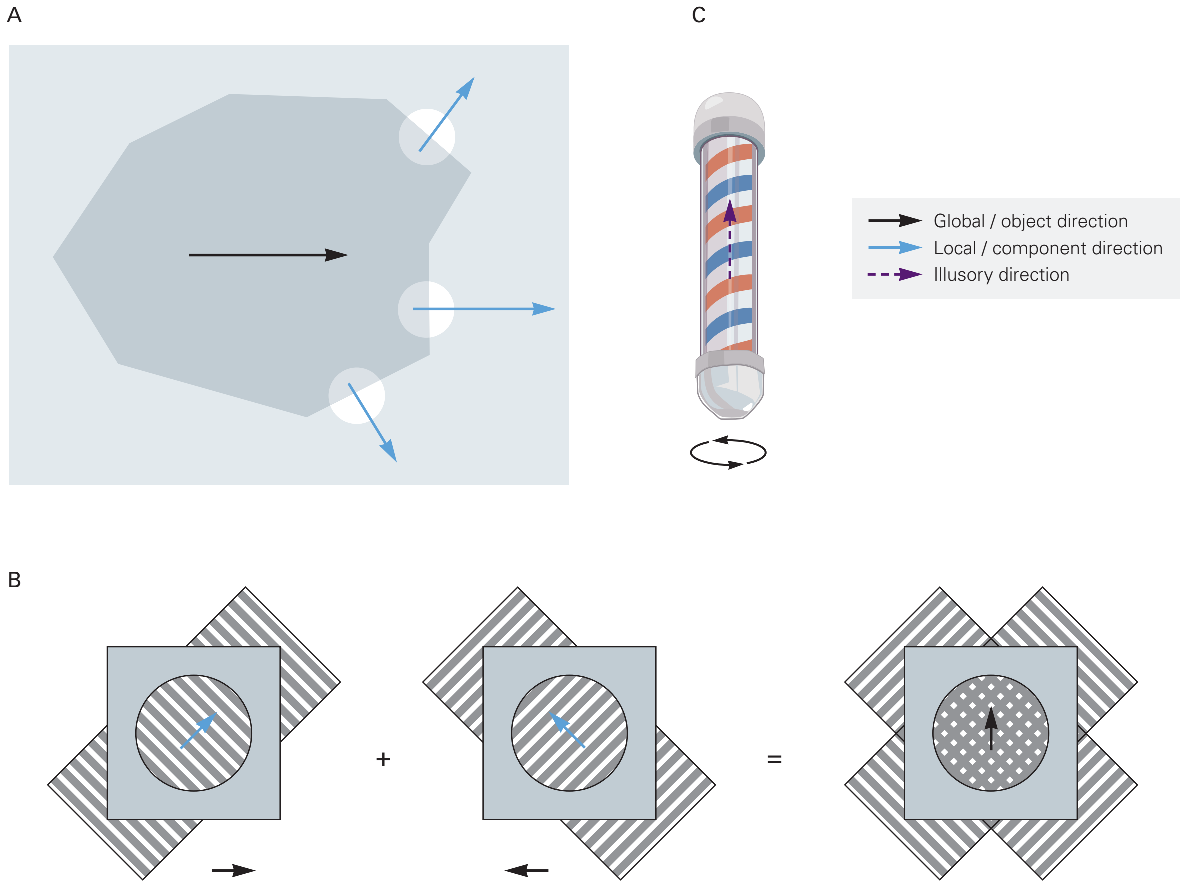

Figure 23–11 The aperture problem and barber-pole illusion.
A. Although an object moves in one direction, each component edge when viewed through a small aperture appears to move in a direction perpendicular to its orientation. The visual system must integrate such local motion signals into a unified 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.

that illuminates them changes from natural to artificial light, from sunlight to shadow, or from dawn to midday (Figure 23–12A).

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

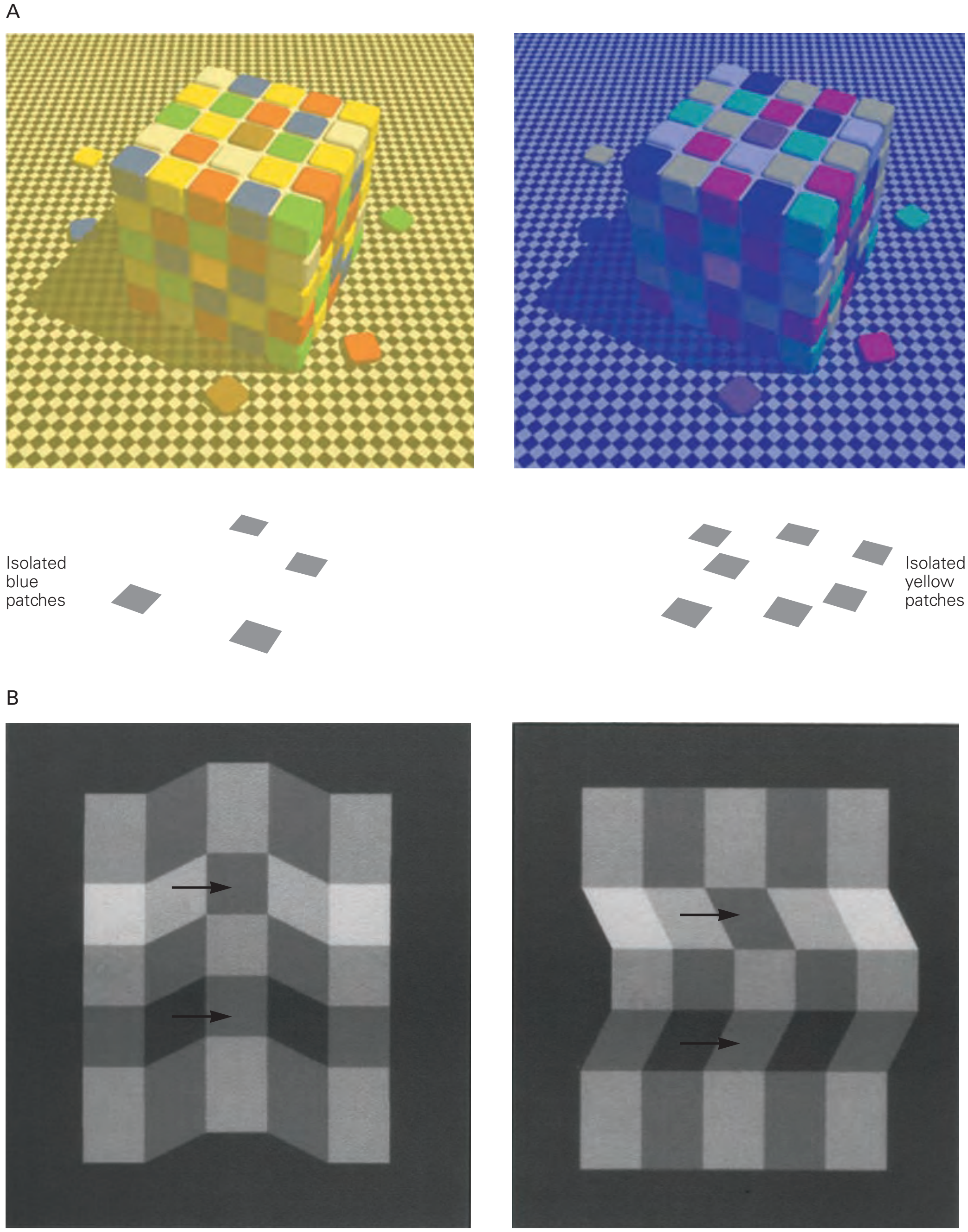

Figure 23–12 Color and brightness perception depend on contextual cues.

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.)

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).

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).

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"

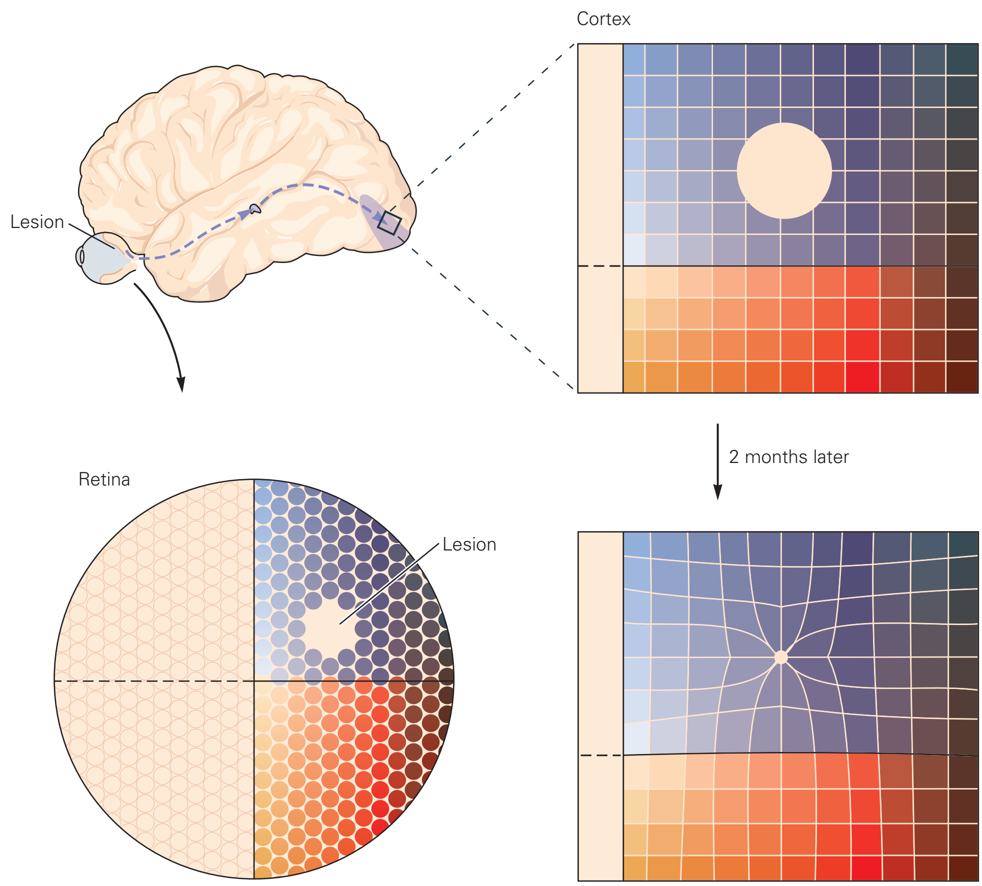

Figure 23–13 **Adult cortical plasticity.** When corresponding positions in both eyes are lesioned, the cortical area receiving input from the lesioned areas—the lesion projection zone—is initially silenced. The receptive fields of neurons in the lesion projection zone eventually shift from the area 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.

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.

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

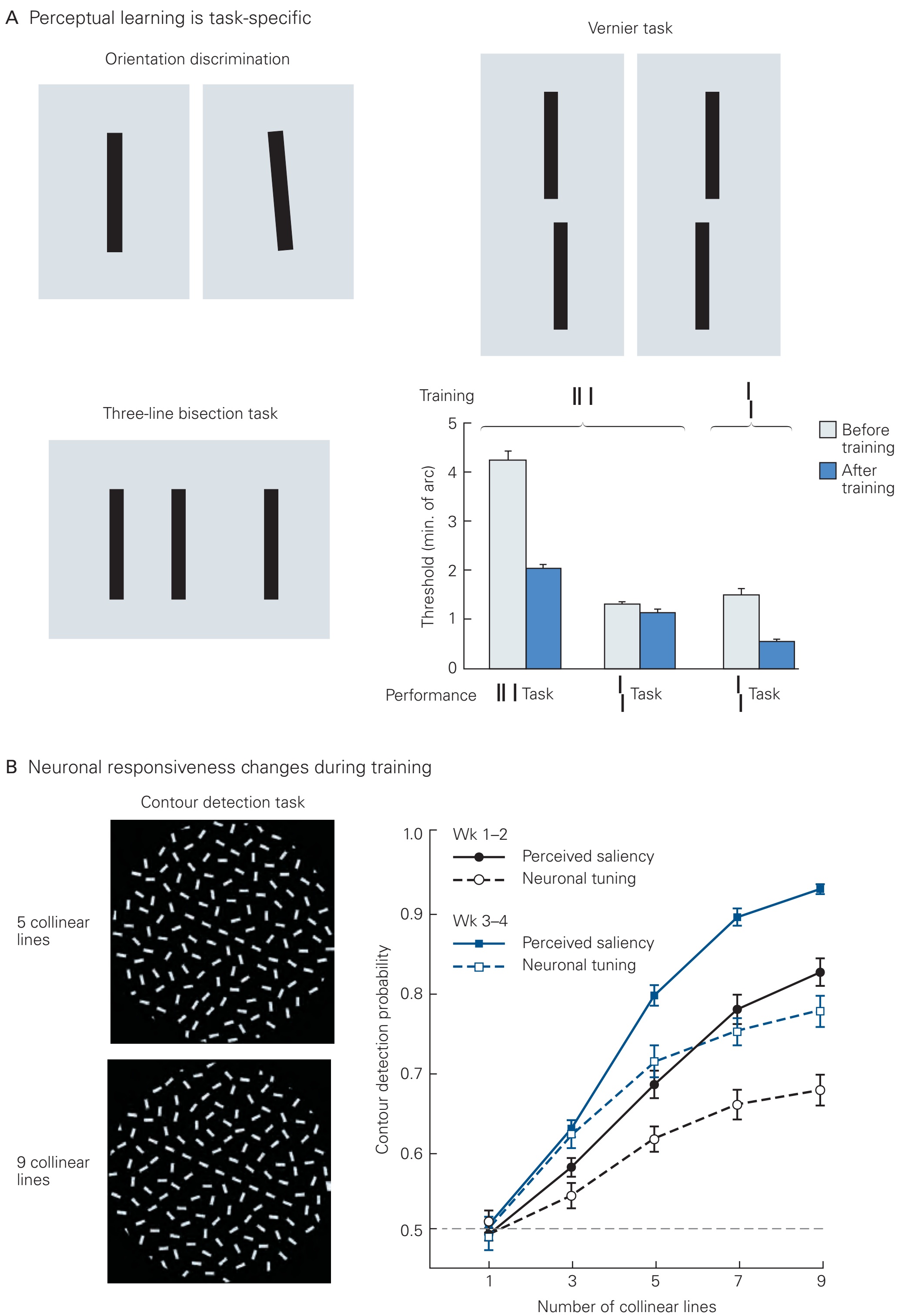

Figure 23–14 Perceptual learning. Perceptual learning is a form of implicit learning. With practice, one can learn to discriminate smaller differences in orientation, position, depth, and direction of movement of objects.

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.)

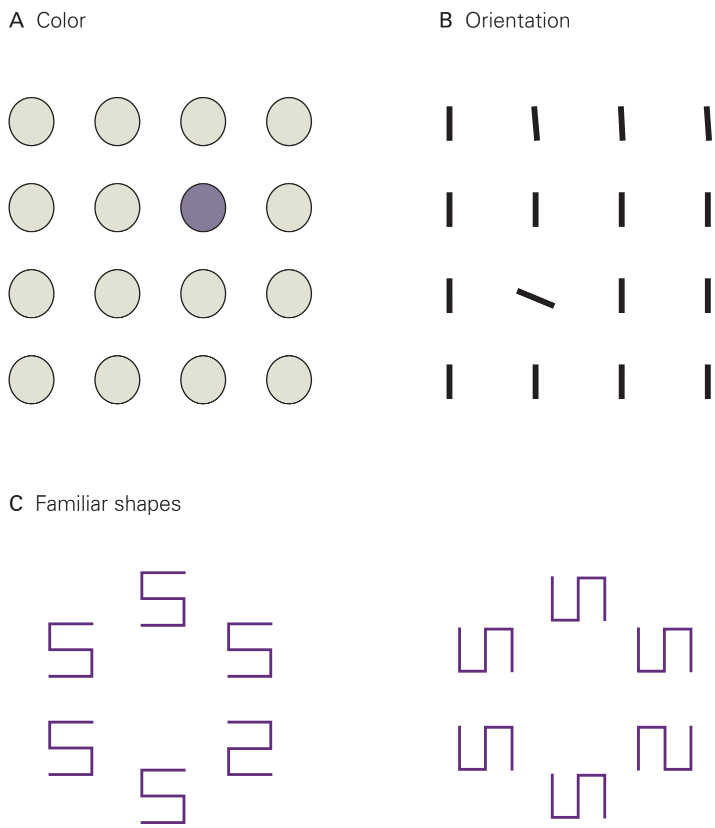

Figure 23–15 **One object in a complex image stands out under certain conditions.**

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.)

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.

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

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

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

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

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

References

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

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.

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

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

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

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

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

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

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

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.

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.

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.

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

24

High-Level Visual Processing: From Vision to Cognition

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).