VISUAL OBJECT RECOGNITION

Nikos K. Logothetis and David L. Sheinberg

Division of Neuroscience, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

ABSTRACT

Visual object recognition is of fundamental importance to most animals. The diversity of tasks that any biological recognition system must solve suggests that object recognition is not a single, general purpose process. In this review, we consider evidence from the fields of psychology, neuropsychology, and neurophysiology, all of which supports the idea that there are multiple systems for recognition. Data from normal adults, infants, animals, and brain-damaged patients reveal a major distinction between the classification of objects at a basic category level and the identification of individual objects from a homogeneous object class. An additional distinction between object representations used for visual perception and those used for visually guided movements provides further support for a multiplicity of visual recognition systems. Recent evidence from psychophysical and neurophysiological studies indicates that one system may represent objects by combinations of multiple views, or aspects, and another may represent objects by structural primitives and their spatial interrelationships.

INTRODUCTION

An essential behavior of animals is the visual recognition of objects that are important for their survival. Human activity, for instance, relies heavily on the classification or identification of a large variety of visual objects. We rapidly and effortlessly recognize these objects even when they are encountered in unusual orientations, under different illumination conditions, or partially occluded by other objects in a visually complicated environment.

How is this performance accomplished by the brain? What kind of information does the visual system derive from the retinal image to construct

577

descriptions of sets of object features that capture the invariant properties of objects? How are such descriptions stored, and how are they activated by the viewed object? Are object representations general, or are they specific to an action or to a cognitive process, such as learning, planning, or reasoning?

These questions have historically been addressed by scientists in a variety of disciplines, including cognitive psychology (Pinker 1985, Biederman 1987, Banks & Krajicek 1991), neurobiology (Gross 1973, Gross et al 1993, Miyashita 1993, Rolls 1994), neuropsychology (Humphreys & Riddoch 1987a, 1987b; Damasio et al 1990; Farah 1990; Grüsser & Landis 1991), and computation and engineering (Marr 1982, Ullman 1989, Koenderink 1990, Aloimonos 1993). In this chapter we review selected work from each of the aforementioned fields that, in combination, shed increasing light on the internal workings of this system.

Our aim is to provide evidence that a multipurpose general recognition system does not actually exist. Instead, in the process of biological recognition, multiple representations of an object are formed, each specific to the transformations required by either perception or action. The reviewed literature suggests that the recognition of prototypical members of an object category, the encoding of dynamic and plastic transformations of objects or object parts, the identification of individual members of a homogeneous object class, and the planning of movements habitually made when interacting with familiar objects rely on different representations that are formed in different neural sites or by different interconnectivity patterns.

We start with an overview of the basic capacities and limitations of the primate recognition system. After a brief description of some general principles of object categorization, we discuss the performance of human and nonhuman primates in different recognition tasks and relate this performance to relevant theoretical models. We then survey a number of human neuropsychological and animal lesion studies showing that damage in different regions of the brain often results in a selective disruption of different recognition processes. In the final section, we discuss findings from psychophysical and electrophysiological experiments in the monkey that examine the role of single neurons of cortical areas thought to be essential in the formation of object representations.

CATEGORIZATION

The world has an infinite number of stimuli that can be discriminated from one another, to an arbitrary degree of detail. Which discriminations are essential for a given recognition system, and what is the basis for organizing information into equivalence classes? In this section, we examine how humans classify and recognize objects, and then we provide evidence that the same principles most likely underlie categorization performed by other biological

recognition systems. Specifically, we show that the generalizations about the world that allow us to categorize objects are not the product of the development of language, but are instead of perceptual origin. Perceptual categorizations, in turn, reflect the redundant, correlational structure of the environment and occur most often at the level at which individual members of categories are most similar to each other and maximally different from members of other categories.

Object Classes and Taxonomies

In human cultures, object categories are usually designated by words that capture the common functional properties of the category's members. Brown (1958) considered the question of why everyday "things" (e.g. pineapples and dimes) are referred to by the same name by most members of a society. He concluded that "[t]he most common name for each of these categorizes them as they need to be categorized for the community's nonlinguistic purposes. The most common name is at the level of usual utility" (Brown 1958, p. 16). The idea that categories may actually reflect more than just linguistic constructs was examined systematically by Rosch and her colleagues (1976a). They showed that human conceptual categories have a perceptual basis and are determined by the high correlational structure of the real world, in which certain combinations of attributes are more probable than others. For instance, attributes such as "feathers" and "wings" co-occur often, while combinations such as "feathers" and "wheels" generally do not. Bundles of such co-occurring attributes form the basis of a natural classification for objects.

Rosch et al (1976a) argued that the world contains "intrinsically separate things," and that there exists a taxonomy for objects within which categories are related to each other by class inclusion. Such categories form natural groupings of stimuli with different perceivable characteristics, or "cues." Cues with high frequency within a given category and low frequency in all other categories are valid category predictors. For example, the cue "long neck" has extremely high validity for the category giraffe because it reliably predicts the presence of a giraffe. The cue "hoofed foot," on the other hand, has low validity for the same category because all ungulate mammals, in addition to giraffes, have hooves. The notion of cue validity extends to categories and is conceived of as the sum of the cue validities of each of the category's features (Reed 1972, Rosch et al 1976a).

General categories, such as mammals, are highly inclusive but have low cue validity, since few perceivable characteristics are shared among their members. Categories such as doberman, on the other hand, are very specific but also have low cue validity because many properties are shared with other categories at the same level of abstraction (e.g. setter, pointer, golden retriever). The most

inclusive category within which attributes are common to most category members is what Rosch and colleagues called the basic-level category, e.g. dog, and it is the category that has the maximum cue validity. Classifications more general than the basic level are called superordinate categories, while those that are more specific are called subordinate categories (Rosch et al 1976a).

When human subjects are asked to list as many attributes that apply to certain objects as they can, they report the greatest increase in the number of characteristic features when describing objects at the basic level (Rosch et al 1976a, Tversky & Hemenway 1984). Objects in the same basic-level category are also manipulated using common motor sequences and share considerable shape similarity with each other but not with objects of most other groups. The similarity of basic objects is such that shape-based averaging of two members of the category will often yield a new object that can also be recognized as a category member. In fact, in the case of a highly homogeneous group of objects like faces, photographic averages from two separate 20-person groups, selected according to gender and age, have been shown to yield two "average" faces remarkably similar to each other (Katz 1953).

Empirically, recognition of objects at the basic level often occurs more rapidly and more accurately than the recognition of objects at any other taxonomic level. Exceptions to this rule are atypical exemplars of basic categories that have pronounced shape differences from the prototype, or central tendency, of the class. For example, humans usually identify "penguins" or "racing cars" faster as such than as "birds" or "cars," which are the basic-level classifications for these objects (Jolicoeur et al 1984, Murphy & Brownell 1985). To describe the level at which specific objects are first accessed irrespective of inclusiveness or cue validity, Jolicoeur et al (1984) coined the term "entry point" of recognition.

Interestingly, the entry point of individual objects, which usually coincides with the basic level of classification, can shift to the subordinate level when perceivers become especially sensitive to subtle differences between objects of the same class. For example, Rosch et al (1976a) noted major differences in the descriptions of object attributes between experts in a field and unspecialized subjects. An airplane mechanic, for instance, when asked to list airplane attributes, spontaneously reported a large number of attributes of airplanes that are potentially available to the casual observer but that are usually ignored. In a systematic study of this phenomenon, Tanaka & Taylor (1991) showed that for experts in a field (such as bird watching), subordinate categories become as differentiated as basic-level categories. This entry point change is also evident in naming latencies, which become as short as those of the basic classifications as expertise increases.

In summary, humans systematically categorize objects in the world based on natural groupings of attributes. Do, however, such categorization principles

also apply in situations where the observer has no prior conceptual information about the objects to be classified? Most importantly, do they apply for the nonverbal observer? If natural categories do develop independently of preexisting conceptual or linguistic labels, then the same principles of categorization may underlie the recognition skills of other animals, in which the neural representations of objects can be studied directly using neurophysiological techniques. At least three lines of evidence, discussed briefly below, suggest that categorization may indeed rely on principles applying to any recognition system.

Perceptual Categorization

LEARNING NOVEL STIMULI   Evidence suggesting some universal principles in the formation of categories comes from recognition experiments with visually novel objects that are unrelated to any previously experienced verbal codes or abstract concepts. Such experiments show that in the process of learning basic objects, humans can detect consistent features of minimum interindividual variability, e.g. features of high cue validity, that characterize most exemplars of an object class, thereby extracting class invariances. In their seminal work, Posner & Keele (1968) probed the representations stored by humans when learning to classify patterns with individual variance around a common abstract structure. They used dot patterns (Figure 1a) as prototypes, and created individual category instances (exemplars) at specified distances, or deviations, from the original pattern by applying statistical distortion rules (Posner et al 1967). Their subjects were taught to classify distorted patterns constructed from three different prototypes, and they were subsequently tested in a recognition task in which they were exposed to the previously viewed patterns, the prototype pattern, and new, distorted patterns. Interestingly, subjects recognized the prototype pattern almost as quickly and as accurately as they recognized previously memorized patterns, even though they had never been directly exposed to it, suggesting that the prototype is a main constituent in the category's memorial representation (see also Franks & Bransford 1971, Strange et al 1970). Moreover, information about the central tendency of such sets of exemplars was found to be extracted and stored during learning and not during the process of recognition (Posner & Keele 1970, Homa et al 1973, Strange et al 1970).

Categorizations are not based upon a recognition threshold that, once exceeded, definitively endows a stimulus with class membership. In other words, class boundaries are not formed by sharp transition hypersurfaces in a multidimensional feature space. Instead, a familiarity continuum exists, according to which the probability of a correct classification depends on the structural typicality of the stimulus, determined by the closeness to the class prototype
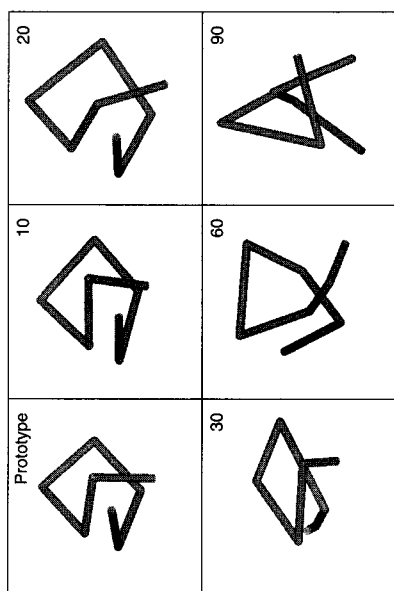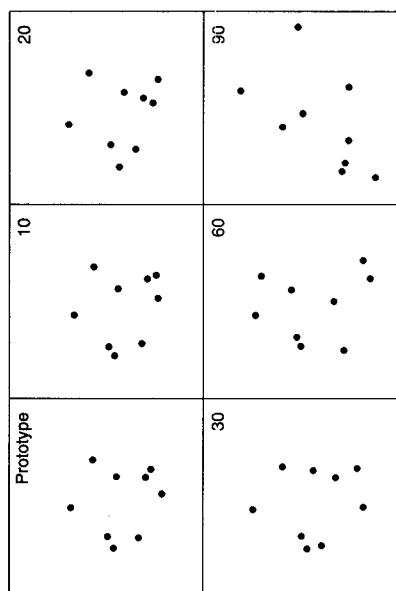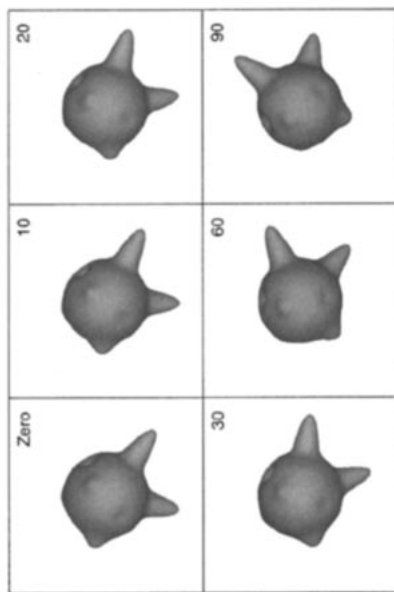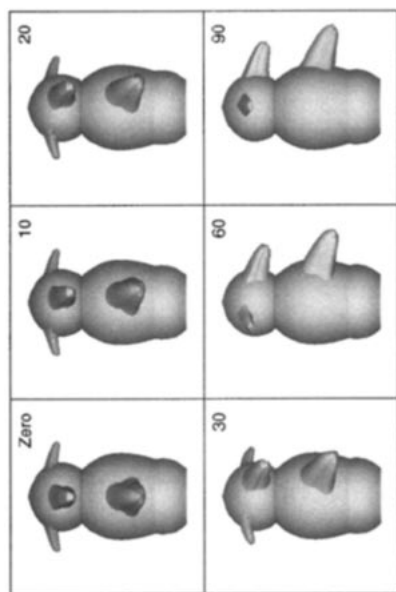
(Rosch et al 1976b). Atypical exemplars greatly differing from the prototype are recognized as individual entities rather than as class members, i.e. they themselves become the entry point of recognition.

What follows demonstrates that the principles emerging from experiments with novel objects appear to underlie the complex and sophisticated processes that infants possess for categorizing and representing their experiences with the environment.

CATEGORIZATION IN INFANTS   Piaget (1969), after carefully observing the development of perception and cognition in children, concluded that these capacities are rooted in prelinguistic constructs that are only later enriched through the use of language. A large number of studies using behavioral paradigms developed for research in preverbal observers—such as the preferential looking paradigm, a technique that capitalizes on infants' innate preference for novel stimuli (Fantz 1964)—have shown that infants as young as three or four months old can form categorical representations based on visual and auditory stimuli (for review, see Quinn & Eimas 1986).

Infants, for instance, can form categorical representations for animals from different basic-level categories that are sufficiently distinct (Quinn et al 1993, Eimas & Quinn 1994). Specifically, they categorized various horses as different from cats, zebras, and giraffes, and perceived cats as different from tigers and horses, but not female lions. The inclusion of female lions in the category of cats was found to disappear, however, by 6 to 7 months of age (Eimas & Quinn 1994), as more subordinate-level recognition skills develop with increasing demands for finer-level discriminations. On the other hand, it is remarkable that some types of subordinate recognition, such as the identification of familiar faces, appear to begin extremely early in life, as neonates can visually discriminate between their mother's face and the face of a stranger (Bushnell et al 1989). In contrast, superordinate-level classifications of object

←——————————————————————————————————

Figure 1   Example stimuli used in object recognition experiments. (*a*) Random dot patterns formed by distorting a prototype (*upper left*) by increasing levels of dot-position perturbation. Each pattern consists of nine dots placed randomly in a 512 × 512 unit area. The number in the top right corner indicates the average distance each dot was displaced from its position in the prototype. Stimuli of this sort were first used by Posner et al (1967) to investigate how humans form abstract visual categories. [Figure adapted from Knapp & Anderson (1984).] (*b*) Prototype distortions of a three-dimensional (3D) wireframe object, similar to those used by Edelman & Bülthoff (1992). Distortions of the prototype were created by randomly displacing each of the vertices by a percentage of the original segment length. (*c*) Example of a "greeble" object (Gauthier 1995) used to study mechanisms underlying recognition performance of experts. The zero view of the object (*upper left*) is shown rotated around the vertical axis in five different poses. Degrees of rotation are indicated by the number in the upper right-hand corner. (*d*) Six views of a spheroidal object shown rotated in the image plane and similar to those used by Edelman & Bülthoff (1992) and Logothetis et al (1994).

pictures appear later and improve with age, usually in close relation to linguistic developments (Rosch et al 1976a).

Most interestingly, the processes involved in forming perceptual categories appear to be very similar to those involved in categorization in adults. By combining the preferential looking technique with the random dot stimuli introduced by Posner & Keele (1968), Bomba & Siqueland (1983) investigated the processes underlying the ability of infants to abstract a prototype from sets of novel stimuli.

The infants were first familiarized with distortions of one of three dot pattern prototypes: a square, a triangle, or a diamond. Once habituated to the exemplars, they were presented with the prototype of the learned exemplars paired with the prototype of one of the other categories. When the number of exemplars used was small (six stimuli) and the recognition test immediately followed the familiarization period, infants learned the individual examples but failed to extract a class representative prototype. In contrast, infants reliably associated the familiarized exemplars with their prototype when the number of examples was increased (12 stimuli) or when a delay of 3 min was introduced between training and testing (Posner & Keele 1970, Homa & Vosburgh 1976). In other words, infants, like human adults, tend to abstract the central tendency of a category when exposed to a sufficient number of exemplars, but they learn individual entities when presented with only a few exemplars. The reported "delay effects," whereby individual exemplars are remembered if testing immediately follows familiarization but the prototype is extracted when testing occurs after a delay, have also been observed in adult subjects and have been taken as evidence for different memory decay times for specific exemplars and for the category prototype.

Taken together these studies suggest that structure in the environment is more critical for categorization than are the linguistic labels assigned to stimulus classes. Although the transition between perceptual and conceptual categorization is by no means a settled issue, categorical representations of infants below 15 months are predominantly perceptual, and conceptual representations begin to slowly emerge only later in infancy (Eimas & Quinn 1994). In Lorenz's (1971) words, "a young child which is already capable of referring to all dogs as 'bow-wow' and all cats as 'miaow-miaow' has quite definitely not abstracted the zoological identification formula for *Canis familiaris* and *Felis ocreata*" (p. 306).

CATEGORIZATION IN ANIMALS   Finally, the generality of classification rules is perhaps best demonstrated in experiments examining concept formation and categorization performance in animals other than humans. Nonhuman primates clearly are capable of making various categorizations at different abstraction levels, of associating meaning or purpose to different objects, and of possessing

natural concepts (e.g. see Lorenz 1971, Davis 1974). Monkeys, for instance, can learn to perform various types of complex classification tasks in the laboratory (e.g. see Davis 1974). They are obviously capable of making basic-level categorizations, but they can also easily learn to discriminate individual human or monkey faces (Rosenfeld & Van Hoesen 1979, CJ Bruce 1982) and novel artificial object classes, even generalizing learning across basic image transformations (Logothetis et al 1994, 1995).

Category formation is not exclusive to primates. For instance, the ability to discriminate between basic classes has been demonstrated in the goldfish (Bowman & Sutherland 1970) and in many different bird species, which have been shown to recognize even impoverished stimuli (Watanabe et al 1993). Herrnstein & Loveland (1964) showed that pigeons can easily learn to peck a key in the presence of a color slide containing people and to withhold pecks for slides not containing people. Similarly, Herrnstein et al (1976) found that pigeons can reliably classify novel photographs of either trees, water, or a particular woman.

Cerella (1979) used a similar procedure and found that pigeons could also learn to classify novel silhouettes of oak leaf patterns from other leaves, although he had great difficulty training the pigeons to respond selectively to a single, specific oak leaf. Based on this inability, he concluded that "the pigeon is most strongly disposed to code class (i.e. generic) descriptions of visual input. This tendency can be countered only to a limited extent, to achieve stimulus-specific descriptions" (Cerella 1979, p. 75).

However, certain subordinate-level discriminations are commonly performed by many animals. For instance, the development of personal recognition of specific individuals is essential in the closed societies of birds and mammals, as it allows recognition of nonmembers and of the internal rank order prevailing between group members. Ryan (1982) has shown that chickens can discriminate slides of one bird in a variety of poses from slides of other birds, and they can transfer this discrimination to novel sets of slides. A notable fact is that subordinate recognition in some species depends on the relevance of the objects to the animal. For example, although birds can recognize scrambled parts of the Charlie Brown cartoon figure as Charlie Brown, they fail to recognize a pigeon's head as such if it is not presented as a full, unscrambled face (Watanabe & Ito 1991).

Task-Specific Representations

The studies reviewed above strongly suggest that categorizations made by humans and other animals have a general, perceptual basis, reflecting the structure of the world. Classifying objects at the basic level is a fundamental recognition task, and it is likely to be the only task that simple recognition

systems perform. Nonetheless, as mentioned above, in primates and other mammals, subordinate-level recognition is also essential in various social and cognitive tasks. Do categorizations at different abstraction levels rely on the same type of stored representations? And, are similar representations used when the perceived object elicits a visually guided motor action?

REPRESENTING PROTOTYPES VS REPRESENTING EXEMPLARS   An interesting observation, pertinent to this question, was made by Homa and his colleagues, who investigated the abstraction of prototypical or exemplar information in categories having either uniform or mixed (low, medium, and high) distortion levels (Homa & Vosburgh 1976). They found that the breadth of a given category, in terms of mixing groups of patterns with different degrees of distortion, has a profound effect on both prototype abstraction and classification performance. In their experiments, recognition was little affected by retention delays (up to 10 weeks) as long as an adequate number of training exemplars were provided. However, when only a few training exemplars were learned, transfer to new stimuli was better if the original training set was not mixed but rather contained uniformly low distortion exemplars (Homa & Chambliss 1975, Homa & Vosburgh 1976). Importantly, training with a few, high distortion exemplars resulted in a form of generalization in which subjects appeared to store information about the individual exemplars and delay the extraction of any prototype information until the onset of the recognition testing. Furthermore, by systematically varying the similarity of old and new exemplars, Homa et al (1981) found that although in general the accuracy of classification of new instances depends on their similarity to old exemplars, the effect of old-new similarity is much greater for small categories (5 items) than for large ones (20 items), implying that for large categories, individual members are not specifically encoded.

At least two models can account for the effects of category breadth on categorization performance. One hypothesizes that all forms of categorization depend on the workings of a single system, based on distributed memory storage (e.g. Anderson et al 1977, Knapp & Anderson 1984, McClelland & Rumelhart 1985). In such a model, when the number of exemplars of the stimulus pattern is small, the new patterns are classified according to their similarity to the learned patterns, but as the size of the training set increases, accuracy depends on the new pattern's similarity to the category's prototype. Alternatively, category breadth–dependent performance can be explained by assuming the existence of two representation systems, the contribution of each of which depends on the level of classification. Support for the latter comes from a recent experiment by Marsolek (1995) that suggests that at least two separate visual form systems may exist in humans: (a) one that is used to classify different instances of an object as belonging to the same

abstract category and that involves the left hemisphere and (*b*) another involving the right hemisphere that appears to preserve visual details of objects in order to distinguish specific exemplars of particular object classes (Marsolek 1995).

ACTION-RELATED REPRESENTATIONS   Recognition of a stimulus is often signaled by the ability of the subject to respond appropriately to that stimulus. In many instances such responses involve visually guided reaching and grasping. In humans and old-world monkeys, prehensile movements almost always require visually acquired shape information. Recent experiments with human subjects show that such action-relevant representations may differ from those used when performing various categorization tasks. For example, when normal human subjects reach for an object, they move their fingers into a certain spatial configuration appropriate for grasping the object. If, however, the perceived size of an object is different from its actual size, as may occur with some form illusions, then a dissociation is observed between the perceived size and the size of the object represented in the systems mediating the grasping behavior (Vishton & Cutting 1995). For instance, when subjects are asked to give verbal estimates of the size of a small horizontal line intercepted by a vertical line of the same length, and then close their eyes and reach for it, their verbal estimate reveals the expected biases from veridicality—in this case overestimating the length of the vertical segment—but their thumb-to-finger distance during grasping shows no such bias, providing reliable estimates for both the vertical and the horizontal lines.

Together, these findings indicate that multiple recognition systems may be employed during the categorization of stimuli at different levels of abstraction and during visuomotor activities. The following sections provide further evidence supporting this point by surveying studies on object constancy in the primate.

RECOGNITION PERFORMANCE IN HUMANS AND MONKEYS

Introspection indicates that the recognition of familiar objects is largely insensitive to changes in their retinal image. Nonetheless, careful examination shows that image transformations, even simple scaling and translation, can sometimes affect recognition. Moreover, invariance to some transformations, such as rotations in depth, appears to depend strongly on familiarity, as well as on the nature of the object and task.

Effects of Size and Position

The effects of scale changes on recognition have been examined in experiments in which subjects classify shapes as being the same or different, disregarding

changes in size. In these tasks, recognition performance, typically assessed by measuring response latencies, varies as a function of the size ratio between the two stimuli, with increasing size discrepancy between the two shapes resulting in elevated reaction times (Jolicoeur 1987, Ellis et al 1989). Studies in which the viewing distance and the size of novel objects were manipulated have found that the perceived, and not the retinal, size of the objects determines the size ratio effects (Milliken & Jolicoeur 1992).

The partial dependency of recognition on object size has led to the belief that shapes are stored at a particular size and that their sensory representation has first to be scaled before recognition occurs (Ullman 1989). As both Jolicoeur (1987) and Biederman & Cooper (1992) have suggested, however, the effects of size obtained in these experiments may reflect the processes of memory-based comparisons rather than the perceptual representation of the objects. To isolate the perceptual effects from those of episodic discrimi- nations, Biederman & Cooper (1992) used a picture-priming task in which objects viewed on one occasion are more quickly and accurately perceived when presented on a second occasion. In such priming experiments, reaction times were found to be independent of whether the primed object was presented at the same or a different size from when originally viewed (Biederman & Cooper 1992). In contrast, explicit memory tasks using the same stimuli showed clear size effects on recognition, suggesting that dif- ferential results obtained from priming and episodic memory experiments might reflect the differential functioning of two representation systems: one underlying the description of an object's shape, and the other its metric attributes, such as its size, orientation, or position (Biederman & Cooper 1992, see also Cooper et al 1992).

Somewhat less pronounced are the effects of stimulus position on recog- nition performance. Response latencies in visual priming tasks are affected very little by stimulus translation (Biederman & Cooper 1991). However, translational disparity between study and test has been shown to reduce recognition accuracy in a successive presentation same-different task (Foster & Kahn 1985) and in a memory task in which subjects were trained to recognize small novel dot stimuli and thin lines presented at only a single retinal location during the learning phase, and then tested with the same stimuli translated to two new positions (Nazir & O'Regan 1990). Although recognition accuracy initially decreased at the new positions, criterion per- formance was restored after only a few presentations. Interestingly, no effects of translation were found when the stimuli were either very simple or very complex patterns, both of which usually contain salient diagnostic features that are themselves translation invariant (O'Regan 1992), a finding that again suggests different recognition strategies for different tasks or object classes.

Effects of Rotation in the Picture Plane

Studies specifically directed at assessing the effects of image-plane rotation on recognition of familiar shapes, such as letters and digits (Corballis et al 1978, Simion et al 1982), or of shapes with pronounced diagnostic features, such as line drawings of natural objects (Eley 1982, Jolicoeur 1985), have found relatively small costs—in terms of error rates or reaction times—associated with the misorientation of the stimuli. Moreover, after practice, even these small costs were found to generally disappear (Shinar & Owen 1973, Jolicoeur 1985, McMullen & Jolicoeur 1992). However, when the familiarity of test objects was more closely controlled by presenting subjects with novel, letter-like shapes possessing no diagnostic features, Tarr & Pinker (1989) found that responses to stimuli rotated away from the training view were slower and less accurate. Continued practice reduced the effects of rotation for the newly familiarized views, but this practice did not transfer to "surprise" views presented later. In contrast, transfer of practice to never-experienced views was found to occur for line drawings of everyday objects, which presumably possess rotation invariant, diagnostic features (Murray et al 1993).

Image plane rotations have been also studied in the monkey by using novel objects in exemplar identification tasks (Logothetis et al 1995). In the early phases of testing, the monkeys exhibited orientation dependency in their recognition performance. However, over time, their ability to generalize across rotations in the picture plane improved, even in the absence of feedback. Initial view-dependent performance often progressed rapidly, over the course of a few test sessions, to view-invariant performance.

Of particular interest are the effects of inversion on object recognition. In humans, inversion strongly affects the processing of distinct classes of objects, most notably faces (Valentine 1988). At present, whether faces, per se, are special or simply represent the most common class of objects that must be identified based on subtle shape differences is controversial (see below). Nevertheless, evidence showing an inversion effect for other overlearned but highly similar stimuli, such as different dog breeds (Diamond & Carey 1986) and artificial stimuli designed to mimic animate objects (Gauthier 1995; also see Figure 1c), indicates that recognition of individual members of homogeneous object classes relies predominately on the processing of configurational information, and very little, if at all, on the discrimination of features. This configuration-based recognition, on the other hand, reveals exactly the same inversion effects observed when recognizing faces. The importance of configurational processing for intra-class recognition has also been shown by asking subjects to identify the individual halves of composite face stimuli. Under these conditions, two unmatched halves interfere with the recognition of the upright but not the inverted composite face stimuli (Young et al 1987).

Interestingly, in monkeys, who commonly encounter faces from many viewpoints (not just the upright), the effects of inversion are not found (CJ Bruce 1982), highlighting the importance of experience in the development of configurational processing. Similarly, young children, who may not yet have developed appropriate sensitivity for holistic stimuli, are also less affected by face inversion than adults (Carey & Diamond 1977) and are likely to rely on individual facial features that can be recognized from any viewpoint.

Effects of Rotations in Depth

One of the most active areas of recognition research in the past decade has concentrated on the effects of rotations in depth on the recognition of 3D objects. Experiments demonstrating viewpoint-invariant recognition have been interpreted as evidence that the visual system employs an object-centered reference frame for representing stimuli. Conversely, experiments demonstrating viewpoint-dependent recognition performance have been taken as evidence for viewer-centered representations. We propose here that the visual system actually uses both types of representation, depending on the classification task and the subject's familiarity with the test objects (see Ullman 1989, Tarr 1995).

DEPTH ROTATION OF COMMON OBJECTS   Similar to the effects of position and size changes, the effects of viewpoint changes on recognition are subtle for common familiar objects. Bartram (1974), using a visual priming paradigm, found that naming objects in pictures was facilitated most by a previous presentation of the identical view of the object, compared to presentations of either a different view of the same object or a view of a different object with the same name. In a sequential same-different matching task using photographs of objects with high- and low-frequency names, Bartram (1976) found little effect of viewpoint change for the pictures with high-frequency names, but significant effects on judgments of rotated low-frequency named objects and for different exemplars with the same name. These results suggest that frequently encountered everyday objects can be accessed equally well from multiple views, while the activation of less common object representations is subject to greater viewpoint dependence, perhaps because such objects are coded by storing specific object views (see below). Furthermore, the priming effects are at least partly visual, because the recognition of objects with the same name as the prime, but that are visually dissimilar, are not equally speeded. Biederman & Gerhardstein (1993) directly studied the effects of rotation in depth on recognition by using a naming task in which subjects identified line drawings of rapidly presented familiar objects, some of which had been seen during a prior priming block. During the test block, the naming latencies for objects that had appeared during the priming block were speeded,

but the magnitude of the priming effect varied only slightly with changes in orientation, indicating again that at least under some circumstances, recognition of everyday objects seems invariant to rotations in depth.

Nevertheless, although recognition of highly familiar objects seems to be viewpoint invariant, it is still "view biased" in the sense that human subjects consistently label one or two views of common objects as subjectively better than all other views (Palmer et al 1981). Naming of objects occurs fastest when the stimulus is shown from such a better view, designated a canonical view, with response times increasing monotonically with increasing angular disparity between a test view and the canonical view.

In a recent experiment, Srinivas (1993) selected 42 common objects that were found to have both usual and unusual views for use in a name priming task. She found that recognition of all objects improved if they had been previously presented but also that the cost of switching viewpoints was greater for tests of unusual views. In other words, seeing the unusual view during study provided the same benefits for recognizing the usual view during test as did a previous encounter with the usual view. Seeing a usual view during study, however, did not facilitate recognition of unusual views as much as did the familiarization with the visually identical unusual view. Thus, while exposure to a usual view of an object may prime only nearby views of the same object, unusual views may activate both similar unusual views and automatically prime the object's canonical view (see also Warrington & Taylor 1973).

The largely viewpoint-invariant recognition of familiar objects can be explained by a number of different theoretical models (Ullman 1989). For example, recognition has been proposed to occur by detecting properties of objects that are invariant to all image transformations, such as the ratio between an object's apparent area and its volume, a compactness measure, or certain parametric shape descriptions, such as Fourier descriptors or object moments.

Alternatives to the invariant-properties approach include theories relying on the decomposition of objects into natural parts. A horse, for instance, can be thought of as a set of components, such as the torso, the legs, the head, and the tail, each of which can be recognized on its own. One can, therefore, assume that the process of recognition of a complex object can be reduced to the recognition of its parts and their relationships (Palmer 1977, Marr & Nishihara 1978). Recursive decomposition may lead to simple volumetric primitives, the combination of which can represent any complex object. Such descriptions obviously do not constitute a true theory of recognition, rather they simply provide one adequate type of representation with which a recognition theory could be constructed. The human body, for instance, can be described as a collection of points of different brightnesses, as a collection of lines and curves, as a group of planes, or as a set of 3D block structures.

Marr & Nishihara (1978) proposed a representational scheme for objects,

using axis-based structural descriptions that can be decomposed into sets of generalized cones. The class includes simple shapes such as a pyramid or a sphere, as well as natural forms such as arms and legs. In a similar vein, Pentland (1986) suggested that most complex natural objects may be described by combinations of superquadric components, such as spheres, wedges, etc, which might be the basic components that the recognition system recovers while analyzing images.

A recent example of a structural description theory is the recognition by components (RBC) theory (Biederman 1987). The underlying assumption of this theory is that an object can be decomposed into volumetric parts called geons. These parts have simple spatial relationships to each other that remain invariant for all object views, and recognition involves the indexing of these parts and the detection of their structural relationships. A solution to the formidable inverse-optics problem for complex objects is reduced, in this approach, to an inverse-optics problem for simple volumes and their two-dimensional (2D) arrangements.

Biederman and colleagues suggest that recognition of most common objects can be accomplished by indexing structural descriptions based on geons, provided that the following three principal conditions are met: (a) Objects are decomposable, (b) they have different part descriptions, and (c) different viewpoints lead to the same configuration of geons. Such conditions are indeed often met when recognizing entry point objects. These conditions are clearly not met, however, when recognizing objects at the subordinate level, that is, when specific exemplars must be identified or when the objects to be recognized cannot be meaningfully decomposed into simpler parts. In these cases, recognition is often view dependent, and object constancy can only be achieved through perceptual learning.

Finally, viewpoint-invariant performance can also be explained by a viewer-centered recognition system that stores a limited number of object views or aspects and is capable of combining information from these views to recognize any view of the object (Seibert & Waxman 1990, Poggio 1990).

DEPTH ROTATION OF NOVEL OBJECTS   The study of recognition of unfamiliar, novel stimuli has provided important insights into the representation of visual objects. Familiar objects can often confound the issues of object constancy, since a recognition system based on 3D descriptions cannot easily be distinguished from a view-based system exposed to a sufficient number of object views.

The first demonstration of strong viewpoint dependence in the recognition of novel objects was that of Rock and his collaborators (Rock et al 1981, Rock & DiVita 1987). These investigators examined the ability of human subjects to recognize 3D, smoothly curved wire objects seen from one viewpoint, when

encountered from a different attitude and thus having a different 2D projection on the retina. Although their stimuli were real objects (made from 2.5-mm wire) and provided the subject with full 3D information, there was a sharp drop in recognition for view disparities larger than approximately 30°. In fact, as subsequent investigations showed, subjects had difficulty even imagining how wire objects would appear when rotated, even when instructed to visualize the object from another viewpoint (Rock et al 1989).

A number of recent experiments have further studied subordinate-level recognition by using computer-rendered novel 3D stimuli, including wire or spheroidal objects (Bülthoff & Edelman 1992, Edelman & Bülthoff 1992), cube-composed stick figures (Tarr 1995), and novel clay shapes (Humphrey & Khan 1992). With these stimuli, recognition was again found to be strongly view dependent, and generalization could only be accomplished by familiarizing the subject with multiple views of target objects.

Such results may present an extreme case in terms of performance. Farah et al (1994) observed that when Rock's wire-forms were interpolated with a smooth clay surface (creating "potato chip" objects), subjects' recognition accuracy was less affected by the same changes in viewpoint tested by Rock. They concluded that object shape and structure plays a significant role in the ability of humans to compensate for variations in viewpoint. One possibility is that as the structure of objects becomes more regular (in terms of properties such as spatial relations and symmetries), the ability to compensate efficiently for changes in viewpoint is enhanced because the resultant image structure is more predictable (Vetter et al 1994, Liu et al 1995). Under these conditions, recognition may be faster than it is for less regular objects, although it is possible that mixed strategies or verification procedures will yield response times that are still dependent on viewpoint, even for familiar objects (Palmer et al 1981) and faces (V Bruce 1982).

The effects of rotation in depth on recognition were recently studied systematically in monkeys by using computer-rendered objects in both basic- and subordinate-level tasks (Logothetis et al 1994). In these experiments, the animals easily learned to generalize recognition to novel views of objects, such as those illustrated in Figures 1b and 1d, when the set of distractors included objects from other basic-level categories. Their ability to identify individual exemplars of either the wire or the spheroidal objects, however, was found to depend strongly on the viewpoint from which the object was encountered during the training phases. The monkeys were unable to recognize objects rotated more than approximately 40° from a familiar view. However, when two views of the target were presented in the training phase, 75 to 120° apart, the animals interpolated between them, often reaching perfect levels of performance for any novel view between the two trained views, as has been shown in human experiments (Bülthoff & Edelman 1992). For all the monkeys tested,

training with a limited number of views (about 10 views for the entire viewing sphere) was usually sufficient to achieve view-independent performance. This ability of humans and monkeys to interpolate between different familiar views of a novel object is consistent with a recent viewer-centered theory of recognition based on regularization networks. In short this theory assumes the existence of units acting as blurred templates, the receptive field of which develops selectivity to the views seen in the training phase. Generalization occurs by linearly combining the output of such units (for details, see review by Poggio 1990).

NEUROPSYCHOLOGICAL STUDIES

Behavioral studies of humans and animals with brain damage provide important insights about the organization of neural modules that participate in visual object recognition. While a detailed discussion of this literature is well beyond the scope of this review (see Farah 1990, Grüsser & Landis 1991), we present evidence here that dissociable clinical deficits in the visual processing of objects, for purposes of categorization, identification, and goal-directed action, strongly suggest a multiple systems architecture for recognition.

Category-Specific Breakdowns

An essential subordinate classification task for both humans and nonhuman primates is the recognition of facial identity. Ischaemic infarcts in the inferomedial occipito-temporal region of the right hemisphere have been shown to disrupt the recognition of familiar faces (Landis et al 1988). The lesions typically involve the fusiform and lingual gyri or their interconnections and are caused by strokes of the right posterior cerebral artery that typically extends from the level of the splenium of the corpus callosum to the occipital pole (but see Damasio et al 1982).

Prosopagnosia, as this disorder was called by Bodamer (1947), was traditionally considered a specific agnosia, which renders human patients incapable of recognizing the faces of familiar or famous persons but spares their ability to recognize common objects. Prosopagnosic patients can recognize individuals by their voice or even by watching their gait but fail to do so by looking at their face (Damasio et al 1982). They can also recognize a face as the object "face" and name and point to its parts (Lhermitte et al 1972, Whiteley & Warrington 1977, Damasio et al 1982). Furthermore, configurational complexity does not appear to be the cause of this disorder, since the patients do recognize visual objects that may be structurally more complex than faces (Benton 1980, Damasio 1985).

The notion that prosopagnosia is the disruption of the function of a recog-

nition system specialized for the processing of faces was questioned, however, by Faust (1955), who suggested, instead, that the major deficit of prosopagnosic patients is their inability to evaluate the structural signs in any figure that confers individuality. In support of this claim were later findings showing that prosopagnosics indeed have problems distinguishing individuals within other objects classes, such as fruits, playing cards, housefronts, and automobiles (De Renzi et al 1968, Macrae & Trolle 1956, Lhermitte & Pillon 1975). Bernstein et al (1969) reported a bird-watcher who had lost the ability to differentiate visually between birds and another patient who could no longer recognize his cows.

An appealing explanation of the various deficits observed in prosopagnosic patients was offered by Damasio and his collaborators (Damasio et al 1982). After systematically examining three such patients, these investigators noticed that, as previously described, the subjects were unable to identify various objects, including automobiles, articles of clothing, cooking utensils, and food ingredients. Recognition tests using a carefully chosen stimulus set—consisting of photographs of animals, abstract symbols, and motor vehicles—revealed that failure to identify any of these objects was caused primarily by visuostructural similarity and was not due to complexity. The patients were able to recognize animals such as horses, owls, and elephants, but they confused simpler abstract objects such as a dollar sign, a British pound sign, and a musical clef. Like the infants described above, they also failed to discriminate animals such as cats, tigers, or panthers, which, despite their great size differences, do share many similarities in shape. Damasio et al (1982) suggested that focal brain damage, such as that observed in prosopagnosics, interferes with the patient's ability to perform within-category discriminations, or identifications, without affecting the recognition of the generic class to which the stimulus belongs (see also Gaffan & Heywood 1993).

The hypothesis that the focal brain damage causing agnosia of faces may be interfering with specific, subordinate-level recognition processes and that the preserved representations may be those general representations that are the easiest to access also receives support from the finding that a patient with car agnosia could still identify special-purpose vehicles, such as an ambulance or a fire engine (Lhermitte et al 1972). Given their very atypical appearance, such objects are likely to have their own entry point attributes.

Evidence against this hypothesis, on the other hand, comes from those clinical reports showing that prosopagnosia does occasionally occur without any other subordinate recognition deficits (De Renzi 1986). Moreover, Farah et al (1991) reported that even after factors of visual complexity, such as inter-item similarity or specificity of identification, were accounted for in the analysis of visual recognition performance of two visual agnosic patients, the recognition of "living things" was still disproportionately impaired compared

to the recognition of "nonliving things." Finally, a recent study found that the well-known face-inversion effect found in normal subjects (see above) was reversed in a prosopagnosic patient, who could more accurately match inverted faces than upright ones (Farah et al 1995). Normal subjects in this task are especially proficient at matching upright faces and, unlike the prosopagnosic patient, perform much worse when the face stimuli are inverted. These clinical data support the view that specialized neural modules for recognition may coexist and may thus be selectively impaired.

Whether or not prosopagnosia is the result of failing to detect individuality, the aforementioned investigations show that at least two separate representation systems may be involved in the recognition of visual objects: one that represents prototypical objects and a second one that is employed when subordinate classifications are required. The latter system is often selectively affected by the infero-medial occipito-temporal lesions that also disturb the recognition of faces. It remains to be seen whether the capacity to perform subordinate-level recognition tasks can be selectively impaired in the absence of any deficits in face recognition, which would suggest that there are actually at least three different systems underlying recognition: one for basic-level classifications, one for subordinate-level identification, and one for recognizing the class of animate objects.

Deficits in the Recognition of Facial Expressions

Another dissociation in the recognition of objects is that observed between the identification of faces and the recognition of facial expressions—an especially challenging task for a recognition system. Facial expressions, such as smiles and frowns, are nonrigid, stretching, or bulging nonaffine transforms that have to be "discarded" when recognizing the same face under different emotional conditions. At the same time, however, the dynamic configuration of a face is endowed with a number of different meanings that are essential for social interaction in many species. Recent evidence shows this type of facial expression analysis proceeds independently of face identification. Specifically, some prosopagnosics fail to recognize the identity of a face or another unique, individual object of a category, although they retain their ability to recognize facial expressions (Tranel et al 1988).

Recently, Damasio's group (Adolphs et al 1994) reported a case of a 30-year-old woman with a confined amygdala lesion. She suffered from Urbach-Wiethe disease, which led to almost complete bilateral destruction of the amygdala, while sparing her hippocampus and all neocortical structures. The patient was tested in rating facial expressions—such as happiness, surprise, fear, anger, disgust, sadness, as well as neutral faces—and was found to be much poorer at this task than age-matched controls. Her ability to identify

individual faces, however, was completely preserved. A similar condition is found in monkeys with bilateral amygdalectomy, in which stimuli that would normally induce a fearful response fail to do so (Weiskrantz 1956, Blanchard & Blanchard 1972, Davis 1992).

Evidence for category-specific representations also comes from the description of a double dissociation between reported forms of metamorphopsia (Bodamer 1947). Metamorphopsias are severe pseudohallucinations in the perception of visual stimuli. One of three patients reported on by Bodamer (1947) experienced severe distortions while looking at faces [a condition later termed prosopo-metamorphopsia by Critchley (1953)] but did not lose the capacity to recognize the faces themselves. In contrast, there are also reports of prosopagnosic patients who are impaired in their ability to identify individual faces, but who also experience metamorphopsias with all visual stimuli except faces (Bodamer 1947).

Selective Damage of Visuomotor Representations

The posterior parietal area is known to play an important role in sensory-motor integration, as lesions in this brain region in either humans or monkeys produce a variety of spatio-perceptual or spatio-behavioral disorders. The clinical neurologist Balint was the first to describe three characteristic deficits—now known together as Balint's syndrome—in a patient with bilateral posterior parietal lesions: (*a*) psychic paralysis of the gaze; (*b*) optic ataxia, i.e. impairment of object-bound movements of the hand performed under visual guidance; and (*c*) a form of simultanagnosia, which is the inability to perceive more than one object at a time, irrespective of the object's angular extent and despite preserved visual fields (for further discussion, see Farah 1990).

Optic ataxia, which was initially considered to be simply a deficit in reaching for objects (Damasio & Benton 1979), is now known to also be a disturbance of visually guided grasping, since patients appear not only to misreach for objects, but also to demonstrate impaired execution of finger movements and show a remarkable disturbance in the formation of finger grip and hand orientation before reaching for target objects (for review, see Jeannerod et al 1995). These deficits in "preshaping" occur despite the facts that the patients can correctly perceive the shape of the objects for which they attempt to reach and that movements that do not require visual guidance, such as those directed to the body, are executed correctly (Damasio & Benton 1979). Similarly, in the monkey, a unilateral parietal leukotomy, damaging a portion of the parieto-occipital white matter but sparing the optic radiations, causes a severe contralateral impairment of fine finger-movement control and misreaching (Haaxma & Kuypers 1975).

Recent studies demonstrated that the disturbance of preshaping can be

dissociated from visuospatial perception, suggesting a double dissociation between the representations used for action and those that may underlie the perception of an object. For example, a recently reported agnosic patient was unable to perceive the orientation or size of objects, although she could still accurately use orientation and size information for visuomotor actions (Goodale et al 1991). The patient, who suffered an episode of carbon monoxide poisoning, had diffuse lesions in areas 18 and 19, with her primary visual cortex being largely intact. She was severely incapacitated in her ability to recognize visual objects based on shape information, being unable to perform elementary shape discrimination tasks (Milner & Heywood 1989). Nonetheless, she had adequate visual acuity, preserved central visual fields, and impaired but clear stereopsis, motion, and color vision (Milner & Heywood 1989).

When she was presented with a pair of rectangular plaques of the same or different dimensions, she was unable to distinguish between them. Similarly, her estimates of the width of a single plaque were only randomly related to the actual dimensions of the object, showing a considerable trial-to-trial variability (Goodale et al 1991). When the same patient, however, was asked to reach out and pick up the plaques, she did so with an index finger and thumb aperture that accurately reflected the dimensions of the objects. Based on these and similar findings, Goodale and colleagues suggested a dissociation between the representations used for apprehension and those used for action, in particular for prehensile hand movements that require accurate visual guidance.

ELECTROPHYSIOLOGICAL STUDIES

Inferotemporal Cortex

In order to discuss the physiological studies pertinent to object representation, it is important to explore, in some detail, the anatomical framework underlying visual processing in the primate brain. A survey of the anatomical organization of the visual system is also important, since the extreme diversity and complexity of areas and connections that make up the primate visual system themselves lend strong support to the idea that there are, indeed, multiple systems for recognition.

Sensory information from the primary visual cortex reaches the temporal and parietal lobes by a number of cortico-cortical stages that form two relatively separate pathways (Ungerleider & Mishkin 1982, Desimone & Ungerleider 1989). One pathway, roughly corresponding to the superior longitudinal fasciculus, passes dorsally in extrastriate cortex to end in the posterior parietal lobule and the frontal lobe; the other, corresponding to the inferior longitudinal fasciculus, passes ventrally in extrastriate cortex to reach the inferior temporal or inferotemporal cortex (IT).

IT is generally considered a large region of cortex extending approximately from just anterior to the inferior occipital sulcus to a couple of millimeters posterior to the temporal pole, and from the fundus of the superior temporal sulcus (STS) to the fundus of the occipito-temporal sulcus (Figure 2). It is roughly coextensive with Brodmann areas 20 and 21, or area TE of Von Bonin & Bailey (1947), which was later subdivided into the areas TE anteriorly and TEO posteriorly (Von Bonin & Bailey 1950, Iwai & Mishkin 1969). Area TEO forms a band extending from the lip of the STS to a few millimeters medial to the occipito-temporal sulcus. Its posterior border is close to the lip of the ascending portion of the inferior occipital sulcus, and its posterior-anterior extent is 10 to 15 mm (Boussaoud et al 1991). Area TE extends further anteriorly to about the sphenoid. Studies based on the deficits that follow focal lesions in IT—where TEO lesions lead to simple pattern deficits, while TE lesions result in associative and visual memory deficits—suggest two functional subdivisions—one posterior and one anterior—that are roughly coextensive with, but not identical to, the previously defined cytoarchitectonic TE and TEO subdivisions (Iwai 1978, 1981, 1985).

Based on topography and the laminar organization of projections, Felleman & Van Essen (1991) subdivided IT into PIT, CIT, and AIT (Figure 2), each having a ventral and dorsal portion. Based on cyto- and myeloarchitectonic criteria, as well as on the pattern of afferent cortical connections, the temporal cortex has been further subdivided into a large number of separate visual areas (Figure 3) (Seltzer & Pandya 1978, 1994), several of which have distinct physiological characteristics (Baylis et al 1987).

Area TEO receives feedforward, topographically organized cortical inputs from areas V2, V3, and V4 and has interhemispheric connections mediated mainly via the corpus callosum. Sparser inputs arise from areas V3A, V4t, and MT. Each of these areas receives a feedback connection from TEO (Distler et al 1993, Rockland et al 1994). TEO projects feedforwardly to the areas TEm, TEa, and IPa, all of which send feedback projections back to TEO. Feedback projections to this area also arise from the parahippocampal area TH and the areas TG and 36 (Distler et al 1993). Cortical projections of area TE include those to TH, TF, STP, frontal eye fields (FEF), and area 46 (Barbas & Mesulam 1981, Barbas 1988, Shiwa 1987). Area TE has both direct and indirect connections to the limbic structures. Direct connections have been reported to the amygdaloid complex (Amaral & Price 1984, Herzog & Van Hoesen 1976, Iwai & Yukie 1987, Turner et al 1980) and to the hippocampus (Yukie & Iwai 1988), which also receives an indirect projection via the parahippocampal gyrus (Van Hoesen 1982). TE does not project directly to entorhinal cortex (Insausti et al 1987), the cortical inputs of which arise primarily in the perirhinal and parahippocampal cortices (Suzuki & Amaral 1994).

Areas TEO and TE are also connected to a large number of subcortical

structures. Both areas receive nonreciprocal inputs from several nuclei of the thalamus and from the hypothalamus, locus coeruleus, reticular formation, basal nucleus of Meynert, and the dorsal and median raphe nuclei. Both are also reciprocally connected with the pulvinar and the ventral portion of the claustrum (Webster et al 1993). The main nonreciprocal output of both areas is a projection to the striatum, while TEO alone projects to superior colliculus and TE to the medial dorsal magnocellular nucleus of the thalamus (Webster et al 1993).

The pathway that begins in striate cortex, passes through the extrastriate and inferotemporal cortices, and reaches these subcortical areas is thought to underlie a variety of cognitive and visuomotor functions, such as recognition, habit formation, associative recall, and formation of visuomotor associations (Mishkin & Appenzeller 1987, Zola-Morgan & Squire 1993, Brothers & Ring 1993, Wilson & Rolls 1993). The diverse subcortical connections of many extrastriate areas, however, indicate that object-related information does not necessarily have to pass through IT to reach the striatum or the limbic structures. Different areas in the ventral, but also in the dorsal, pathway have reciprocal connections to structures such as the caudate, claustrum, amygdala, or hippocampal complex, areas which subserve different types of memory (Yeterian & Van Hoesen 1978, Webster et al 1993, Baizer et al 1993). Therefore some of the higher visual functions mentioned above may also be adequately accomplished by virtue of information derived from earlier processing levels. For example, categorizations based on the detection of some object features with high cue validity could, in principle, be possible even with damage to area TE, if these features are encoded in the activity of neurons in any of the earlier areas that project to limbic or striatal structures. Accordingly, the diversity and specificity of deficits observed in brain-damaged, agnosic patients may reflect an organization in which each processing stage copes with increasingly abstract representations and is capable, on its own, of supporting some types of categorization performance.

Physiological Properties of Inferotemporal Neurons

Gross and colleagues were the first to obtain visually evoked responses in IT by using both macro- and microelectrodes in anesthetized and unanesthetized

Figure 2   Von Bonin & Bailey's (1947) map of the (a) lateral and (b) medial surface of the *Macaca mulatta* brain. Superimposed are the major visual areas as described by Felleman & Van Essen (1991). The names from the Von Bonin & Bailey (1947) parcellation are depicted on the brain, and the labels currently used by most investigators appear adjacent to the relevant areas: visual areas 1, 2, and 4 (V1, V2, V4), ventral posterior (VP), posterior inferotemporal (PIT), central inferotemporal (CIT), anterior inferotemporal (AIT), and dorsal parietal (DP). (c) Lateral and (d) medial surface of the human brain with Brodmann's areas numbered. [Adapted from Nieuwenhuys et al (1980).] Note that the relative sizes of the macaque and human brains are not to scale.
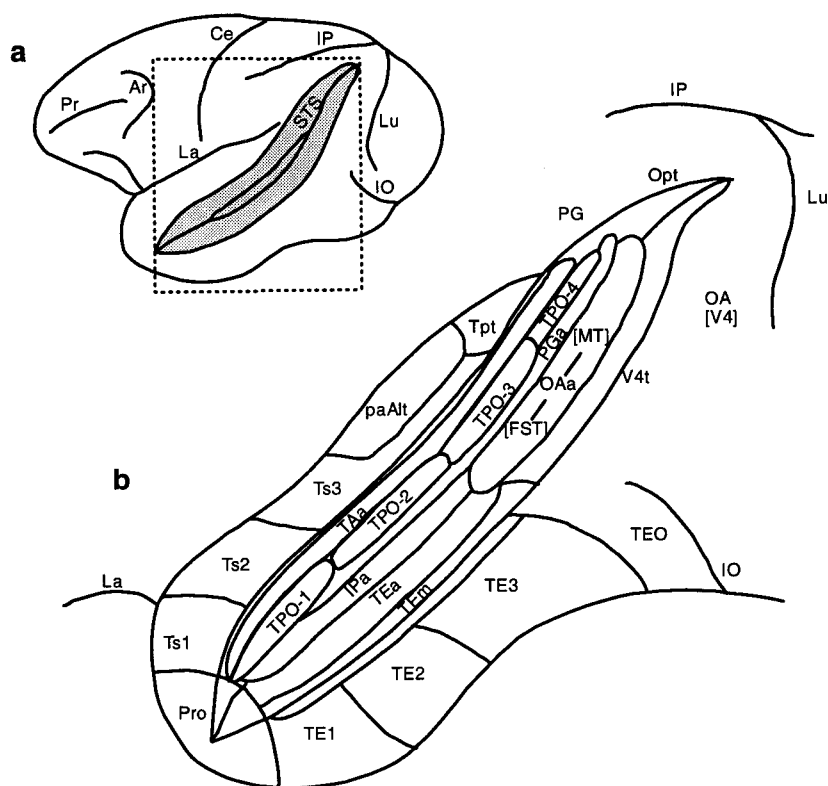
Figure 3 Subdivision of monkey inferior temporal lobe centered around the superior temporal sulcus (STS). (*a*) Lateral view of the cortical surface with major visible sulci labeled: inferior occipital (IO), lunate (Lu), intraparietal (IP), central (Ce), lateral (Sylvian) fissure (La), arcuate (Ar), and principal (Pr). (*b*) Expanded view of the inferior temporal areas surrounding the STS. [Adapted from Seltzer & Pandya (1994).]

monkeys (for review, see Gross 1994). A large number of investigations confirmed and extended the initial findings, establishing the IT as the last exclusively visual area in the ventral pathway. More than 85% of the neurons in this area are excited or inhibited by different simple or complex visual patterns (Desimone et al 1984).

The observed properties of IT cells change significantly as one moves from the most posterior part of TEO, where cells have similar properties to those observed in area V4, to the most anterior part of TE, where neurons rarely respond to such simple stimuli. Among the changing characteristics are topography, receptive field size, and stimulus selectivity. Area TEO has a coarse

visuotopic organization. It has an almost complete representation of the con-
tralateral visual field, with receptive fields that are larger than those of the V4
neurons (Boussaoud et al 1991). In contrast, area TE is not visuotopically
organized. Cells have large, ipsilateral, contralateral, or bilateral receptive
fields, almost always including the fovea (Gross et al 1972). The response of
the cells to stimuli presented at the fovea is usually more vigorous than
elsewhere in the receptive field, possibly due to the fact that IT cortex receives
a strong projection from those parts of the extrastriate cortex in which the
central visual field is overrepresented, with a smaller contribution from those
areas that process peripheral visual stimuli (Desimone et al 1980, Seltzer &
Pandya 1978).

There is a systematic increase in the receptive field size along the posterior-
anterior length of IT, with receptive field diameters in TEO as small as 1.5 to
2.5°, and those in TE reaching diameters of 30 to 50° (Boussaoud et al 1991,
Tanaka 1993). The responsiveness of TE neurons to stimuli presented in the
ipsilateral hemifield depends on the massive projections received from the
opposite hemisphere via the splenium of the corpus callosum and the anterior
commissure (Zeki 1973, Gross et al 1977). Sectioning of the splenium reduces
the incidence of ipsilateral activation by about 50%, while combined splenium
and anterior commissure sections entirely eliminate ipsilateral activation, sug-
gesting that interhemispheric connections do play an essential role in the
positional invariance observed in the response of many neurons in this area.

Many IT neurons are selective for a variety of stimulus attributes, such as
color, orientation, texture, direction of movement, or shape (Gross et al 1972,
Desimone & Gross 1979, Mikami & Kubota 1980). Of particular interest is
the sensitivity of IT neurons to stimulus shape. Although shape selectivity has
also been reported in earlier areas such as V4 (Desimone & Schein 1987,
Tanaka et al 1991, Gallant et al 1993, Kobatake & Tanaka 1994), only in IT
is this selectivity extensively encountered. Neurons in this area respond selec-
tively to a variety of natural or synthetic objects (Desimone et al 1984), to
parametric shape descriptors (Schwartz et al 1983), or to mathematically cre-
ated 2D patterns, e.g. Walsh functions, that can be used to synthesize any
arbitrary image with a given resolution (Richmond et al 1987). Groups of cells
in IT have also been found that respond to the sight of biologically important
objects such as faces or hands (see below). Face cells, which have been
reported in monkeys as young as six weeks old (Rodman et al 1993), are two
to ten times more sensitive to faces than to simple geometrical stimuli or 3D
objects (Perrett et al 1979, 1982).

Interestingly, many IT neurons show various degrees of invariance to image
transformations. The absolute response of the cells only rarely exhibits size or
position constancy (e.g. see Logothetis et al 1995). However, their selectivity
for shape, i.e. their relative preference for the optimal stimulus over several

suboptimal stimuli, is preserved over large changes in stimulus size and position (Sato et al 1980, Schwartz et al 1983, Logothetis et al 1995, Ito et al 1995). In this sense, more than half of the IT neurons can be thought of as demonstrating size and position invariance. The response of the rest of the neurons indicates some degree of size specificity, suggesting that at least some object representations might be stored in a size-specific manner (Ito et al 1995). Selectivity for shape has also been found to be cue invariant, in the sense that cell responses to an optimal stimulus remain unchanged regardless of the cues (motion, texture, or luminance) determining the object's shape (Sáry et al 1993). Contrast polarity, on the other hand, appears to have large effects on the response of IT neurons (Ito et al 1994). The effects of contrast polarity corroborate the proposed role this area may play in shape processing, since the recovery of surface structure relies partly on shading information, which in turn depends on luminance contrast polarity (Cavanagh 1987, Ramachandran 1990).

In summary, IT appears to have all the machinery requisite for the formation of object descriptions. Cells respond selectively to stimulus attributes such as color and texture, to simple and complex patterns, and to complex natural objects such as faces. They also show a certain degree of translation and scale invariance. An obvious question, then, is, What is the encoding scheme used to represent visual objects? Are they represented explicitly by the firing of a few gnostic units? Are they represented by the firing of a small population of neurons, each encoding some features, aspects, or single views? Or, are they only implicitly represented by a large population of cells each acting as a specialized pattern filter that combines certain shapes with different surface properties of objects, such as their texture, color, or lightness?

Electrophysiological findings suggest the existence of at least two possible neural mechanisms for object representation. One system may code the prototypes of objects that can be decomposed into parts and recognized by indexing these parts and their metric or spatial relationships. A second, separate, system may be used when holistic configuration rather than individual features is important and may rely primarily on small populations of neurons with strong configurational selectivity.

Combination Encoding

Recent careful studies of the properties of inferotemporal neurons have revealed a systematic organization in the temporal lobe, wherein neurons with relatively similar response properties are clustered in modules spanning the entire thickness of cortex (Fujita et al 1992, Tanaka 1993, Young 1993, Gawne & Richmond 1993, Kobatake & Tanaka 1994).

Columnar organization is a well-established cortical property in many different areas. In the early visual system, clustering is found for neurons re-

sponding selectively to simple stimulus attributes, such as position in the visual field, ocular preference, orientation, or direction of movement. In area TE, modular organization is less related to retinotopic organization and, instead, reflects similar preferences for combinations of shapes and other stimulus attributes. Details of this work are described elsewhere in this volume by Tanaka (1996). Of particular interest for this discussion are the "elaborate" cells reported by Tanaka et al (1991), which responded only to composite shapes. These cells were studied extensively by reducing the complexity of an effective visual stimulus in a systematic manner until the simplest pattern that would drive the cell maximally was determined. The degree of complexity required to drive an elaborate cell was found to increase, in general, from area TEO to area TE. In addition, cells of different modules showed greater differences in shape selectivity than cells within a single module. Based on these and related findings, it has been argued that the general class of an object could be represented by the activity across different IT modules, while detailed discriminations could, in principle, be accomplished by detecting small differences in the activity of neurons within single modules (Fujita et al 1992, Tanaka 1993, Young 1993, Gawne & Richmond 1993).

At present, it is unclear whether the critical features of the elaborate cells form a complete set of general shape descriptors that could represent any complex object or scene. Such a scheme would, in many ways, be similar to the RBC theory proposed by Biederman (1987), although there is not, as yet, evidence for cells in IT that code for the spatial relations between individual primitives. Nonetheless, the idea that prototypes may be represented by a relatively small number (estimated to be around 1300) of modules is both theoretically appealing and biologically plausible.

Is, however, such a system sufficient for representing individual exemplars of a given object category or when holistic configuration information is necessary to disambiguate individual objects? The perception of overall configuration is often crucial for subordinate-level discriminations, as in the case of face recognition mentioned earlier. Could the representation of holistic configuration be accomplished by the combination encoding scheme described above? While not excluding this possibility, two lines of evidence show that in the monkey visual system, alternative strategies are probably used when configuration or metric information is the determining factor for a categorization task. First, a large number of neurons in TE and STS seem to encode the overall shape of biologically important objects—not specific features or parts (Rolls 1994, Oram & Perrett 1994). Further, recordings from the IT of monkeys trained to identify individual objects from two novel object classes have shown that neurons can be found in TE that respond to a limited subset of views of the objects, as face-selective neurons do for views of faces (Logothetis & Pauls 1995).

Selectivity for Biologically Important Stimuli

Responses of cells to biologically relevant stimuli—such as faces, face features, hands, and other body parts—have been reported by several investigators (Gross et al 1972; Perrett et al 1982, 1985, 1989; Rolls 1984; Desimone et al 1984; Yamane et al 1988; Hasselmo et al 1989; Young & Yamane 1992). Face cells, which seem to be the most prominent class, are found in the STS and IT (areas TPO, TEa, and TEm) (Baylis et al 1987); the amygdala (Sanghera et al 1979; Leonard et al 1985; Rolls 1992a, 1992b); the ventral striatum, which receives a projection from the amygdala (Williams et al 1993, Rolls 1994); and the inferior convexity of the prefrontal cortex (Wilson et al 1994). The large number of areas containing cells responsive to faces is consistent with the hypothesis that object representations are manifest in multiple parallel sites.

There is a considerable differentiation among face-selective neurons. One subgroup appears to be very similar to the elaborate neurons described above. Some cells are selective for particular features of the head and face, e.g. the eyes (Perrett et al 1982, 1992), whereas another population of cells can only be driven by simultaneous presentation of multiple parts of a face (Perrett & Oram 1993, Wachsmuth et al 1994). Yet other face cells require the entire face-view configuration, or even a combination of information such as eye gaze angle, head direction, and body posture (Perrett et al 1992), and in this sense, they encode holistic information about shape and not information about the existence of individual features.

The selectivity of these neurons for faces is maintained over changes in stimulus size and position but less over changes in orientation. Face cells are sensitive to rotations in the picture plane, with a strong bias for upright faces (Tanaka et al 1991), and most of them show selectivity for a specific vantage point. In particular, some cells are maximally sensitive to the front view of a face, and their response falls off as the head is rotated into the profile view, while some others are sensitive to the profile view with no response to the front view of the face (see also Figures 4c and 4d). A detailed investigation of these types of cells by Perrett et al (1985) reported a total of five types of cells in the STS, each maximally responsive to one view of the head: full face, profile, back of the head, head up, and head down. In addition, two subtypes have been discovered that respond only to left profile or only to right profile, confirming that these cells are involved in visual analysis rather than in representing specific behavioral or emotional responses.

Interestingly, a study using correlation analysis between quantified facial features and neural responses has shown that face neurons can detect combinations of the distances between facial parts, such as eyes, mouth, eyebrows, and hair (Yamane et al 1988). These cells show a remarkable redundancy of coding characteristics, as only two facial dimensions seem to be necessary for

explaining most of the variance in the population. For example, all of the width measurements—such as the width of the eyes, the mouth, and the interocular distance—covary with the general width of the face. Young & Yamane (1992) also found that each face cell typically exhibits graded responses to a wide variety of face stimuli and therefore presumably participates in the representation of many different faces.

In general, face-selective neurons responsive explicitly to the identity of faces are found in IT, while cells that respond to facial expressions, gaze direction, and vantage point are mostly located in the STS (Hasselmo et al 1989, Perrett et al 1992). Such functional localization is in agreement with lesion experiments in monkeys (Heywood & Cowey 1992), showing that removal of the cortex in the banks and floor of the rostral STS of monkeys results in deficits in the perception of gaze directions and the facial expression, but not face identification (Heywood & Cowey 1992). It is also in agreement with preoperative electrophysiological recordings in epileptic patients that also suggest separate processing of facial identity and facial expressions (Ojemann et al 1993). In fact, PET studies suggest that the posterior fusiform gyrus is activated during face matching or gender discrimination testing (Haxby et al 1991, Sergent et al 1992), while the presentation of a unique face activates the mid-fusiform gyrus (Sergent et al 1992).

Cells in IT also respond to the sight of the entire human body or of body parts (Wachsmuth et al 1994). About 90% of these neurons responded to the human body in a viewer-centered fashion, whereas the rest responded equally well to any view of the stimulus. Of particular interest here is the observation that about one fifth of the neurons studied responded only to the entire body and not to the sight of any of the constituent body parts alone (Wachsmuth et al 1994).

In summary, the evidence presented here suggests that at least some objects are represented by neurons with a complex configurational selectivity that cannot be reduced to selectivity of individual features or even constellations of such features. One obvious question is whether such a configurational selectivity is specific for animate objects, such as faces and body forms. Clinical observations, as mentioned above, have shown that the recognition of living things can be selectively impaired (see e.g. Farah et al 1991). Thus one possibility is that the perception of these shapes is mediated by specialized neural populations. If so, then the complex-pattern selectivity for faces and body parts reported above may be unique to the representation of the class of living things, with different encoding mechanisms responsible for the recognition of other objects. In support of this hypothesis are the observations that face cells appear very early in the ontogenesis of monkeys (Rodman et al 1993) and that newborn human infants show a special affinity for the sight of faces (Goren et al 1975, Johnson et al 1991). Alternatively, a system based on neurons that are selective for complex

configurations may provide a general mechanism for encoding any object that cannot undergo useful decomposition in the process of recognition.

Configurational Selectivity for Novel Objects

Recently, Logothetis and his colleagues (Logothetis et al 1994, 1995; Logothetis & Pauls 1995) set out to determine whether the configurational selectivity found for IT neurons is specific for faces or body parts or whether it can be generated for any novel object as a result of extensive training. In these combined psychophysical and neurophysiological experiments, macaque monkeys were trained to become experts at identifying novel computer-generated wire and spheroidal objects, similar to those shown in Figures 1*b* and 1*d*. These objects had never been experienced by the monkeys, nor did they possess any inherent biological relevance. Nonetheless, after training, the animals learned to discriminate individual objects from a set of highly similar distractors (Logothetis et al 1994), a task not unlike the problem of identifying a specific face or a particular bird species. Because all of the objects used in testing were composed of the same basic parts, good performance in this task relied upon the detection of subtle shape differences. These experiments were thus directed at understanding how objects are neurally represented when they are encountered in the context of a subordinate-level recognition task.

Physiological recordings from individual neurons in the inferior temporal lobe, near the anterior medial temporal sulcus (AMTS), revealed a subpopulation of cells that were activated selectively by views of previously unfamiliar objects (Logothetis et al 1995, Logothetis & Pauls 1995). Many neurons fired selectively for a small set of views of a spheroidal or wire object that the monkey had learned to recognize from all viewpoints. The cells were most active when the target was presented from one particular view (Figure 4*a*), and their activity declined as the object was rotated in depth.

The neurons found in the temporal lobe of the expert monkeys bear interesting similarities to face-selective cells found in the banks of the rostral STS (NK Logothetis & DL Sheinberg, unpublished observations). Cells from both populations exhibit object-specific as well as view-specific selectivity. Figure 4 illustrates response profiles of four different cells selective for specific views of wire objects (Figures 4*a* and 4*b*) and faces (Figures 4*c* and 4*d*). The neurons depicted in Figures 4*a* and 4*c* show a marked preference for a single view of the test object and a steady decrease in mean activity for increasing object rotations. Such neurons seem to act like blurred templates, with their tolerance for small rotations in depth representing a form of limited generalization. The cells shown in Figures 4*b* and 4*d* reveal a broader form of generalization by their selective response for pseudo-reflected object views, i.e. object views 180° apart that appear as mirror images. Psychological studies have shown
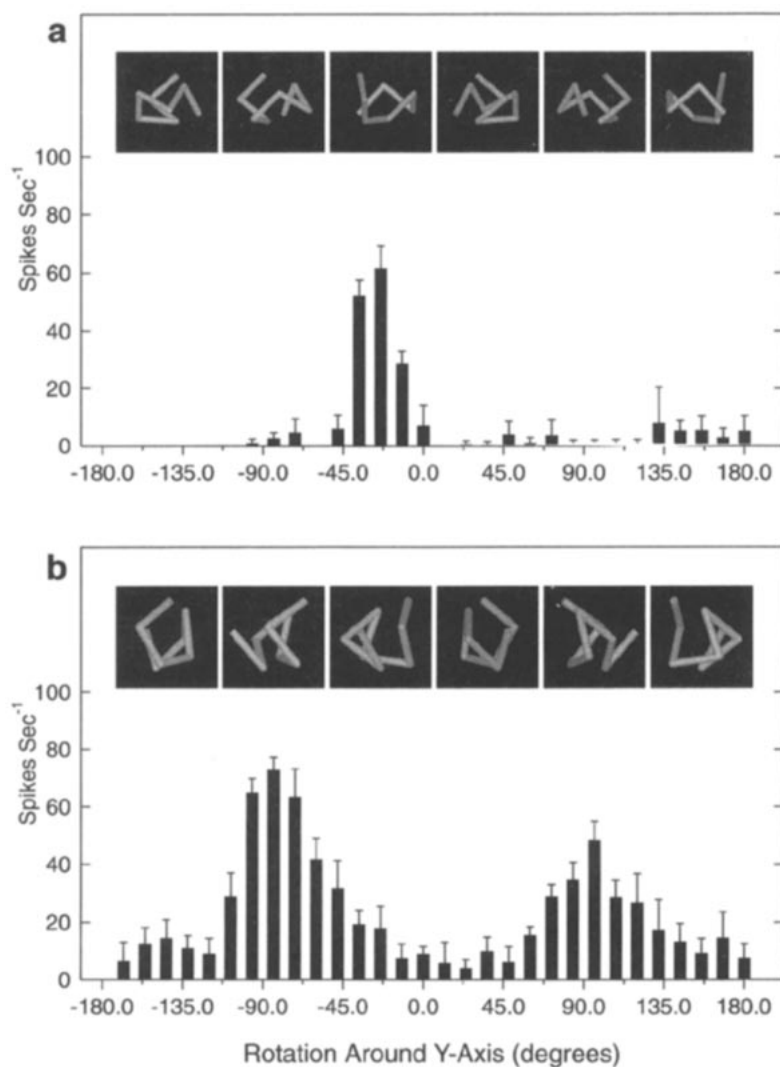
that enantiomorphic views of objects are not easily discriminated by children (Bernstein et al 1978, Corballis & McLaren 1984) and seem to be categorized equivalently by adults (Biederman & Cooper 1991, Cooper et al 1992). Generalization across enantiomorphs, at both the cellular and behavioral level, may be evidence of the primate recognition system's ability to extract a more global representation of a shape, ignoring local deviations in cases where differentiating between the two views is almost always unnecessary for the purpose of object recognition.

The narrow tuning curves encountered for views of stimuli from the two novel object classes clearly shows that high stimulus selectivity is not limited to faces or other biological forms. Thus the ability to make subordinate-level judgments about novel objects may rely on some of the same mechanisms, or perhaps even the same population of cells, that are involved in the recognition of faces.

That the stimulus selectivity of cells in IT can be altered as a result of experience has also been suggested by Kobatake et al (1993). In their experiment, a monkey was first trained for more than two months in a discrimination task, using 28 shapes composed of two or three geometric primitives. Following training, cells in IT were isolated and tested using a battery of visual stimuli while the monkey was anesthetized and immobilized. Interestingly, a much higher proportion of cells in the trained monkey were strongly activated by stimuli in the test set, compared to untrained monkeys. These results imply that long-term changes in tuning characteristics can be induced by experience and that these changes can be observed in the anesthetized animal. These data lend support to the claim that the selectivity of the view-selective cells reported by Logothetis & Pauls (1995) was, in fact, tuned throughout the course of the animals' training. In addition, they further emphasize the importance of characterizing the properties of cells in IT in the context of a behaviorally relevant task.

Action-Related Representations

As mentioned above, damage to the parietal cortex can cause a severe impairment in spatial perception (Lynch 1980, Andersen 1989). In the monkey, this cortical region covers Brodmann's areas 5 (superior parietal lobule) and 7 (inferior parietal lobule), or approximately, areas PE, PG, and PF, as defined by Von Bonin & Bailey (see Figure 2). Physiological investigations have supported the clinical and lesion studies. A major route from the occipital lobe into the parietal lobe is via the middle temporal area (MT) that is located on the posterior bank and floor of the caudal third of the STS (see references in Snowden 1994, Logothetis 1994). Area MT has neurons that are highly selective for binocular disparity, speed, and direction of stimulus motion. Informa-

tion from MT is routed to areas MST and FST, which in turn project to parietal areas, such as LIP, VIP, and 7a (Ungerleider & Desimone 1986, Felleman & Van Essen 1991). The responsiveness of many parietal neurons is strongly modulated by attention, and many neurons are related to visuomotor activity (Robinson et al 1978, Lynch 1980, Mountcastle et al 1984).

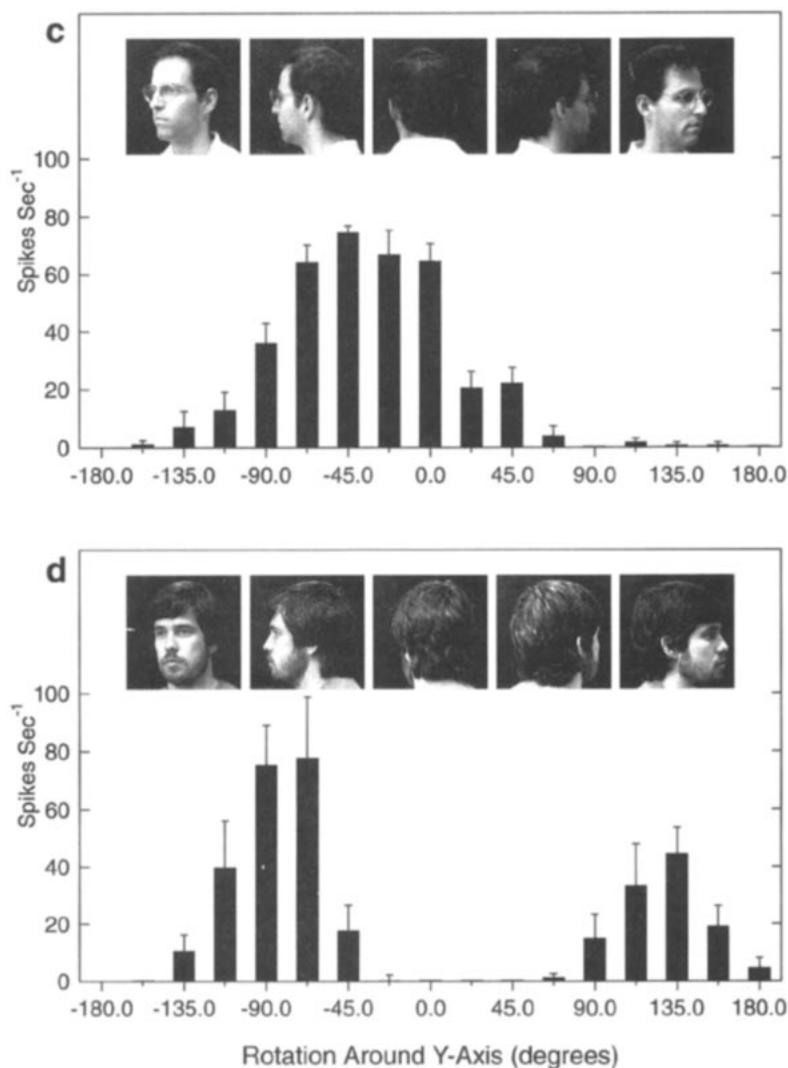Parietal cells are also sensitive to those visual qualities of an object that determine the posture of the hand and fingers during a grasping movement. In

Figure 4   Four different IT neurons selective for views of wires and faces. (*a–d*) Two of the neurons shown here responded maximally for a single view of an object (*a, c*), and response magnitude decreased gradually as the object was rotated in depth away from the preferred view. Figure 4*b* shows an example of a cell responding to two views of a wire object separated by 180°, and Figure 4*d* shows data from a cell that exhibited its maximum response for the left-facing profile of a head and nearly the same response for the right-facing profile. (Error bars indicate standard deviations of mean response rates.)

particular, for reaching and grasping tasks, neurons in area 7 of the monkey have been found to be selectively activated depending upon the configuration and orientation of the target object (Taira et al 1990, Sakata & Taira 1994). Taira et al (1990) reported a class of motor-dominant neurons that fired during hand movements in either the light or the dark but not during the visual fixation of the manipulandum. Another class of visual-and-motor neurons were found to be active in all three conditions, and a third class of visual-dominant cells fired during hand-movement or visual fixations but not in the dark. The first two classes were closely related to hand manipulation and fired more consistently in the light than in the dark, suggesting that the visual-dominant cells recorded in this area may be providing the visual input to these neurons controlling the hand configuration.

Although these experiments clearly show that cells in parietal cortex are responsive to visual stimuli, they do not directly address the question of whether some parietal neurons are truly sensitive to an object's shape. Sereno & Maunsell (1995) recently reported such evidence, however, in a study of single-unit activity in the LIP of monkeys who were trained to perform a short-term memory task. About one third of the units recorded in these experiments did indeed show significant response differences, dependent upon stimulus shape, and about a third of the units showed significant differences in delay period activity, dependent on the shape of the sample. These results suggest that parietal cells may contribute to the memory of shape features, as well as participate in the execution of visually guided actions.

CONCLUSIONS

The research reviewed in this paper suggests that recognition of visual objects relies on different types of stored representations, each employed according to the requirements of the task under study.

A fundamental problem is the classification of objects at the basic categorization level, a task that may encompass all that simple recognition systems are capable of performing. In humans, basic-level object names are the first to enter a child's vocabulary and are used to a much greater extent than any other term to describe categories. Recognition at the basic object level appears largely invariant to surface illumination, hue, or image transformations, such as scaling, translation, or rotation around any axis. The representations used in prototype recognition may rely on structural decomposition of the objects into parts and on indexing these parts and their relationships. In agreement with this notion are electrophysiological studies suggesting that basic object forms may be represented by the activity of neurons across different modules of the inferior temporal cortex, each encoding combinations of various complex forms with surface properties such as texture or color.

Categorization at the subordinate level appears to involve different types of representation, each of which may rely on different neural mechanisms than those used for the recognition of objects at the basic level. Subordinate-level recognition is initially strongly view dependent, with generalization accomplished through perceptual learning. The extent to which the different types of representations are used in recognition is likely to vary depending on the object type or familiarity. Recognition at the subordinate level is differentially affected by brain damage in the absence of any deficits in the recognition of objects at the basic level. Agnosic patients can occasionally recognize natural or synthetic objects with distinct shapes that belong to different classes, but they fail to do so when identification of individual entities is required, as when recognizing personal items such as their own wallets, cars, or articles of clothing. Identification of items of a structurally homogeneous class may rely on the activity of neurons with high configurational selectivity, such as the cells found in the experiments described above.

Neuropsychological evidence suggests a further dissociation between the representations used for the recognition of living things and nonliving things. We have discussed evidence regarding the specificity of prosopagnosia that can occur in the absence of any other associative failure. Similarly, facial memory, i.e. the matching of unfamiliar but previously presented faces, is also a somewhat different task than recognition of familiar faces. Even more striking, however, is the specific loss of humans' or monkeys' ability to recognize facial expressions and the associated emotions, despite the preserved capacity to recognize face identity. While the failure to recognize the identity of a face may simply be due to a general inability to detect individuality within a given homogeneous object class, the agnosia of facial expressions strongly suggests a specialized mechanism for the processing of biologically important configurations, as does the double dissociation between prosopo-metamorphopsia and the metamorphopsias for other visual stimuli. Evidence for brain mechanisms specialized for recognizing biological forms, especially faces, also comes from the electrophysiological findings reviewed in the previous section, showing that neurons in the temporal lobe respond selectively to faces, hands, or other body parts. Finally, both physiological and clinical work suggests a dissociation between the representations underlying perceptual categorization and those used for the action of object grasping.

In reviewing the current state of research in the field of visual object recognition, we were interested to find that the field of memory research has, in the last decade, been characterized in much the same way that we have suggested here (Tulving & Schacter 1990). Insofar as visual recognition necessarily requires memory, similar ideas should hold in both fields. In closing, then, we borrow a conclusion from a recent review on memory research that,

after substituting the term object recognition for memory, seems to apply remarkably well to the field of visual object recognition:

> [I]t is becoming increasingly clear that there are no universal principles of memory and that facts discovered about one form of memory need not hold for other forms. This is why systematic classification of memory systems, both psychological and physiological, is an essential prerequisite for the successful pursuit of the empirical and theoretical understanding of memory processes and mechanisms. The systems approach combined with appropriate processing theories seems to provide the most direct route to the future (Tulving & Schacter 1990, p. 305).

Literature Cited

Adolphs R, Tranel D, Damasio H, Damasio A. 1994. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372:669–72

Aloimonos Y, ed. 1993. *Active Perception.* Hillsdale, NJ: Erlbaum

Amaral DG, Price JL. 1984. Amygdalo-cortical projections in the monkey (*macaca fascicularis*). *J. Comp. Neurol.* 230:465–96

Andersen RA. 1989. Visual and eye movement functions of the posterior parietal cortex. *Annu. Rev. Neurosci.* 12:377–403

Anderson JA, Silverstein JW, Ritz SA, Jones RS. 1977. Distinctive features, categorical perception, and probability learning: some applications of a neural model. *Psychol. Rev.* 84:413–51

Baizer JS, Desimone R, Ungerleider LG. 1993. Comparison of subcortical connections of inferior temporal and posterior parietal cortex in monkeys. *Vis. Neurosci.* 10:59–72

Banks WP, Krajicek D. 1991. Perception. *Annu. Rev. Psychol.* 42:305–31

Barbas H. 1988. Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 276:313–42

Barbas H, Mesulam MM. 1981. Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *J. Comp. Neurol.* 200: 407–31

Bartram DJ. 1974. The role of visual and semantic codes in object naming. *Cogn. Psychol.* 6:325–56

Bartram DJ. 1976. Levels of coding in picture-picture comparison tasks. *Mem. Cogn.* 4: 593–602

Baylis GC, Rolls ET, Leonard CM. 1987. Functional subdivisions of the temporal lobe neocortex. *J. Neurosci.* 7:330–42

Benton AL. 1980. The neuropsychology of facial recognition. *Am. Psychol.* 35:176–86

Biederman I. 1987. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94:115–47

Biederman I, Cooper EE. 1991. Evidence for

complete translational and reflectional invariance in visual object priming. *Perception* 20:585–95

Biederman I, Cooper EE. 1992. Size invariance in visual object priming. *J. Exp. Psychol.: Hum. Percept. Perform.* 18:121–33

Biederman I, Gerhardstein PC. 1993. Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *J. Exp. Psychol.: Hum. Percept. Perform.* 19:1162–82

Blanchard DC, Blanchard RJ. 1972. Innate and conditioned reactions to threat in rats with amygdaloid lesions. *J. Comp. Physiol. Psychol.* 81:281–90

Bodamer J. 1947. Die Prosop-Agnosie (Die Agnosie des Physiognomieerkennens). *Arch. Psychiatr. Nervenkr.* 179:6–54

Bomba PC, Siqueland ER. 1983. The nature and structure of infant form categories. *J. Exp. Child. Psychol.* 35:294–328

Bornstein MH, Gross CG, Wolf JZ. 1978. Perceptual similarity of mirror images in infancy. *Cognition* 6:89–116

Bornstein MH, Sroka H, Munitz H. 1969. Prosopagnosia with animal face agnosia. *Cortex* 5:164–69

Boussaoud D, Desimone R, Ungerleider LG. 1991. Visual topography of area TEO in the macaque. *J. Comp. Neurol.* 306:554–75

Bowman RS, Sutherland NS. 1970. Shape discrimination by goldfish: coding of irregularities. *J. Comp. Physiol. Psychol.* 72:90–97

Brothers L, Ring B. 1993. Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behav. Brain Res.* 57:53–61

Brown R. 1958. How shall a thing be called? *Psychol. Rev.* 65:14–21

Bruce CJ. 1982. Face recognition by monkeys: absence of an inversion effect. *Neuropsychologia* 20:515–21

Bruce V. 1982. Changing faces: visual and nonvisual coding processes in face recognition. *Br. J. Psychol.* 73:105–16

Bülthoff HH, Edelman S. 1992. Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl. Acad. Sci. USA* 89:60–64

Bushnell IWR, Sai F, Mullin JT. 1989. Neonatal recognition of the mother's face. *J. Dev. Psychol.* 7:3–15

Carey S, Diamond R. 1977. From piecemeal to configuration representation of faces. *Science* 195:312–13

Cavanagh P. 1987. Reconstructing the third dimension: interactions between color, texture, motion, binocular disparity and shape. *Comp. Vis. Graphics Image Proc.* 37:171–95

Cerella J. 1979. Visual classes and natural categories in the pigeon. *J. Exp. Psychol.: Hum. Percept. Perform.* 5:68–77

Cooper LA, Schacter DL, Delaney SM,

Ballesteros S, Moore C. 1992. Priming and recognition of transformed three-dimensional objects: effects of size and rotation. *J. Exp. Psychol.: Learn. Mem. Cogn.* 18:43–57

Corballis MC, McLaren R. 1984. Winding one's ps and qs: mental rotation and mirror-image discrimination. *J. Exp. Psychol.: Hum. Percept. Perform.* 10:318–27

Corballis MC, Zbrodoff NJ, Shetzer LI, Butler PB. 1978. Decisions about identity and orientation of rotated letters and digits. *Mem. Cogn.* 6:98–107

Critchley M, ed. 1953. *The Parietal Lobes.* New York: Hafner

Damasio AR. 1985. Disorders of complex visual processing: agnosia, achromatopsia, Balint's syndrome, and related difficulties of orientation and construction. In *Principles of Behavioural Neurology,* ed. MM Mesulam, pp. 259–88. Philadelphia: Davis Co.

Damasio AR, Benton AL. 1979. Impairment of hand movements under visual guidance. *Neurology* 29:170–78

Damasio AR, Damasio H, Van Hoesen GW. 1982. Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology* 32:331–41

Damasio AR, Tranel D, Damasio H. 1990. Face agnosia and the neural substrate of memory. *Annu. Rev. Neurosci.* 13:89–109

Davis M. 1992. The role of amygdala in fear and anxiety. *Annu. Rev. Neurosci.* 15:352–75

Davis RT. 1974. *Monkeys as Perceivers.* New York: Academic

De Renzi E. 1986. Slowly progressive visual agnosia or apraxia without dementia. *Cortex* 22:171–80

De Renzi E, Faglioni P, Spinnler H. 1968. The performance of patients with unilateral brain damage on face recognition tasks. *Cortex* 4:17–34

Desimone R, Albright TD, Gross CG, Bruce CJ. 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4:2051–62

Desimone R, Fleming JFR, Gross CG. 1980. Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Res.* 184:41–55

Desimone R, Gross CG. 1979. Visual areas in the temporal cortex of the macaque. *Brain Res.* 178:363–80

Desimone R, Schein SJ. 1987. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J. Neurophysiol.* 57:835–67

Desimone R, Ungerleider LG. 1989. Neural mechanisms of visual processing in monkeys. In *Handbook of Neuropsychology,* ed. F Boller, J Grafman, 2:267–99. Amsterdam: Elsevier

Diamond R, Carey S. 1986. Why faces are and are not special: an effect of expertise. *J. Exp. Psychol.* 115:107–17

Distler C, Boussaoud D, Desimone R, Unger-

leider LG. 1993. Cortical connections of inferior temporal area TEO in macaque monkeys. *J. Comp. Neurol.* 334:125–50

Edelman S, Bülthoff HH. 1992. Orientation dependence in the recognition of familiar and novel views of 3D objects. *Vis. Res.* 32: 2385–400

Eimas PD, Quinn PC. 1994. Studies on the formation of perceptually based basic-level categories in young infants. *Child Devel.* 65: 903–17

Eley MG. 1982. Identifying rotated letter-like symbols. *Mem. Cogn.* 10:25–32

Ellis R, Allport DA, Humphreys GW, Collis J. 1989. Varieties of object constancy. *Q. J. Exp. Psychol.* 41A:775–96

Fantz RL. 1964. Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science* 146:668–70

Farah MJ. 1990. *Visual Agnosia.* Cambridge, MA: MIT Press

Farah MJ, McMullen PA, Meyer MM. 1991. Can recognition of living things be selectively impaired? *Neuropsychologia* 29:185–93

Farah MJ, Rochlin R, Klein KL. 1994. Orientation invariance and geometric primitives in shape recognition. *Cogn. Sci.* 18:325–44

Farah MJ, Wilson KD, Drain HM, Tanaka JR. 1995. The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific mechanisms. *Vis. Res.* 35: 2089–93

Faust C, ed. 1955. *Die zerebralen Herdstürungen nach Hinterhauptverletzungen und ihre Beurteilung.* Stuttgart: Thieme

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in primate cerebral cortex. *Cereb. Cortex* 1:1–47

Foster DH, Kahn JI. 1985. Internal representations and operations in the visual comparison of transformed patterns: effects of pattern point-inversion, positional symmetry, and separation. *Biol. Cybern.* 51:305–12

Franks JJ, Bransford JD. 1971. Abstraction of visual patterns. *J. Exp. Psychol.* 90:65–74

Fujita I, Tanaka K, Ito M, Cheng K. 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360: 343–46

Gaffan D, Heywood CA. 1993. A spurious category-specific visual agnosia for living things in human and nonhuman primates. *J. Cogn. Neurosci.* 5:118–28

Gallant JL, Braun J, Van Essen DC. 1993. Selectivity for polar, hyperbolic, and cartesian gratings in macaque visual cortex. *Science* 259:100–3

Gauthier I. 1995. *Becoming a "greeble" expert: exploring the face recognition mechanism.* MS thesis. Yale Univ.

Gawne TJ, Richmond BJ. 1993. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J. Neurosci.* 13:2758–71

Goodale MA, Milner AD, Jakobson LS, Carey DP. 1991. A neurological dissociation between perceiving objects and grasping them. *Nature* 349:154–56

Goren CC, Sarty M, Wu RWK. 1975. Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56:544–49

Gross CG. 1973. Visual functions of inferotemporal cortex. In *Handbook of Sensory Physiology,* ed. R Jung, 7/3B:451–82. Berlin: Springer-Verlag

Gross CG. 1994. How inferior temporal cortex became a visual area. *Cereb. Cortex* 4:455–69

Gross CG, Bender DB, Mishkin M. 1977. Contributions of the corpus callosum and the anterior commissure to visual activation of inferior temporal neurons. *Brain Res.* 131: 227–39

Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.* 35:96–111

Gross CG, Rodman HR, Gochin PM, Colombo MW. 1993. Inferior temporal cortex as a pattern recognition device. In *Computational Learning and Cognition: Proc. 3rd NEC Res. Symp.,* ed. E Baum, p. 44. Slam: NEC Res.

Grüsser OJ, Landis T, eds. 1991. *Visual Agnosias and Other Disturbances of Visual Perception and Cognition.* London: Macmillan

Haaxma R, Kuypers HGJM. 1975. Intrahemispheric cortical connexions and visual guidance of hand and finger movements in the rhesus monkey. *Brain* 98:239–60

Hasselmo ME, Rolls ET, Baylis GC, Nalwa V. 1989. Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Exp. Brain Res.* 75:417–29

Haxby JV, Grady CL, Horwitz B, Ungerleider LG, Mishkin M, et al. 1991. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc. Natl. Acad. Sci. USA* 88:1621–25

Herrnstein RJ, Loveland DH. 1964. Complex visual concept in the pigeon. *Science* 146: 549–51

Herrnstein RJ, Loveland DH, Cable C. 1976. Natural concepts in pigeons. *J. Exp. Psychol.: Anim. Behav. Proc.* 2:285–302

Herzog AG, Van Hoesen GW. 1976. Temporal neocortical afferent connections to the amygdala in the rhesus monkey. *Brain Res.* 115:57–69

Heywood CA, Cowey A. 1992. The role of the 'face-cell' area in the discrimination and recognition of faces by monkeys. *Philos. Trans. R. Soc. London Ser. B* 335:31–38

Homa D, Chambliss D. 1975. The relative contributions of common and distinctive information on the abstraction from ill-defined categories. *J. Exp. Psychol.: Hum. Learn. Mem.* 1:351–59

Homa D, Cross J, Cornell D, Goldman D, Shwartz S. 1973. Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *J. Exp. Psychol.* 101:116–22

Homa D, Sterling S, Trepel L. 1981. Limitations of exemplar-based generalization and the abstraction of categorical information. *J. Exp. Psychol.: Hum. Learn. Mem.* 7:418–39

Homa D, Vosburgh R. 1976. Category breadth and the abstraction of prototypical information. *J. Exp. Psychol.: Hum. Learn. Mem.* 2:322–30

Humphrey GK, Khan SC. 1992. Recognizing novel views of three-dimensional objects. *Can. J. Psychol.* 46:170–90

Humphreys GW, Riddoch MJ. 1987a. *To See But Not To See: A Case Study of Visual Agnosia.* Hillsdale, NJ.: Erlbaum

Humphreys GW, Riddoch MJ, eds. 1987b. *Visual Object Processing: A Cognitive Neuropsychological Approach.* Hillsdale, NJ.: Erlbaum

Insausti R, Amaral DG, Cowan WM. 1987. The entorhinal cortex of the monkey. II. Cortical afferents. *J. Comp. Neurol.* 264:356–95

Ito M, Fujita I, Tamura H, Tanaka K. 1994. Processing of contrast polarity of visual images in inferotemporal cortex of the macaque monkey. *Cereb. Cortex* 4:499–508

Ito M, Tamura H, Fujita I, Tanaka K. 1995. Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J. Neurophysiol.* 73:218–26

Iwai E. 1978. The visual learning area in the inferotemporal cortex of monkeys. In *Integrative Control Functions of the Brain,* ed. M Ito, pp. 419–27. Tokyo: Kodansha

Iwai E. 1981. Visual mechanisms in the temporal and prestriate association cortices of the monkey. *Adv. Physiol. Sci.* 17:279–86

Iwai E. 1985. Neuropsychological basis of pattern vision in macaque monkeys. *Vis. Res.* 25:425–39

Iwai E, Mishkin M. 1969. Further evidence on the locus of the visual area in the temporal lobe of the monkey. *Exp. Neurol.* 25:585–94

Iwai E, Yukie M. 1987. Amygdalofugal and amygdalopetal connections with modality-specific visual cortical areas in macaques (*macaca fuscata, m. mulatta* and *m. fascicularis*). *J. Comp. Neurol.* 261:362–87

Jeannerod M, Arbib MA, Rizzolatti G, Sakata H. 1995. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 7:314–20

Johnson MH, Dziurawiec S, Ellis H, Morton J. 1991. Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40:1–19

Jolicoeur P. 1985. The time to name disoriented natural objects. *Mem. Cogn.* 13:289–303

Jolicoeur P. 1987. A size-congruency effect in memory for visual shape. *Mem. Cogn.* 15:531–43

Jolicoeur P, Gluck MA, Kosslyn SM. 1984. Pictures and names: making the connection. *Cogn. Psychol.* 16:243–75

Katz D, ed. 1953. *Studien zur experimentellen Psychologie.* Basel: Schwabe

Knapp AG, Anderson JA. 1984. Theory of categorization based on distributed memory storage. *J. Exp. Psychol.* 10:616–37

Kobatake E, Tanaka K. 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* 71:856–67

Kobatake E, Tanaka K, Wang G, Tamori Y. 1993. Effects of adult learning on the stimulus selectivity of cells in the inferotemporal cortex. *Soc. Neurosci. Abstr.* 19:975

Koenderink JJ. 1990. *Solid Shape.* Cambridge, MA: MIT Press

Landis T, Regard M, Bliestle A, Kleihues P. 1988. Prosopagnosia and agnosia for non-canonical views. An autopsied case. *Brain* 111:1287–97

Leonard CM, Rolls ET, Wilson FA, Baylis GC. 1985. Neurons in the amygdala of the monkey with responses selective for faces. *Behav. Brain. Res.* 15:159–76

Lhermitte F, Chain F, Escourolle R, Ducarne B, Pillon B. 1972. Etude anotomo-clinique d'un cas de prosopagnosie. *Rev. Neurol.* 126:329–46

Lhermitte F, Pillon B. 1975. La prosopagnosie. Role de l'hemisphere droit dans la perception visuelle. *Rev. Neurol.* 131:791–812

Liu Z, Knill DC, Kersten D. 1995. Object classification for human and ideal observers. *Vis. Res.* 35:549–68

Logothetis NK. 1994. Physiological studies of motion inputs. In *Visual Detection of Motion.* ed. AT Smith, RJ Snowden, pp. 177–216. New York: Academic

Logothetis NK, Pauls J. 1995. Psychophysical and physiological evidence for viewer-centered representations in the primate. *Cereb. Cortex* 5:270–88

Logothetis NK, Pauls J, Bülthoff HH, Poggio T. 1994. View-dependent object recognition by monkeys. *Curr. Biol.* 4:401–14

Logothetis NK, Pauls J, Poggio T. 1995. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5:552–63

Lorenz K. 1971. *Studies in Animal and Human Behaviour.* Vol. 2. Cambridge, MA: Harvard Univ. Press

Lynch JC. 1980. The functional organization of posterior parietal association cortex. *Behav. Brain Sci.* 3:485–534

Macrae D, Trolle W. 1956. The defect of function in visual agnosia. *Brain* 79:94–110

Marr D. 1982. *Vision.* New York: Freeman

Marr D, Nishihara HK. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. London Ser. B* 200:269–94

Marsolek CJ. 1995. Abstract visual-form representation in the left cerebral hemisphere. *J. Exp. Psychol.: Hum. Percept. Perform.* 21:375–86

McClelland JL, Rumelhart DE. 1985. Distributed memory and the representation of general and specific information. *J. Exp. Psychol.* 114:159–88

McMullen B, Jolicoeur P. 1992. The reference frame and effects of orientation on finding the top of rotated objects. *J. Exp. Psychol.: Hum. Percept. Perform.* 18:802–20

Mikami A, Kubota K. 1980. Inferotemporal neuron activities and color discrimination with delay. *Brain Res.* 182:65–78

Milliken B, Jolicoeur P. 1992. Size effects in visual recognition memory are determined by perceived size. *Mem. Cogn.* 20:83–95

Milner AD, Heywood CA. 1989. A disorder of lightness discrimination in a case of visual form agnosia. *Cortex* 25:489–94

Mishkin M, Appenzeller T. 1987. The anatomy of memory. *Sci. Am.* 256:80–89

Miyashita Y. 1993. Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* 16:245–63

Mountcastle VB, Motter BC, Steinmetz MA, Duffy CJ. 1984. Looking and seeing: the visual functions of the parietal lobe. In *Dynamic Aspects of Neocortical Function.* ed. GM Edelman, WE Call, WM Cowan, pp. 159–93. New York: Wiley

Murphy GL, Brownell HH. 1985. Category differentiation in object recognition: typicality constraints on the basic category advantage. *J. Exp. Psychol.: Learn. Mem. Cogn.* 11:70–84

Murray JE, Jolicoeur P, McMullen PA, Ingleton M. 1993. Orientation-invariant transfer of training in the identification of rotated natural objects. *Mem. Cogn.* 21:604–10

Nazir TA, O'Regan JK. 1990. Some results on translation invariance in the human visual system. *Spatial Vis.* 5:81–100

Nieuwenhuys R, Voogd J, van Huijzen C. 1980. *Das Zentrainervensystem des Menschen.* Berlin: Springer-Verlag

Ojemann GA, Ojemann JG, Haglund M, Holmes M, Lettich E. 1993. Visually related activity in human temporal cortical neurons. In *Functional Organisation of the Human Cortex.* ed. B Gulyaas, D Ottoson, PE Roland, pp. 279–89. Oxford: Pergamon

Oram MW, Perrett DI. 1994. Modeling visual recognition from neurobiological constraints. *Neural Networks* 7:945–72

O'Regan JK. 1992. Solving the 'real' mysteries of visual perception: the world as an outside memory. *Can. J. Psych.* 46:461–88

Palmer SE. 1977. Hierarchical structure in perceptual representation. *Cogn. Psychol.* 9:441–74

Palmer SE, Rosch E, Chase P. 1981. Canonical perspective and the perception of objects. In *Attention and Performance IX,* ed. J Long, A Baddeley, pp. 135–51. Hillsdale: Erlbaum

Pentland AP. 1986. Perceptual organization and the representation of natural form. *Artif. Intell.* 28:293–331

Perrett DI, Harries MH, Bevan R, Thomas S, Benson PJ, et al. 1989. Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* 146:87–113

Perrett DI, Hietanen JK, Oram MW, Benson PJ. 1992. Organization and functions of cells responsive to faces in the temporal cortex. *Philos. Trans. R. Soc. London Ser. B* 335:23–30

Perrett DI, Oram MW. 1993. Neurophysiology of shape processing. *Image Vis. Comput.* 11:317–33

Perrett DI, Rolls ET, Caan W. 1979. Temporal lobe cells of the monkey with visual responses selective for faces. *Neurosci. Lett. Suppl.* S3:S358

Perrett DI, Rolls ET, Caan W. 1982. Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* 47:329–42

Perrett DI, Smith PAJ, Potter DD, Mistlin AJ, Head AS, et al. 1985. Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proc. R. Soc. London Ser. B* 223:293–317

Piaget J. 1969. *The Mechanisms of Perception.* Transl. M. Cook. New York: Basic Books

Pinker S. 1985. Visual cognition: an introduction. In *Visual Cognition.* ed. S Pinker, pp. 1–63. Cambridge, MA: MIT Press

Poggio T. 1990. A theory of how the brain might work. *Cold Spring Harbor Symp. Quant. Biol.* 55:899–910

Posner MI, Goldsmith R, Welton KE. 1967. Perceived distance and the classification of distorted patterns. *J. Exp. Psychol.* 73:28–38

Posner MI, Keele SW. 1968. On the genesis of abstract ideas. *J. Exp. Psychol.* 77:353–63

Posner MI, Keele SW. 1970. Retention of abstract ideas. *J. Exp. Psychol.* 83:304–8

Quinn PC, Eimas PD. 1986. Categorization in early infancy. *Merrill-Palmer Q.* 32:331–63

Quinn PC, Eimas PD, Rosenkrantz SL. 1993. Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception* 22:463–75

Ramachandran VS. 1990. Perceiving shape from shading. In *The Perceptual World.* ed. I Rock, pp. 127–38. New York: Freeman

Reed SK. 1972. Pattern recognition and categorization. *Cogn. Psychol.* 3:382–407

Richmond BJ, Optican LM, Podell M, Spitzer H. 1987. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. 1. Response characteristics. *J. Neurophysiol.* 57:132–46

Robinson DL, Goldberg ME, Stanton GB. 1978. Parietal association cortex in the primate: sensory mechanisms and behavioral modulations. *J. Neurophysiol.* 41:910–32

Rock I, DiVita J. 1987. A case of viewer-centered object perception. *Cogn. Psychol.* 19: 280–93

Rock I, DiVita J, Barbeito R. 1981. The effect on form perception of change of orientation in the third dimension. *J. Exp. Psychol.* 7: 719–32

Rock I, Wheeler D, Tudor L. 1989. Can we imagine how objects look from other viewpoints? *Cogn. Psychol.* 21:185–210

Rockland KS, Saleem KS, Tanaka K. 1994. Divergent feedback connections from areas V4 and TEO in the macaque. *Vis. Neurosci.* 11:579–600

Rodman HR, Ò Scalaidhe SP, Gross CG. 1993. Response properties of neurons in temporal cortical visual areas of infant monkeys. *J. Neurophysiol.* 70:1115–36

Rolls ET. 1984. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum. Neurobiol.* 3:209–22

Rolls ET. 1992a. Neurophysiolgy and functions of the primate amygdala. In *The Amygdala*, ed. JP Aggleton, pp. 143–65. New York: Wiley-Liss

Rolls ET. 1992b. Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philos. Trans. R. Soc. London Ser. B* 35:11–21

Rolls ET. 1994. Brain mechanisms for invariant visual recognition and learning. *Behav. Process.* 22:113–38

Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P. 1976a. Basic objects in natural categories. *Cogn. Psychol.* 8:382–439

Rosch E, Simpson C, Miller RS. 1976b. Structural bases of typicality effects. *J. Exp. Psychol.: Hum. Percept. Perform.* 2:491–502

Rosenfeld SA, Van Hoesen GW. 1979. Face recognition in the rhesus monkey. *Neuropsychologia* 17:503–9

Ryan CME. 1982. Concept formation and individual recognition in the domestic chicken (*gallus gallus*). *Behav. Anal. Lett.* 2:213–20

Sakata H, Taira M. 1994. Parietal control of hand action. *Curr. Opin. Neurobiol.* 4:847–56

Sanghera MK, Rolls ET, Roper-Hall A. 1979. Visual responses of neurons in the dorsolateral amygdala of the alert monkey. *Exp. Neurol.* 63:610–26

Sáry C, Vogels R, Orban GA. 1993. Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995–97

Sato T, Kawamura T, Iwai E. 1980. Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. *Exp. Brain Res.* 38:313–19

Schwartz EL, Desimone R, Albright TD, Gross CG. 1983. Shape recognition and inferior temporal neurons. *Proc. Natl. Acad. Sci. USA* 80:5776–78

Seltzer B, Pandya DN. 1978. Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res.* 149:1–24

Seltzer B, Pandya DN. 1994. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J. Comp. Neurol.* 343:445–63

Sereno A, Maunsell JHR. 1995. Spatial and shape selective sensory and attentional effects in neurons in the macaque lateral intraparietal cortex (LIP). *Invest. Ophthalmol. Vis. Sci. Suppl.* 36:S692

Sergent J, Ohta S, MacDonald B. 1992. Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115:15–36

Shinar D, Owen DH. 1973. Effects of form rotation on the speed of classification: the development of shape constancy. *Percept. Psychophys.* 14:149–54

Shiwa T. 1987. Corticocortical projections to the monkey temporal lobe with particular reference to the visual processing pathways. *Arch. Ital. Biol.* 125:139–54

Simion F, Bagnara S, Roncato S, Umilta C. 1982. Transformation processes upon the visual code. *Percept. Psychophys.* 31:13–25

Snowden RJ. 1994. Motion processing in the primate cerebral cortex. In *Visual Detection of Motion*, ed. AT Smith, RJ Snowden, pp. 51–83. New York: Academic

Srinivas K. 1993. Perceptual specificity in nonverbal priming. *J. Exp. Psychol.: Learn. Mem. Cogn.* 19:582–602

Strange W, Kenney T, Kessel F, Jenkins J. 1970. Abstraction over time of prototypes from distortions of random dot patterns. *J. Exp. Psychol.* 83:508–10

Suzuki WA, Amaral DG. 1994. Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J. Neurosci.* 14:1856–77

Taira M, Mine S, Georgopoulos AP, Murata A, Sakata H. 1990. Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp. Brain Res.* 83:29–36

Tanaka JW, Taylor M. 1991. Object categories

and expertise: Is the basic level in the eye of the beholder? *Cogn. Psychol.* 23:457–82

Tanaka K. 1993. Neuronal mechanisms of object recognition. *Science* 262:685–88

Tanaka K. 1996. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19:109–39

Tanaka K, Saito HA, Fukada Y, Moriya M. 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* 66:170–89

Tarr M, Pinker S. 1989. Mental rotation and orientation-dependence in shape recognition. *Cogn. Psychol.* 21:233–82

Tarr MJ. 1995. Rotating objects to recognize them: a case study of the role of mental transformations in the recognition of three-dimensional objects. *Psychol. Bull. Rev.* 2: 55–82

Tranel D, Damasio AR, Damasio H. 1988. Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology* 38:690–96

Tulving E, Schacter DL. 1990. Priming and human memory systems. *Science* 247:301–6

Turner BH, Mishkin M, Knapp M. 1980. Organization of the amygdalopetal projections from modality specific cortical association areas in the monkey. *J. Comp. Neurol.* 191: 515–43

Tversky B, Hemenway K. 1984. Objects, parts, and categories. *J. Exp. Psychol.* 113:169–93

Ullman S. 1989. Aligning pictorial descriptions: an approach to object recognition. *Cognition* 32:193–254

Ungerleider LG, Desimone R. 1986. Projections to the superior temporal sulcus from the central and peripheral field representations of VI and V2. *J. Comp. Neurol.* 248:147–63

Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In *Analysis of Visual Behavior*, ed. DJ Ingle, pp. 549–86. Cambridge, MA: MIT Press

Valentine T. 1988. Upside-down faces: a review of the effect of inversion upon face recognition. *Br. J. Psychol.* 79:471–92

Van Hoesen GW. 1982. The parahippocampal gyrus: new observations regarding its cortical connections in the monkey. *Trends Neurosci.* 52:345–50

Vetter T, Poggio T, Bülthoff HH. 1994. The importance of symmetry and virtual views in three-dimensional object recognition. *Curr. Biol.* 4:18–23

Vishton PM, Cutting JE. 1995. Veridical size perception for action reaching vs. estimating. *Invest. Ophthalmol. Vis. Sci. Suppl.* 36:S358

Von Bonin G, Bailey P. 1947. *The Neocortex of Macaca Mulatta.* Urbana, IL: Univ. Ill. Press. 4th ed.

Von Bonin G, Bailey P. 1950. *The Neocortex of the Chimpanzee.* Urbana, IL: Univ. Ill. Press

Wachsmuth E, Oram MW, Perrett DI. 1994.

Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb. Cortex* 4: 509–22

Warrington EK, Taylor AM. 1973. The contribution of the right parietal lobe to object recognition. *Cortex* 9:152–64

Watanabe S, Ito Y. 1991. Discrimination of individuals in pigeons. *Bird Behav.* 9:20–29

Watanabe S, Lea SEG, Dittrich WH. 1993. What can we learn from experiments on pigeon concept discrimination? In *Vision, Brain, and Behavior in Birds,* ed. HP Zeigler, HJ Bischof, pp. 351–76. Cambridge, MA: MIT Press

Waxman AM, Seibert M, Bernardon AM, Fay DA. 1993. Neural systems for automatic target learning and recognition. *Lincoln Lab J.* 6:77–116

Webster MJ, Bachevalier J, Ungerleider LG. 1993. Subcortical connections of inferior temporal areas TE and TEO in macaque monkeys. *J. Comp. Neurol.* 335:73–91

Weiskrantz L. 1956. Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *J. Comp. Physiol. Psychol.* 49:381–91

Whiteley AM, Warrington EK. 1977. Prosopagnosia: a clinical, psychological, and anatomical study of three patients. *J. Neurol. Neurosurg. Psychiatry* 40:395–403

Williams GV, Rolls ET, Leonard CM, Stern C. 1993. Neuronal responses in the ventral striatum of the behaving macaque. *Behav. Brain Res.* 55:243–52

Wilson FA, Rolls ET. 1993. The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Exp. Brain Res.* 93:367–82

Wilson FAW, Ó Scalaidhe SP, Goldman-Rakic PS. 1994. Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 91:4009–13

Yamane S, Kaji S, Kawano K. 1988. What facial features activate face neurons in the inferotemporal cortex of the monkey? *Exp. Brain Res.* 73:209–14

Yeterian EH, Van Hoesen GW. 1978. Corticostriate projections in the rhesus monkey: the organization of certain cortico-caudate connections. *Brain Res.* 139:43–63

Young AW, Hellawell D, Hay DC. 1987. Configurational information in face perception. *Perception* 16:747–60

Young MP. 1993. Visual cortex: modules for pattern recognition. *Curr. Biol.* 3:44–46

Young MP, Yamane S. 1992. An analysis at the population level of the processing of faces in the inferotemporal cortex. In *Brain Mechanisms of Perception and Memory: From Neuron to Behaviour,* ed. L Squire, T

Ono, M Fukuda, D Perrett, pp. 47–71. New York: Oxford Univ. Press

Yukie M, Iwai E. 1988. Direct projections from the ventral TE area of the inferotemporal cortex to hippocampal field CA1 in the monkey. *Neurosci. Lett.* 88:6–10

Zeki SM. 1973. Comparison of the cortical de-generation in the visual regions of the temporal lobe of the monkey following section of the anterior commissure and the splenium. *J. Comp. Neurol.* 148:167–76

Zola-Morgan S, Squire LR. 1993. Neuro-anatomy of memory. *Annu. Rev. Neurosci.* 16:547–63