

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.

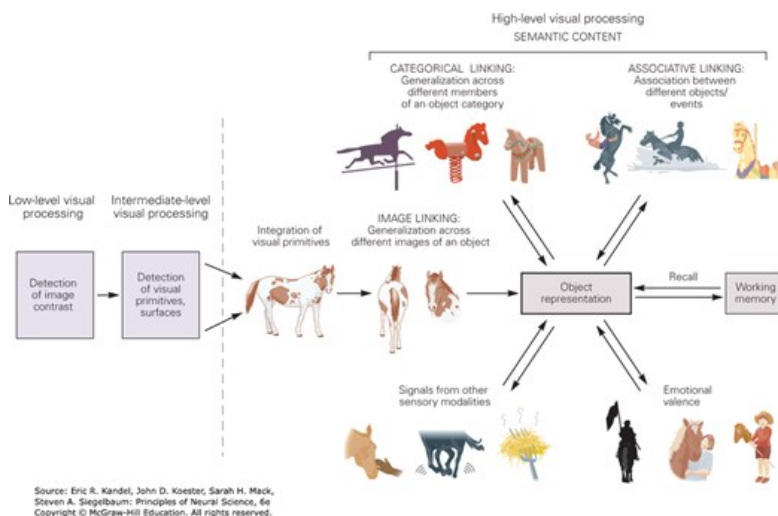
High-Level Visual Processing Is Concerned With Object Recognition

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.



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.

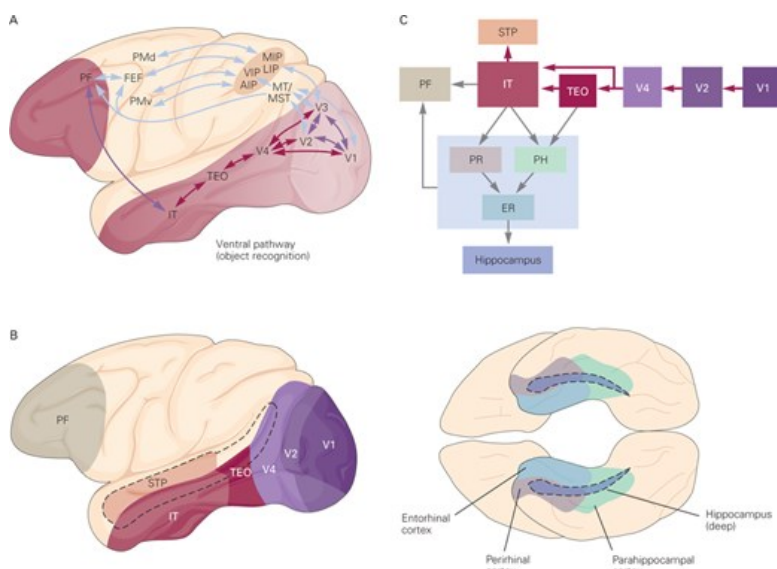
Figure 24–2

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.)



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 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.

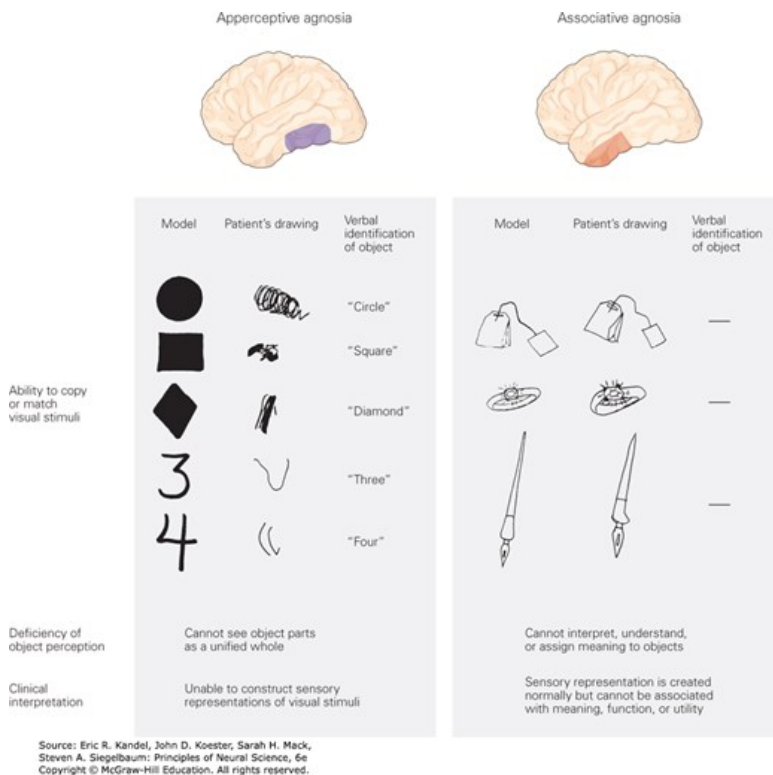
Clinical Evidence Identifies the Inferior Temporal Cortex as Essential for Object Recognition

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.)



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.

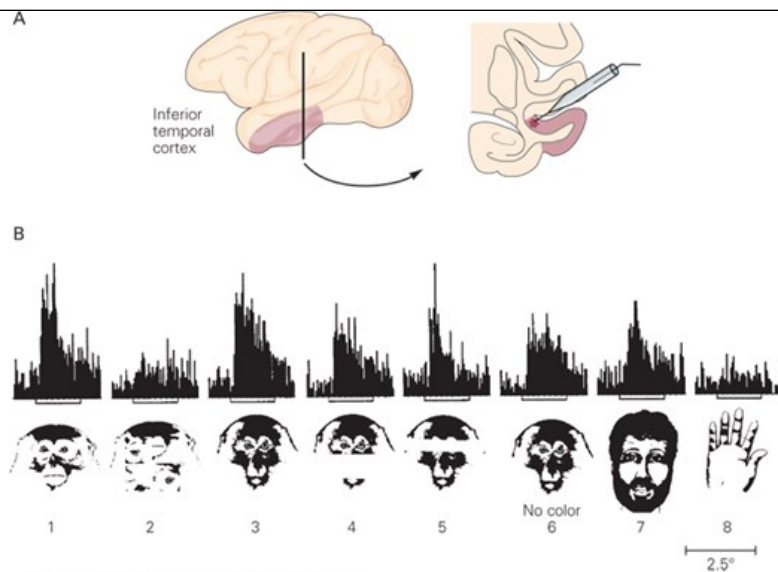
In fact, within the inferior temporal cortex, several small subpopulations of neurons are activated by highly meaningful objects, such as faces and hands (Figure 24–4), as Charles Gross discovered. For cells that respond to the sight of a hand, individual fingers are particularly critical. Among cells 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.

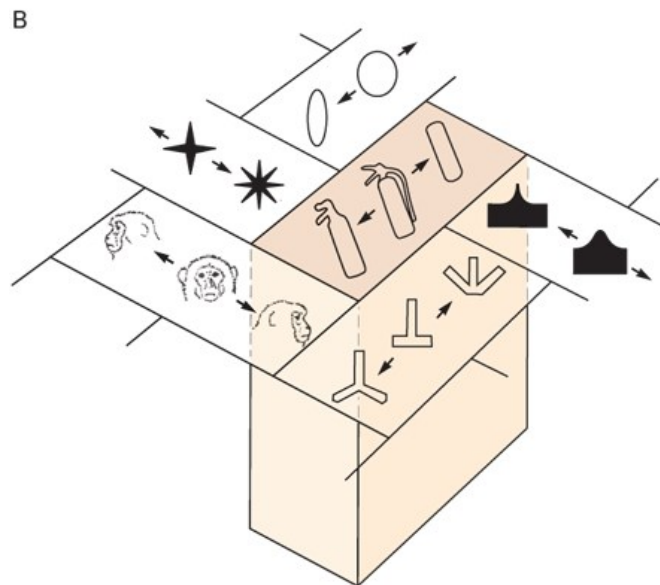
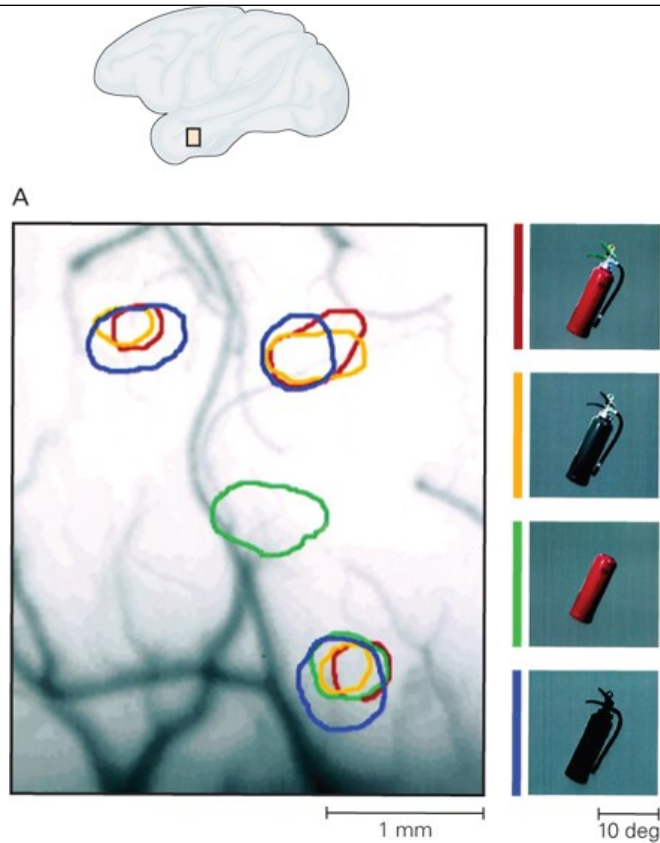
In initial relays in the cortical visual system, neurons that respond to the same stimulus features, such as orientation or direction of motion, but from different parts of the visual field are organized in columns. Cells within the inferior temporal cortex are similarly organized. Columns of neurons 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 Copyright © McGraw-Hill Education. All rights reserved.

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.

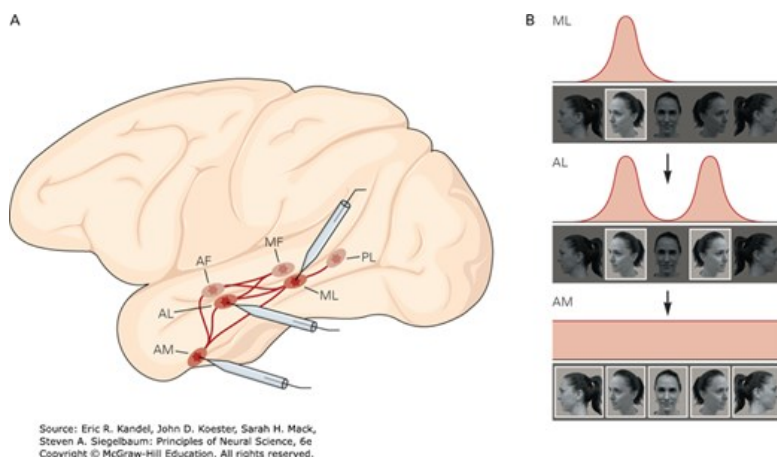
Subsequently, several more face-selective areas were found, primarily in temporal but also in prefrontal cortex. Early studies of these areas provided circumstantial evidence for clustering of face-selective neurons. In later studies, Doris Tsao, Winrich Freiwald, and colleagues directly demonstrated 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**).



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.

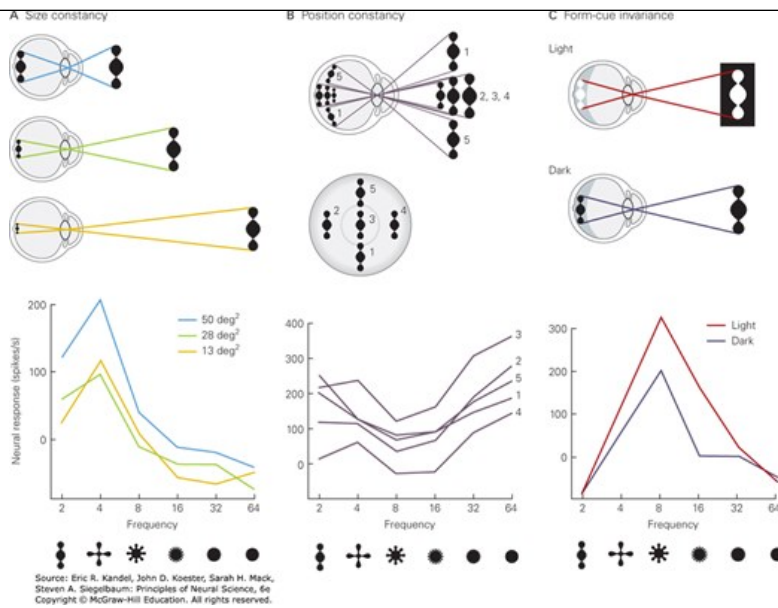
For object recognition to take place, these invariant attributes must be represented independently of other image properties. The visual system does this with proficiency, and its behavioral manifestation is termed *perceptual constancy*. Perceptual constancy has many forms ranging from invariance across 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.



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.

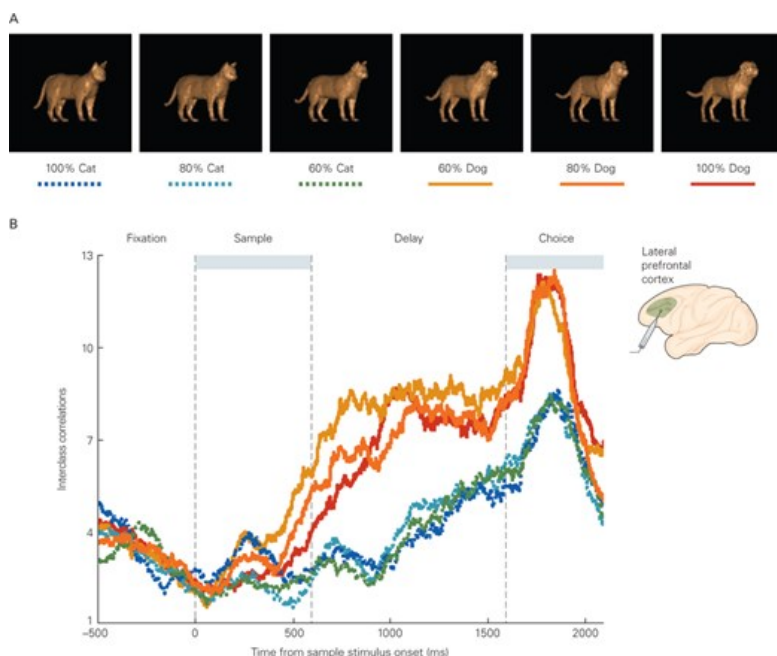
Is there a population of neurons that respond uniformly to objects within a category and differentially to objects of different categories? To test this, David Freedman and Earl Miller and colleagues created a set of images in which features of dogs and cats were merged; the proportions of dog and cat 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 Copyright © McGraw-Hill Education. All rights reserved.

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.

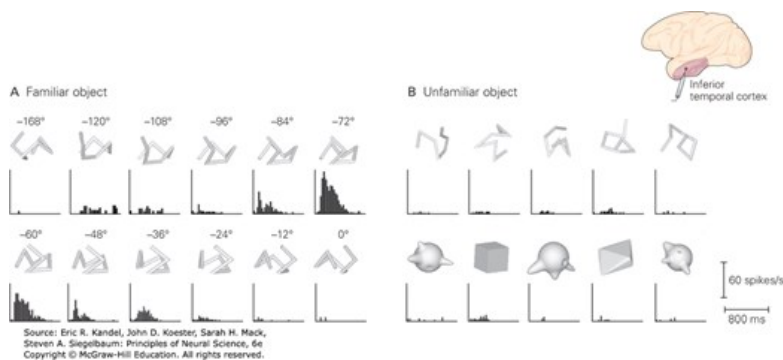
In the inferior temporal cortex, neuronal selectivity for complex objects can undergo change that parallels change in the ability to distinguish objects. For example, in a study by Logothetis and colleagues, monkeys were trained to identify novel three-dimensional objects, such as randomly bent wire 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.



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.

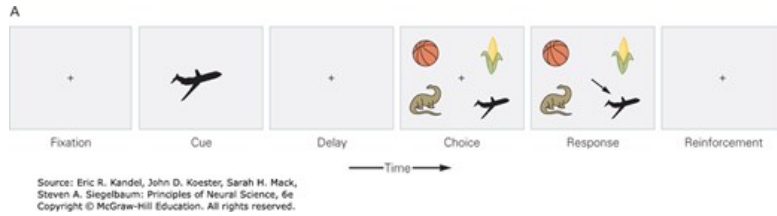


Figure 24–10B

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.

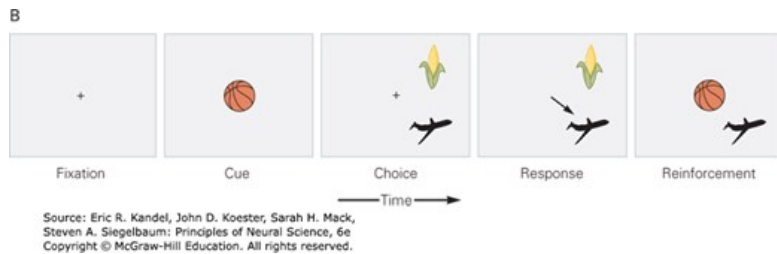
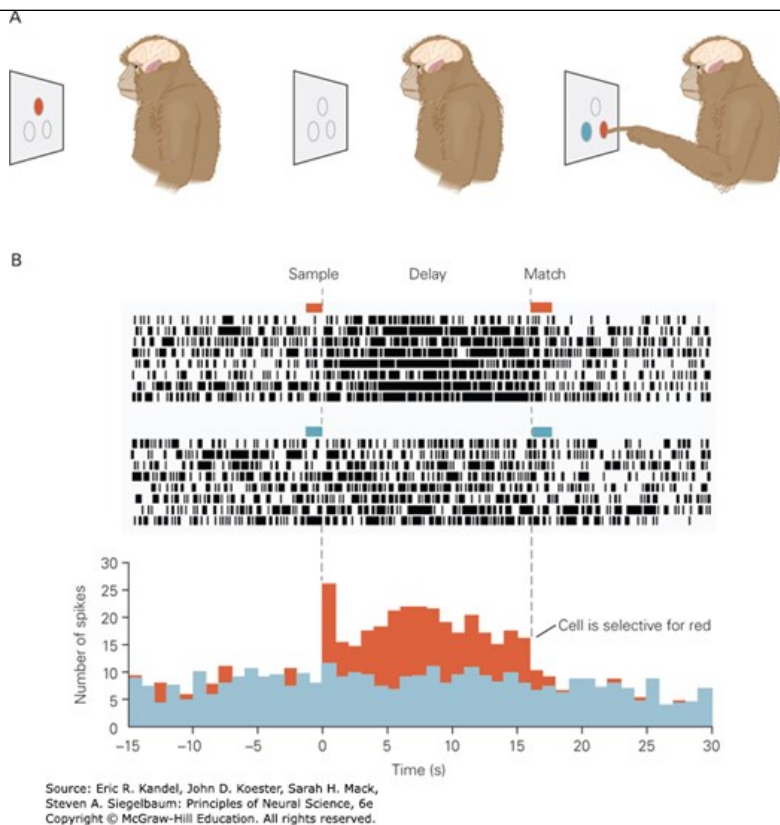


Figure 24–11

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.



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.

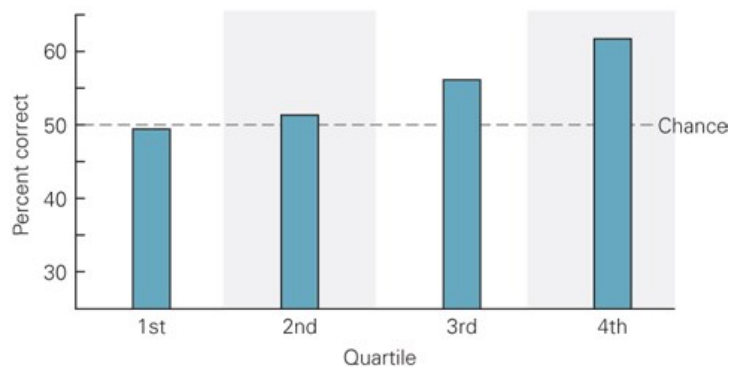
Figure 24–12

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.)

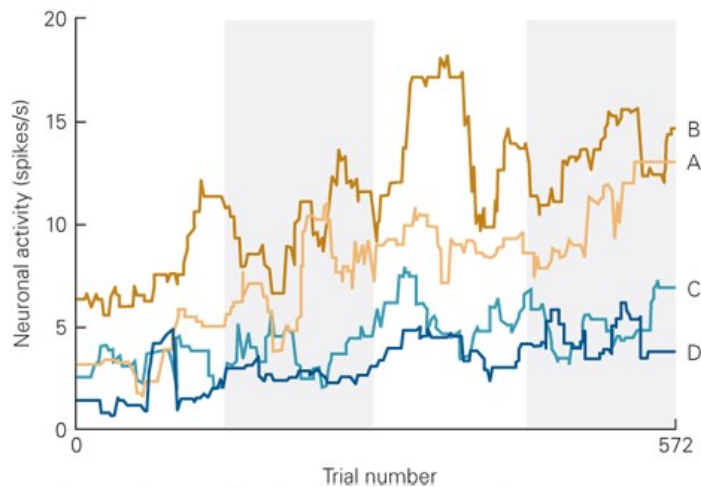
A. Behavioral performance on a paired-association task is plotted for each quartile of a single training session (572 trials). The animal was presented with four novel stimuli (A, B, C, D) and was required to learn two paired associations (A–B, C–D). As expected, performance began at chance (50% correct)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

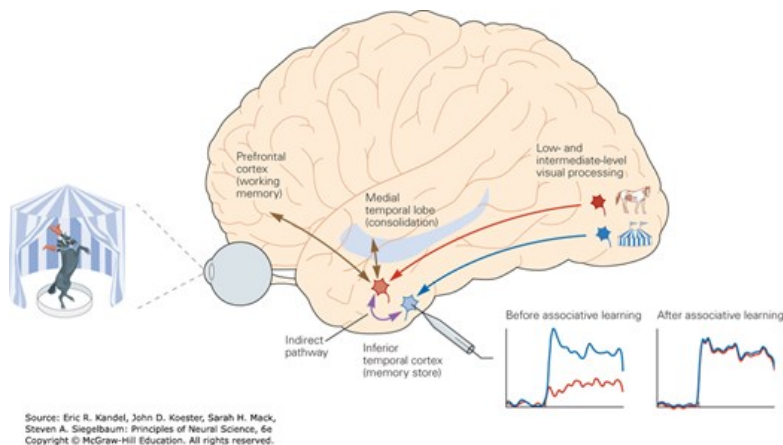
Associative Recall of Visual Memories Depends on Top-Down Activation of the Cortical Neurons That Process Visual Stimuli

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.



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]