Review

Uncovering the visual "alphabet": Advances in our understanding of object perception

Leslie G. Ungerleider *, Andrew H. Bell

Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA

ABSTRACT

The ability to rapidly and accurately recognize visual stimuli represents a significant computational challenge. Yet, despite such complexity, the primate brain manages this task effortlessly. How it does so remains largely a mystery. The study of visual perception and object recognition was once limited to investigations of brain-damaged individuals or lesion experiments in animals. However, in the last 25 years, new methodologies, such as functional neuroimaging and advances in electrophysiological approaches, have provided scientists with the opportunity to examine this problem from new perspectives. This review highlights how some of these recent technological advances have contributed to the study of visual processing and where we now stand with respect to our understanding of neural mechanisms underlying object recognition.

Published by Elsevier Ltd.

1. Introduction

In many ways, the neural processes associated with object recognition are not unlike those associated with language. Just as words are formed from a combination of letters, visual objects are formed from a combination of individual features, such as lines or textures. By themselves, these features reveal little regarding the identity of the specific object to which they belong, much like individual letters do not convey the meaning of a word. The features must be combined in specific ways, into "syllables" and "words", to create a unified percept of the object of interest. Moreover, this must be done irrespective of changes in stimulus orientation, illumination, and position to yield object invariance – the ability to recognize an object under changing conditions. Yet despite this complexity, the brain accomplishes the task effortlessly. How it does so remains largely a mystery.

Visual processing is thought to consist of two stages. The first stage involves transforming the visual stimulus into neural impulses that are transmitted via the retina and the lateral geniculate nucleus to the primary visual cortex (V1), where the process of analyzing the individual features begins. From V1, the visual information is distributed to a number of extrastriate visual areas, including areas V2, V3, V4, and MT, which process shape, color, motion, and other visual features. These features are then combined to create a complete representation of the image under

scrutiny. The second stage involves binding this representation with its categorical (i.e., recognizing the collection of features as a face, a car, etc.) and ultimately its specific identity (i.e., mother, father, Ford Taurus, etc.). Historically, these two stages were believed to occur simultaneously, dependent on the same neural mechanism (i.e., seeing is recognizing). However, with the discovery of perceptual deficits in recognition in the absence of any significant loss of vision (i.e., visual agnosias; see below), it has become clear that they are separable and likely mediated by different anatomical substrates.

Lesion studies in patients have revealed a number of regions in the occipital and temporal cortices that are intimately linked to visual processing. Subsequent electrophysiological studies have uncovered individual neurons that selectively respond to a variety of visual stimuli, ranging from isolated bars of light to intact faces. Despite these advances, it is still unclear how object recognition takes place. How are visual features, such as shapes, colors and textures encoded by individual neurons (i.e., what is the visual alphabet)? How are these features combined to produce complete representations of complex objects (i.e., visual "words")? And finally, how are these representations anatomically organized (i.e., what is the visual "dictionary")?

Over the last 25 years, there have been a number of technological advances that have led to major strides in our understanding of the role of the inferior temporal (IT) cortex in visual processing and object recognition. This review will discuss how these technological advances have contributed to the study of visual processing and where we now stand with respect to our understanding of object recognition. We begin with a brief overview of the early studies that identified IT cortex as the center of object recognition in the

* Corresponding author. Address: Laboratory of Brain and Cognition, National Institute of Mental Health, 10 Center Drive, 4C104, Bethesda, MD 20892, USA. Fax:+1 301 402 0921.

E-mail address: ungerlel@mail.nih.gov (L.G. Ungerleider).

brain. We next discuss how advances in optical imaging and physiological recording have revealed much about the complex properties of IT neurons. We then highlight how functional imaging revolutionized the study of cognitive neuroscience, and visual cognition in particular. Finally, we conclude by discussing how, in the last decade, efforts to relate findings from human and non-human primates have led to more integrated theories regarding the processing of faces, which are believed to represent a special category of visual stimuli.

2. Historical overview

2.1. Deficits in perception – "Seeing" is not "Recognizing"

Despite a long history of research into brain function, it is only relatively recently that different cognitive functions have been attributed to specific locations within the brain. Prior to the 19th century, it was believed that cognitive functions were equally distributed throughout the cerebral hemispheres, subcortical structures, and even the ventricles (see Gross (1999) for review). The first widely acknowledged theory to attribute specific functions to certain brain areas was the field of phrenology, pioneered by Gall, Spurzheim, and others (Gall & Spurzheim, 1809, 1810). Later, more objective examples of localized cortical function began to emerge. Paul Broca, a French pathology professor, noted that damage to the left frontal cortex was often associated with speech impairments (Broca, 1861). John Hughlings Jackson proposed the existence of a motor area based on his studies of epileptic patients (Jackson, 1870). Such demonstrations led to the establishment of cerebral localization as the standard model of brain function, and prompted the hunt for other examples of specialized regions within the brain.

One of the more active areas of research at this time was the search for visual centers of the brain. During the late 19th and early 20th century, a number of experiments were conducted that ultimately confirmed the occipital cortex as the primary visual center in primates. However, these experiments also highlighted an important dissociation in visual processing between "seeing" a visual stimulus and "recognizing" it. For example, Hermann Munk (Munk, 1881) and Brown and Schafer (1888) independently found that monkeys and dogs with temporal lobe lesions would ignore food and water, even if hungry and thirsty. However, they could still navigate rooms and avoid obstacles placed in front of them, indicating that their visual acuity was, at least partially, intact. Munk referred to this syndrome as "psychic blindness" to describe the apparent inability of the animal to understand the meaning of the visual stimuli presented. This was in contrast to "cortical blindness", which referred to the complete loss of vision that followed extensive occipital lobe lesions.

Psychic blindness, later renamed "visual agnosia", has since been demonstrated in humans who suffered damage to the occipital and temporal cortices (see Farah (2004) for review). Two separate forms of visual agnosia have been characterized. Apperceptive agnosia, which occurs after diffuse damage to the occipital and temporal cortices, affects both early and late visual processing. Patients suffering from apperceptive agnosia are impaired at simple discrimination tasks, have difficulty copying pictures, and cannot name or categorize objects. Patients suffering from associative agnosia, on the other hand, can successfully identify and copy shapes. However, their ability to associate any meaning with the images is severely impaired. Associative agnosia often involves damage to the inferior portions of the posterior cortex, the ventral occipital cortex and/or the lingual and fusiform gyri in the posterior temporal lobe.

The detailed studies of perceptual deficits in patients suffering from associative agnosia foreshadowed much of what we now know about higher-order visual processing and the cortical mechanisms associated with object recognition. First, it was shown that the inability of patients to categorize or identify a given stimulus was unaffected by changes in viewpoint, size, orientation, or any other "low-level" manipulation (e.g., McCarthy & Warrington, 1986). This finding confirmed that the disorder was not associated with an impairment in visual acuity per se, but in linking the physical features of a stimulus with its meaning. In turn, this implied that the brain areas associated with object recognition should be equally insensitive to such changes (i.e., show object invariance). Second, it was shown that perceptual deficits could be restricted to specific visual categories. Joachim Bodamer (1947) reported findings from several cases of patients who lacked the ability to recognize familiar faces, which he termed prosopagnosia. In cases of pure prosopagnosia, the ability to classify and categorize non-face stimuli is unimpaired (see Damasio, Damasio, and Van Hoesen (1982) for review). In other instances, a patient may be more impaired for certain categories compared to others (e.g., a greater impairment for images of living objects such as plants and animals compared to images of inanimate objects; Warrington & Shallice, 1984). The selective nature of these deficits suggests not only that the areas responsible for "seeing" a visual stimulus are distinct from those responsible for "recognizing" it, but that the anatomical substrates for object recognition might be arranged according to categorical distinctions.

2.2. Identifying the contribution of the inferior temporal cortex to object recognition

While the involvement of the occipital cortex in vision was firmly established by the early 20th century, the role of the temporal cortex in visual perception was not fully appreciated until several decades later. In 1938, Heinrich Klüver and Paul Bucy conducted a series of lesion experiments in monkeys that produced a characteristic set of behavioral symptoms (see Klüver, 1948; Klüver & Bucy, 1938). They noted that monkeys with bilateral removal of the temporal lobe showed not only decreased emotional reactivity and increased sexual activity, but also a significant impairment in their ability to recognize and/or discriminate visual stimuli. This collection of symptoms later became known as the Klüver–Bucy syndrome. However, these changes in visual ability were not accompanied by any appreciable loss of vision. It was later revealed that the visual deficits of the Klüver–Bucy syndrome were due to selective damage of the inferior temporal neocortex (Blum, Chow, & Pribram, 1950; Chow, 1961; Iwai & Mishkin, 1969; Mishkin, 1954; Mishkin & Pribram, 1954). Damage to medially adjacent temporal-lobe regions (i.e., the amygdala) failed to produce visual deficits, but instead led to the social and sexual abnormalities.

The inferior temporal (IT) cortex of monkeys was further divided into architectonically distinct regions, most notably area TE, located anteriorly along the inferior temporal gyrus, and area TEO, located posterior to area TE near the posterior middle temporal sulcus (Fig. 1) (Boussaoud, Desimone, & Ungerleider, 1991; Seltzer & Pandya, 1978; von Bonin & Bailey, 1947). Later work showed that damage to area TE resulted in greater deficits in visual recognition whereas damage to area TEO resulted in greater deficits in visual discrimination (Cowey & Gross, 1970; Ettlinger, Iwai, Mishkin, & Rosvold, 1968; Iwai & Mishkin, 1968).

It was known at this time that more posterior visual areas, such as striate (V1) and extrastriate visual areas, receive visual inputs from several subcortical structures such as the pulvinar and superior colliculus (Felleman & Van Essen, 1991). However, lesions of the pulvinar and superior colliculus repeatedly failed to produce deficits in visual discrimination (Chow, 1951; Ungerleider & Pribram, 1977), as one might expect if the region involved in visual discrimination, namely IT cortex, relied on inputs from these
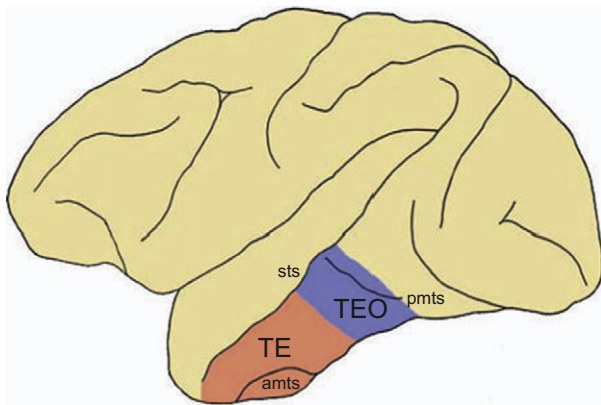
Adapted from Webster et al., 1991

Fig. 1. Lateral view of the monkey brain showing the two architectonic subdivisions of the inferior temporal cortex: area TE and area TEO. amts: anterior middle temporal sulcus; pmts: posterior middle temporal sulcus; sts: superior temporal sulcus (from Webster, Ungerleider, & Bachevalier, 1991).

subcortical structures. In 1975, Rocha-Miranda and colleagues confirmed that activation of IT cortex depends exclusively on inputs received from corticocortical connections originating in V1 (Desimone, Fleming, & Gross, 1980; Rockland & Pandya, 1979; Zeki, 1971). They found that bilateral removal of V1 obliterated all visual responses in IT neurons. However, unilateral lesions of V1 or transection of the forebrain commissures resulted in IT neurons that only responded to stimuli in the contralateral visual field. The behavioral significance of this corticocortical pathway could be seen in the crossed-lesion experiments conducted by Mishkin and colleagues in the 1960s (see Mishkin, 1966). They combined V1 lesions in one hemisphere with lesions of IT cortex in the opposite hemisphere and found no visual impairment. Presumably, visual inputs from V1 of the intact hemisphere could still reach the contralateral IT cortex via the forebrain commissures. However, if the forebrain commissures were cut, thus completely isolating the intact IT cortex from its V1 input, the animals would exhibit profound visual deficits. One would not expect such a deficit if IT cortex could still obtain sufficient visual information via subcortical pathways.

2.3. Neurons in IT cortex are selective for complex stimuli

In 1969, Gross and colleagues reported on the first electrophysiological recordings from individual neurons in IT cortex. They found that the majority of neurons in IT cortex (approximately 80%) are visually responsive, displaying both excitatory and suppressed responses to visual stimuli (Gross, Bender, & Rocha-Miranda, 1969; Gross, Rocha-Miranda, & Bender, 1972). Neurons in IT cortex were found to have very large receptive fields relative to earlier visual areas, often spanning both hemifields and almost always including the center of gaze (Gross et al., 1972). However, what was perhaps most noteworthy about neurons in IT cortex is their strong selectivity for complex shapes, including faces and objects (e.g., Fig. 2) (Bruce, Desimone, & Gross, 1981; Desimone, Albright, Gross, & Bruce, 1984; Gross et al., 1972; Kobatake & Tanaka, 1994; Perrett, Rolls, & Caan, 1982; Rolls, Judge, & Sanghera, 1977; Tanaka, Saito, Fukada, & Moriya, 1991). Moreover, this selectivity was shown to be insensitive to changes in size and location (Desimone et al., 1984), as predicted by the earlier neuropsychological evidence. The majority of these early electrophysiological studies were done in anesthetized animals, demonstrating that the unique sensitivity of IT neurons to complex shapes is not a

product of top-down or attentional factors but is intrinsic to the neurons themselves.

The discovery of neurons selective for specific objects in IT cortex had profound repercussions on theories concerning object recognition. Some years earlier, Konorski (1967) had postulated the existence of "gnostic" units, which are, in theory, individual neurons that signal the presence of a particular object. Such a neuron would presumably code for all iterations of this object – across changes in size, viewpoint, location, etc. The concept of gnostic units was later reworked into the "The Grandmother Cell Hypothesis", a term first credited to Jerry Lettvin (see Gross, 2002). If object recognition proceeds along a strictly hierarchical pathway in which selectivity among neurons gets progressively more complex, one should theoretically be able to locate a particular neuron or small population of neurons that responds selectivity to a particular object or person, such as one's grandmother. Taking this one step further, if we were to destroy this neuron (or population of neurons), we would lose the ability to recognize said grandmother.

2.4. Summary

The progressive accumulation of information regarding the properties of visual cortex ultimately led to the hypothesis of a visual recognition pathway, stretching from V1 to area TE (the ventral stream or "what" pathway; Ungerleider & Mishkin, 1982). A second visual pathway, stretching from V1 to parietal cortex was postulated to process visuospatial information (the dorsal stream or "where" pathway). As one moves along both pathways, there is a progressive loss of retinotopic information and a progressive increase in convergence among neuronal projections (Shiwa, 1987; Weller & Kaas, 1987; Weller & Steele, 1992). For the ventral visual pathway, this could account for the insensitivity of IT neurons to spatial location, size, and viewpoint: invariance necessary for object recognition.

One assumption based on this evidence is that object recognition is a serial process that proceeds through a chain of hierarchically organized areas beginning with early visual areas responsible for analyzing simple visual features (e.g., lines, color, etc.), to highly selective areas that respond to individual categorical exemplars (e.g., faces). Neurons within these end-stage areas (i.e., a "gnostic" or "grandmother" neuron) would presumably encode a complete representation of a given object. It was eventually pointed out, however, that if objects were indeed encoded by individual neurons, there would be an insufficient number of neurons to encode *all possible* objects, and thus the concept of such a sparse-encoding scheme fell out of favor. Nonetheless, it did leave the community with two critical, and as of yet, unanswered questions. First, how are objects represented, if not by "grandmother" or "gnostic" neurons? And second, what is the relationship of such object representations to visual agnosias? In the remainder of this review, we will highlight several research avenues followed over the last 25 years that have attempted to address these fundamental questions.

3. Building object representations in the inferior temporal cortex

3.1. "Combination coding": representing complex objects through combinations of simpler elements

The 1980s and 1990s saw a number of electrophysiological studies attempting to characterize the basic response properties of IT neurons and their relationship to object recognition. What rapidly became clear was that IT neurons show a wide range of stimulus preferences that do not immediately lend themselves to an obvious organizational scheme. Unlike neurons in earlier visual

Example 1
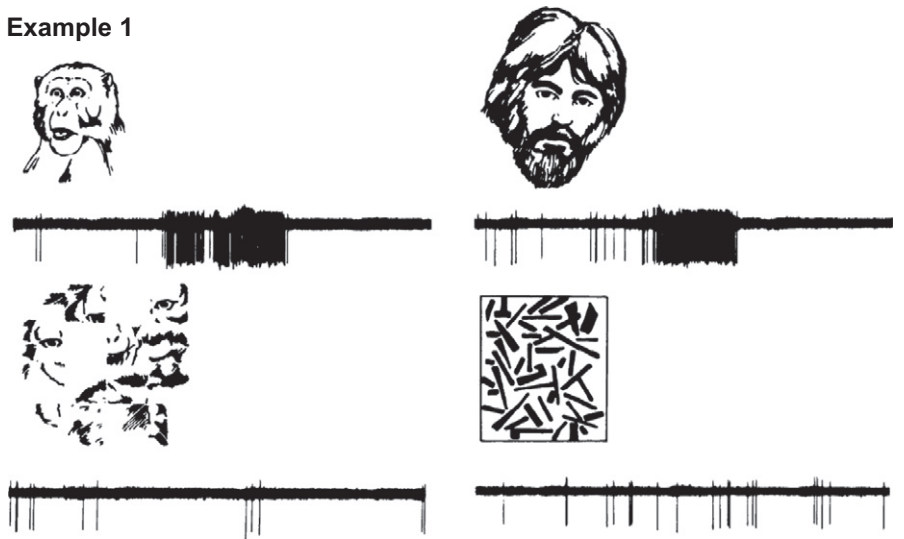
Adapted from Bruce et al., 1981
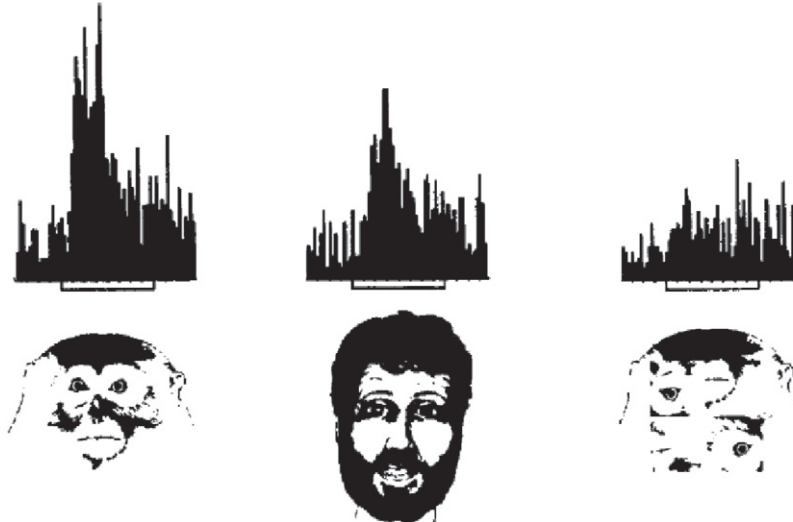
Example 2

Adapted from Desimone et al., 1984

Fig. 2. Two examples of face-selective neurons recorded in the temporal cortex of anaesthetized monkeys. Example 1 was recorded from the superior bank of the superior temporal sulcus (from Bruce et al., 1981). Example 2 was recorded from the inferior temporal cortex (from Desimone et al., 1984). Both neurons responded vigorously to images of human and monkey faces. However, their activity decreased significantly when presented with scrambled versions of those images, or other non-face stimuli.

areas such as V1, which are organized retinotopically, or V2, which are organized into modules according to functional properties (e.g., color), neurons in IT cortex did not, at first, seem to be organized according to any recognizable scheme.

In 1984, Desimone and colleagues performed one of the first systematic studies of the neuronal properties of IT neurons. They compared the responses of IT neurons in anesthetized monkeys to a wide array of stimuli, ranging from simple shapes, such as bars and circles, to complex stimuli, such as faces and snakes. They found that while a small subset of neurons responded selectively to faces and hands, the majority responded to a variety of different simple stimuli. In the latter case, they attempted to identify the critical features of the stimulus that were responsible for driving individual neurons by presenting cutouts or otherwise altered versions of the stimuli (see also Schwartz, Desimone, Albright, & Gross, 1983). They found that the majority of visually responsive neurons required complex shapes, textures, colors or some combination thereof to be maximally activated.

Several years later, Tanaka and colleagues developed a more sophisticated stimulus reduction algorithm that allowed them to better characterize the critical features that activated each sampled neuron (Fujita, Tanaka, Ito, & Cheng, 1992; Ito, Tamura, Fujita, & Tanaka, 1995; Kobatake & Tanaka, 1994; Tanaka et al., 1991). They found that neurons in the posterior one-third of IT cortex (area TEO) could be maximally activated by bars or circles, whereas those in the anterior two-thirds of IT cortex (area TE) required more complex arrangements of features to elicit a maximal response (e.g., Fig. 3). These critical features were often small groups of elements derived from intact real-world objects. Tanaka and colleagues proposed that objects might be represented by a small number of neurons, each sensitive to the various features of that object ("combination coding"; Fig. 4). The benefit of combination over sparse coding is that the former is much more flexible, allowing for limitless representations through the integration of the responses from different sets of feature-selective neurons.

Reprinted from Tanaka, 1996
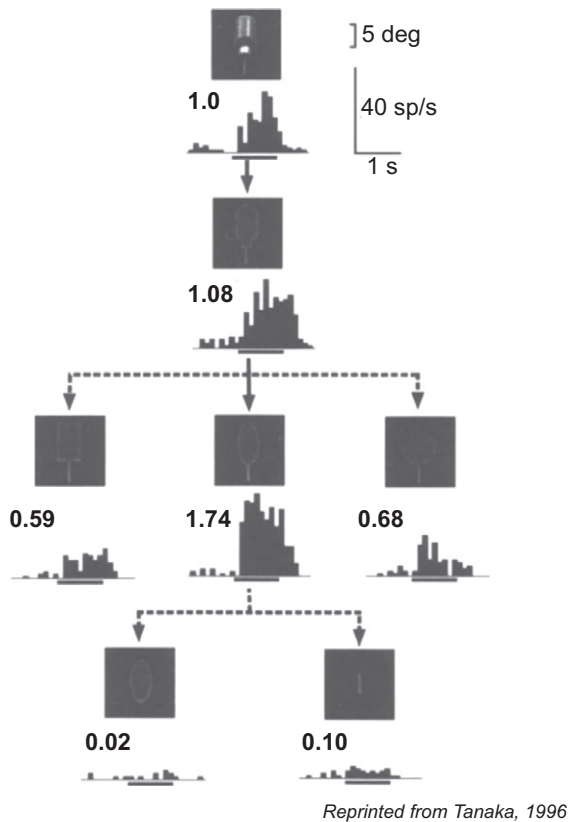
Fig. 3. Systematic reduction of real-world object reveals feature selectivity of IT neurons. Responses of an IT neuron to intact (top panel) and systematically reduced versions of a water bottle. Inset numbers above the histograms represent the normalized response magnitudes relative to the response to the intact image. Stimulus presentation window is indicated by horizontal line below each histogram (from Tanaka, 1996).
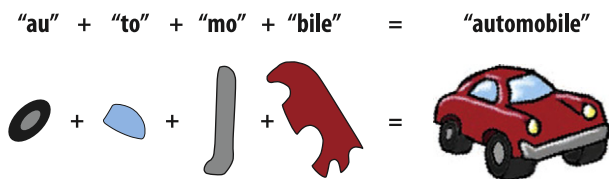
Fig. 4. Real-world objects represented by combinations of different features. Similar to how individual syllables are combined to create words, "combination coding" suggests that objects are represented by neurons (or small populations of neurons) each coding for the different complex features that comprise the object (see Tanaka (1996) for review).

Around the same time, psychologists were developing similar "object-based" models of object recognition, which were all, more or less, based on the principle of simpler shapes being combined to produce complex representations. Perhaps the most well known of these is the theory of "Recognition-by-Components" (RBC), first proposed by Biederman (1987). Similar to combination coding, RBC proposed that objects are represented through various combinations of more primitive shapes. What was particularly intriguing about RBC is that it constrained these primitives to a set of no more than 36 geometric icons or "geons" (Fig. 5A). These geons were thought to be distinguished from one another on the basis of five properties: curvature, collinearity, symmetry, parallelism, and cotermination.

Object-based models of object recognition make certain predictions about what one might expect to find at the neuronal level. For example, object-based models predict the existence of neurons
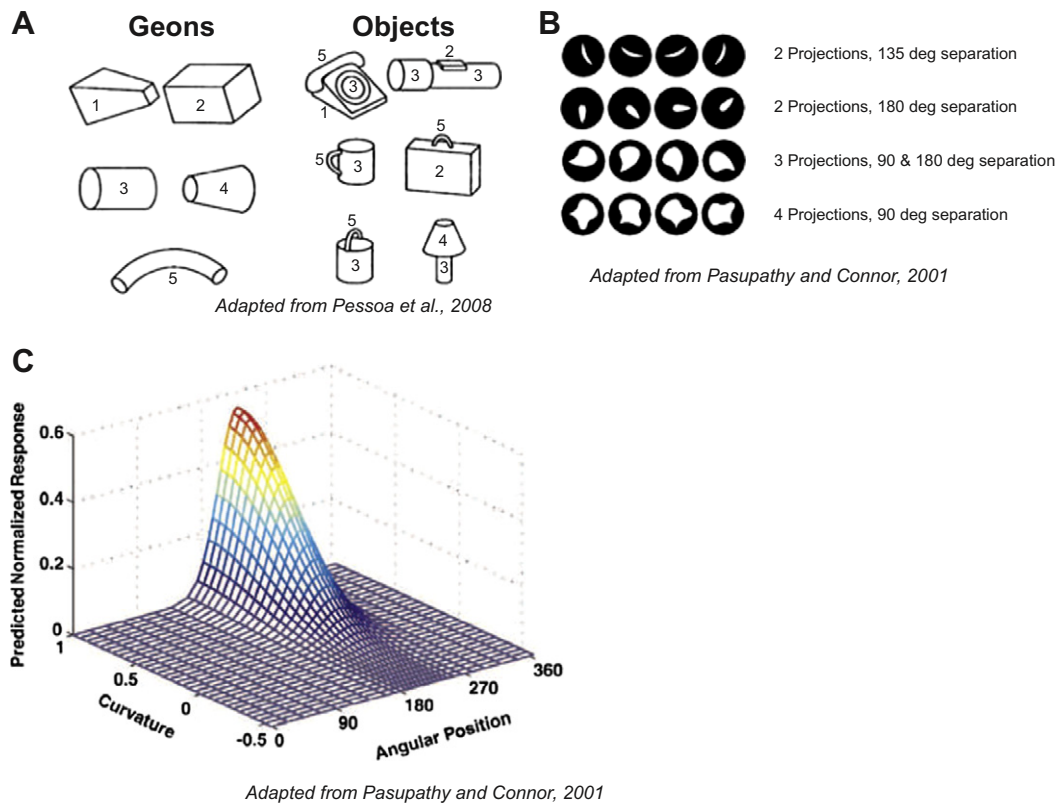
sensitive to intermediate shapes (e.g., geons). Tanaka and colleagues were among the first to provide direct neurophysiological evidence for object-based models of object recognition – specifically, neurons that respond selectively and most strongly to decomposed versions of real-world objects. This was followed up by an extensive examination of an intermediate stage along the visual pathway by Connor and colleagues, which provided further support for object-based models.

Pasupathy and Connor (1999, 2001, 2002) measured the responses of neurons in area V4 to a large set of contoured shapes, each with a different number of convex projections (2–4), different orientations, and degrees of curvature (Fig. 5B). They found that neurons in V4 responded most strongly to single contours or combinations of contours, arranged in particular configurations (e.g., acute convexity located in the lower right of the stimulus, acute convexity immediately adjacent to a shallow concavity, etc.; see example in Fig. 5C). Notably, this tuning remained constant when the particular configuration was placed within a variety of more complex shapes, thus illustrating how a single neuron might participate in the coding of many different complex objects.

Later, Brincat and Connor (2004, 2006) reported similar findings in posterior IT cortex (i.e., area TEO). They further found that some neurons within this region encoded shape information at multiple levels: an initial visual response that represented the simple contour fragments and a second, later response that differed according to how the various shape fragments were combined. This biphasic property is very similar to that found among neurons in anterior IT cortex as described by Sugase and colleagues (1999) and others (e.g., Tamura & Tanaka, 2001; see below). Brincat and Connor inferred that the second phase of the response represented the outcome of a recursive network process responsible for constructing selective responses of increased complexity.

These data suggest that neurons in V4 and posterior IT cortex might serve to identify and discriminate between the primitive "geons" described in Biederman's RBC theory. At the very least, these data show how neuronal selectivity progresses from simple line segments (in V1) to simple curves (in V2), to complex curves or combination of curves (in V4 and posterior IT cortex), and finally to complex shapes such as faces and hands (in anterior IT cortex).

Object-based models are considered to be "view-independent" because they rely on hard-coded representations of geons (or similar primitives) encompassing all possible orientations. As such, the orientation of the object being examined is irrelevant provided the various components can be recognized and distinguished from one another. An alternative to such models are the so-called "view-dependent" or "image-based" models of object recognition (Logothetis & Sheinberg, 1996; Tarr & Bulthoff, 1998). These models propose that objects are represented by multiple independent views of that object. Immediate criticisms of these image-based (view-dependent) models include the unlikelihood that we could store all possible viewpoints for all objects in our environment (i.e., "combinatorial explosion") or that we are so limited in our recognition abilities that we can only recognize a given object if we have previously seen it in that precise configuration. Therefore, to overcome these criticisms, image-based models assume that the brain is somehow able to extrapolate a more complete representation of a given object from an incomplete collection of experienced viewpoints. Despite this assumption, such models do predict that subjects should be faster at recognizing objects when presented in familiar vs. unfamiliar viewpoints – which is indeed the case (e.g., Humphrey & Khan, 1992). Bulthoff and Edelman (1992) evaluated the ability of human subjects to correctly identify previously unfamiliar objects when presented in unfamiliar orientations. Subjects were first presented a series of 2D images of an unfamiliar object from a limited number of viewpoints. Later, subjects were presented with new images of the same object from different

A Geons Objects

Adapted from Pessoa et al., 2008

B

2 Projections, 135 deg separation

2 Projections, 180 deg separation

3 Projections, 90 & 180 deg separation

4 Projections, 90 deg separation

Adapted from Pasupathy and Connor, 2001

C

Adapted from Pasupathy and Connor, 2001

Fig. 5. Object-based (view-independent) models of object recognition. (A) Examples of geometric icons ("geons") as originally proposed by Biederman (1987), and how they could be combined to produce representations of real-world objects. (B) Stimuli used by Pasupathy and Connor (2001) to investigate tuning properties of V4 neurons. Stimuli contained between 2 and 4 projections, separated by 90–180°, presented at various orientations. (C) Example of V4 neuron exhibiting tuning to a specific feature composed of a particular curvature and angular position.

perspectives. The ability of subjects to correctly recognize the objects systematically decreased with greater deviations from the familiar viewpoints.

Image-based models also predict that neuronal responses among IT neurons should show evidence of view-dependence, which has again been demonstrated. Using a paradigm similar to that of Bulthoff and Edelman (1992), Logothetis and colleagues demonstrated that monkeys were better at recognizing the objects when presented in a previously presented viewpoint (as compared to a novel viewpoint) (Logothetis, Pauls, Bulthoff, & Poggio, 1994). When they sampled activity from neurons in IT cortex while monkeys performed this task, they found that neurons in IT cortex responsive to the objects were viewpoint dependent, responding most strongly to one viewpoint and decreasing activity systematically as the object was rotated away from its preferred orientation (Logothetis, Pauls, & Poggio, 1995). They proposed that a view-invariant representation of a given object might be generated by a small population of *view-dependent* neurons that can, through changes in their individual firing patterns, signal the presence of a given object as well as extrapolate its current orientation.

Evidence for both object-based and image-based models of object recognition continues to accumulate, leading most to believe that the underlying mechanism incorporates aspects of both. Exemplar level recognition (e.g., discriminating between a Toyota and a Ford) is addressed particularly well by image-based models, which are essentially based on storing multiple snapshots of specific exemplars. On the other hand, how these snapshots are grouped according to semantic relationships to perform categorical distinctions (e.g., cars from chairs) is unclear. For that matter, it is unclear how object-based models construct categorical groupings, highlighting one of the many challenges in understanding object recognition.

3.2. IT neurons are organized into columns according to feature selectivity

The studies cited above, and others like them, detailed the functional role of individual IT neurons in object recognition. Around the same time, evidence began to emerge that described how neurons in IT cortex might be spatially organized. Early experiments showed that neuronal activity within IT cortex is better correlated among pairs of neighboring neurons (within 100 μm) and weakens significantly when the two neurons are separated by more than 250 μm (Gochin, Miller, Gross, & Gerstein, 1991). Similarly, neurons along the same recording track perpendicular to the cortical surface exhibit similar preferences for visual features (Fujita et al., 1992). These findings suggested that neurons might be clustered according to their stimulus preferences.

Critical evidence for a clustered arrangement of IT neurons came from a landmark study that used optical imaging to investigate the preferences of neurons in area TE. Optical imaging measures differences in light absorption on the cortical surface based on the ratio between deoxygenated to oxygenated blood. The extent to which deoxygenated compared to oxygenated blood absorbs light varies as a function of wavelength. If a particular region of cortex is more active than another, a change will occur in the ratio of deoxygenated vs. oxygenated blood, thereby causing a change in light absorption that can be measured using sensitive cameras (Grinvald, Lieke, Frostig, Gilbert, & Wiesel, 1986). Wang and colleagues (1996) used this technology to show that neurons in area TE are arranged into patches, spanning approximately 500 μm, which are spatially distributed according to their selectivity for complex features. By combining this approach with electrophysiological recordings, Tsunoda and colleagues (2001) demonstrated that as real-world objects are systematically reduced into their

component elements, the patches (and corresponding neurons) that would normally respond to the elements removed are effectively silenced.

Studies such as these have shown us how responses from small populations of neurons, each sensitive to a particular set of features that might belong to many different objects, can be combined to produce a complete representation of a given object. Returning to the analogy of language, it is as if individual neurons in IT cortex code for individual syllables that, when combined, spell out different words – the number of neurons (or populations of neurons) involved being a function of the complexity of the word/object. Further, these neurons are organized not according to their receptive fields, as in earlier visual areas, but according to the similarity of their preferred features.

3.3. Dynamic properties of IT neurons

In the preceding section, we outlined how electrophysiological studies combined with optical imaging revealed a potential mechanism through which objects might be encoded by small populations of IT neurons. The sophisticated stimulus selectivity of IT neurons follows a logical progression of increasing complexity originating with simple neurons in V1. However, the majority of the aforementioned studies were performed in anaesthetized animals, which ignores any potential impact of behavioral factors on the neural responses. Once investigators began to conduct studies in awake, behaving animals, it was immediately clear that IT neurons were more flexible in their response properties than had been originally shown. In the following section, we describe a few of the many properties of IT neurons that illustrate how their responses reflect far more than the visual features presented to the retina.

Presenting a novel object to a subject naturally captures attention. New stimuli in the environment could represent a potential threat, or a new potential for reward. In either case, a new stimulus warrants further investigation. One way in which IT neurons denote familiarity is through repetition suppression, a phenomenon whereby response magnitudes decrease over repeated presentations of the same stimulus. Miller and colleagues (1991) trained monkeys to perform a delayed match-to-sample task in which a sample image was first presented followed by a series of test images, one after the other. The monkey was required to release a bar if the test image matched the original sample. The sample and the match could be separated by as many as four intervening non-match distractor stimuli, thus requiring the monkey to hold the identity of the sample image in working memory throughout the duration of the trial. It was found that responses to the matching test image were attenuated compared to that to the initial sample. Thus, when the same stimulus was presented more than once and became increasingly familiar to the animal, the responses to that stimulus decreased (i.e., "repetition suppression"). This same effect was later shown to persist over the course of an entire recording session (>1 h), after hundreds of intervening stimuli (Li, Miller, & Desimone, 1993). In essence, repetition suppression enables IT neurons to act as filters for novel stimuli. Repetition suppression is found throughout the brain (e.g., ventral temporal cortex, Fahy, Riches, & Brown, 1993; Riches, Wilson, & Brown, 1991; prefrontal cortex, Dobbins, Schnyer, Verfaellie, & Schacter, 2004; Mayo & Sommer, 2008; parietal cortex, Lehky & Sereno, 2007) and is believed to play a critical role in behavioral priming (McMahon & Olson, 2007; Wiggs & Martin, 1998).

Another example of the flexible nature of IT neurons is how they can adjust their tuning profiles and selectivity according to the task at hand. For example, through associative learning, it is possible to dynamically "retune" IT neurons to respond to additional stimuli. Sakai and Miyashita (1991) trained monkeys to memorize pairs of stimuli by first presenting them with a cue

stimulus followed shortly by two test stimuli. The monkeys were rewarded for selecting one of the latter two stimuli that made up an arbitrary pair with the cue stimulus. Initially, neurons in IT cortex responded only to one stimulus of the pair. However, over the course of the experiment, as the monkeys learned to associate one stimulus with another, a subset of neurons responded to both elements of the pair. The IT neurons had "learned" to signal the presence of either element of the pair. Similarly, Sigala and Logothetis (2002) found that IT neurons could shift their tuning according to the demands of the current behavioral task. They trained monkeys to discriminate images of fish and faces based on a limited set of diagnostic features (e.g., fin shape, nose length). They found that IT neurons became tuned to the specific features necessary to form these discriminations (Fig. 6).

These studies are a few of many showing how activity among IT neurons reflects more than the presence or absence of certain visual features in the external environment. Activity in IT cortex has been shown to be affected not only by stimulus novelty and experience, associative learning, and task demands but also by spatial attention (e.g., Buffalo, Bertini, Ungerleider, & Desimone, 2005; Moran & Desimone, 1985) and stimulus value (e.g., Mogami & Tanaka, 2006). This evidence tells us that IT neurons are not hard-wired to respond to a particular set of features, but can adapt dynamically to the current behavioral objectives, making them ideally suited as the neural substrate for object recognition.

3.4. Generating invariant representations through experience

Above, we described how responses among IT neurons decrease as stimuli become more familiar to the subject (i.e., repetition
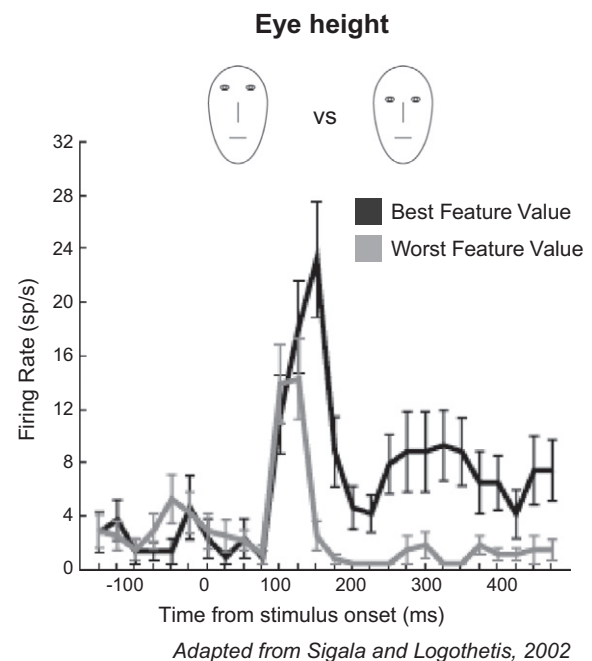
Fig. 6. Example of an individual IT neuron whose selectivity was shaped by current task demands. Monkeys were trained to discriminate between different face caricatures on the basis of four diagnostic features: eye height, eye separation, nose length, and mouth height. The black trace indicates the average response to the best feature value (i.e., the exemplar that elicited the strongest response). The gray trace indicates the average response for the worst feature value (i.e., the exemplar that elicited the weakest response). When the monkey was required to discriminate faces on the basis of eye height, for example, the selectivity of the neuron was shaped according to the appropriate feature (from Sigala & Logothetis, 2002).

suppression). It has been argued that experience might also be responsible for one of the hallmark characteristics of IT neurons, position invariance. Position invariance refers to the ability of IT neurons to maintain their stimulus selectivity regardless of where the stimulus appears in the visual field. For example, if an IT neuron shows a preference for a particular stimulus (e.g., an image of a face), that preference will be maintained regardless of where the face appears in the visual field or how large the retinal image is. This would be true regardless of the absolute magnitude of the responses. That is to say while the response to a preferred stimulus might be 50 spikes/s when presented to one location and 30 spikes/s when presented to another; the response to a non-preferred stimulus will always be less in both locations (e.g., 35 and 15 spikes/s, respectively). Expanding on this point, if a given neuron prefers faces to houses, and houses to fruit, as long as this ordering of stimulus preference is maintained, the absolute magnitude of the individual responses (which might change in response to changes in size, position, etc.) to the different stimuli is irrelevant (Li, Cox, Zoccolan, & DiCarlo, 2009). This property, possibly unique to IT cortex, is extremely advantageous to object recognition because it allows neurons to signal the presence of a particular object regardless of the particular viewing conditions.

Li and DiCarlo (2008) proposed that position invariant representations could be constructed based on the assumption that temporally contiguous shifts of stimuli correspond to the same object. In other words, if a given set of features is, at one moment, 10° to your left and you generate a saccade 10° to the right, then the set of features that are now located 20° to your left likely belong to the same object. They argued that over many exposures, IT neurons could learn to associate a given set of features in one retinal location with the same set of features in another. As such, the same IT neuron would eventually respond to the same set of features, regardless of where they appear. Li and DiCarlo tested this theory by determining whether it would be possible to "fool" a neuron into responding to two different stimuli as if they were the same stimulus presented to two different locations (Fig. 7A). While monkeys fixated centrally, a preferred stimulus (i.e., a stimulus that produced a robust response in the neuron currently being recorded) was presented in the periphery. When the monkeys generated a saccade to that stimulus, its identity immediately changed to a different, non-preferred, stimulus. Initially, the two stimuli evoked different response magnitudes, which would effectively

allow neurons further downstream to discriminate between them. However, over repeated exposures, the difference in response magnitude between the preferred and the non-preferred stimuli (i.e., the measure of object selectivity) decreased systematically (Fig. 7B). One interpretation of this result is that, based on the new similarity between the magnitudes of the neuronal responses, the two stimuli were "seen" as the same object by neurons further downstream: the ability of such downstream neurons to discriminate between the two stimuli had been abolished. In the case of this particular experiment, the effect was limited to one particular location but, in a natural setting over many such exposures, this mechanism could result in a single neuron responding to the same stimulus, regardless of its position in the visual field. Presumably, a similar mechanism could explain not only position invariance but also size invariance. For example, as we approach or move away from a given object, the retinal image of that object will increase or decrease in size, thus recruiting different populations of neurons in earlier areas of visual cortex. However, the visual features will remain the same and, over time, neurons in IT cortex could learn to respond to the different retinal images as the same object. To our knowledge, this adaptability is unique to neurons within IT cortex, further emphasizing its particular relevance for object recognition.

3.5. Summary

Therefore, in contrast to the sparse-encoding scheme originally suggested by gnostic neurons (i.e., "Grandmother Cells"), it is now hypothesized that objects are represented by a population of neurons, each sensitive to different visual features of the stimulus, moving towards a more distributed encoding perspective of object representations in IT cortex. In essence, researchers had discovered the visual alphabet, which appears to be organized into syllables – syllables that do not necessarily convey meaning, but allow for great flexibility in the number of objects they can encode. Additional evidence of the flexibility of IT neurons is their ability to maintain their high degree of stimulus selectivity over different image sizes and positions (i.e., size/position invariance), to signal familiarity vs. novelty, and to adapt to current task demands. Interestingly, these characteristics are all remarkably reminiscent of the predications regarding the neural correlates of object recognition made on the basis of the early neuropsychological and behavioral
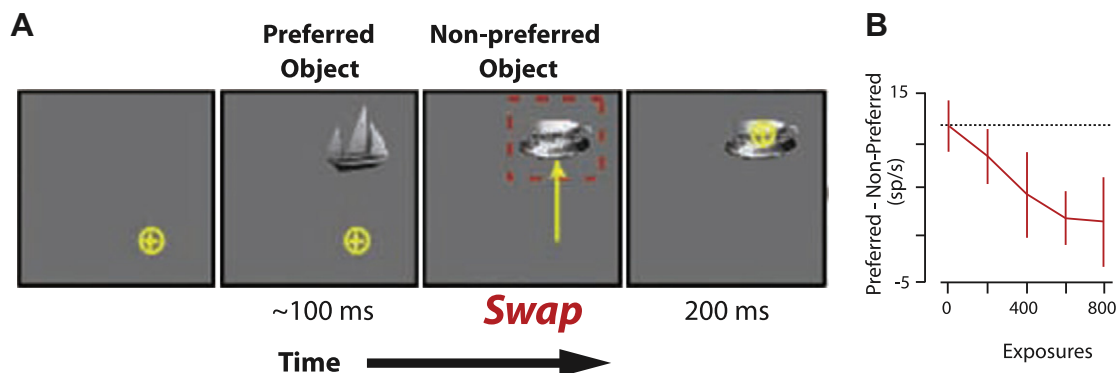
Adapted from Li and Dicarlo, 2008

Fig. 7. Example of how altered experience reduces selectivity of IT neurons to stimuli presented to different retinal positions. (A) While their fixation was stable, monkeys were presented with a preferred stimulus (i.e., a stimulus that evoked a strong response in the neuron currently being recorded). When the monkey initiated a saccade to that stimulus, its identity changed to that of a non-preferred stimulus (i.e., one that evoked a weak response). (B) Difference between the responses to preferred vs. non-preferred stimuli for 10 multiunit recording sites as a function of the number of swap exposures. Initially, the two stimuli evoked significantly different responses (i.e., strong object selectivity). However, as the number of swap exposures increased, the difference between the two stimuli decreased to the point where the neurons responded with approximately the same magnitude of response to the two different stimuli (i.e., weak object selectivity). In effect, the neurons responded as if the same stimulus was present in the two different locations. The effect was specific to the swap exposure location. This mechanism shows how experience can lead to position invariant responses within individual IT neurons (from Li & DiCarlo, 2008).

studies (see above). However, this model is not as easily reconciled with the neuropsychological observations in human patients with agnosias: How can we have a highly specific deficit for a particular category of stimuli, which might include all manner of visual features, if the neurons encoding those features are distributed throughout IT cortex? The answer to this question arose as a result of approaching the problem from a completely different perspective using an innovative new technology.

4. Functional neuroimaging investigations of visual processing

The use of functional neuroimaging has revolutionized the study of cognitive neuroscience. Prior to the availability of neuroimaging techniques, the only method for researchers to examine the correlation between brain structures and behavior in humans was to study brain-damaged individuals. Functional neuroimaging techniques, and functional magnetic resonance imaging (fMRI) in particular, provided researchers with the opportunity to examine brain function in healthy subjects as well as in patient populations using a non-invasive approach. Interestingly, fMRI does not measure neuronal activity *per se*, but instead the hemodynamic responses that occur as a result of changes in neuronal activity. Specifically, fMRI provides an indirect measure of neuronal activity by taking advantage of differences in the magnetic susceptibility between oxygenated and deoxygenated hemoglobin. As brain regions become more active, their metabolic demands increase and local oxygen delivery is enhanced. Oxygenated hemoglobin is diamagnetic and produces a strong magnetic signal as compared to deoxygenated hemoglobin, which is paramagnetic and produces a weaker signal. Using an MRI scanner, one can obtain a measure of the relative dependence on oxygenated blood within a given brain region (i.e., the "blood oxygen level dependent" or "BOLD" signal). Although it is generally accepted that the BOLD signal provides some indication of the neuronal activity in the underlying brain tissue, the precise physiological basis of the BOLD signal is still poorly understood.

4.1. A first glimpse at the intact human brain

By the time functional neuroimaging methods became available, much was known about the hierarchical organization and properties of the different visual regions in the monkey brain. We knew, for example, that a series of interconnected visual areas, many with independent retinotopic maps, were organized in a pathway stretching from V1 to area TE in the anterior temporal lobe (i.e., the ventral visual pathway; see above). Many of the early neuroimaging studies attempted to determine if humans showed a similar organization. Haxby and colleagues (1994) used positron emission tomography (PET) to demonstrate that humans, like monkeys, show two visual pathways: a dorsal pathway, stretching from V1 through extrastriate to parietal cortex, which is primarily concerned with visuospatial processing, and a ventral pathway, stretching from V1 through extrastriate to temporal cortex, which is primarily concerned with object processing. Functionally distinct regions in human cortex along both the dorsal and ventral pathways were delineated on the basis of retinotopic fMRI maps (e.g., Fig. 8) (DeYoe, Bandettini, Neitz, Miller, & Winans, 1994; Engel et al., 1994; Huk, Dougherty, & Heeger, 2002; Sereno et al., 1995; Tootell, Hadjikhani, Hall, et al., 1998; Tootell, Hadjikhani, Vanduffel, et al., 1998; Tootell, Mendola, Hadjikhani, Liu, & Dale, 1998; Tootell et al., 1997; Wade, Brewer, Rieger, & Wandell, 2002). In monkeys, we know these maps become increasingly less well defined as one moves anteriorly along the inferior temporal cortex (see above). Grill-Spector and colleagues (1998) used fMRI to show a similar trend in humans.
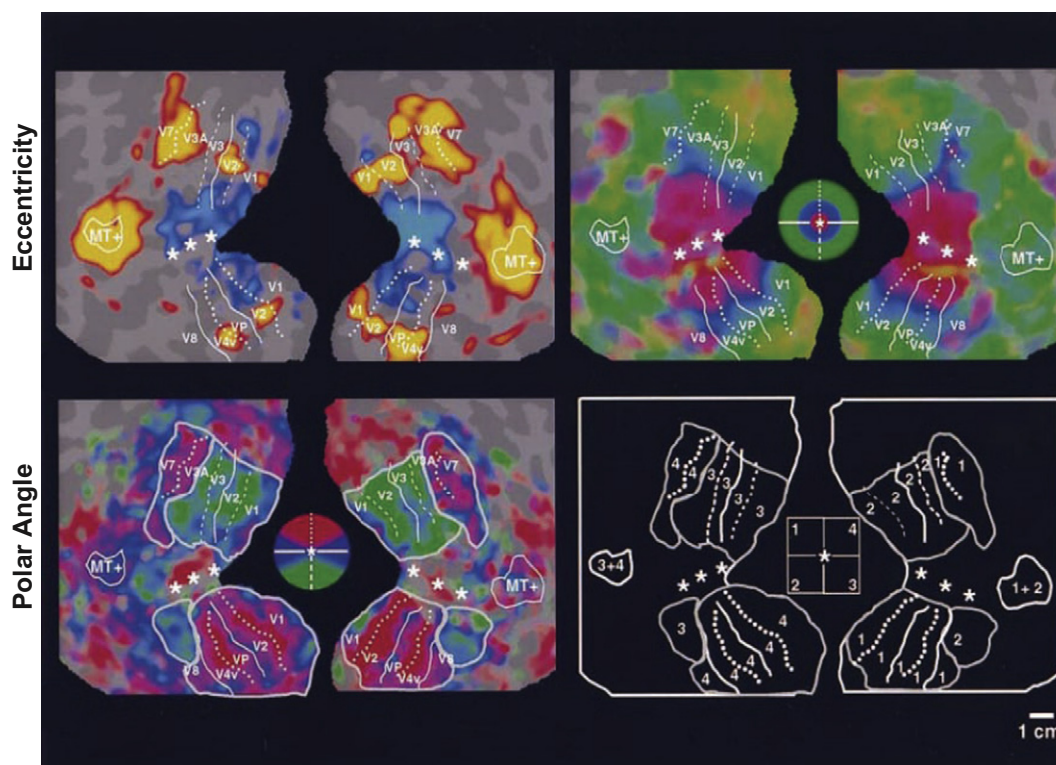
Reprinted from Tootell et al., 1998

Fig. 8. Retinotopic organization of visual cortex revealed by fMRI. Inflated and flattened views of the human brain showing eccentricity and polar angle maps, stretching from V1 to V4. Maps were produced using expanding and rotating checkerboard patterns, respectively (from Tootell et al., 1998).

In addition to its use as an anatomical tool, fMRI has also allowed researchers to identify regions involved in object recognition on the basis of their functional properties. According to the monkey literature, regions involved in object recognition not only show weak retinotopic organization and become increasingly more selective for complex objects, but also are increasingly less responsive to scrambled images (see above; Bruce et al., 1981; Vogels, 1999). Furthermore, they often show position and size invariance (i.e., the representations of individual objects remain constant regardless of their retinal position or size; see above; Ito et al., 1995). When the sensitivity of the ventral visual pathway in humans to scrambled images was assessed, the first stage at which regions no longer responded strongly to scrambled images was located near the occipitotemporal junction (Grill-Spector et al., 1998; Kanwisher, Chun, McDermott, & Ledden, 1996; Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001; Malach et al., 1995). The underlying collection of areas, termed the lateral occipital complex (LOC; see Grill-Spector, Kourtzi, & Kanwisher, 2001), is believed to be where regions associated with processing simple visual features (e.g., lines, simple shapes) transition to those responsible for processing more complex visual stimuli (e.g., complex shapes, faces, scenes, etc.). Shortly thereafter, it was shown that LOC is less sensitive to changes in stimulus size and position as compared to changes in viewpoint or illumination (Grill-Spector et al., 1999). This result is consistent with the properties of neurons in IT cortex, but not of those in earlier visual areas of monkeys (such as V1–V4), supporting the idea that LOC may be a functional homolog of IT cortex in monkeys (Grill-Spector et al., 1998; Malach et al., 1995).

4.2. Functional neuroimaging studies reveal neural substrates of category-selective agnosias

Although LOC exhibited many of the features one associates with a higher-order object recognition area, including weak retinotopy, sensitivity to image scrambling, and size/position invariance, it did not appear to possess the degree of selectivity necessary to identify specific objects. Namely, LOC responds strongly to *any object* but does not show selective responses to specific objects or features. However, anterior to LOC, a number of

other brain areas were identified that were found to be very selective for specific categories of objects (Fig. 9). For example, Sergent and colleagues (1992) identified regions within the right fusiform gyrus and anterior superior temporal sulcus selectively recruited during a face discrimination task. Puce and colleagues (Puce, Allison, Asgari, Gore, & McCarthy, 1996; Puce, Allison, Gore, & McCarthy, 1995) and Kanwisher and colleagues (1997) later replicated these findings using fMRI. The latter group called the fusiform activation the "fusiform face area" (Fig. 9) (FFA; Kanwisher et al., 1997). Other face-selective regions were also identified: one in the ventral occipital cortex (the "occipital face area", OFA) (Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999), another in the superior temporal sulcus (Haxby et al., 1999), and yet another in the anterior portion of the temporal lobe (Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Rajimehr, Young, & Tootell, 2009; Tsao, Moeller, & Freiwald, 2008). The significance of these regions for the processing of face stimuli is best illustrated by the observation that damage that includes the occipitotemporal cortex and/or the fusiform gyrus produces acquired prosopagnosia (e.g., Landis, Regard, Bliestle, & Kleihues, 1988; Rossion et al., 2003).

In addition to face-selective regions, regions selective for other visual categories were identified. An area selectively activated by locations, places, and landmarks was identified in the parahippocampal gyrus, and termed the "parahippocampal place area" (PPA) (Fig. 9) (Epstein & Kanwisher, 1998). Activation of the PPA requires the presentation of surfaces that convey information about the local environment; random displays of objects (e.g., furniture) produce minimal responses, but if those same objects are arranged in such a way as to give the illusion of a room, this area becomes significantly more responsive. As with face-selective regions, damage to the parahippocampal gyrus produces difficulties in navigating familiar locations ("topographical disorientation"; Epstein, DeYoe, Press, & Kanwisher, 2001).

Along the lateral extrastriate cortex, an area selectively activated by images of human body-parts was found, the so-called extrastriate body-part area (EBA) (Fig. 9) (Downing, Jiang, Shuman, & Kanwisher, 2001). This region was later shown to also respond to limb and goal-directed movements (Astafiev, Stanley, Shulman, & Corbetta, 2004). There was even the discovery of a region selective
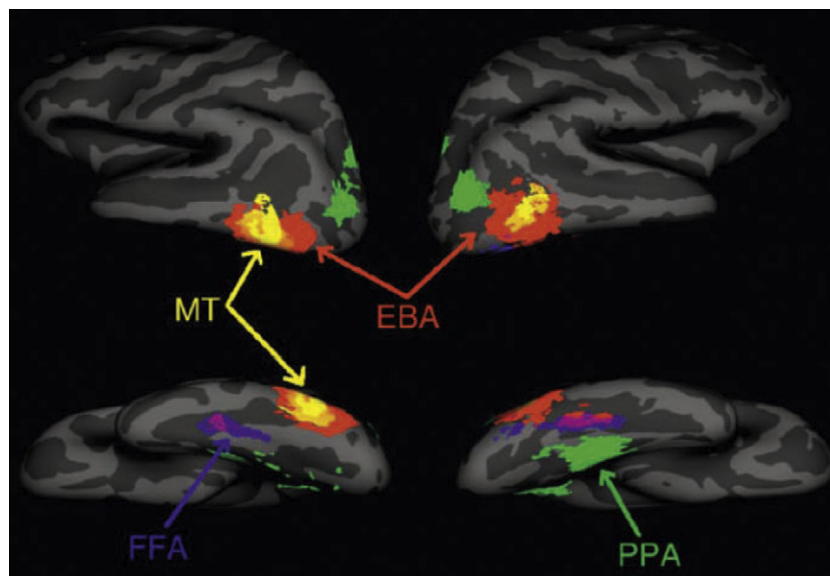
Reprinted from Spiridon et al., 2006

Fig. 9. Inflated views of the human brain showing regions selective for images of faces in the fusiform gyrus ("fusiform face area, FFA"), images of places in the adjacent parahippocampal gyrus ("parahippocampal place area, PPA"), and images of human body-parts near the lateral occipitotemporal cortex ("extrastriate body-part area, EBA"). MT: motion-selective region. (from Spiridon et al., 2006).

for letters and nonsense letterstrings (Puce et al., 1996). This so-called "visual word form area" (VWFA) was localized to a region within/adjacent to the fusiform gyrus and was further shown to respond preferentially to real words compared to nonsense letter-strings (Cohen et al., 2002). The VWFA is found almost exclusively in the left hemisphere – the same hemisphere as other language-related areas, such as Broca's and Wernicke's areas (Binder et al., 1997). Importantly, because we are not born with knowledge about written words, the existence of VWFA indicates that at least some category-specific cortex must arise as a result of experience.

Recently, neuroimaging experiments in monkeys have revealed similar category-selective regions throughout IT cortex (Fig. 10). Tsao and colleagues (2003) and later others (Bell, Hadj-Bouziane, Frihauf, Tootell, & Ungerleider, 2009; Pinsk, DeSimone, Moore, Gross, & Kastner, 2005; Pinsk et al., 2009) identified several regions along the superior temporal sulcus (STS) selectively responsive to faces (of both monkeys and humans). Recordings from one of these face-selective regions has demonstrated high proportions of face-selective neurons (Tsao, Freiwald, Tootell, & Livingstone, 2006). Located adjacent to these face-selective regions are regions selective for monkey body-parts, and regions selective for places have been found within the STS and along ventral IT cortex (Bell et al., 2009; Pinsk et al., 2005). Establishing precise homologies between the two species will require adherence to both anatomical and functional criteria (Rajimehr et al., 2009; Tsao et al., 2008) and is still debated. Nonetheless, the existence of fMRI-identified category-selective regions in the monkey brain provides a clear bridge between experiments involving fMRI and electrophysiology, greatly facilitating cross-species comparisons.

In humans, these category-selective regions were shown to be sensitive to the conscious percept of a stimulus. Tong and colleagues (1998) presented subjects with an image of a face to one eye and an image of a house to the other, thereby creating a bi-stable percept that shifted back and forth between faces and houses every few seconds. They found that FFA activation increased when subjects reported perceiving a face whereas PPA activation increased when subjects reported perceiving a house, despite both images being present at all times. Moreover, these regions respond to the percept of the stimulus, even in its absence. Summerfield and colleagues (2006) asked subjects to discriminate between noisy images of faces and houses. When subjects incorrectly reported seeing a face when actually presented with a house, activation in FFA increased, indicating that the activation was related more to the percept than to the physical stimulus itself.

How (and when during development) these category-selective regions originate is still unclear, although there is mounting evidence to suggest that they are not innate but arise through learning and experience. For example, a number of studies have demonstrated training and experience-related changes in activation of temporal cortex. Op de Beeck and colleagues (2006) scanned subjects before and after discrimination training on a set of novel objects and found that activation in areas responsive to the objects increased after training. This increase in signal strength due to training was not uniform across temporal cortex (i.e., was not, for example, a non-specific attentional effect) but rather was greater in some visual cortical areas (e.g., right fusiform gyrus) compared to others. Wong and colleagues (2009) offered further insights into the spatial distribution of training-related increases in activation. They found that training subjects to discriminate a set of novel objects at an exemplar (subordinate) level produced greater and more focal activation in right fusiform gyrus than in other areas of visual cortex (see also van der Linden, Murre, & van Turennout, 2008). By contrast, training subjects to discriminate objects at a categorical level produced more diffuse increases in activation. Changes in activation may reflect an underlying increase in neuronal selectivity among IT neurons that is observed following discrimination learning (Baker, Behrmann, & Olson, 2002). In some cases, experience-related changes in activation and selectivity may take place over several years. For example, although humans are relatively adept at face discrimination at a very early age (Fagan, 1972; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995), face-selective regions (i.e., FFA, OFA) do not appear until early to late adolescence (Golarai et al., 2007; Scherf, Behrmann, Humphreys, & Luna, 2007).

To summarize, the existence of brain regions selectively activated by images from specific categories offered, for the first time, a clear correlate to some of the category-selective agnosias first characterized over half a century earlier (e.g., prosopagnosia). Damage to these regions produces perceptual deficits specific for these categories. What is perhaps most interesting about these regions is how they respond to all conceptually related stimuli, regardless of their individual visual features. Images of body-parts, for example, might include many different shapes and textures but activate the same small, circumscribed region. This coding mechanism, based on conceptual similarity, is in direct contrast with the combination coding mechanism, based on visual similarity, currently proposed in monkeys. The former mechanism appears well suited for making rapid category-level identifications but may not have the selectivity necessary to dissociate individual exemplars from within individual categories. On the other hand, combination coding is well suited to dissociate individual exemplars based on subtle differences in visual appearance, but has no
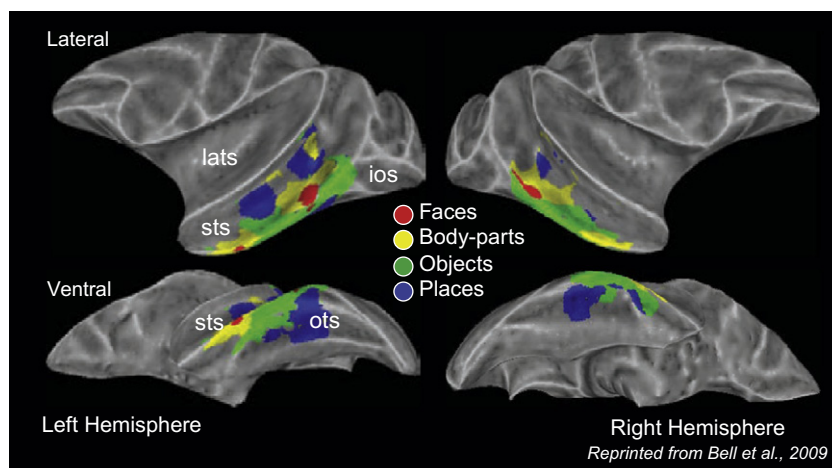
Fig. 10. Inflated views of the monkey brain showing regions selective for images of faces, body-parts, objects, and places. ios: inferior occipital sulcus; lats: lateral sulcus; ots: occipitotemporal sulcus; sts: superior temporal sulcus (from Bell et al., 2009).

obvious way to organize or link representations based on perceptual or semantic relationships. Reconciling these two theories will reveal much about how the brain recognizes both categorical distinctions and individual exemplars from within categories.

4.3. Advancements in neuroimaging methods and analyses: category-selective regions form interconnected networks

Most categories tested have been shown to evoke multiple category-selective regions within the human and monkey brains: Images of faces evoke four in the human occipital and temporal cortex (FFA, OFA, STS, anterior temporal cortex) and between two and six in the monkey temporal cortex (Bell et al., 2009; Moeller, Freiwald, & Tsao, 2008; Pinsk et al., 2005). Places and scenes evoke activation in the human PPA as well as in a region of the retrosplenial cortex (Epstein, Parker, & Feiler, 2008; Park & Chun, 2009; Walther, Caddigan, Fei-Fei, & Beck, 2009). Similarly, several place-selective regions in the monkey temporal cortex have been described (Bell et al., 2009). The existence of multiple regions for each category strongly suggests the existence of cortical networks specialized for the processing of different stimulus categories. Evaluating the architecture and input–output anatomical relationships among regions in distributed networks was once only possible through terminal tracer experiments in animals. However, new analytic techniques in fMRI offer the possibility of characterizing cortical networks using non-invasive approaches (see Bandettini (2009) for review). Diffusion tensor imaging (DTI), for example, measures the diffusion of water along axonal fibers, providing both the location and relative size of white matter tracts between structures and thus allows investigators to quantify the degree of connectivity between functionally similar regions (Basser, 1995; Basser, Mattiello, & LeBihan, 1994). DTI has been used to investigate how changes in the connectivity among face-selective regions in the ventral temporal cortex of humans affect face processing. For example, patients suffering from congenital prosopagnosia show reduced volume in the inferior longitudinal fasciculus and inferior fronto-occipito fasciculus, two white matter pathways that pass through the fusiform gyrus (Fig. 11) (Thomas et al., 2009). Incidentally, the latter fasciculus has been shown to degrade with age, which correlates with a reduction in face discrimination abilities among older subjects (Thomas et al., 2008).

Another analytic technique, called "functional connectivity", estimates connections between brain regions based on the relationship between activity patterns in different seed regions. These can be assessed at rest (e.g., resting state connectivity; Greicius, Krasnow, Reiss, & Menon, 2003; Raichle et al., 2001) or during the performance of a task (e.g., dynamic causal modeling; DCM; Friston, Harrison, & Penny, 2003). Resting state connectivity has been used extensively in the study of neuropsychiatric disorders (Greicius, 2008). Patients suffering from Alzheimer's disease show altered connectivity between the hippocampus, cingulate and prefrontal cortices (Wang et al., 2006). Assessing functional connectivity during task performance can provide insights as to how connectivity is altered during different behavioral states. Using DCM, Nummenmaa and colleagues (2009) examined how face-selective regions in the ventral temporal cortex interact with attention and oculomotor structures in the frontal cortex to mediate shifts in attention induced by changes in gaze direction.

In monkeys, it is possible to use electrical microstimulation in combination with fMRI to visualize the connections between areas (Field, Johnston, Gati, Menon, & Everling, 2008; Tolias et al., 2005). Moeller and colleagues (2008) used this technique to map the connections among six discrete face-selective regions in IT cortex, which were found to be strongly interconnected with one another but not with adjacent regions (Fig. 12), thereby demonstrating that


Congenital Prosopagnosia   **Matched Control**

ILF

IFOF
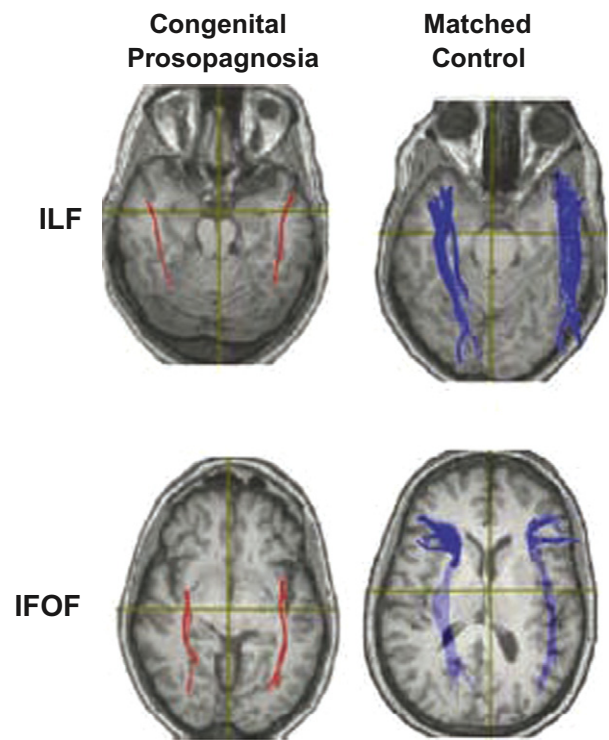
Adapted from Thomas et al., 2009

Fig. 11. Diffusion tensor images of patient with congenital prosopagnosia (left column) and age/gender matched control (right column), plotted on axial slices. Patients with congenital prosopagnosia show reduced white matter volume in inferior lateral and inferior fronto-occipito fasciculi (ILF and IFOF), two white matter tracts that pass through the fusiform gyrus (from Thomas et al., 2009).

monkeys show similar specialized networks to those found in humans.

Finally, in addition to new experimental methodologies, the last several years have seen a number of advancements in how imaging data are analyzed. Traditionally, fMRI data are analyzed by first using a localizer contrast (e.g., faces vs. places) to locate a region or regions of interest (ROIs; e.g., face-selective regions), and then restricting subsequent analyses to these isolated voxels. Alternatively, ROIs might be defined anatomically (e.g., hippocampus). Examining ROIs as isolated entities is highly advantageous because it restricts the analysis to the regions most likely to show an effect. However, it also ignores any information that might be contained within the *pattern* of activation among voxels outside the ROIs. Recently, investigators have begun to examine these patterns of activation (e.g., Haxby et al., 2001; Kriegeskorte et al., 2006; Sapountzis et al., 2010; Oosterhof et al., 2010; Norman et al., 2006; Naselaris et al., 2010; Kamitani and Tong, 2005; see Bandettini (2009) for review). For example, Kay and colleagues (2008) used patterns of activation within early visual areas to identify which of 120 novel natural images the subject was viewing. Subjects were first shown a set of 1750 natural images while fMRI data were collected from V1 to V3. These data were used to create a receptive field model for each voxel. The model was then used to predict the pattern of activation that would result when the subjects were later shown the 120 novel images. Accuracy exceeded 80% – well above chance performance. Using fMRI data in this way is often referred to as "brain decoding" and has been used in a number of recent fMRI studies looking at a variety of cognitive functions (e.g., Brouwer & Heeger, 2009; Esterman, Chiu, Tamber-Rosenau, & Yantis, 2009; Macevoy & Epstein, 2009; Reddy, Tsuchiya, & Serre, 2010; Rodriguez, 2010).
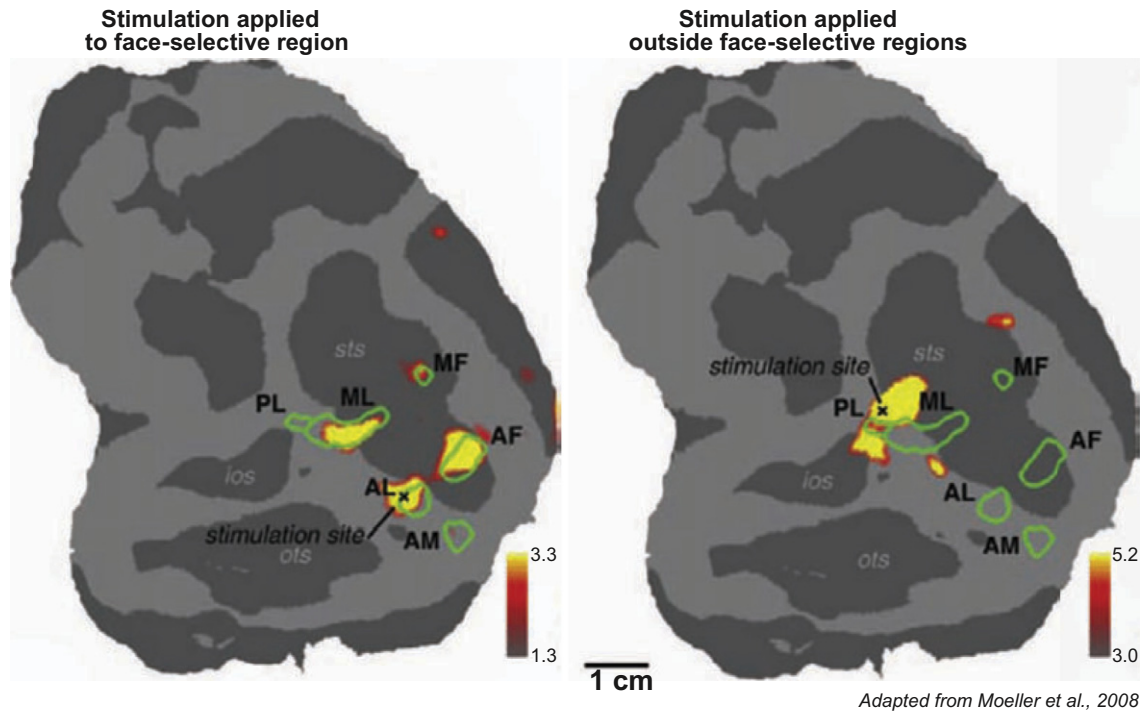
Fig. 12. Interconnections between face-selective regions in the monkey temporal cortex revealed using electrical microstimulation and fMRI. Microstimulation applied to a face-selective region in area TE evoked activation in several nearby face-selective regions, indicating functional connections among them (left panel). Microstimulation applied outside all face-selective regions failed to evoke activation within any of the face-selective regions (right panel). Similar experiments were conducted to show that all face-selective regions within the temporal cortex of the monkey are interconnected. AF: anterior fundus; AL: anterior lateral; AM: anterior medial; MF: middle fundus; ML: middle lateral; PL: posterior lateral (from Moeller et al., 2008).

Pattern analysis and brain decoding may challenge some pre-conceived ideas about how representations are organized in the brain. For instance, it is easy to assume that only those regions identified using ROI-based analyses process stimuli from a specific category and yet pattern analysis might subsequently reveal valuable information about that category outside of these ROIs. However, this leads to the problem of how does one interpret the results of a pattern-analysis? That is to say – just because an analysis might provide accurate information, we cannot assume that this is how the brain actually accomplishes the task. This is a difficult question to address as the answer may change from one study to another, or from one brain region to another – requiring caution when applying pattern analysis to fMRI data. Nonetheless, the positive results obtained thus far from pattern analysis emphasize the importance of considering activations outside isolated regions of interest when examining fMRI data.

4.4. Summary

Neuroimaging has allowed researchers to learn much about the neural substrates of perceptual representations in the human brain. The demonstration of category-selective regions in ventral occipital and temporal cortex offers clear correlates to the category-selective agnosias observed in humans. Activation within these category-selective regions appears to be associated with the conscious perception of stimuli from a given category. Finally, these regions form interconnected networks specialized for processing stimuli from a given category. These data suggest that as one moves anteriorly along the occipito-temporal pathway, the neuronal representation of a given stimulus is transformed from a strict reflection of its visual appearance (i.e., a sphere, a rectangle) to one that incorporates its categorical identity (i.e., a face, a car). As new analytic techniques evolve, potential for greater insights

into understanding mechanisms for categorical and exemplar distinctions will increase.

5. A special case of visual perception – face processing

Given their unique importance to social communication, faces undoubtedly represent a special category of stimuli with respect to visual processing. We rely on our ability to distinguish individuals and read their facial expressions in order to understand our relationships with others and to interpret their intentions. Face perception is multifaceted, involving face detection, identity discrimination, and the analysis of the changeable aspects of the face relevant for social context (e.g., facial expressions). In the following sections, we will describe how studies conducted over the last two decades have greatly expanded our knowledge of how the brain codes and interprets face stimuli.

5.1. Face-selective neurons in monkey IT cortex sensitive to socially relevant aspects of facial stimuli

Since the earliest recording studies, researchers have found neurons in IT cortex that are highly selective for face stimuli. Moreover, unlike the majority of IT neurons in the monkey that respond systematically to reduced versions of real-world objects, face-selective neurons typically respond only to intact images of faces (see above; Desimone et al., 1984). If the individual elements of the face are rearranged or parts removed (e.g., eyes, mouth), the firing rates of these neurons drop significantly (Desimone et al., 1984; Hasselmo, Rolls, & Baylis, 1989; Perrett et al., 1982). Thus, face-selective neurons in monkey IT cortex appear to be sensitive to the holistic percept of an intact face. Given this degree of specificity, face-selective neurons likely represent the primary neural substrate for face processing in the primate brain. As such, one

would expect to see evidence of all aspects of face processing (i.e., face detection, discrimination, etc.) within the neuronal responses, which is indeed the case. For example, in addition to signaling the presence of a face, face-selective neurons also appear to be sensitive to a number of socially relevant cues, such as head orientation (Desimone et al., 1984; Perrett et al., 1985) and gaze direction (Perrett et al., 1985). Neurons responsive to different facial expressions (e.g., open-mouth threat) have also been identified (Hasselmo et al., 1989) and activation throughout IT cortex has been shown, using fMRI, to be modulated in response to different facial expressions (Hadj-Bouziane, Bell, Knusten, Ungerleider, & Tootell, 2008; Hoffman, Gothard, Schmid, & Logothetis, 2007). Face-selective neurons can therefore encode both the presence of a face as well as the social content of the face image – but how might this be accomplished by a single neuron?

Sugase and colleagues (1999) offered a potential solution. In their study, monkeys were trained to maintain central fixation while visual stimuli were presented foveally. The stimuli included images of monkey and human faces, each with a different identity and facial expression. They found that responses among IT neurons often consisted of an early phase followed by a later phase (Fig. 13). The early phase, which occurred approximately 100–150 ms following stimulus onset, encoded the global features of the stimulus that were relevant for face detection or simple discrimination (e.g., between a face and a non-face, or between a monkey face and a human face). The later phase, which occurred approximately 50 ms later, encoded the fine features of the stimulus, such as those necessary to discriminate between individual faces or between different facial expressions.

The encoding of individual facial identities has also been hypothesized to occur at the level of individual neurons. Leopold and colleagues (2006) have suggested that facial identity is encoded by single neurons using a norm-based encoding scheme. According to this theory, neurons in IT cortex signal the distance of a given face from a prototypical or "mean" face along multiple dimensions, each representing different diagnostic features. This

scheme, in contrast to the earlier "Grandmother Cell Hypothesis", allows for infinite flexibility because it does not rely on previous viewings or many stored representations. A similar norm-based encoding scheme has been proposed in humans (Loffler, Yourganov, Wilkinson, & Wilson, 2005).

The studies cited above have highlighted the capabilities of individual neurons to represent different facial features. Another important, and as of yet unanswered, question is how do these neurons contribute to behavior? Surprisingly, only a few studies have addressed this question directly. Eifuku and colleagues (2004) trained monkeys to perform a face identity discrimination task in which monkeys were first presented with a sample face stimulus followed by several potential target faces. The monkeys were required to push a lever when the target stimulus matched the identity of the sample stimulus. The two matching stimuli were pictures of the same familiar person, but not necessarily from the same viewpoint, thus requiring the monkey to make true face identifications and not simply exemplar matches. It was found that face-selective neurons in the anterior IT cortex showed correlations between the latency of the neuronal response and the latency of the behavioral response, suggesting that these neurons directly contributed to the perceptual decision. More recently, Afraz and colleagues (2006) provided direct evidence linking the responses of face neurons to the perception of a face. After locating a cluster of face-selective neurons, they applied electrical microstimulation while monkeys performed a difficult face vs. non-face discrimination task. When stimulation was applied, monkeys were significantly more likely to report a face, regardless of the identity of the visual stimulus.

Individual face-selective neurons have a remarkable capacity to encode different aspects of the facial image. A single neuron not only signals the presence of a face within the visual field, but also conveys information about the identity of the face as well as its emotional content. Furthermore, the activity of these neurons correlates with performance in face detection and face discrimination tasks, suggesting they contribute directly to perceptual decisions involving faces.
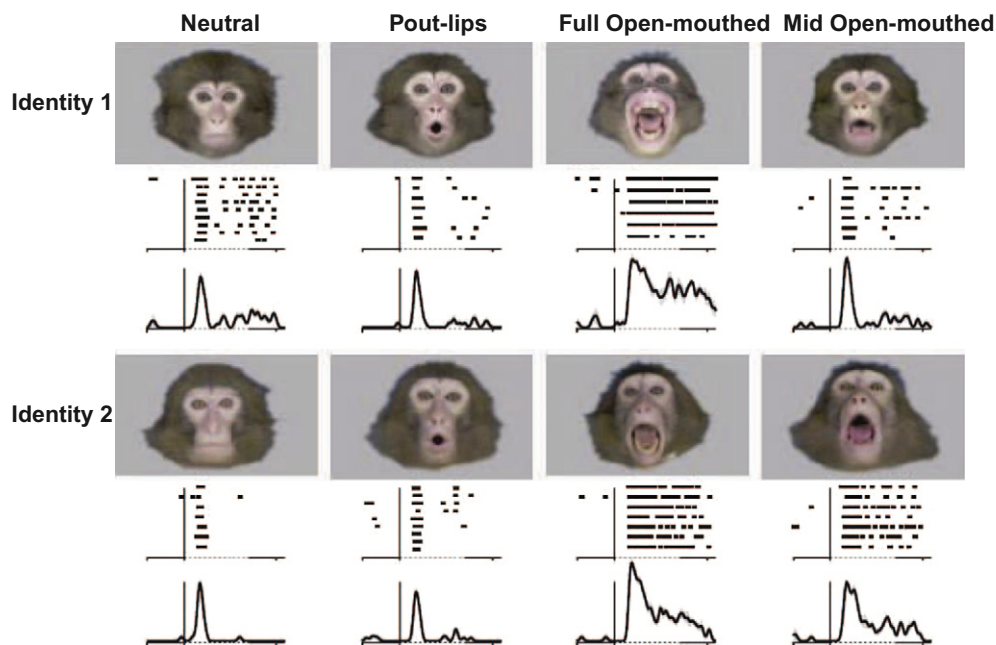
Adapted from Sugase et al., 1999

Fig. 13. Example of an IT neuron encoding multiple levels of information related to face perception. Each face evokes a sharp but brief response, beginning approximately 100–150 ms following stimulus onset. The magnitude of the second phase of the response, beginning approximately 50 ms later, differs according to the facial expression and identity of the facial image (from Sugase et al., 1999).

5.2. Two systems in the human brain specialized for different aspects of face processing

In the human, fMRI has been used extensively to investigate not only the sensitivity of the different face-selective regions but also how they might be functionally organized. Evaluating the identity and socially relevant information of a face (e.g., facial expression and gaze direction), for example, requires greater attention to individual facial features. Therefore, areas showing greater sensitivity to individual facial features are likely involved in extracting this information. Hoffman and Haxby (2000) probed the relative influence of facial identity vs. gaze direction in face-selective regions by manipulating selective attention to one or the other. They hypothesized that areas more sensitive to facial identity would show greater activation when subjects were cued to attend to the identity of a face stimulus as compared to when subjects were cued to attend to the gaze direction. They found that selective attention had a greater effect on identity in face-selective regions in the lateral and ventral temporal regions and a greater effect on gaze direction in regions located in the STS. Similar findings linking identity discrimination with face-selective regions in the lateral and ventral temporal cortex have been reported by others (e.g., Eger, Schyns, & Kleinschmidt, 2004; Fox, Moon, Iaria, & Barton, 2009; Rhodes, Michie, Hughes, & Byatt, 2009; Winston, Henson, Fine-Goulden, & Dolan, 2004). Expanding on this research, Haxby and colleagues proposed that the face-processing network consists of two separate systems. The core system, which includes the occipito-temporal face-selective regions (e.g., FFA, OFA, STS), is primarily responsible for the visual analysis of facial features (for identity discrimination, processing of emotional expressions, etc.). The extended system, which includes the amygdala, anterior temporal regions, and the prefrontal cortex, is primarily responsible for the further analysis of the changeable aspects of the face and for interacting with other systems that rely on face processing (e.g., speech perception). Using DCM (see above), Fairhall and Ishai (2007) evaluated the connectivity between these two systems. They found that correlations between areas in both the core and extended systems changed depending on the type of face presented. Emotional faces, for example, increased the strength of functional connections between the FFA and the amygdala, inferior frontal gyrus, and orbitofrontal cortex – all structures linked to the processing of emotional expressions (Adolphs, Damasio, & Tranel, 2002).

5.3. Summary

The study of face processing in the primate brain is an ideal example of the benefits of combining several approaches across multiple species to address a single problem. The first suggestions of a specialized face-processing network came from lesion work in humans, which was later supported by electrophysiological recordings in monkeys. Later still, neuroimaging studies bridged the gap between these two approaches, greatly increasing our understanding of this network at multiple levels. We now know that individual neurons can encode multiple levels of information regarding face stimuli and that these neurons are concentrated into several patches distributed throughout the temporal cortex. Based on work obtained in humans, it appears that these patches are functionally organized into different pathways, each specialized for different aspects of face processing. One of the next steps will be to determine how the information encoded by the face-processing networks in the temporal cortex is used by downstream structures, such as those within the prefrontal cortex, to form decisions about face discrimination and social behavior.

6. General summary

In this review, we have described a few of the many discoveries in object and face processing made over the last 25 years. Technological advances, such as neuroimaging, combined with new approaches, such as physiological studies in behaving animals, have provided many insights that were not possible in the past. We have learned how the temporal cortex contains networks for processing different categories of visual stimuli, each of which consists of different regions selective for these categories. We have learned about the relative contributions of these regions to the different aspects of object processing, particularly face processing. Finally, we have identified a combination coding mechanism that could underlie how individual objects might be represented.

Yet, despite this progress, many unanswered questions remain. How precisely does the brain integrate individual visual features into a single percept and, perhaps more importantly, how does it bind that percept to the object's semantic identity? Does the existence of a category-selective region in cortex indicate that *all* processing for said category takes place within that region? For that matter, how are individual exemplars from within a given category recognized and what about stimuli from categories that do not have a specialized region (e.g., fruit)?

How will we answer these questions in the future? Much like the progress made in the last 25 years, discovering new information about the brain during the next 25 years will likely require new approaches. For example, many laboratories have already started to apply genetics to the study of visual perception. Two recent studies have compared the ability of monozygotic and dizygotic twins to recognize and discriminate faces in order to evaluate the genetic basis of face recognition (Wilmer et al., 2010; Zhu et al., 2010). Similarly, others studying congenital prosopagnosia, a condition in which patients are born with an impaired ability to recognize faces, have begun to include genotyping and heritability assays to evaluate the role of genetics in face processing disorders (Gruter, Gruter, & Carbon, 2008; Kennerknecht, Pluempe, & Welling, 2008; Lee, Duchaine, Wilson, & Nakayama, 2009). As new experimental techniques are applied, yielding fresh perspectives, our understanding of the brain and the "visual alphabet" will continue to evolve.

References

Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: A 3-D lesion study. *Emotion, 2*(1), 23–51.

Afraz, S. R., Kiani, R., & Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. *Nature, 442*(7103), 692–695.

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience, 7*(5), 542–548.

Baker, C. I., Behrmann, M., & Olson, C. R. (2002). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience, 5*(11), 1210–1216.

Bandettini, P. A. (2009). What's new in neuroimaging methods? *Annals of the New York Academy of Sciences, 1156*, 260–293.

Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine, 8*(7–8), 333–344.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance Series B, 103*(3), 247–254.

Bell, A. H., Hadj-Bouziane, F., Frihauf, J. B., Tootell, R. B., & Ungerleider, L. G. (2009). Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *Journal of Neurophysiology, 101*(2), 688–700.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115–147.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*(1), 353–362.

Blum, J. S., Chow, K. L., & Pribram, K. H. (1950). A behavioral analysis of the organization of the parieto-temporo-preoccipital cortex. *Journal of Comparative Neurology, 93*(1), 53–100.

Bodamer, J. (1947). Die Prosop-Agnosie (Die Agnosie des Physiognomieerkennens). *Archiv fur Psychiatrie und Nervenkrankheiten, 179*, 48.

Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology, 306*(4), 554–575.

Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience, 7*(8), 880–886.

Brincat, S. L., & Connor, C. E. (2006). Dynamic shape synthesis in posterior inferotemporal cortex. *Neuron, 49*(1), 17–24.

Broca, P. (1861). Remarks on the seat of the faculty of articulate language, followed by an observation of aphemia. In G. Von Bonin (Ed.), *Some papers on the cerebral cortex*. Springfield.

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience, 29*(44), 13992–14003.

Brown, S., & Schafer, E. A. (1888). An investigation into the functions of the occipital and temporal lobes of the monkey's brain. *Philosophical Transactions of the Royal Society of London, 179*, 303–327.

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*(2), 369–384.

Buffalo, E. A., Bertini, G., Ungerleider, L. G., & Desimone, R. (2005). Impaired filtering of distracter stimuli by TE neurons following V4 and TEO lesions in macaques. *Cerebral Cortex, 15*(2), 141–151.

Bulthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences of the United States of America, 89*(1), 60–64.

Chow, K. L. (1951). Effect of partial extirpations of the posterior association cortex on visually mediated behavior. *Comparative Psychological Monographs, 20*, 187–217.

Chow, K. L. (1961). Anatomical and electrographical analysis of temporal neocortex in relation to visual discrimination learning in monkeys. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning*. Oxford: Blackwell Publishing.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain, 125*(Pt 5), 1054–1069.

Cowey, A., & Gross, C. (1970). Effects of foveal prestriate and inferotemporal lesions on visual discrimination by rhesus monkeys. *Experimental Brain Research, 11*(2), 128–144.

Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology, 32*(4), 331–341.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4*(8), 2051–2062.

Desimone, R., Fleming, J., & Gross, C. G. (1980). Prestriate afferents to inferior temporal cortex: An HRP study. *Brain Research, 184*(1), 41–55.

DeYoe, E. A., Bandettini, P., Neitz, J., Miller, D., & Winans, P. (1994). Functional magnetic resonance imaging (FMRI) of the human brain. *Journal of Neuroscience Methods, 54*(2), 171–187.

Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature, 428*(6980), 316–319.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470–2473.

Eger, E., Schyns, P. G., & Kleinschmidt, A. (2004). Scale invariant adaptation in fusiform face-responsive regions. *Neuroimage, 22*(1), 232–242.

Eifuku, S., De Souza, W. C., Tamura, R., Nishijo, H., & Ono, T. (2004). Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *Journal of Neurophysiology, 91*(1), 358–371.

Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., et al. (1994). FMRI of human visual cortex. *Nature, 369*(6481), 525.

Epstein, R., DeYoe, E. A., Press, D., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Psychology, 18*(6), 481–508.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*(6676), 598–601.

Epstein, R. A., Parker, W. E., & Feiler, A. M. (2008). Two kinds of FMRI repetition suppression? Evidence for dissociable neural mechanisms. *Journal of Neurophysiology, 99*(6), 2877–2886.

Esterman, M., Chiu, Y. C., Tamber-Rosenau, B. J., & Yantis, S. (2009). Decoding cognitive control in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 106*(42), 17974–17979.

Ettlinger, G., Iwai, E., Mishkin, M., & Rosvold, H. E. (1968). Visual discrimination in the monkey following serial ablation of inferotemporal and preoccipital cortex. *Journal of Comparative and Physiological Psychology, 65*(1), 110–117.

Fagan, J. F. 3rd, (1972). Infants' recognition memory for faces. *Journal of Experimental Child Psychology, 14*(3), 453–476.

Fahy, F. L., Riches, I. P., & Brown, M. W. (1993). Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research, 96*(3), 457–472.

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex, 17*(10), 2400–2406.

Farah, M. J. (2004). *Visual agnosia*. Cambridge: MIT Press.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.

Field, C. B., Johnston, K., Gati, J. S., Menon, R. S., & Everling, S. (2008). Connectivity of the primate superior colliculus mapped by concurrent microstimulation and event-related FMRI. *PLoS ONE, 3*(12), e3928.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *Neuroimage, 44*(2), 569–580.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage, 19*(4), 1273–1302.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature, 360*(6402), 343–346.

Gall, F. J., & Spurzheim, J. C. (1809). *Recherches sur le Sysstème Nerveux en generale et celui du Cerveau en particulier*. Paris: Schoell.

Gall, F. J., & Spurzheim, J. C. (1810). *Anatomie et physiologie du systeme nerveux en generale, et du cerveau en particulier avec des observations sur la possibilitie de reconnaitre plusiers dispositions intellectuelles et morales de l'homme et des animaux par la configuration de leurs tetes*. Paris: Schoell.

Gochin, P. M., Miller, E. K., Gross, C. G., & Gerstein, G. L. (1991). Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Experimental Brain Research, 84*(3), 505–516.

Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., et al. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience, 10*(4), 512–522.

Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Current Opinion in Neurology, 21*(4), 424–430.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 100*(1), 253–258.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*(10–11), 1409–1422.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*(1), 187–203.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6*(4), 316–328.

Grinvald, A., Lieke, E., Frostig, R. D., Gilbert, C. D., & Wiesel, T. N. (1986). Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature, 324*(6095), 361–364.

Gross, C. (1999). *Brain, vision, and memory: Tales in the history of neuroscience*. Cambridge: MIT Press.

Gross, C. G. (2002). Genealogy of the "grandmother cell". *Neuroscientist, 8*(5), 512–518.

Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science, 166*(910), 1303–1306.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology, 35*(1), 96–111.

Gruter, T., Gruter, M., & Carbon, C. C. (2008). Neural and genetic foundations of face recognition and prosopagnosia. *Journal of Neuropsychology, 2*(Pt 1), 79–97.

Hadj-Bouziane, F., Bell, A. H., Knusten, T. A., Ungerleider, L. G., & Tootell, R. B. (2008). Perception of emotional expressions is independent of face selectivity in monkey inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 105*(14), 5591–5596.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32*(3), 203–218.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience, 14*(11 Pt 1), 6336–6353.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron, 22*(1), 189–199.

Hoffman, K. L., Gothard, K. M., Schmid, M. C., & Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology, 17*(9), 766–772.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3*(1), 80–84.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience, 22*(16), 7195–7205.

Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Canadian Journal of Psychology, 46*(2), 170–190.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences of the United States of America, 96*(16), 9379–9384.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology, 73*(1), 218–226.

Iwai, E., & Mishkin, M. (1968). Visual discrimination in the monkey following serial ablation of inferotemporal and preoccipital cortex. *Journal of Comparative and Physiological Psychology, 65*(1), 110–117.

Iwai, E., & Mishkin, M. (1969). Further evidence on the locus of the visual area in the temporal lobe of the monkey. *Experimental Neurology, 25*(4), 585–594.

Jackson, J. H. (1870). A study of convulsions. *Trans. St. Andrews. Med. Grad. Assoc, 3*, 162–204.

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Brain Research. Cognitive Brain Research, 5*(1–2), 55–67.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311.

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature, 452*(7185), 352–355.

Kennerknecht, I., Pluempe, N., & Welling, B. (2008). Congenital prosopagnosia – A common hereditary cognitive dysfunction in humans. *Frontiers in Bioscience, 13*, 3150–3158.

Klüver, H. (1948). Functional differences between the occipital and temporal lobes with special reference to the interrelations of behavior and extracerebral mechanisms. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley Inc.

Klüver, H., & Bucy, P. C. (1938). An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey, with special reference to "psychic blindness". *Journal of Psychology, 5*, 33–54.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71*(3), 856–867.

Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 104*(51), 20600–20605.

Landis, T., Regard, M., Bliestle, A., & Kleihues, P. (1988). Prosopagnosia and agnosia for noncanonical views. An autopsied case. *Brain, 111*(Pt 6), 1287–1297.

Lee, Y., Duchaine, B., Wilson, H. R., & Nakayama, K. (2009). Three cases of developmental prosopagnosia from one family: Detailed neuropsychological and psychophysical investigation of face processing. *Cortex*.

Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology, 97*(1), 307–319.

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature, 442*(7102), 572–575.

Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex, 11*(4), 287–297.

Li, N., Cox, D. D., Zoccolan, D., & DiCarlo, J. J. (2009). What response properties do individual neurons need to underlie position and clutter "invariant" object recognition? *Journal of Neurophysiology, 102*(1), 360–376.

Li, N., & DiCarlo, J. J. (2008). Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science, 321*(5895), 1502–1507.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*(6), 1918–1929.

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). FMRI evidence for the neural representation of faces. *Nature Neuroscience, 8*(10), 1386–1390.

Logothetis, N. K., Pauls, J., Bulthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology, 4*(5), 401–414.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology, 5*(5), 552–563.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience, 19*, 577–621.

Macevoy, S. P., & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology, 19*(11), 943–947.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*(18), 8135–8139.

Mayo, J. P., & Sommer, M. A. (2008). Neuronal adaptation caused by sequential visual stimulation in the frontal eye field. *Journal of Neurophysiology, 100*(4), 1923–1935.

McCarthy, R. A., & Warrington, E. K. (1986). Visual associative agnosia: A clinico-anatomical study of a single case. *Journal of Neurology, Neurosurgery and Psychiatry, 49*(11), 1233–1240.

McMahon, D. B., & Olson, C. R. (2007). Repetition suppression in monkey inferotemporal cortex: Relation to behavioral priming. *Journal of Neurophysiology, 97*(5), 3532–3543.

Miller, E. K., Gochin, P. M., & Gross, C. G. (1991). Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Visual Neuroscience, 7*(4), 357–362.

Mishkin, M. (1954). Visual discrimination performance following partial ablations of the temporal lobe. II. Ventral surface vs. hippocampus. *Journal of Comparative and Physiological Psychology, 47*(3), 187–193.

Mishkin, M. (1966). Visual mechanisms beyond the striate cortex. In R. Russell (Ed.), *Frontiers of physiological psychology* (pp. 93–119). New York: Academic Press.

Mishkin, M., & Pribram, K. H. (1954). Visual discrimination performance following partial ablations of the temporal lobe. I. Ventral vs. lateral. *Journal of Comparative and Physiological Psychology, 47*(1), 14–20.

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science, 320*(5881), 1355–1359.

Mogami, T., & Tanaka, K. (2006). Reward association affects neuronal responses to visual stimuli in macaque te and perirhinal cortices. *Journal of Neuroscience, 26*(25), 6761–6770.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*(4715), 782–784.

Munk, H. (1881). On the functions of the cortex. In G. Von Bonin (Ed.), *Uber die funktionen der Grosshirnrinde, 3te Millteilung. The cerebral cortex* (pp. 97–117). Springfield: Thomas.

Nummenmaa, L., Passamonti, L., Rowe, J., Engell, A. D., & Calder, A. J. (2009). Connectivity analysis reveals a cortical network for eye gaze perception. *Cerebral Cortex*.

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience, 26*(50), 13025–13036.

Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage, 47*(4), 1747–1756.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and extension. *Infant Behavior & Development, 18*, 79–85.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology, 82*(5), 2490–2502.

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology, 86*(5), 2505–2519.

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience, 5*(12), 1332–1338.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47*(3), 329–342.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 223*(1232), 293–317.

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., et al. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative FMRI study. *Journal of Neurophysiology, 101*(5), 2581–2600.

Pinsk, M. A., DeSimone, K., Moore, T., Gross, C. G., & Kastner, S. (2005). Representations of faces and body parts in macaque temporal cortex: A functional MRI study. *Proceedings of the National Academy of Sciences of the United States of America, 102*(19), 6996–7001.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16*(16), 5205–5215.

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology, 74*(3), 1192–1199.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2), 676–682.

Rajimehr, R., Young, J. C., & Tootell, R. B. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences of the United States of America, 106*(6), 1995–2000.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information during mental imagery. *Neuroimage, 50*(2), 818–825.

Rhodes, G., Michie, P. T., Hughes, M. E., & Byatt, G. (2009). The fusiform face area and occipital face area show sensitivity to spatial relations in faces. *European Journal of Neuroscience, 30*(4), 721–733.

Riches, I. P., Wilson, F. A., & Brown, M. W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience, 11*(6), 1763–1779.

Rocha-Miranda, C. E., Bender, D. B., Gross, C. G., & Mishkin, M. (1975). Visual activation of neurons in inferotemporal cortex depends on striate cortex and forebrain commissures. *Journal of Neurophysiology, 38*(3), 475–491.

Rockland, K. S., & Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Research, 179*(1), 3–20.

Rodriguez, P. F. (2010). Neural decoding of goal locations in spatial navigation in humans with fMRI. *Human Brain Mapping, 31*(3), 391–397.

Rolls, E. T., Judge, S. J., & Sanghera, M. K. (1977). Activity of neurones in the inferotemporal cortex of the alert monkey. *Brain Research, 130*(2), 229–238.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain, 126*(Pt 11), 2381–2395.

Sakai, K., & Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature, 354*(6349), 152–155.

Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science, 10*(4), F15–30.

Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences of the United States of America, 80*(18), 5776–5778.

Seltzer, B., & Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Research, 149*(1), 1–24.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*(5212), 889–893.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain, 115*(Pt 1), 15–36.

Shiwa, T. (1987). Corticocortical projections to the monkey temporal lobe with particular reference to the visual processing pathways. *Archives Italiennes de Biologie, 125*(2), 139–154.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*(6869), 318–320.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*(6747), 869–873.

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2006). Mistaking a house for a face: Neural correlates of misperception in healthy humans. *Cerebral Cortex, 16*(4), 500–508.

Tamura, H., & Tanaka, K. (2001). Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cerebral Cortex, 11*(5), 384–399.

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology, 66*(1), 170–189.

Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition, 67*(1–2), 1–20.

Thomas, C., Avidan, G., Humphreys, K., Jung, K. J., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature Neuroscience, 12*(1), 29–31.

Thomas, C., Moya, L., Avidan, G., Humphreys, K., Jung, K. J., Peterson, M. A., et al. (2008). Reduction in white matter connectivity, revealed by diffusion tensor imaging, may account for age-related changes in face perception. *Journal of Cognitive Neuroscience, 20*(2), 268–284.

Tolias, A. S., Sultan, F., Augath, M., Oeltermann, A., Tehovnik, E. J., Schiller, P. H., et al. (2005). Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. *Neuron, 48*(6), 901–911.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*(4), 753–759.

Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron, 21*(6), 1409–1422.

Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., et al. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 811–817.

Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., et al. (1997). Functional analysis of V3A and related areas in human visual cortex. *Journal of Neuroscience, 17*(18), 7060–7078.

Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Liu, A. K., & Dale, A. M. (1998). The representation of the ipsilateral visual field in human cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 818–824.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature Neuroscience, 6*(9), 989–995.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670–674.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America, 105*(49), 19514–19519.

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience, 4*(8), 832–838.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. G. Ingle (Ed.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Ungerleider, L. G., & Pribram, K. H. (1977). Inferotemporal versus combined pulvinar–prestriate lesions in the rhesus monkey: Effects on color, object and pattern discrimination. *Neuropsychologia, 15*(4–5), 481–498.

van der Linden, M., Murre, J. M., & van Turennout, M. (2008). Birds of a feather flock together: Experience-driven formation of visual object categories in human ventral temporal cortex. *PLoS One, 3*(12), e3995.

Vogels, R. (1999). Effect of image scrambling on inferior temporal cortical responses. *NeuroReport, 10*(9), 1811–1816.

von Bonin, G., & Bailey, P. (1947). *The neocortex of Macaca mulatta*. Urbana: University of Illinois Press.

Wade, A. R., Brewer, A. A., Rieger, J. W., & Wandell, B. A. (2002). Functional measurements of human ventral occipital cortex: Retinotopy and colour. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 357*(1424), 963–973.

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience, 29*(34), 10573–10581.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science, 272*(5268), 1665–1668.

Wang, L., Zang, Y., He, Y., Liang, M., Zhang, X., Tian, L., et al. (2006). Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI. *Neuroimage, 31*(2), 496–504.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(Pt 3), 829–854.

Webster, M. J., Ungerleider, L. G., & Bachevalier, J. (1991). Connections of inferior temporal areas TE and TEO with medial temporal-lobe structures in infant and adult monkeys. *Journal of Neuroscience, 11*(4), 1095–1116.

Weller, R. E., & Kaas, J. H. (1987). Subdivisions and connections of inferior temporal cortex in owl monkeys. *Journal of Comparative Neurology, 256*(1), 137–172.

Weller, R. E., & Steele, G. E. (1992). Cortical connections of subdivisions of inferior temporal cortex in squirrel monkeys. *Journal of Comparative Neurology, 324*(1), 37–66.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8*(2), 227–233.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences of the United States of America, 107*(14), 5238–5241.

Winston, J. S., Henson, R. N., Fine-Goulden, M. R., & Dolan, R. J. (2004). FMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92*(3), 1830–1839.

Wong, A. C., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PLoS One, 4*(12), e8405.

Zeki, S. M. (1971). Convergent input from the striate cortex (area 17) to the cortex of the superior temporal sulcus in the rhesus monkey. *Brain Research, 28*(2), 338–340.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20*(2), 137–142.