

## High-Level Visual Processing: Cognitive Influences

### High-Level Visual Processing Is Concerned with Object Identification

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

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

Neurons in the Inferior Temporal Cortex Encode Complex Visual Stimuli

Neurons in the Inferior Temporal Cortex Are Functionally Organized in Columns

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

#### Object Recognition Relies on Perceptual Constancy

#### Categorical Perception of Objects Simplifies Behavior

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

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

Explicit Visual Learning Depends on Linkage of the Visual System and Declarative Memory Formation

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

#### An Overall View

**T**HE IMAGES PROJECTED ONTO THE retina are generally complex dynamic patterns of light of varying intensity and color. As we have seen, low-level visual processing is responsible for detection of various types of contrast in these images (see

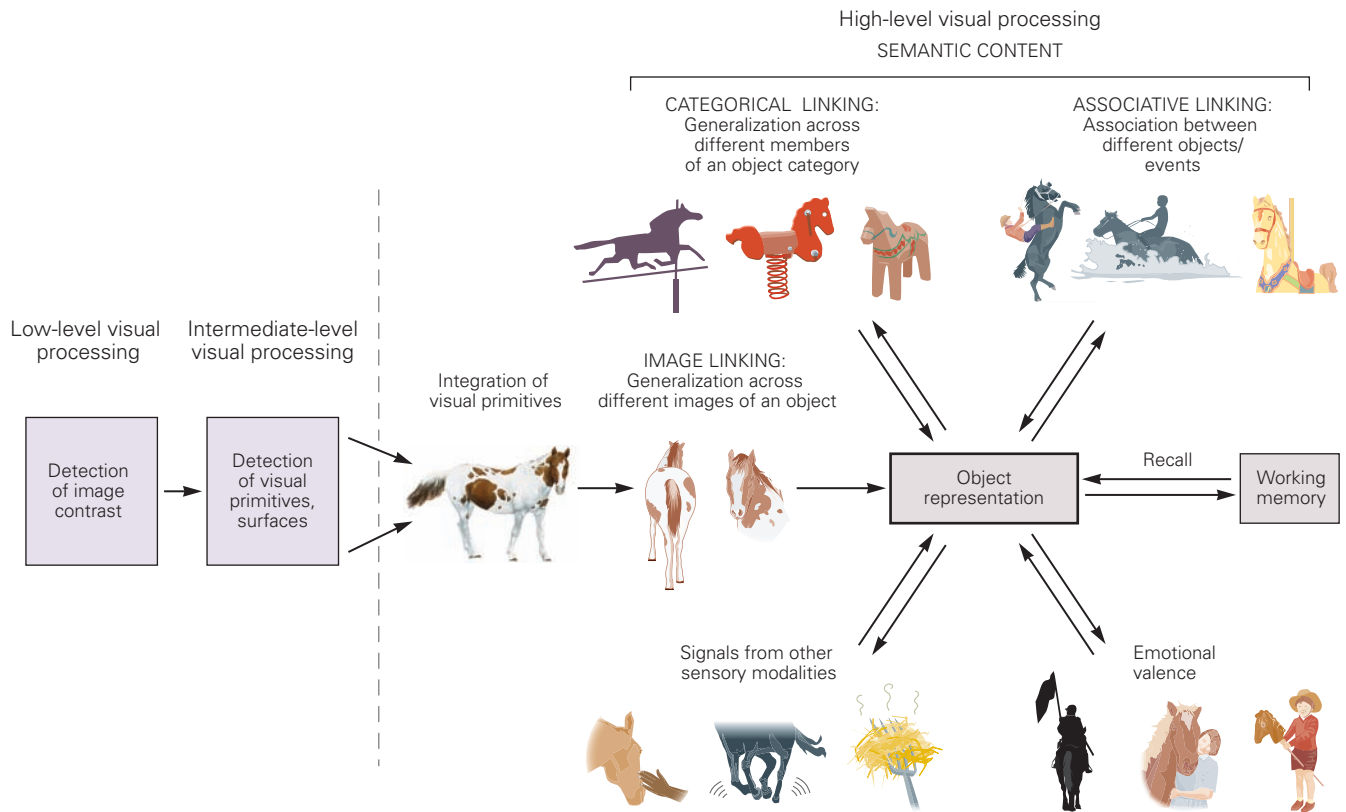
Chapters 25 and 26), whereas intermediate-level processing is involved in the identification of so-called visual primitives, such as contours and fields of motion, and the representation of surfaces (see Chapter 27). High-level visual processing integrates information from a variety of sources and is the final stage in the visual pathway leading to conscious visual experience.

In practice high-level visual processing depends on top-down signals that imbue bottom-up (afferent) sensory representations with semantic significance, such as that arising from short-term working memory, long-term memory, and behavioral goals. High-level visual processing thus selects behaviorally meaningful attributes of the visual environment (Figure 28–1).

### High-Level Visual Processing Is Concerned with Object Identification

Our visual experience of the world is fundamentally object-centered. Objects are often visually complex, being composed of a large number of conjoined visual features. In addition, the features projected on the retina by an object vary greatly under different viewing conditions, such as lighting, angle, position, and distance.

Moreover, objects are commonly associated with specific experiences, other remembered objects, other sensations—such as the hum of the coffee grinder or the aroma of a lover’s perfume—and a variety of emotions. Animate beings, which are objects to the visual



**Figure 28–1** The neuronal representation of entire objects is central to high-level visual processing. Object representation involves integration of visual features extracted at earlier stages in the visual pathways. Ideally the resulting representation 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.

system, also direct intentions, desires, and actions at others and ourselves. In conjunction with our own behavioral goals, it is the behavioral saliency of individual objects, memories, and emotional valences as well as the real or implied actions of others that enables us to take action based on visual information. Object perception is thus the nexus between vision and cognition.

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

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.

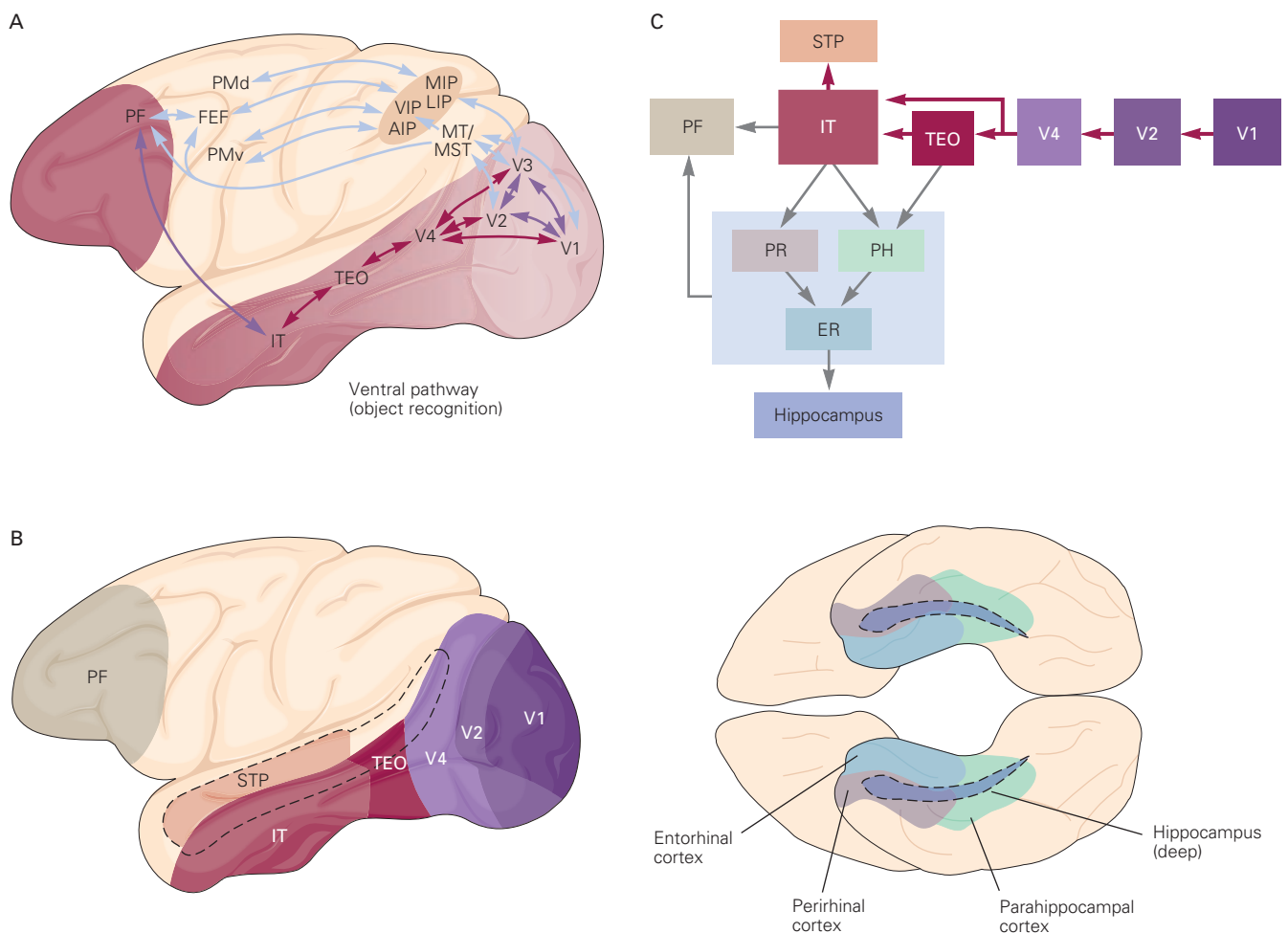
As we shall later see, neuropsychological studies have found that damage to the inferior temporal cortex can produce specific failures of object recognition. Neurophysiological and brain-imaging studies have in turn yielded remarkable insights into the ways in which the activity of inferior temporal 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). Thereafter ascending visual pathways follow two parallel and hierarchically organized streams: the ventral and dorsal streams (see Chapter 25). The ventral stream extends ventrally and anteriorly from V1 through

V2, V4, and the temporal-occipital junction before reaching the inferior temporal cortex, which comprises the lower bank of the superior temporal sulcus and the ventrolateral convexity of the temporal lobe (Figure 28–2). This pathway makes the inferior temporal cortex the seat of the highest stage of cortical visual processing. Neurons at each synaptic relay in this ventral stream receive convergent input from the preceding stage. Inferior temporal neurons are thus in a position to

integrate a large and diverse quantity of visual information over a vast region of visual space.

The inferior temporal cortex is a large brain region. The patterns of anatomical connections to and from this area indicate that it comprises at least two main functional subdivisions: the posterior and anterior inferior temporal cortex. Anatomical evidence identifies the anterior subdivision as a higher processing stage than the posterior subdivision. As we shall see,



**Figure 28–2** Cortical pathway for object recognition.

**A.** The pathway for object recognition (red) is identified in a lateral view of the brain showing the major pathways involved in visual processing. (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.** Cortical areas involved in object recognition are shown on lateral and ventral views of the monkey brain.

**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. (ER, entorhinal cortex; PF, prefrontal cortex; PH, parahippocampal cortex; PR, perirhinal cortex; STP, superior temporal polysensory area; TEO, temporo-occipital cortex.)

this distinction 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 resulted in loss of the ability to recognize objects. This impairment is distinct from the deficits that accompany lesions of occipital cortical areas in that sensitivity to basic visual attributes, such as color, motion, and distance, remains intact. 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 (Figure 28–3). With apperceptive agnosia the ability to match or copy complex visual shapes or objects is impaired. This impairment is perceptual in nature, resulting from disruption of the first stage of object recognition: integration of visual features into sensory representations of entire objects. In contrast, patients with associative agnosia can match or copy complex objects, but their ability to identify the 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.

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, a later stage in the functional hierarchy. Neurons in the anterior subdivision exhibit a variety of memory-related properties not seen in the posterior area.

### **Neurons in the Inferior Temporal Cortex Encode Complex Visual Stimuli**

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 yet higher-order neuronal representations of meaningful objects.

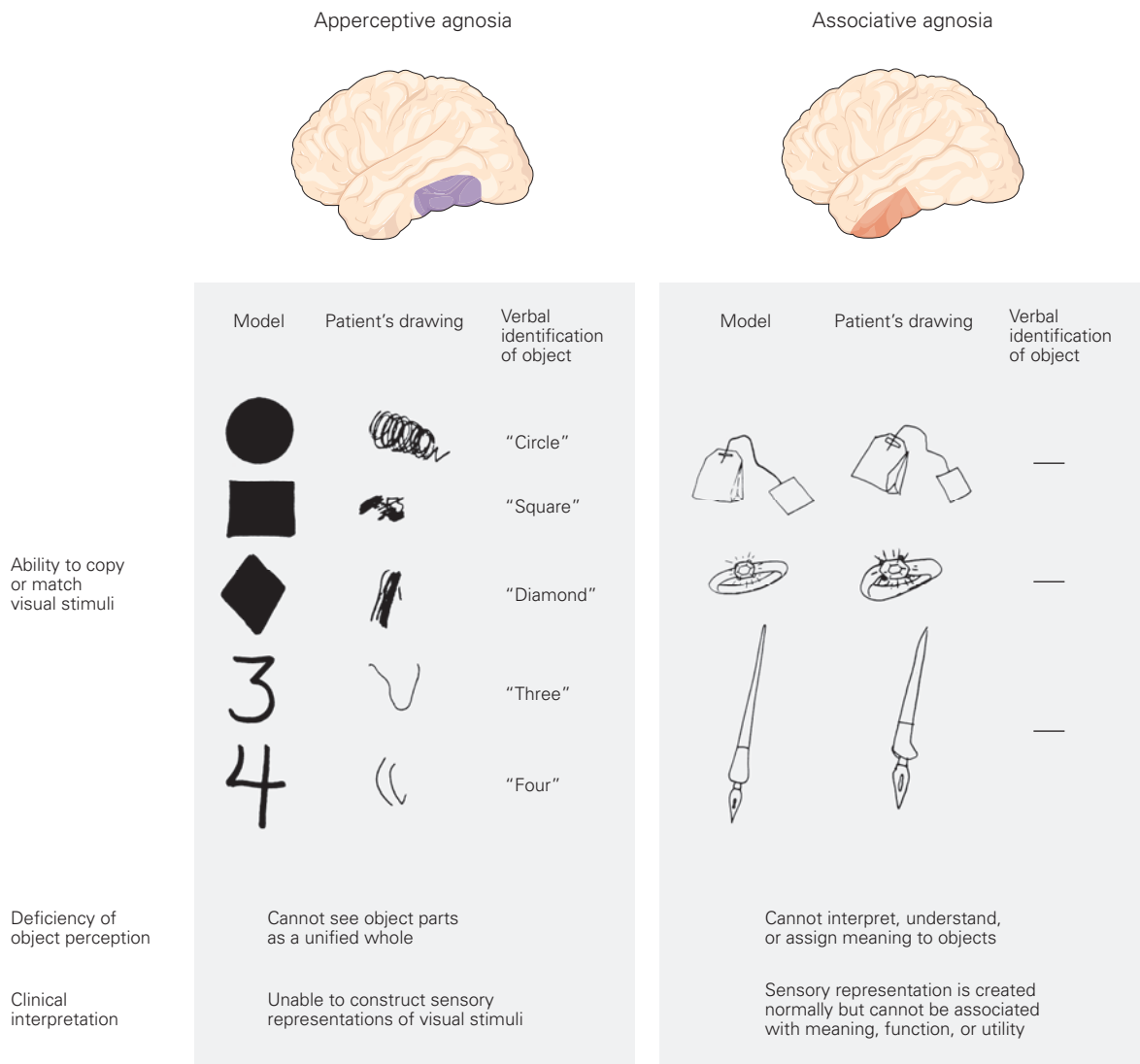
Indeed, several small subpopulations of neurons encode objects that convey to the observer highly meaningful information, such as faces and hands (Figure 28–4). For cells that respond to the sight of a hand, individual fingers are particularly critical. Among cells that respond to faces, the most effective stimulus for some cells is the frontal view of the face, whereas for others it is the side view. Although some neurons respond preferentially to faces, others respond to facial expressions. It seems likely that such cells contribute directly to face recognition.

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 one example of “category-specific” agnosia, in which patients with temporal-lobe damage fail to recognize items within a specific semantic category. There are reported cases of category-specific agnosias for living things, fruits, vegetables, tools, or animals. Owing to the pronounced behavioral significance of faces and the normal ability of people to recognize an extraordinarily large number of items from this category, prosopagnosia may simply be the most common variety of category-specific agnosia.

### **Neurons in the Inferior Temporal Cortex Are Functionally Organized in Columns**

Early relays in the cortical visual system are organized in columns of neurons that represent the same stimulus features, such as orientation or direction of motion, in different parts of the visual field. Cells within the inferior temporal cortex are also organized in columns of neurons representing the same or similar stimulus properties (Figure 28–5). These columns commonly extend throughout the cortical thickness and over a range of approximately 400  $\mu\text{m}$ . Columnar patches in the inferior temporal cortex are arranged such that different stimuli that possess some similar features are



**Figure 28–3** Neuropsychological evidence for the neuronal correlates of object recognition in the temporal lobe. Damage to inferior temporal cortex (IT) impairs the ability to recognize visual objects, a condition known as visual agnosia.

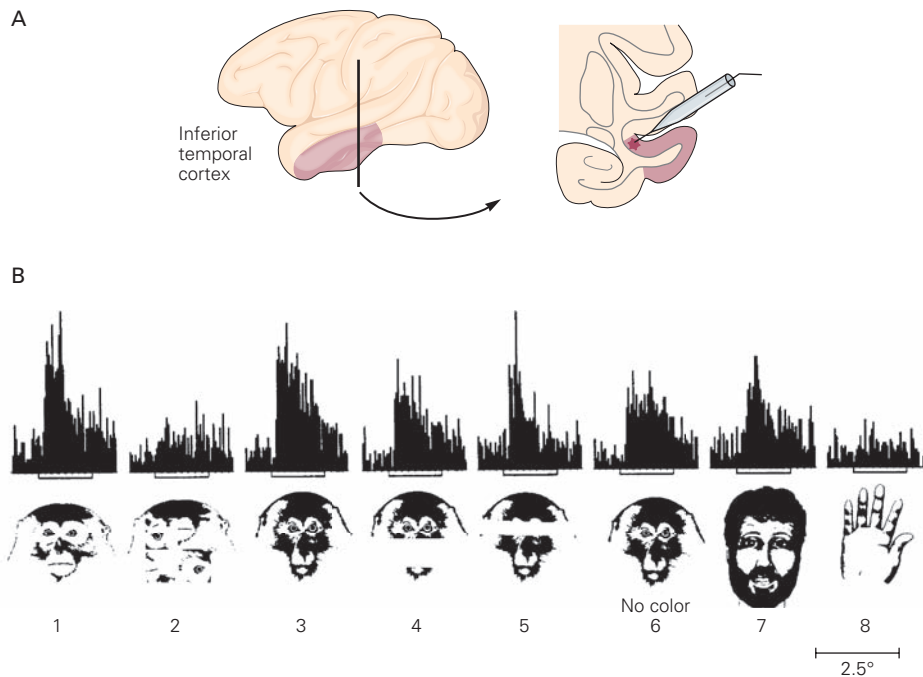
There are two major categories of visual agnosia: apperceptive, a result of damage to the posterior region, and associative, resulting from damage of the anterior region. (Reproduced, with permission, from Farah 1990.)

represented in partially overlapping columns (Figure 28–5). Thus one stimulus can activate multiple patches within the cortex. Long-range horizontal connections within the cortex may serve to connect patches into distributed networks for object representation.

Face-selective cells constitute a highly specialized class of neurons. Indeed, the fact that prosopagnosia often occurs in the absence of any other form of agnosia suggests that face-selective neurons of the temporal lobe may be located in exclusive clusters. While many early studies of neuronal response properties offered circumstantial evidence for such clustering, in

2006 Doris Tsao and Margaret Livingstone obtained dramatic support for this hypothesis. Functional magnetic resonance images of monkeys that were viewing faces revealed large active zones in a region of cortex in the lower bank of the superior temporal sulcus. Neurophysiological recordings of neurons in these zones confirmed that face-recognition cells formed large, dense clusters (Figure 28–6). Winrich Freiwald and Tsao later found that the five face-representation areas in monkeys interconnect with one another and form a processing system, with each node apparently concerned with a different aspect of face recognition.





**Figure 28-4** Neurons in the inferior temporal cortex of the monkey are involved in face recognition. (Reproduced, with permission, from Desimone et al. 1984.)

**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 histograms illustrate the frequency of action potentials in a single neuron in response to the different images illustrated below. 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.

### 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 28-2).

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 (see Figure 28-2C). 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, which is increasingly recognized as an important contributor to high-level vision. As we shall see, prefrontal neurons play important roles in categorical visual perception, visual working memory, and recall of stored memories.

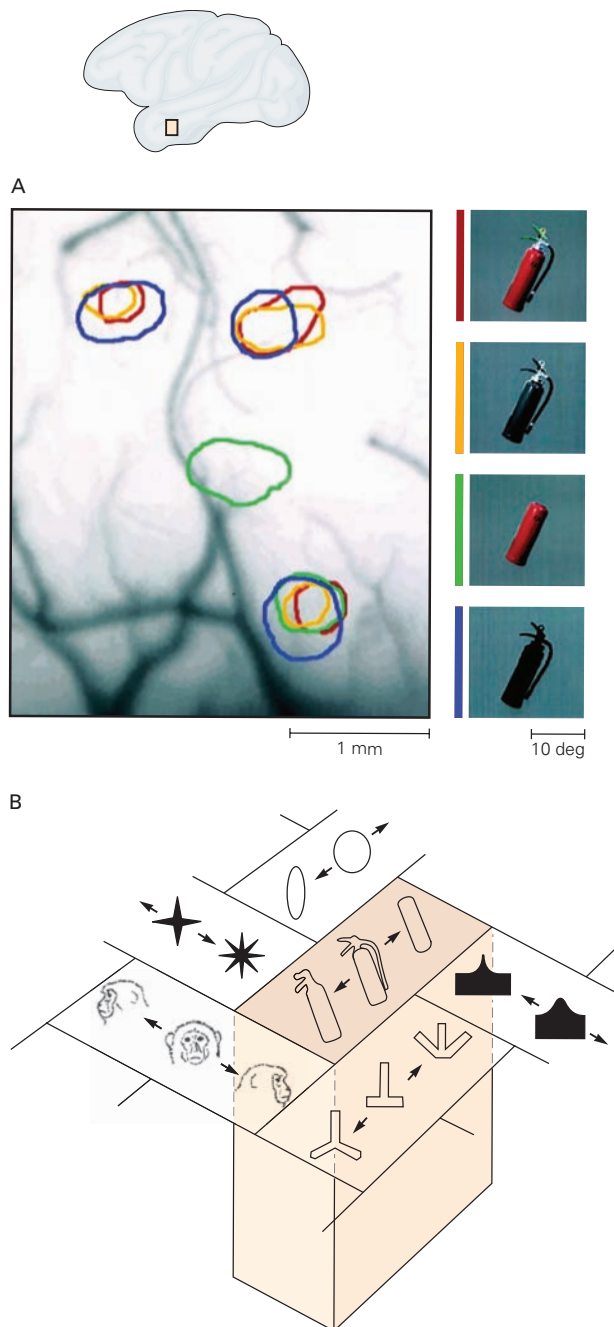
The inferior temporal cortex also provides input—directly and indirectly via the perirhinal cortex—to the amygdala, which is believed to apply emotional

valence to sensory stimuli and to engage the cognitive and visceral components of emotion (see Chapter 48). Finally, the inferior temporal cortex is a major source of input to multimodal sensory areas of cortex such as the superior temporal polysensory area.

### 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, the 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



**Figure 28-5** Neurons in the anterior portion of the inferior temporal cortex that represent complex visual stimuli are organized into columns. (Reproduced, with permission, from Tanaka 2003.)

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

B. In this schematic depiction of the columnar structure of the inferior temporal cortex the vertical axis represents cortical depth. According to this model each column includes neurons that represent a distinct complex pattern. Columns of neurons that represent variations of a pattern, such as the different faces or the different fire extinguishers, constitute a hypercolumn.

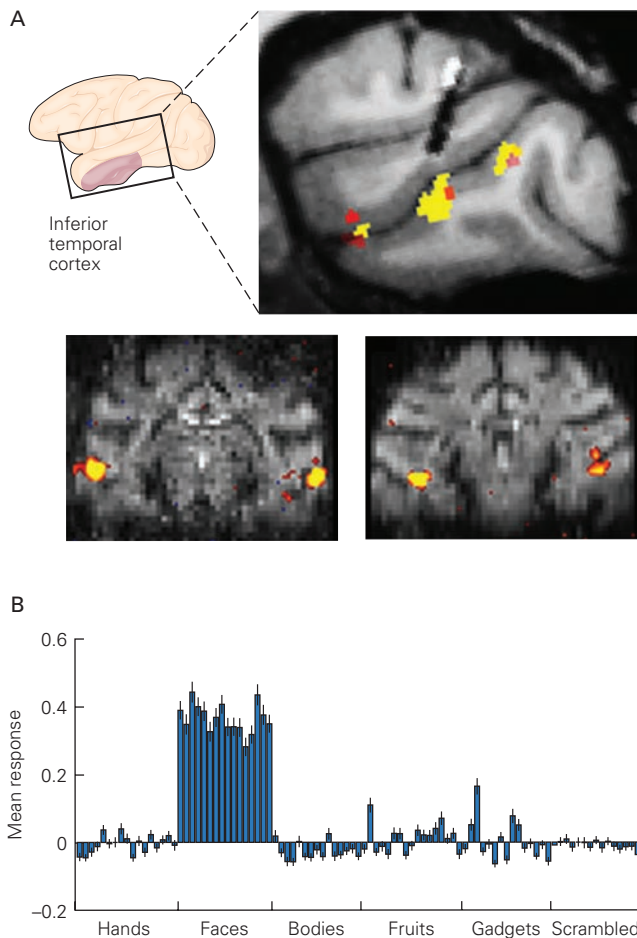
termed *perceptual constancy*. Perceptual constancy has many forms ranging from invariance across simple transformations of an object, such as size, position, and rotation, to the sameness of objects within a common category: All zebras look alike.

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 and best-documented features of the response properties of individual inferior temporal neurons is the invariance of their pattern of selective responses to changes in stimulus size (Figure 28-7A).

Another relatively simple 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 responses of many inferior temporal neurons does not vary when the position of an object within their large receptive fields is changed (Figure 28-7B). *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 28-7C), color, or texture.

*Viewpoint invariance* refers to the perceptual constancy of three-dimensional objects observed from different angles. Despite the limitless range of retinal images that might be cast by a familiar object, an observer can readily recognize the 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.

Investigators have looked for neurons whose response properties would account for viewpoint invariance but have found surprisingly little evidence. On the contrary, most neurons are tuned for specific viewing angles of a three-dimensional object. Although this tuning is often broad, thus reflecting partial viewpoint invariance, it appears that individual neurons do not generalize across inputs sufficiently to account for viewpoint invariance. Another possibility is that viewpoint invariance is the product of population coding by an ensemble of neurons each tuned to a different viewing angle. Finally, viewpoint invariance may be



**Figure 28-6** The inferior temporal cortex contains dense clusters of face-selective neurons. (Reproduced, with permission, from Tsao et al. 2006.)

**A.** Functional magnetic resonance imaging (fMRI) identifies three regions of the inferior temporal cortex that are selectively activated by faces. The upper image, a sagittal section, shows the three active zones along the lower bank of the superior temporal sulcus in one monkey. The two lower images are coronal sections through the face-representation areas in two monkeys.

**B.** Neurophysiological recordings reveal a preponderance of face-selective neurons in the middle face area identified by fMRI. The histogram plots the mean normalized response rate (minus baseline) of 182 neurons in the middle face area of one monkey. The monkey was shown 96 visual stimuli in six categories. Only faces elicited consistently vigorous responses.

achieved at a higher stage of cortical processing, such as the prefrontal cortex, through convergent inputs from neurons selective for specific viewpoints.

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 the same, a confusion reflecting a false-positive identification by the system for viewpoint invariance. Carl Olson and colleagues examined the responses of neurons in the inferior temporal cortex to stimuli that were mirror images. Consistent with the perceptual confusion, many inferior temporal neurons responded similarly to both images. This result reinforces the conclusion that activity in the inferior temporal cortex reflects perceptual invariance, albeit incorrectly in this case, rather than the actual sensory information.

### Categorical Perception of Objects Simplifies Behavior

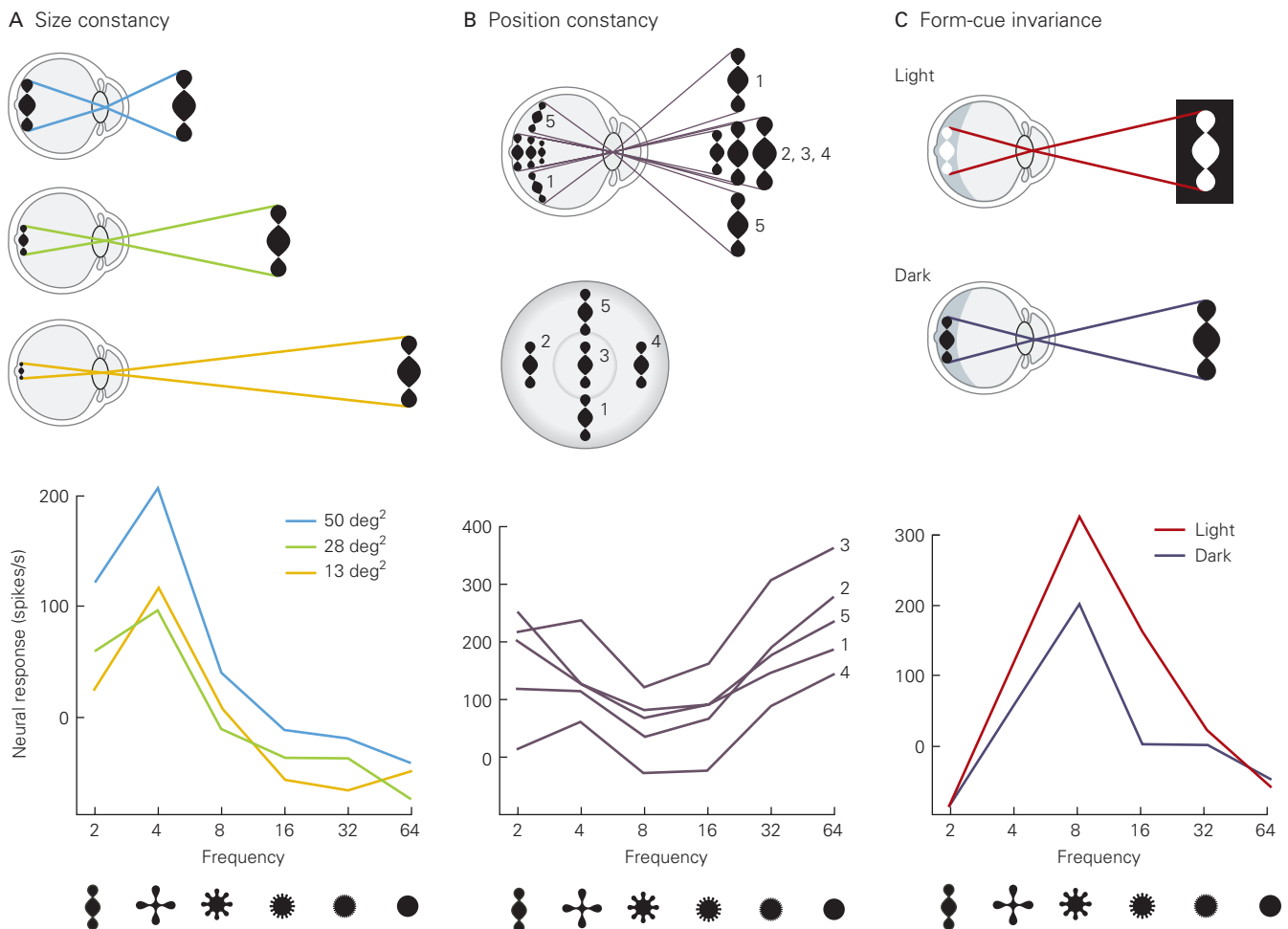
All forms of perceptual constancy are the product of the visual system's attempts to generalize across different retinal images generated by a single object. A still more general type of constancy is the perception of individual objects as belonging to the same semantic category. The apples in a basket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

Categorical perception is classically defined by the ability to distinguish objects of different categories even when objects of the same category cannot be distinguished. 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 reflects sensitivity to invariant visual attributes.

Is there a population of neurons that respond uniformly to objects within a category and differentially to objects of different categories? To test this 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 lateral prefrontal cortex, a region that receives direct input from the inferior





**Figure 28-7** Perpetual 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 particular numbers or frequencies of lobes but invariant with regard 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.

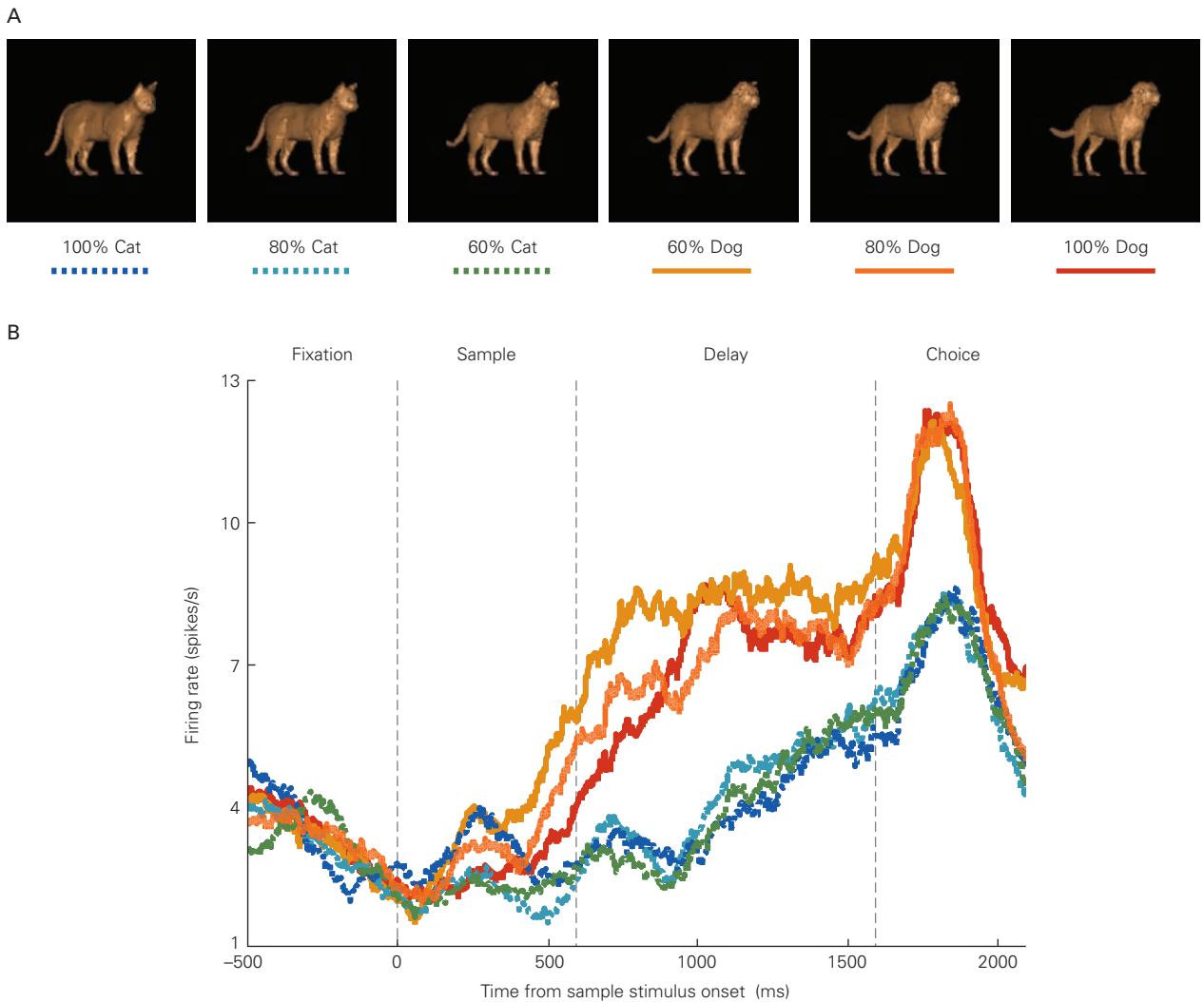
**B. Position constancy.** An object is perceived to be the same despite changes in position in the retinal image. Almost all inferior temporal neurons respond similarly to the same stimulus in different positions in the visual field, as illustrated here by the record of a single neuron.

**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 viewing conditions illustrated, as shown in the record of the individual neuron.

temporal cortex. Not only did these neurons exhibit the predicted category-specific responses—responding well to cat but not dog, or *vice versa*—but the neuronal category boundary also corresponded to the behaviorally learned category boundary (Figure 28-8).

The fact that category-specific agnosias sometimes follow damage to the temporal lobe suggests there are neurons in the inferior temporal cortex that have

category-specific responses similar to those of neurons in the prefrontal cortex. Face-recognition cells appear to meet this criterion, for their responses to a range of faces are often similar. Face-recognition cells may constitute a special case, however, for learned category-specific responses of the sort tested by Miller in the prefrontal cortex are rarely seen in the inferior temporal cortex. For most stimulus conditions category-specific



**Figure 28–8 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 identify an image as cat or dog if it had 50% or more features of that animal.

**B.** Peristimulus histograms illustrate the responses of a prefrontal cortex neuron to the images shown in part A. The

neuron was selectively responsive to images of dogs. Despite the different retinal images, the responses to images within each category (dog or cat) are similar. By contrast, the responses to images in different categories (cat *versus* dog) differ significantly. Category-specific responses are common among visual neurons of the prefrontal cortex.

representations may be generated 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 in the visual system. 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 (see Chapter 27). The other occurs in connection with the storage of explicit learning, the learning of facts or events that can be recalled consciously (see Chapter 67).

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

The ability to resolve differences between complex visual stimuli is highly modifiable by experience. For example, individuals who attend to and examine fine differences between different models of automobiles or eyeglasses become far better at discriminating and recognizing such differences.

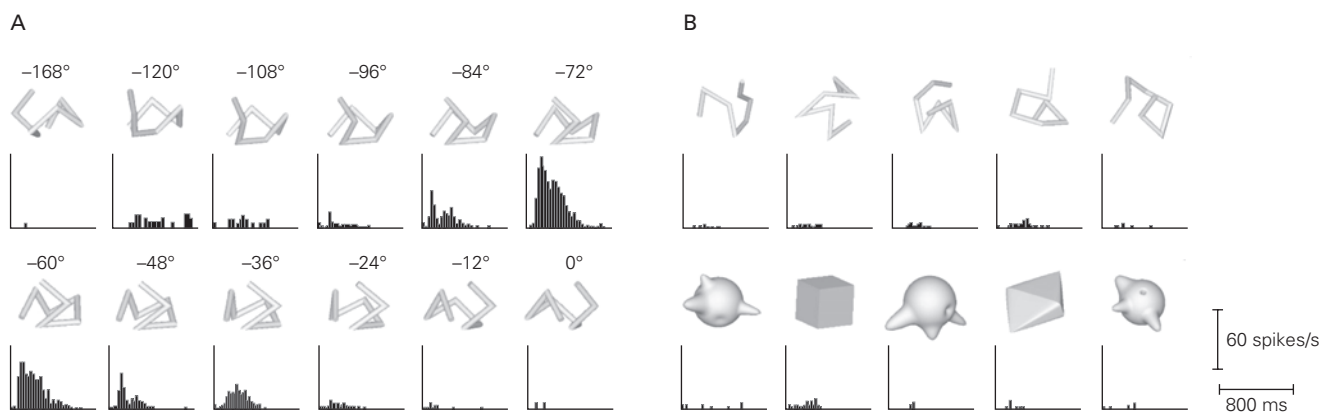
In the inferior temporal cortex neuronal selectivity for complex objects can undergo change that parallels changes in the ability to distinguish objects. For example, in one study monkeys were trained to identify novel three-dimensional objects, such as randomly bent wire forms, from two-dimensional views of the objects. Extensive training led to pronounced improvements in the ability to recognize the objects from two-dimensional views. Extracellular recordings from the inferior temporal cortex after training revealed a population of neurons that exhibited marked selectivity for the views seen earlier but not for other two-dimensional views of the same objects (Figure 28–9).

Other studies with monkeys have shown that familiarity with novel faces alters the tuning of many 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 can result from either active discrimination or passive viewing and are often manifested as a sharpening of stimulus 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.

### Explicit Visual Learning Depends on Linkage of the Visual System and Declarative Memory Formation

Progress has been made in understanding the neurobiology of interaction between vision and memory, specifically in relation to two issues. First, how is visual sensory information maintained in short-term working memory? Working memory has a limited capacity, acting like a buffer in a computer operating system, and is susceptible to interference as when trying to remember the face of a person you have just met. Second, how are long-term visual memories and the associations between them stored and recalled?

Visual neurons in both the inferior temporal and prefrontal cortices continue firing during the delay in a visual delayed-response task (Box 28–1).



**Figure 28–9** Familiarity with particular complex objects leads inferior temporal neurons to respond selectively for those objects. (Reproduced, with permission, from Logothetis and Pauls 1995.)

**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 degrees in each successive view. Once performance was stable at a high level, extracellular recordings

were made from neurons in the inferior temporal cortex while each view was presented. The responses of a typical neuron to each view are plotted in the form of peristimulus histograms. This neuron responded selectively to views that represented a small range of rotation of the object.

**B.** The same neuron was tested with two sets of “distractor” stimuli that were unfamiliar to the monkey. It failed to respond to any of these stimuli.

## Box 28–1 Investigating Interactions Between Vision and 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*. An animal 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 sample (Figure 28–10A).

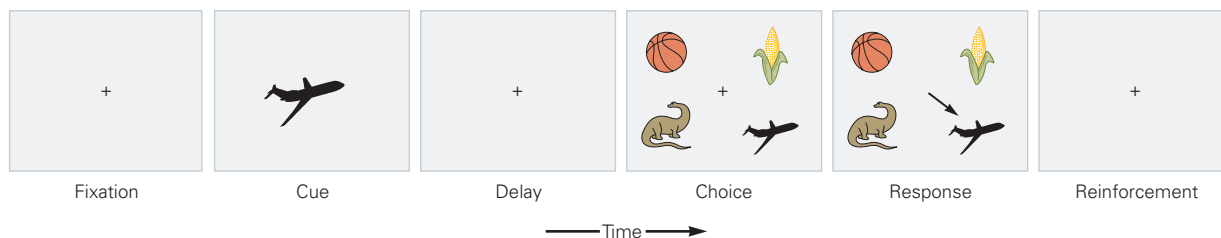
For example, the subject is shown a photograph of a tractor and then, after a brief delay, is shown several photographs of tractors, only one of which is identical to the sample tractor previously viewed. The task is to identify the tractor that matches the sample.

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 sample stimulus;

(2) the short-term or working-memory component, the response that occurs during the delay between the sample 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 sample stimulus.

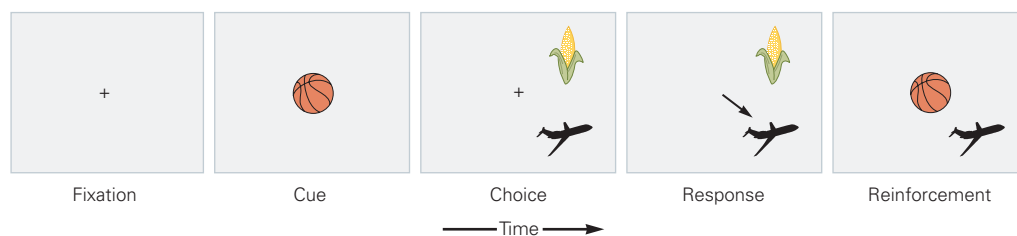
A second behavioral paradigm, known as the *visual paired-association task*, has been used in conjunction with electrophysiology to explore the cellular mechanisms underlying the long-term storage and recall of associations between visual stimuli (Figure 28–10B).

This task differs from the delayed match-to-sample task in that the match and sample are two different stimuli. 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.



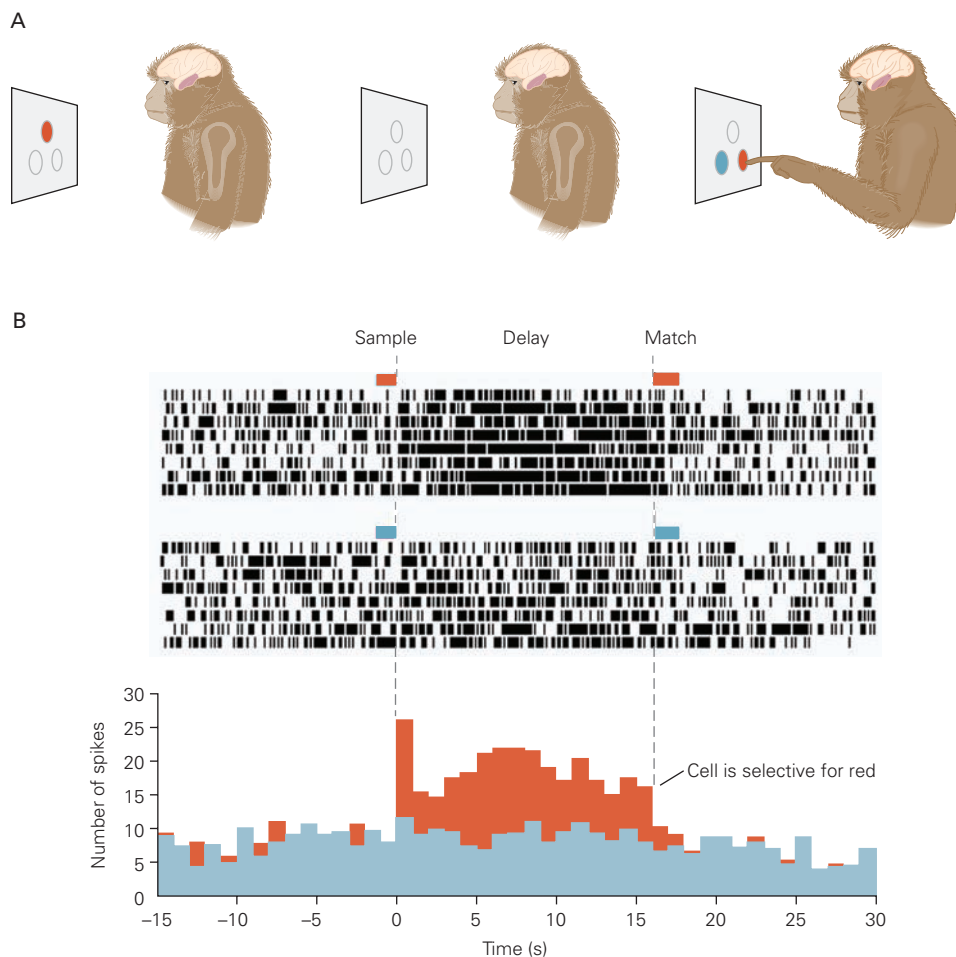
**Figure 28–10A** Delayed match-to-sample task. In this paradigm a trial begins with the appearance of a fixation spot that directs the subject's attention and gaze to the center of the computer screen. A sample image then appears briefly, typically for 500 ms, followed by a delay in which the display is blank. The delay can be varied to fit the experimental goals. Following the delay several test images, including the sample, are displayed. The monkey must choose the sample, typically either by pressing

a button or by a saccade to the stimulus. If the animal chooses the sample, it receives a small juice reward. 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.



**Figure 28–10B** The paired-association task. This paradigm resembles the match-to-sample paradigm except that the sample and test stimuli are not the same. The subject must discover the correct association by trial-and-error learning. The task thus serves to build an association

between 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.



**Figure 28–11** Neural activity representing an object is sustained while the object is held in working memory. (Reproduced, with permission, from Fuster and Jervey 1982.)

**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 inferior temporal cortex is shown.

**B.** Peristimulus 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 from trials in which it was green. 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.

This delay-period activity is thought to maintain information in short-term working memory. Delay-period activity in the inferior temporal and in prefrontal cortices differ in a number of ways. For one, the activity in the inferior temporal cortex is associated with the short-term storage of visual patterns and color information, whereas the activity in the prefrontal cortex encodes not only visual spatial information but also information about other sensory modalities. Delay-period activity in

the inferior temporal cortex also appears to be closely attuned to visual experience, for it encodes the sample image and can be eliminated by the appearance of another image (Figure 28–11).

In the prefrontal cortex, by contrast, delay-period activity is more likely to depend on task requirements and is not terminated by later 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 and then tested. Each behavioral test began with presentation of a single sample object. After a brief delay a monkey was shown a test object and asked to indicate whether it was the object paired with the sample during training.

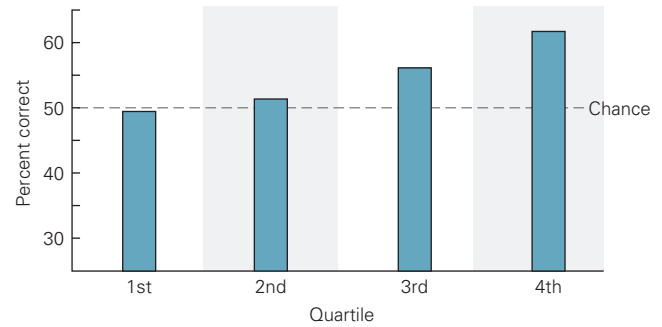
There are two possible ways to solve this task. During the delay the animal could remember the sample object by retaining a sensory code or thinking ahead to the expected object—the one associated with the sample during training—using a “prospective code.” Remarkably, neuronal activity appears to transition 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. A century ago William James, a founder of the American school of experimental psychology, suggested that learning of visual associations might be mediated by enhanced connectivity between the neurons encoding individual stimuli. To test this hypothesis monkeys were trained 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.

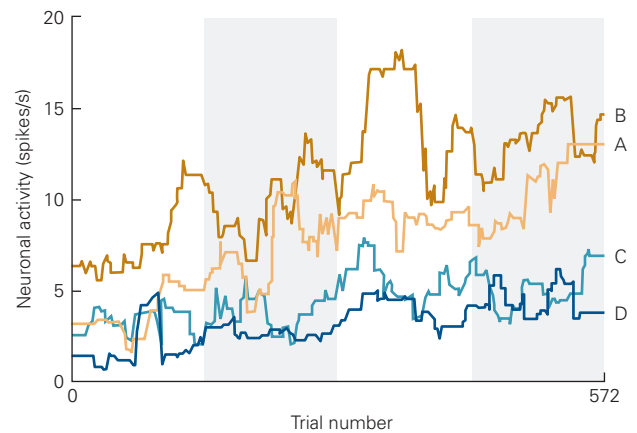
To determine whether this pattern of selectivity was indeed temporally and conditionally tied to learning, Thomas Albright and colleagues recorded from individual inferior temporal neurons while monkeys were learning new visual associations. They found that responses to paired objects became more similar over the course of training (Figure 28–12). Most importantly, the changes in neuronal activity occurred on the same timescale as the behavioral changes and were dependent on successful learning.

The 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 visual associative memories. The results also support the view that learned associations are implemented rapidly by highly specific changes in

#### A Animals learn to associate pairs of stimuli



#### B After training neurons respond similarly to paired stimuli



**Figure 28–12** 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.)

**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 and was required to learn two paired associations. 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). Although 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.

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 visual associative memories and for the functional plasticity of the inferior temporal cortex. The hippocampus and medial temporal lobe may facilitate the reorganization of local neuronal circuitry in the inferior temporal cortex as required to store visual associative memories. The reorganization itself may reflect a form of Hebbian plasticity, initiated by the temporal coincidence of the associated visual stimuli.

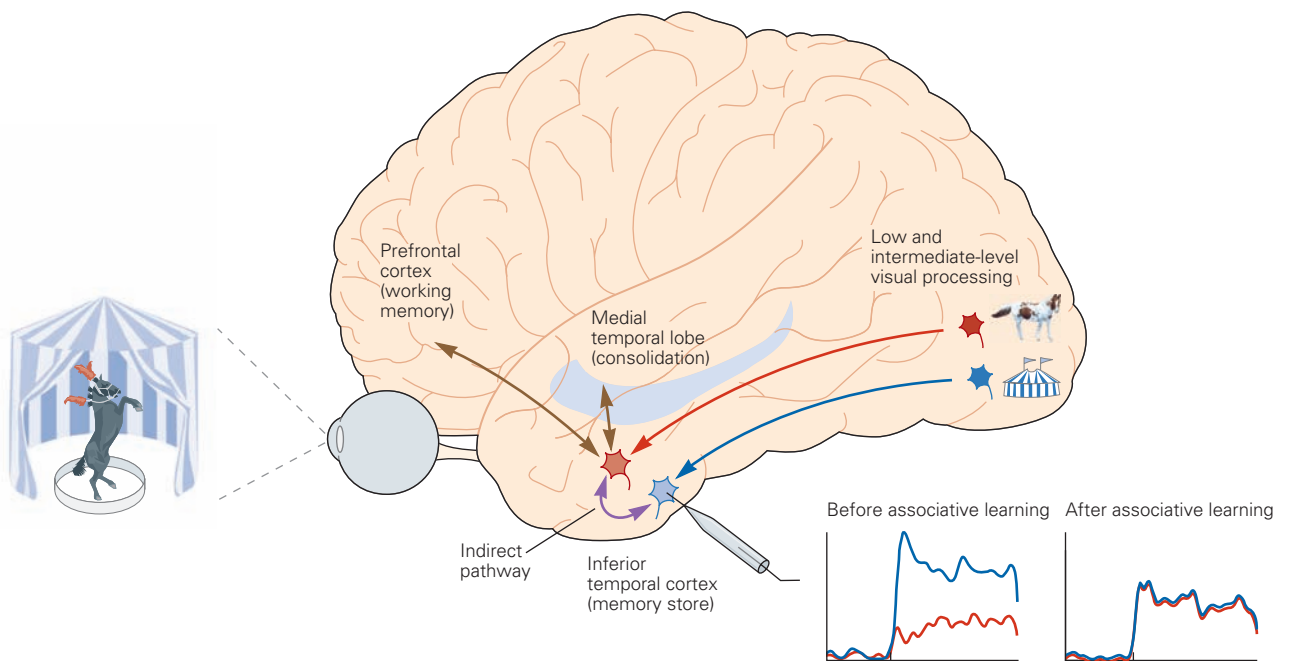
### Associative Recall of Visual Memories Depends on Top-Down Activation of the Cortical Neurons That Process Visual Stimuli

One of the most intriguing features of high-level visual processing is the fact that the sensory experience 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 visual associative 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 stimulus *A* and stimulus *B* after these stimuli have been associated (Figure 28–13). Activation of an *A*-responsive neuron by the associated stimulus *B* can be viewed as the neuronal correlate of top-down recall of stimulus *A*.

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



**Figure 28–13** Circuits for visual association and recall. Bottom-up signals—afferent signals initiated by objects in the observer's visual field—lead to representation of those objects in the inferior temporal cortex. Before associative learning, a neuron (light blue) responds well to the circus tent but not to the horse. Learned associations between objects are consolidated in the inferior temporal cortex by strengthening connections between neurons representing each of the paired objects

(the indirect pathway in the figure). That consolidation is mediated by memory structures of the medial temporal lobe. 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 experience is the product of a combination of direct and indirect inputs to the inferior temporal cortex.

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, the prefrontal cortex is essential for activating these circuits for conscious recall. The afferent signal for one of a pair of images might be received by the inferior temporal cortex and relayed to the prefrontal cortex, whereupon the information would be maintained in working memory. As we have seen, the signaling of many prefrontal neurons during the delay period of a delayed match-to-sample task initially encodes the sample image but changes to encode 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 visual recall.

## An Overall View

The eminent neuropsychologist Hans-Lukas Teuber once wrote that failure of object recognition “would appear in its purest form as a normal percept that has somehow been stripped of its meaning.” Indeed, the assignment of meaning is one of the most important processes in vision and forms the core of the high-level stage of visual processing.

Although meaning is itself difficult to define, it has generally acknowledged connotations. For example, meaning suggests the ability to identify things that are the same. One of the most striking features of object recognition is the ability to identify an object as the same despite an infinite variety of retinal images. This occurs because a neuron in the inferior temporal cortex is activated by the various retinal images of the same object. Similarly, visual neurons in the prefrontal cortex fire in response to objects that are physically different but semantically related.

Meaning may also connote function, utility, or intention. In the case of visual recognition meaning is formed by the observer’s prior sensory experiences and the acquired associations between these experiences. These attributes are fundamental in high-level visual processing and include enhanced perceptual and neuronal selectivity for objects that are commonplace as well as associative links between neuronal representations of objects.

Although much is now known about the neuronal correlates of object recognition, very little is yet known

about the circuits that *cause* these neuronal representations. Even less is known about the cellular and molecular mechanisms by which these circuits are modified by visual experience. Thus future experiments need to address a number of important questions.

How are categorical representations formed? What is the mechanism by which incoming sensory stimuli are compared with stored representations to achieve object recognition? If we accept that associative memories are stored as patterns of connections between neocortical neurons, what then are the specific contributions of the hippocampus and neocortical structures of the medial temporal lobe, and by what cellular mechanisms do they exert their influences? And how does reinforcement cement changes in the strength of the interconnections between neurons that are presumed to underlie associative memories?

The confluence of molecular-genetic, cellular, neurophysiological, and behavioral approaches in solving these and other problems promises a bright future for understanding of high-level visual processing.

---

Thomas D. Albright

## Selected Readings

- Freedman DJ, Miller EK. 2008. Neural mechanisms of visual categorization: insights from neurophysiology. *Neurosci Biobehav Rev* 32:311–329.
- Gross CG. 1999. *Brain, Vision, Memory: Tales in the History of Neuroscience*. Cambridge, MA: MIT Press.
- Logothetis NK, Sheinberg DL. 1996. Visual object recognition. *Annu Rev Neurosci* 19:577–621.
- 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.
- Miller EK, Li L, Desimone R. 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254:1377–1379.
- Miyashita Y. 1993. Inferior temporal cortex: where visual perception meets memory. *Annu Rev Neurosci* 16:245–263.
- Schlack A, Albright TD. 2007. Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* 53:881–890.
- Squire LR, Zola-Morgan S. 1991. The medial temporal lobe memory system. *Science* 253:1380–1386.
- 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.

## 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.
- 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.
- 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.
- 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.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2002. Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *J Neurophysiol* 88:929–941.
- Fujita I, Tanaka K, Ito M, Cheng K. 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360:343–346.
- Fuster JM, Jervey JP. 1982. Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *J Neurosci* 2:361–375.
- Gross CG, Bender DB, Rocha-Miranda CE. 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* 166:1303–1306.
- Kosslyn SM. 1994. *Image and Brain*. Cambridge, MA: MIT Press.
- Logothetis NK, Pauls J. 1995. Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex* 5:270–288.
- 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.
- Miyashita Y, Chang HS. 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70.
- Rainer G, Rao SC, Miller EK. 1999. Prospective coding for objects in primate prefrontal cortex. *J Neurosci* 19:5493–5505.
- Rollenhagen JE, Olson CR. 2000. Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science* 287:1506–1508.
- Sakai K, Miyashita Y. 1991. Neural organization for the long-term memory of paired associates. *Nature* 354:152–155.
- 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.
- Suzuki WA, Amaral DG. 2004. Functional neuroanatomy of the medial temporal lobe memory system. *Cortex* 40:220–222.
- 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.
- Teuber HL. 1968. Disorders of memory following penetrating missile wounds of the brain. *Neurology* 18:287–288.
- 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.
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS. 2006. A cortical region consisting entirely of face-selective cells. *Science* 311:670–674.
- 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.