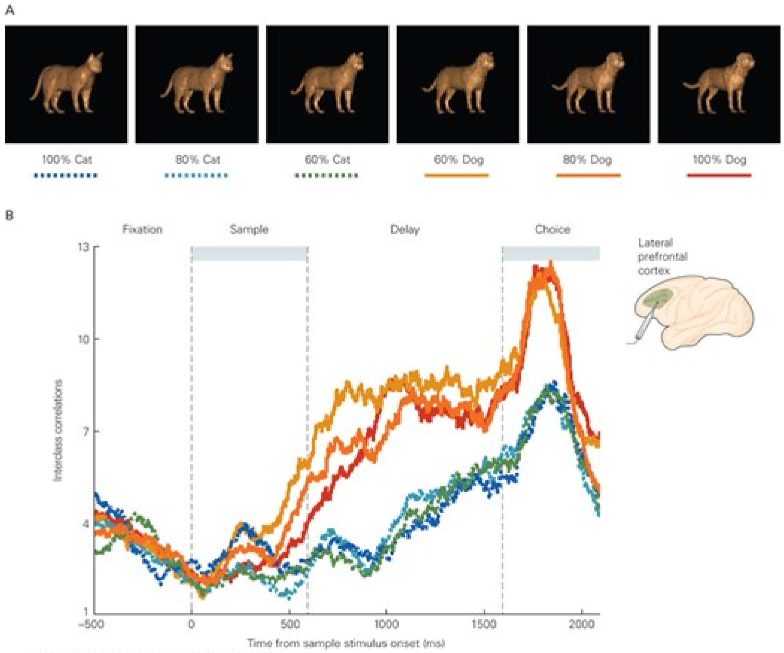
Is there a population of neurons that respond uniformly to objects within a category and differentially to objects of different categories? To test this, David Freedman and Earl Miller and colleagues created a set of images in which features of dogs and cats were merged; the proportions of dog and cat in the composite images varied continuously from one extreme to the other. Monkeys were trained to identify these stimuli reliably as either dog or cat. Miller and colleagues then recorded from visually responsive neurons in the dorsolateral prefrontal cortex, a region that receives direct input from the inferior temporal cortex. Not only did these neurons exhibit the predicted category-selective responses—responding well to cat but not dog, or vice versa—but the neuronal category boundary also corresponded to the behaviorally learned boundary (Figure 24–8). By contrast, neurons in inferior temporal cortex represented similarity of features, not categories.

Figure 24–8

**Neural coding for categorical perception.** (Reproduced, with permission, from Freedman et al. 2002.)

**A.** The images combine cat and dog features in varying proportions. Monkeys were trained to categorize an image as cat or dog if it had 50% or more features of that animal.

**B.** Peristimulus time histograms illustrate the responses of a prefrontal cortex neuron to the images shown in part A. The neuron responded much more weakly to images of cats (100%, 80%, and 60%) than to images of dogs (60%, 80%, and 100%). Responses to images from the same category were very similar despite variations in retinal images that were as large as or even larger than the differences in retinal images between categories. Thus, the cell was category-specific. Such category-specific responses were common among visual neurons of the lateral prefrontal cortex.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

The fact that category-specific agnosias sometimes follow damage to the temporal lobe suggests there are neurons in the inferior temporal cortex that are category-selective similar to those of neurons in the prefrontal cortex. Face-selective cells in the temporal cortex appear to meet this criterion, because their responses to a range of faces are often similar. Yet, these may constitute a special case, whereas for most stimulus conditions, category-selective responses may be characteristic of neurons in the prefrontal cortex, where visual responses are more commonly linked to the behavioral significance of the stimuli.

## Visual Memory Is a Component of High-Level Visual Processing

Visual experience can be stored as memory, and visual memory influences the processing of incoming visual information. Object recognition in particular relies on the observer's previous experiences with objects. Thus, the contributions of the inferior temporal cortex to object recognition must be modifiable by experience.

Studies of the role of experience in visual perception have focused on two distinct types of experience-dependent plasticity. One stems from repeated exposure or practice, which leads to improvements in visual discrimination and object recognition ability. These experience-dependent changes constitute a form of implicit learning known as perceptual learning (Chapter 23). The other occurs in connection with the storage of explicit learning, the learning of facts or events that can be recalled consciously (Chapter 54).

### Implicit Visual Learning Leads to Changes in the Selectivity of Neuronal Responses

The ability to discriminate complex visual stimuli is highly modifiable by experience. For example, individuals who attend to fine differences between automobile models improve their ability to recognize such differences.

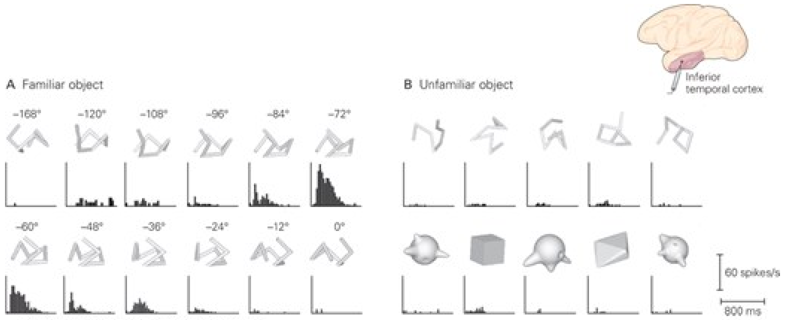
In the inferior temporal cortex, neuronal selectivity for complex objects can undergo change that parallels change in the ability to distinguish objects. For example, in a study by Logothetis and colleagues, monkeys were trained to identify novel three-dimensional objects, such as randomly bent wire forms, from two-dimensional views of the objects. Extensive training led to pronounced improvements in the ability to recognize the objects from two-dimensional views. After training, a population of neurons was found that exhibited marked selectivity for the views seen earlier but not for other two-dimensional views of the same objects (Figure 24–9).

Figure 24–9

**Familiarity with particular complex objects leads inferior temporal neurons to respond selectively for those objects.** (Reproduced, with permission, from Logothetis and Pauls 1995. Copyright © 1995 Oxford University Press.)

**A.** Monkeys were trained to recognize a randomly bent wire from a set of two-dimensional views of the wire. The wire form was rotated 12° in successive views. Once recognition performance was stable at a high level, recordings were made from neurons in the inferior temporal cortex while each view was presented. Peristimulus time histograms show the responses of a typical neuron to each view. This neuron responded selectively to views that represented a small range of rotation of the object.

**B.** When the same neuron was tested with two sets of stimuli that were unfamiliar to the monkey, it failed to respond to any of these stimuli.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Other studies with monkeys have shown that familiarization with novel faces alters the tuning of face-selective neurons in the inferior temporal cortex. Similarly, when an animal has experience with novel objects formed from simple features, inferior temporal neurons become selective for those objects. Such neuronal changes have been found as a consequence of the animal engaging in active discrimination or simply passive viewing of visual stimuli, and they are often manifested as a sharpening of neural selectivity rather than changes in absolute firing rate. Sharpening is precisely the sort of neuronal change that could underlie improvements in perceptual discrimination of visual stimuli.

## The Visual System Interacts With Working Memory and Long-Term Memory Systems

Object recognition and learning are intricately linked. In fact, learning can generate entire areas of functional specialization within inferior temporal cortex. For example, monkeys who learn at a young age to associate specific shapes (eg, a number symbol) with particular reward magnitudes develop specialized brain areas that process these specific shapes. These brain regions develop close to the temporal lobe face patches discussed earlier.

Two issues concerning interaction between vision and memory have been investigated. First, how is visual information maintained in short-term working memory? Working memory has a limited capacity, acting like a buffer in a computer operating system, and consolidation into long-term memory is susceptible to interference (Chapter 54). Second, how are long-term visual memories and the associations between them stored and recalled?

In a visual delayed-response task requiring access to stimulus information beyond the duration of the stimulus (Box 24–1), many vision-related neurons in both the inferior temporal and prefrontal cortices continue firing during the delay. This delay-period activity is thought to maintain information in short-term working memory (Figure 24–11). Delay-period activity in the inferior temporal and prefrontal cortices differs in a number of ways. For one, activity in the inferior temporal cortex is associated with the short-term storage of visual patterns and color information, whereas activity in the prefrontal cortex encodes visuospatial information as well as information received from other sensory modalities. Delay-period activity in the inferior temporal cortex also appears to be closely attuned to visual perception, for it encodes the sample image, but can be eliminated by the appearance of another image.

### Box 24–1 Investigating Interactions Between Vision and Working Memory

The relationship between vision and memory can be studied by combining a neuropsychological approach with single-cell electrophysiological methods.

One behavioral paradigm used to study memory is the *delayed-response task*. A subject is required to make a specific response based on information remembered during a brief delay. In one form of this task, known as *delayed match-to-sample*, the subject must indicate whether a visual stimulus is the same or different from a previously viewed cue stimulus (sample) (Figure 24–10A).

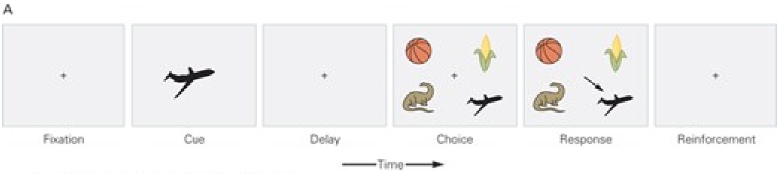
When used in conjunction with single-cell recording, this task allows the experimenter to isolate three key components of a neuronal response: (1) the sensory component, the response elicited by the cue stimulus; (2) the short-term or working-memory component, the response that occurs during the delay between the cue and the match; and (3) the recognition-memory or familiarity component, the difference between the response elicited by the match stimulus and the earlier response to the cue stimulus.

A second behavioral paradigm, the *visual paired-association task*, has been used in conjunction with electrophysiology to explore the cellular mechanisms underlying the long-term storage and recall of associations. This task differs from the delayed match-to-sample task in that the match and cue are two different stimuli (Figure 24–10B).

The sample stimulus might consist of the letter *A* and the match stimulus the letter *B*. Through repeated temporal pairing and conditional reinforcement, subjects learn that *A* and *B* are predictive of one another: They are associated.

Figure 24–10A

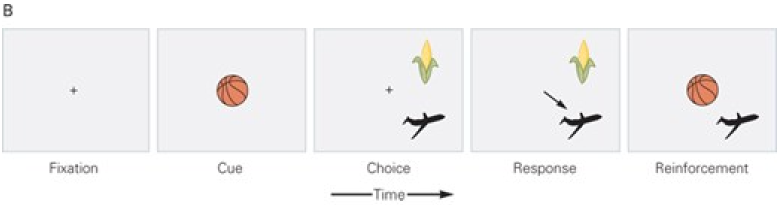
**Delayed match-to-sample task.** In this paradigm, a trial begins with the appearance of a fixation spot that directs the subject's attention and gaze to the center of the computer screen. A cue stimulus (the "sample") then appears briefly, typically for 500 ms, followed by a delay in which the display is blank. The delay can be varied to fit the experimental goals. Following the delay, the choice display appears, which contains several images, one of which is the cue (the "match"). The subject must respond by choosing the cue stimulus, typically either by pressing a button or by a saccade to the stimulus. In the task illustrated here, all of the test images appear at once (a simultaneous match-to-sample task). They can also be presented sequentially (a sequential match-to-sample task). Although the trial's duration may be longer for the sequential task, this paradigm can be advantageous for electrophysiological studies by limiting the visual stimuli present at any time.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

Figure 24–10B

**The paired-association task.** This paradigm resembles the match-to-sample paradigm except that the cue and match are different stimuli. In the illustrated example, the basketball is the cue stimulus and the airplane is the experimenter-designated match stimulus. Because these stimuli have no inherent association, the subject must discover the designated association through trial-and-error learning. The task is thus to establish an association between nonidentical stimuli. The paired-association task can also incorporate a delay between presentation of the sample and test stimuli, and it can be used in both simultaneous (shown) and sequential forms.



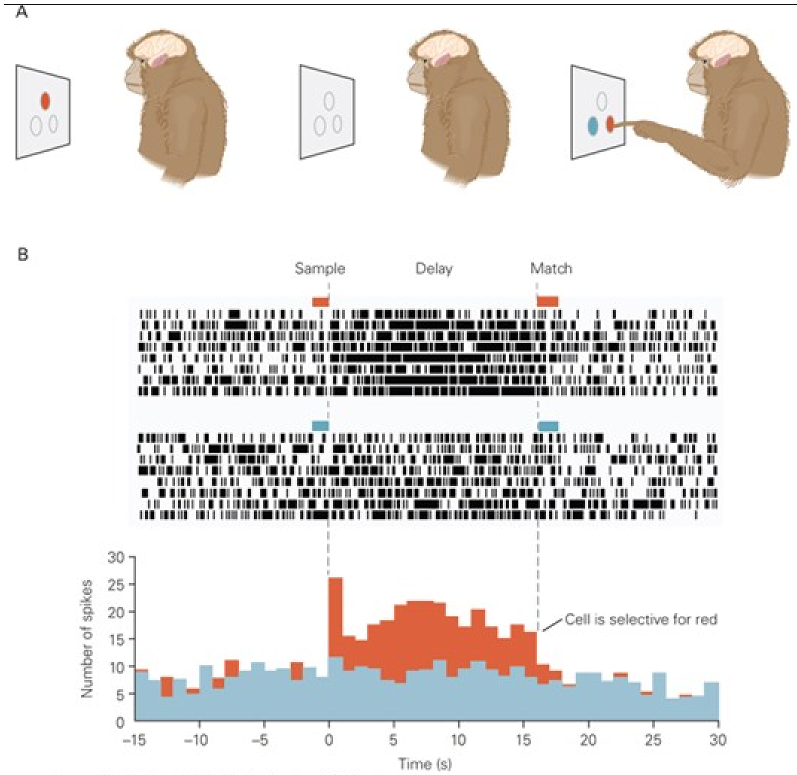
Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

Figure 24–11

**Neural activity representing an object is sustained while the object is held in working memory.** (Reproduced, with permission, from Fuster and Jervey 1982. Copyright © 1982 Society for Neuroscience.)

**A.** Monkeys were trained to perform a color match-to-sample task. For example, a red stimulus was first presented and the animal later had to choose a red stimulus from among many colored stimuli. The task incorporated a brief delay (1–2 seconds) between display of the sample and the match, during which information about the correct target color had to be maintained in working memory. The **purple** area in the monkey's brain indicates the inferior temporal cortex.

**B.** Peristimulus time histograms and raster plots of action potentials illustrate responses of a single neuron in the inferior temporal cortex during the delayed match-to-sample task. The upper record is from trials in which the sample was red, and the lower record is from trials in which it was green (shown here as **blue**). The recordings show that the cell responds preferentially to red stimuli. In trials with a green sample, the activity of the neuron does not change, whereas in trials with a red sample, the cell exhibited a brief burst of activity following presentation of the sample and continued firing throughout the delay. Many visual neurons in the inferior temporal and prefrontal cortices exhibit this kind of behavior.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

In the prefrontal cortex, by contrast, delay-period activity depends more on task requirements and is not terminated by intermittent sensory inputs, suggesting that it may play a role in the recall of long-term memories. Experiments by Earl Miller and colleagues support this view. In these experiments, monkeys were trained to associate multiple pairs of objects. They were then tested on whether they had learned these pairwise associations, using the following procedure. First, a single (sample) object was presented; then, after a brief delay, a second (test) object appeared. The monkey was instructed to indicate whether the test object was the object paired with the sample during previous training.

There are two possible ways to solve this task. During the delay, the animal could use a sensory code and keep a representation of the sample object online until the appearance of the test object, or it could remember the sample object's associate and keep information about the associate object online in a "prospective code" of what might appear as the test object. Remarkably, neuronal activity appears to transition from one to the other during the delay. Neurons in the prefrontal cortex initially encode the sensory properties of the sample object—the one just seen—but later begin to encode the expected (associated) object. As we shall see, such prospective coding in the prefrontal cortex may be the source of top-down signals to the inferior temporal cortex, activating neurons that represent the expected object and thus giving rise to conscious recall of that object.

The relation between long-term declarative memory storage and visual processing has been explored extensively in the context of remembered associations between visual stimuli. Over a century ago, William James, a founder of the American school of experimental psychology, suggested that learning visual associations might be mediated by enhanced connectivity between the neurons encoding individual stimuli. To test this hypothesis, Thomas Albright and colleagues trained monkeys to associate pairs of objects that had no prior physical or semantic relatedness. The monkeys were later tested while extracellular recordings of neurons in the inferior temporal cortex were made. Objects that had been paired often elicited similar neuronal responses, as one would expect if functional connections had been enhanced, whereas responses elicited by unpaired objects were unrelated. Recordings from individual inferior temporal neurons while monkeys were learning new visual associations showed that a cell's responses to paired objects became more similar over the course of training (Figure 24–12). Most importantly, the changes in neuronal activity occurred on the same timescale as the changes in behavior, and the changes in neural activity depended on successful learning.

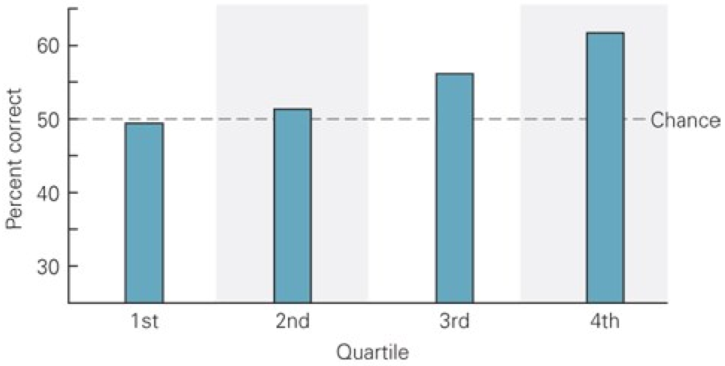
Figure 24–12

**Object recognition is linked to associative memory.** Monkeys learned associations between pairs of visual stimuli while activity was recorded from a neuron in the inferior temporal cortex. (Reproduced, with permission, from Messinger et al. 2001. © 2001 National Academy of Sciences.)

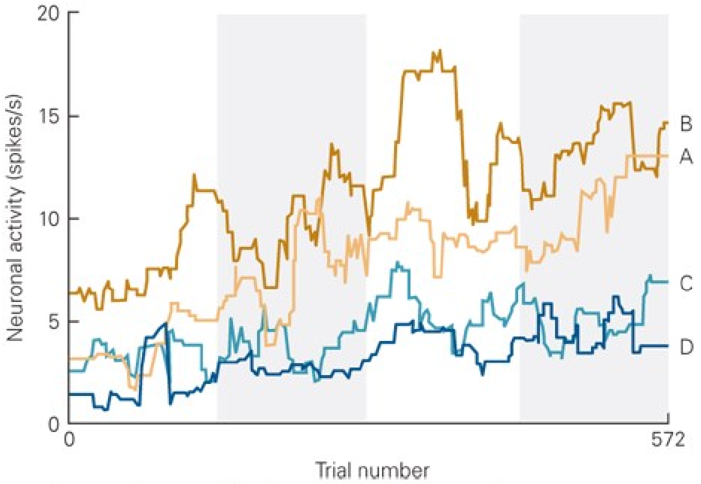
**A.** Behavioral performance on a paired-association task is plotted for each quartile of a single training session (572 trials). The animal was presented with four novel stimuli (A, B, C, D) and was required to learn two paired associations (A–B, C–D). As expected, performance began at chance (50% correct) and gradually climbed as the animal learned the associations.

**B.** Mean firing rates of an inferior temporal neuron recorded during the behavioral task described in part A. Each trace represents the firing rate during presentation of one of the four stimuli (A, B, C, or D). The responses to all stimuli were of similar magnitude at the outset. As the paired associations were learned, the neuronal responses to the paired stimuli A and B began to cluster at a different level from responses to the paired stimuli C and D. The neuron's activity thus corresponded to the learned associations between the two pairs.

A Animals learn to associate pairs of stimuli



B After training neurons respond similarly to paired stimuli



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e
Copyright © McGraw-Hill Education. All rights reserved.

These learning-dependent changes in the stimulus selectivity of inferior temporal cortex neurons are long-lasting, suggesting that this cortical region is part of the neural circuitry for associative visual memories. The experimental results also support the view that learned associations are implemented rapidly by changes in the strength of synaptic connections between neurons representing the associated stimuli.

We know that the hippocampus and neocortical areas of the medial temporal lobe—the perirhinal, entorhinal, and parahippocampal cortices—are essential both for the acquisition of associative visual memories and for the functional plasticity of the inferior temporal cortex. In fact, work by Yasushi Miyashita and colleagues showed that the aforementioned pair-association neurons are much more prevalent in perirhinal cortex than in anterior inferior temporal cortex. Thus, although learning changes the stimulus selectivity of neurons in both areas, the association between visually associated pairs grows stronger from inferior temporal to perirhinal cortex (Figure 24–2C). The hippocampus and medial temporal lobe may facilitate the reorganization of local neuronal circuitry in the inferior temporal cortex necessary to store associative visual memories. The reorganization itself may be a form of Hebbian plasticity (Chapter 49) initiated by the temporal coincidence of the associated visual stimuli.

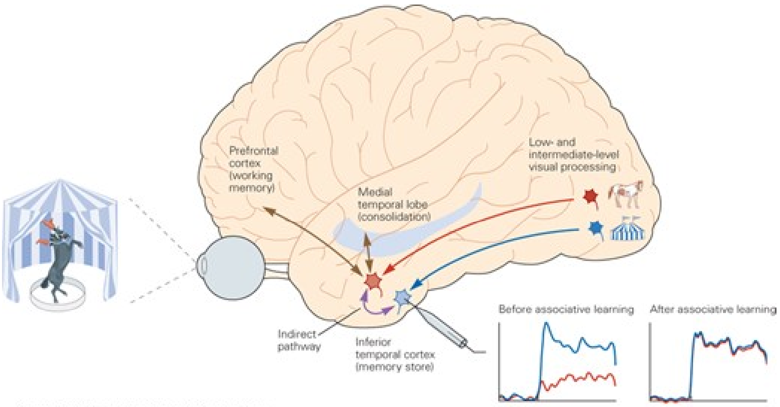
## Associative Recall of Visual Memories Depends on Top-Down Activation of the Cortical Neurons That Process Visual Stimuli

One of the most intriguing features of high-level visual processing is the fact that the detection of an image in one's visual field and the recall of the same image are subjectively similar. The former depends on the bottom-up flow of visual information and is what we traditionally regard as vision. The latter, by contrast, is a product of top-down information flow. This distinction is anatomically accurate but obscures the fact that under normal conditions afferent and descending signals collaborate to yield visual experience.

The study of associative visual memory has provided valuable insights into the cellular mechanisms underlying visual recall. As we have seen, visual associative memories are stored in the visual cortex through changes in the functional connectivity between neurons that independently represent the associated stimuli. The practical consequence of this change is that a neuron that responded only to stimulus *A* prior to learning will respond to both *A* and *B* after these stimuli have been associated (Figure 24–13). Activation of an *A*-responsive neuron by stimulus *B* can be viewed as the neuronal correlate of top-down recall of stimulus *A.*

Figure 24–13

**Circuits for visual association and recall.** Bottom-up signals—afferent signals conveying information about objects in the observer's visual field—are combined into object representations in the inferior temporal cortex. Before associative learning, a neuron (**blue**) responds well to the circus tent but not to the horse. Learned associations between objects are mediated in the inferior temporal cortex by strengthening connections between neurons representing each of the paired objects (the indirect pathway in the figure). Thus, recall of the circus tent following presentation of the horse is achieved by activating the indirect pathway. Indirect activation can also be triggered by the contents of working memory (feedback from the prefrontal cortex). Under normal conditions, visual perception is the product of a combination of direct and indirect inputs to inferior temporal neurons.



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Neurons in the inferior temporal cortex exhibit precisely this behavior. The activity correlated with cued recall is nearly identical to the bottom-up activation by the stimulus. These neurophysiological findings are supported by a number of brain imaging studies that have identified selective activity in the visual cortex during cued and spontaneous recall of objects.

Although learned associations between images are likely to be stored through circuit changes in the inferior temporal cortex, activation of these circuits for conscious recall depends on input from the prefrontal cortex. The afferent signal for one of a pair of images might be received by the inferior temporal cortex and relayed to prefrontal cortex, where the information would be maintained in working memory. As we have seen, the continued firing of many prefrontal neurons during the delay period of a delayed match-to-sample task initially represents information about the sample image but changes to the associated image that is expected to follow. Signals from prefrontal cortex to the inferior temporal cortex would selectively activate neurons representing the associated image, and that activation would constitute the neural correlate of visual recall.

## Highlights

1. A key function of high-level vision is object recognition. Object recognition imbues visual perception with meaning. As the eminent

neuropsychologist Hans-Lukas Teuber once wrote, failure of object recognition “would appear in its purest form as a normal percept that has somehow been stripped of its meaning.”

2. Object recognition is difficult, primarily because of changes in appearance with changes in position, distance, orientation, or lighting conditions, possibly rendering different objects similar in appearance. Building computer models mimicking primate object recognition capabilities is a major challenge for current and future research.

3. Object recognition relies on a region of the temporal lobe called inferior temporal cortex. Visual information reaching inferior temporal cortex has already been processed through mechanisms of low- and mid-level vision.

4. Lesions to inferior temporal cortex cause visual agnosia, a loss in the ability to recognize objects. Apperceptive agnosia, the inability to match or copy complex objects, is distinguished from associative agnosia, the impairment of the ability to recognize an object’s meaning or function. Predicting the exact nature of an agnosia from the pattern of lesioned or inactivated areas, and thus to go from understanding the *correlates* to the *causes* of neural object representations, is a major goal for the field of object recognition and neurology.

5. Individual cells in inferior temporal cortex can be highly shape-selective and respond selectively, eg, to a hand or a face. They can maintain selectivity across position, size, and even rotation—properties that might explain perceptual constancy.

6. Inferotemporal cortex comprises a yet-unknown number of areas with very different functional specializations. While the functional logic of the overall organization remains unknown, we do know that cells with similar selectivity group into cortical columns and that face cells are organized into larger units called face areas.

7. Face recognition is supported by multiple face areas, each with a unique functional specialization. Face areas are selectively coupled to form a face-processing network, which has emerged as a model system for high-level vision.

8. Inferotemporal cortex is interconnected with perirhinal and parahippocampal cortices for memory formation, with the amygdala for the assignment of emotional valence to objects, and with prefrontal cortex for object categorization and visual working memory. If associative memories are stored as patterns of connections between neurons, what then are the specific contributions of hippocampus and neocortical structures of the medial temporal lobe, and by what cellular mechanisms do they exert their influences? The confluence of molecular-genetic, cellular, neurophysiological, and behavioral approaches promises to solve these and other problems.

9. Objects are perceived as members of a category. This simplifies the selection of appropriate behaviors, which often do not depend on stimulus details. Neurons with categorical selectivity are found in dorsolateral prefrontal cortex, a main projection site of inferior temporal cortex.

10. Object recognition relies on past experience. Perceptual learning can improve the ability to discriminate between complex objects and refine neural selectivity in inferior temporal cortex.

11. Visual information can be held in short-term working memory to be available beyond the duration of a sensory stimulus. Neurons in temporal and prefrontal cortex can exhibit delay-period activity after the disappearance of a stimulus. How these networks establish the ability to keep information online is an open question.

12. High-level visual information processing changes with top-down modulation. The sensory experience of an image in view and the recall of the same stimulus from memory are subjectively similar. Neurons in inferior temporal cortex exhibit similar activity during bottom-up activation and cued recall.

**Thomas D. Albright**

**Winrich A. Freiwald**

## Selected Reading

Freedman DJ, Miller EK. 2008. Neural mechanisms of visual categorization: insights from neurophysiology. Neurosci Biobehav Rev 32:311–329. [PubMed: 17950874]

Gross CG. 1999. *Brain, Vision, Memory: Tales in the History of Neuroscience.* Cambridge, MA: MIT Press.

Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302–4311. [PubMed: 9151747]

Logothetis NK, Sheinberg DL. 1996. Visual object recognition. Annu Rev Neurosci 19:577–621. [PubMed: 8833455]

McCarthy G, Puce A, Gore J, Allison T. 1997. Face-specific processing in the human fusiform gyrus. J Cog Neurosci 9:605–610.

Messinger A, Squire LR, Zola SM, Albright TD. 2005. Neural correlates of knowledge: stable representation of stimulus associations across variations in behavioral performance. Neuron 48:359–371. [PubMed: 16242414]

Miller EK, Li L, Desimone R. 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. Science 254:1377–1379. [PubMed: 1962197]

Miyashita Y. 1993. Inferior temporal cortex: where visual perception meets memory. Annu Rev Neurosci 16:245–263. [PubMed: 8460893]

Schlack A, Albright TD. 2007. Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881–890. [PubMed: 17359922]

Squire LR, Zola-Morgan S. 1991. The medial temporal lobe memory system. Science 253:1380–1386. [PubMed: 1896849]

Ungerleider LG, Courtney SM, Haxby JV. 1998. A neural system for human visual working memory. Proc Natl Acad Sci U S A 95:883–890. [PubMed: 9448255]

## References

Baker CI, Behrmann M, Olson CR. 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat Neurosci 5:1210–1216. [PubMed: 12379864]

Brown S, Schafer ES. 1888. An investigation into the functions of the occipital and temporal lobes of the monkey's brain. Philos Trans R Soc Lond B Biol Sci 179:303–327.

Damasio AR, Damasio H, Van Hoesen GW. 1982. Prosopagnosia: anatomic basis and behavioral mechanisms. Neurology 32:331–341. [PubMed: 7199655]

Desimone R, Albright TD, Gross CG, Bruce CJ. 1984. Stimulus selective properties of inferior temporal neurons in the macaque. J Neurosci 8:2051–2062.

Desimone R, Fleming J, Gross CG. 1980. Prestriate afferents to inferior temporal cortex: an HRP study. Brain Res 184:41–55. [PubMed: 6766778]

Farah MJ. 1990. *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*. Cambridge, MA: MIT Press.

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1:1–47. [PubMed: 1822724]

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2002. Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. J Neurophysiol 88:929–941. [PubMed: 12163542]

Freiwald WA, Tsao DY. 2010. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science

330:845–851. [PubMed: 21051642]

Fujita I, Tanaka K, Ito M, Cheng K. 1992. Columns for visual features of objects in monkey inferotemporal cortex. Nature 360:343–346. [PubMed: 1448150]

Fuster JM, Jervey JP. 1982. Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. J Neurosci 2:361–375. [PubMed: 7062115]

Gross CG, Bender DB, Rocha-Miranda CE. 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science 166:1303–1306. [PubMed: 4982685]

Kosslyn SM. 1994. *Image and Brain.* Cambridge, MA: MIT Press.

Leibo JZ, Liao Q, Anselmi F, Freiwald WA, Poggio T. 2017. View-tolerant face recognition and Hebbian learning imply mirror-symmetric neural tuning to head orientation. Curr Biol 27:62–67. [PubMed: 27916522]

Logothetis NK, Pauls J. 1995. Psychophysical and physiological evidence for viewer-centered object representations in the primate. Cereb Cortex 5:270–288. [PubMed: 7613082]

Messinger A, Squire LR, Zola SM, Albright TD. 2001. Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A 98:12239–12244. [PubMed: 11572946]

Miyashita Y, Chang HS. 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature 331:68–70. [PubMed: 3340148]

Rainer G, Rao SC, Miller EK. 1999. Prospective coding for objects in primate prefrontal cortex. J Neurosci 19:5493–5505. [PubMed: 10377358]

Rollenhagen JE, Olson CR. 2000. Mirror-image confusion in single neurons of the macaque inferotemporal cortex. Science 287:1506–1508. [PubMed: 10688803]

Sakai K, Miyashita Y. 1991. Neural organization for the long-term memory of paired associates. Nature 354:152–155. [PubMed: 1944594]

Schwartz EL, Desimone R, Albright TD, Gross CG. 1983. Shape recognition and inferior temporal neurons. Proc Natl Acad Sci U S A 80:5776–5778. [PubMed: 6577453]

Suzuki WA, Amaral DG. 2004. Functional neuroanatomy of the medial temporal lobe memory system. Cortex 40:220–222. [PubMed: 15070014]

Tanaka K. 2003. Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. Cereb Cortex 13:90–99. [PubMed: 12466220]

Teuber HL. 1968. Disorders of memory following penetrating missile wounds of the brain. Neurology 18:287–288. [PubMed: 5690365]

Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y. 1999. Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature 401:699–703. [PubMed: 10537108]

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS. 2006. A cortical region consisting entirely of face-selective cells. Science 311:670–674. [PubMed: 16456083]

Wheeler ME, Petersen SE, Buckner RL. 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc Natl Acad Sci U S A 97:11125–11129. [PubMed: 11005879]