# Neural Mechanisms of High-Level Vision

Jay Hegdé<sup>\* 1,2,3,4</sup>

#### ▶ ABSTRACT

The last three decades have seen major strides in our understanding of neural mechanisms of highlevel vision, or visual cognition of the world around us. Vision has also served as a model system for the study of brain function. Several broad insights, as yet incomplete, have recently emerged. First, visual perception is best understood not as an end unto itself, but as a sensory process that subserves the animal's behavioral goal at hand. Visual perception is likely to be simply a side effect that reflects the readout of visual information processing that leads to behavior. Second, the brain is essentially a probabilistic computational system that produces behaviors by collectively evaluating, not necessarily consciously or always optimally, the available information about the outside world received from the senses, the behavioral goals, prior knowledge about the world, and possible risks and benefits of a given behavior. Vision plays a prominent role in the overall functioning of the brain providing the lion's share of information about the outside world. Third, the visual system does not function in isolation, but rather interacts actively and reciprocally with other brain systems, including other sensory faculties. Finally, various regions of the visual system process information not in a strict hierarchical manner, but as parts of various dynamic brain-wide networks, collectively referred to as the "connectome." Thus, a full understanding of vision will ultimately entail understanding, in granular, quantitative detail, various aspects of dynamic brain networks that use visual sensory information to produce behavior under real-world conditions. © 2018 American Physiological Society. Compr Physiol 8:903-953, 2018.

### **Didactic Synopsis**

#### Major teaching points

- Visual perception is essentially an inferential process, in that the visual system infers the likely interpretation of a given image by evaluating the various underlying statistical (i.e., probabilistic) factors. This is not to say that the inferences are necessarily optimal or that they are made consciously.
- For last several decades, neurophysiological research has been guided by the implicit assumption that the goal of visual processing is to help construct a veridical internal representation of the external visual world. This, along with earlier methodological difficulties in monitoring the activity of large number of neurons in multiple areas, led to a couple of decades of neurophysiological research focused on delineating functional specialization of *individual* visual areas, which has to do with figuring out "which visual area does what" in representing the external visual world.
- More recent research, however, suggests that a "goaloriented connectomic" view is a better framework for understanding visual processing as a sensory process wherein large number of different brain areas act, not as individual areas but as parts of as a larger network, to implement the animal's behavioral goal at hand.
- The networks that implement a given behavior vary dynamically and probabilistically depending on various factors,

including the sensory information, behavioral goal, possible rewards or punishments, etc.

- Vision is not a single process, but a collection of processes.
- Even though vision is the dominant sensory modality for humans and other primates, it does not function in isolation, but interacts actively with other senses and the brain systems that implement behavior.

### Introduction

#### Overview

Few exercises in scientific chauvinism are better grounded in truth than the conceit that vision is the mother of all senses. Vision is the dominant sensory modality of humans and other primates, presumably because it can provide detailed and

\*Correspondence to jay@hegde.us

DOI: 10.1002/cphy.c160035

<sup>&</sup>lt;sup>1</sup>Brain and Behavior Discovery Institute, Augusta University, Augusta, Georgia, USA

<sup>&</sup>lt;sup>2</sup>James and Jean Culver Vision Discovery Institute, Augusta University, Augusta, Georgia, USA

<sup>&</sup>lt;sup>3</sup>Department of Ophthalmology, Medical College of Georgia, Augusta University, Augusta, Georgia, USA

<sup>&</sup>lt;sup>4</sup>The Graduate School, Augusta University, Augusta, Georgia, USA Published online, July 2018 (*comprehensivephysiology.com*)

Copyright © American Physiological Society.

behaviorally relevant information about the outside world, and from safe distances at that. It is far easier to find and flee danger, and gain eating and mating opportunities, with vision than without. It is undoubtedly because of the behavioral utility of vision that a substantial portion of the brain is devoted to the processing of visual information. By any reckoning, we primates are visual animals.

The preeminence of vision as a sensory faculty has meant not only that understanding the neurophysiological mechanisms of vision has been understood to be a worthy goal unto itself but that vision has served as a useful model system for understanding brain function at large, and served as a template for understanding other sensory systems. But, as is frequently the case of accumulating wisdom primarily due to the labors of others, one of the biggest lessons of visual research over the last few decades has been that studying vision for its own sake tends to produce a misleading and warped picture of how the brain works.

# Scope and organization of the review: Broad operational principles of high-level vision for nonexperts

A short review like this cannot provide an exhaustive account of all aspects of high-level vision. For a more exhaustive treatment of the subject, the reader should consult the references listed in the *Further Reading* section below. This review has a much more modest goal. It simply aims to give the reader an intuitive feel for high-level vision by presenting a "highlight reel" of what we currently know and do not know about how high-level vision works. It is worth remembering that while there is much that we do not yet know about high-level vision, it is part of what makes vision so exciting to study.

This review is primarily meant as a pedagogical tool to help provide a critical understanding of the field to intelligent nonexperts, but may also be useful to practitioners in the field as one neurophysiologist's point of view about the current status and, even more importantly, the future direction of the field. This meant focusing not only on what we know about high-level vision but also on some key issues about which we know surprisingly little. I have tried to present both these aspects of the field in a way to incite further thinking and perhaps discussion.

An exercise in capturing the "gist" of a vast field will necessarily entail an arbitrary and subjective selection of topics to cover and studies to cite. For a fuller, more technical understanding of the topics discussed, readers should consult the references cited here, and many excellent books on the topic of vision (a selected few of which are listed below), and the various citations therein. Also, not all the studies cited in this review represent the latest ones on a given topic. Indeed, some of the studies discussed in detail represent some of the *earliest* studies on that topic, especially in cases where I felt it was more helpful to the reader to appreciate the trend-setting nature of the studies.

Most of the psychophysical studies of vision have been carried out in humans, and almost all of the relevant microelectrode recording studies have been carried out in animal models, especially awake, behaving monkeys.<sup>1</sup> Relatively little is known about the neuronal mechanisms of high-level perception in "lower" vertebrates, such as rodents, birds, zebrafish, and invertebrates such as insects. This happens to give us a convenient excuse to largely sidestep these animal models and the attendant thorny questions about the extent to which such animals can be thought to have highlevel visual cognition in the conventional sense of the term (240, 243, 253, 302). It is worth remembering, however, that such animal models probably represent a major future trend in neurophysiological research, both because these models are more amenable to genetic manipulations, and because it is becoming increasingly difficult to carry out neurophysiological research in monkeys and other "higher" mammalian species.

#### What is high-level vision?

"High-level vision" is not a specific process. Rather, it is a term of art that vision scientists use for describing a loose collection of processes that help the viewer to understand, and operate in, the visual environment (95, 337). Low-level vision involves measurements of basic properties of the retinal image. For instance, measuring the basic properties of the abstract painting in Figure 1, such as the color or brightness



Figure 1 The abstract painting *Interchanged* (1955) by Willem de Kooning. At a reported purchase price of US \$300 million, it is one of the most expensive paintings in the world. See text for additional details.



Figure 2 The importance of recurrent processing in visual perception. When viewed for the first time, this two-tone "Mooney" image appears to be an unrecognizable pattern of black and white blobs. That is, it is hard to interpret this image based on sensory information (i.e., bottom-up or feed-forward processing) alone. However, after viewing a full grayscale or color counterpart (for which see Figure 3A—which, being smaller in size, has no pixel-to-pixel correspondence with the above image), the Mooney image becomes easy to interpret. Note that viewing the disambiguating image rapidly, drastically, and enduringly alters our perception of the Mooney image, although the Mooney image itself remains physically unchanged. Feed-forward theories of vision cannot help explain such phenomena. Recurrent (or "reentrant") neural signals bring to bear such top-down influences as prior knowledge and the behavioral context to help constrain the interpretation of the visual image. Learning to interpret Mooney images can be understood as an extreme case of knowledge-mediated disambiguation that is part and parcel of normal visual perception (51, 82, 125, 142).

of various image regions, the contrast, or differences, of color, brightness, etc. between different parts of the image, the orientation and curvature of the various edges and surfaces, etc., are typically thought of as low-level visual processes. By contrast, recognizing the artistic message of the painting involves high-level visual processes. From psychophysical and neurophysiological viewpoints, low-level visual processes have been much better understood so far compared to high-level processes.

Of course, from a visual information processing standpoint, real world is far more complex and dynamic than the image in Figure 1, with multiple objects, often moving and dynamically interacting with each other and the viewer. To operate in the real-world, we must make sense of it, which is domain of high-level vision. Low-level processes of making relatively simple, spatially localized measurements of image properties are necessary, but are not sufficient.

# All our complex visual abilities are functions of high-level vision

Figures 2 through 6 illustrate some operating principles of, and computational issues in, high-level vision. For instance, consider the two-tone image (often referred to as a "Mooney" image) in Figure 2. At first glance, there appears to be nothing more to the image than just a pattern of black-and-white splotches. But take a quick look—as brief a glance as you can make it—at the image in Figure 3A and look back at Figure 2.

The black and white splotches probably will make some sense now. Repeat this process. You will probably see more in Figure 3A with each glance. Note also that your understanding of Figure 2 also evolves over time, and that this happens because your brain brought to bear the information it gleaned from Figure 3A to disambiguating Figure 2. This is a very rapid, "online" form of perceptual learning (or sensory learning; also see below), and it helps illustrate the fact that visual images are fundamentally ambiguous, and we use our prior knowledge of the visual world to help make sense of what is in front our eyes.

For both images, your understanding of the image evolved with time, even though the image itself remains unchanged. This helps illustrate the principle that vision is neither a static nor an instantaneous process, but rather a dynamic process where our understanding of the scene evolves over time (over the course of a few hundred milliseconds, as we will see below). It also straightforwardly demonstrates that vision is not a deterministic process either, where the input—in this case, the visual image—solely determines the outcome. If this were the case, the percept would have remained the same while the stimulus remained unchanged.



Figure 3 Some complexities of natural images and the information processing required to making sense of them. Panels A through D show real-world scenes that help illustrate some of the complexities of such scenes. For instance, there are multiple objects of the same type in each picture. They all vary greatly in image size, illumination, shadows, occlusion, viewpoint, etc. But the visual system recognizes them as objects of the same kind. In other words, the brain must be able to discard a variety of image features as irrelevant to recognizing the images. Note that, in case of each scene in this figure, our understanding of the scene evolves over time. Note also that the objects, their spatial relationships, and even the semantic similarities and differences among them are such that the percept that each scene elicits is more than just the sum of the parts. In addition, each picture contains implicit cues to motion and/or depth that static, 2D pictures such as these do not do justice to. In fact, natural scenes differ from each other and from the relatively simple stimuli used in many a vision study, such as a sinusoidal grating on a neutral gray background, in myriad ways. Studies have shown that the visual system is adapted to contend with the statistics of natural scenes (307), so that neural responses to "artificial" stimuli may provide a substantially different and potentially misleading picture of how the brain works. On the other hand, note that our visual system performs very well in recognizing scenes with "unnatural" statistics, such as the street scene in panel D, even though it evidently did not encounter the "unnatural" statistics of human-made objects until quite recently on the evolutionary time scale. This is because the brain is quite good at adapting to a variety of nonoptimal inputs.

In many real-world visual scenes, information about an object of interest is diminished because of poor lighting or shadows (see, e.g., Fig. 3A) or is simply missing, due to occlusion (Fig. 4A). Indeed, occlusion is ubiquitous in the real world. But the visual system generally has little trouble recognizing familiar objects from just a few tell-tale cues in the image (Fig. 4A-C). This is because the brain can bring to bear what it knows about the visual appearance of objects in the world (Fig. 4A and B) and/or what it remembers about its experiences (Fig. 4C) to make up for what is missing in the image. Obviously, this would not happen if visual perception were purely image-driven.

Visual images can also elicit powerful emotions (Fig. 5). While we are not accustomed to thinking emotions as the business of vision researchers, the inescapable fact is that the ability to experience humor, beauty, love, hatred, or horror are a part of what makes us human and emotions self-evidently influence our behavior. That is, the brain's high-vision faculties are intricately liked to its affective faculties.

Figures 2 through 5 also help illustrate some of the additional complexities of real-world scenes that high-level visual processes must contend with (also see (36)). The visual images in these figures, while themselves static, point to the fact that visual scenes change dynamically, because objects in the scene tend to move, as does the viewer (and the viewer's head and the eyes, all potentially independent of each other), which in turn changes the retinal image. The order in which the viewer's eyes move to the various regions of the image



Figure 4 Visual system can often recognize objects with great precision with very little information. A veridical description of the object, even if it were possible, is not necessary. (A) In the picture on the left, very little of the dog is visible. But we have little trouble recognizing the dog. Many probably can even readily name the breed of the dog. (B) Few would have trouble recognizing the person from the picture. (C) The tell-tale hat. For most readers, this picture simply shows a hat hung on some plumbing. But this famous picture, on the cover of the very first issue of *Physics Today* in May 1948, showed the famous "pork pie" hat of J. Robert Oppenheimer, the father of the atomic bomb. Therefore, the picture was highly topical and was readily understood at the time. But for most of us living today, this picture requires either some historical knowledge or explanation. Thus, our perception of a visual image is influenced by a great many nonvisual factors. Neurophysiological understanding of such influences on cognition represents a major challenge.

adds another level of complexity, because it changes the temporal sequence of information processing. For instance, in case of Figure 5, the eyes are initially likely to move to visually "salient" regions of the image, such as local areas with high luminance contrast (86, 332), and it is likely to take additional scrutiny of the image for the larger import of the vulture to "sink in." Oftentimes, there are multiple objects in the scene, where one object may be of current behavioral interest (i.e., "target") and other objects are of lesser behavioral interest ("distractors"). It is known that, even when the distractors do not occlude or otherwise interact with the target, their mere presence in the vicinity can influence whether or how we perceive the target, a phenomenon known as crowding or



Figure 5 There is more to vision than feed-forward processing. This picture of an unknown little girl was taken by South African photojournalist Kevin Carter in the famine-stricken South Sudan in March 1993. The girl had reportedly collapsed from weakness on her way to a United Nations food center. (Reproduced, with permission, from *The Vulture and the Little Girl* by Kevin Carter. Pulitzer Prize for Feature Photography, 1994.) Note that feed-forward processing by itself would completely miss the import of the picture.

clutter (157, 202, 251, 252, 336, 360).<sup>2</sup> In some cases, the visual appearance of the target object itself varies from time to time, owing to changes in viewpoint, viewing distance, illumination, or other viewing conditions. The visual system can typically recognize a given object regardless such incidental changes in its visual appearance, a phenomenon known variously as object invariance, perceptual constancy, or invariant recognition (26). Since incidental variations in viewing conditions are common and unavoidable, organisms simply cannot survive in the real world without invariant recognition.

# What machine vision reveals about biological high-level vision

The neurophysiological mechanisms of object invariance remain poorly understood, but it is clear that the underlying computations are dauntingly intricate (159, 275). Nonetheless, advances in machine vision have devised computers with nontrivial, real-world object invariance. A notable example of this is self-driving cars (see Fig. 6A). Briefly, such machines learn the properties of the real world by a process called deep learning (DL; for a list of abbreviations, see Table 1). In DL, a computer program, typically implemented as some type of neural network, learns the abstract statistical patterns in real-world scenes from a large set of examples where the various objects in the scene are labeled (such as "parking lot," "traffic cone," "curb," "wheel stop," etc. in Fig. 6A). Note that this process is directly analogous to how we learn to recognize objects and operate in the world during development. Computers trained in this fashion can recognize realworld objects invariantly enough to operate robustly in the real world (173, 193, 199, 275). In doing so, such machines help illustrate the computational problems that biological systems must solve to operate in the real world.

There are a large number of other instances where machines can operate with great success in the real world, that is, such machines have many, although not yet all, of the faculties of high-level vision (199). The current successes of machine vision suggest that we are on the right track. After all, if we can broadly mimic the neural processes in these computers to produce high-level visual performance in the real world akin to that of humans and other visual animals, we must be doing something right.

The current shortcomings of intelligent machines also help highlight some additional principles of biological highlevel vision. Almost all of these shortcomings have to do with one or both of following: First, successful machines tend to be hyperspecialized, and are typically unable to generalize beyond a very limited realm of operation. For instance, while self-learning cars can generalize across a vast array of traffic scenes, they are currently limited to operating in conventional road traffic conditions; for instance, they tend to fare poorly at off-road driving (16). IBM Watson trained to compete in the Jeopardy! Television game show cannot be used to detect anomalies in medical images, and vice versa (53, 247). This inability to generalize is shared by all current machine vision systems, a limitation that humans do not have to the same extent. Many modern websites exploit this fact to distinguish humans from machines using the CAPTCHA login tests (Fig. 6B; see legend for details) (121, 326).



Figure 6 (A) A self-driving car. (B) Two instances of the CAPTCHA (or Completely Automated Public Turing test to tell Computers and Humans Apart) internet device designed to prevent automated logins (121, 326). In the example on the *left*, the website asks the user to perform a straightforward object categorization task, namely distinguish pictures of people wearing glasses from pictures of people without glasses. This tends to be quite successful in preventing machines from logging on to the website. However, this success is not so much because it would be all that difficult nowadays to train a suitably designed computer program to distinguish the two categories of human faces (see, e.g., (2, 191)), but essentially because, at present, such programs tend to be hyperspecialized, and tend not to generalize beyond their training sets to other visual objects or tasks (121, 199). That is, such an "intelligent" program, once trained to tell aforementioned types of faces apart, can be readily stumped at present by rather slight changes in the underlying categories (panel B, *right*), or the task (not shown) (2, 191). However, recent research suggests that the problem of overspecialization of intelligent machines is likely to straightforwardly surmountable, so that CAPTCHA strategies such as the one illustrated in this panel are unlikely to be effective for long. That is, machine vision is getting ever better at mimicking human high-level vision (199).

#### Table 1 List of Abbreviations

2-DG, 2-deoxyglucose	LOC, lateral occipital complex (human)
2 <sup>1</sup> / <sub>2</sub> D, 2 <sup>1</sup> / <sub>2</sub> dimensional	MEG, magnetoencephalography
3D, three-dimensional	MFC, medial frontal cortex
AIT, anterior infereotemporal	MR, magnetic resonance
BM, BrainMapping (database)	MT, middle temporal area
BOLD, blood oxygenation level-dependent	MST, medial superior temporal area
Cal S, calcarine sulcus	MTL, medial temporal lobe
CAPTCHA, Completely Automated Public Turing test to tell Computers and Humans Apart	MVPA, multivoxel pattern analysis
CC, cortico-cortical pathways	
CIT, central inferotemporal area	OFC, orbitofrontal cortex
Col, S, collateral sulcus	PET, positron emission tomography
CP, choice probability	PFC, pretrontal cortex
cpd, cycles per degree	
CRF, classical receptive field	
CS, central sulcus	KGC, refinal ganglion cells
CTC, cortico-thalamo-cortical pathways	rCBF, regional cerebral blood flow
DCM, dynamic causal modeling	rstMKI or K-tMKI, resting state functional magnetic resonance imaging
DF, dorsal foci (human)	KSN, resting state network
dMFC, dorsal medial frontal cortex	IE, temporal area
EEG, electroencephalogram	TEO, temporooccipital area
ERMF, event-related magnetic field	IMS, transcranial magnetic stimulation
ERP, event-related potential	IP, tempoparietal junction
FD <sub>A</sub> , dorsolateral prefrontal region	SNr, substantia nigra pars reticulate
FDv, ventrolateral prefrontal region	SNc, substantia nigra pars compacta
FEF, frontal eve fields	SOA, stimulus onset asynchrony
FFA, fusiform face area (human)	STN, subthalamic nucleus
fMRI, functional magnetic resonance imaging	STS, superior temporal sulcus
GC, Granger causality	STP, superior temporal polysensory area
GPe, Globus pallidus, external portion	SVM, support vector machine
GPi, Globus pallidus, internal portion	IE, temporal area (macaque)
ICA, independent components analysis	IE, temporo-occipital (macaque)
IOG, inferior occipital gyrus	IP, temporo-occipital junction
IPS, intraparietal sulcus	V1,V2, V4, or V5, visual area 1, 2, 4, or 5
IT. inferior temporal/inferotemporal visual area (macaaue)	VEF, visually evoked potential
ITC, inferior temporal/inferotemporal cortex (human)	vMFC, ventral medial trontal cortex
Lateral S. lateral sulcus	VSD, voltage-sensitive dye
LFP, local field potential	VIA, Ventral tegmental area
LGN, lateral geniculate nucleus	TAH, "you are here"
-	

The second main current shortcoming of intelligent machines is that they do not have the same high-level semantic understanding of scenes as we do. This is an important limitation, because as noted above (e.g., Fig. 5), semantic understanding of the scene is intimately related to high-level vision, and vice versa. There is broad agreement that when we recognize an object, we assign it, however implicitly or abstractly, a category label (126, 138, 276). While machine systems do associate objects with the corresponding labels, indeed, they learn from training examples in which objects are suitably labeled—they cannot be said to understand the *meaning* of the labels.

The assertion that intelligent machine do not truly understand the meaning of words may seem surprising, because "intelligent assistants" such as Siri, Alexa, Cortana, etc. respond adequately a large diversity of verbal commands and questions, as though they truly understand language. But this is a mistaken impression; the intelligent machines simply learn to probabilistically associate a given word labels in the training set with a given set of likely responses. In this sense, they have no better understanding of the words than a "talking" parrot.

Parenthetically, it is worth noting that the nature of category labels in animals remains unclear. However, it is obvious enough that visual recognition amounts to some type of categorization in animals too, because they can differentiate predator from prey, preferred food from nonpreferred food, etc.

#### A brief, historical caricature of the conventional, outdated framework of vision: Vision for vision's sake

Conventional theoretical framework: Veridical representation of the visual world through bottom-up processing of the retinal image

To understand what we currently know and do not know about how vision works, it is useful to briefly review how we got here.<sup>3</sup> Until the 1980s, neurophysiological studies of vision focused, by and large, on the mechanisms of "early vision," or the processing of low-level image characteristics in the early parts of the visual pathway. While a rich body of literature existed-some of it more than a century old, characterizing high-level perception of visual objects and scenes at the perceptual level-neither neurophysiological techniques nor a conceptual framework for addressing high-level questions were available at the time. Beginning in the 1980s, advances in neurophysiological recording techniques, a better understanding of the anatomical layout of the visual processing pathway, along with a candidate computational framework for understanding visual perception as a whole, made it possible to pursue questions about high-level visual perception.

With some notable exceptions (see, e.g., (15,22,23)), most of the earlier theories implicitly or explicitly took a "vision for vision's sake" approach, wherein visual cognition was an end



Figure 7 An outline of Marr's theory of visual processing. See text for details.

unto itself. The computational framework that hugely influenced vision research in general, and visual neurophysiological research in particular, was the work of David Marr (1945-1980) and his colleagues, distilled in Marr's 1982 book *Vision* (214). The basic idea behind Marr's model was brilliant, and was well ahead of its time. But it is now clear, in retrospect, that the main thrust of his framework is not only wrong but also misleading. Nonetheless, it is important to understand this framework to have a contextual understanding of visual neurophysiological research since the publication of his book. Marr's framework of vision posits, to a first approximation, that the visual system builds a veridical internal representation of the visual world using a stepwise processing of the information in the retinal image (see Fig. 7).

The model consists of several successive levels of visual information processing, with the output of each level of processing serving as the input to the next higher level of processing. Marr's model can be understood as follows: The visual system first evaluates local, low-level characteristics, or "primitives," of the retinal image. Using these primitives as building blocks, the visual system constructs a surface level representation, or a "2<sup>1</sup>/<sub>2</sub> D sketch," of the objects in the image, that describes the local orientation of the surface (for a list of abbreviations, see Table 1). Using the 2<sup>1</sup>/<sub>2</sub> D sketch, the visual system then builds a more or less veridical 3D representation of the objects in the visual world ("3D sketch").

Marr's model is hierarchical, in the sense that it posits that vision is a more or less stepwise process. Marr's model is essentially a feed-forward or "bottom-up" model of vision, in the sense that it conceives of vision as driven by the sensory input, where the information entering the bottom level of the processing hierarchy (i.e., the retina) travels up the various levels of hierarchy, where a more-or-less veridical representation of the real world is constructed at the highest levels of processing. This "feed-forward thinking" essentially envisions visual processing as a one-way conveyor belt, where information processed at earlier stages of visual processing is passed on to progressively higher stages of processing, so that a veridical representation of the visual world emerges. For all practical purposes, it ignores "top-down" factors, such as attention, memory, learning, etc., which can be thought of arising from the high-level centers of the brain and acting on the visual information processing at lower levels. Visual cognition is conceived as the function of information representation in the highest level of visual processing, and earlier stages of visual processing play no role in visual cognition. It paints a modular picture of visual processing, wherein the visual system initially takes apart different types of visual information, such motion, shape, etc., processes them more or less separately, and then puts it all together (see below). It does not directly contemplate how vision subserves behavior, because it does not suppose vision is meant to subserve behavior. Rather, vision is an end unto itself; its job is to construct a true representation of the outside visual world. Presumably, when the visual information is needed to guide some behavior, the brain systems that subserve the given behavior requisition the visual system for the information.

#### Conventional view of neural mechanisms: Feed-forward processing of visual information

How does the brain actually implement the above scheme of information processing? Two broad sets of neurobiological findings seemed to fit well with the hierarchical, modular nature of Marr's model. Almost all of these insights came from animal research, mostly in macaque monkeys, but also from cats and other animals.

The first set of findings is more neuroanatomical than neurophysiological. Starting in the 1980s, it became clear that the primate visual cortex was not a monolithic expanse of gray matter. Rather, it was organized as a number of distinct visual areas, more or less clearly distinguishable from each other based on a variety of principled criteria (184, 342).<sup>4</sup> By the time a few dozen of these areas were identified, it also became clear that the primate visual system is organized as an anatomical hierarchy (for overviews, see (98, 344)) (Fig. 8).

Briefly, as articulated by Felleman and Van Essen in their highly influential work on the topic (98), visual anatomical hierarchy refers to the fact that when one keeps track of which visual area receives feedforward input from, and in turn sends output to, which other areas, and which areas receive feedback information from which other areas, one can organize the known visual areas in the brain in a stepwise fashion, that is, as a hierarchy (98, 141, 216). Since this anatomical hierarchy seemed to fit well with Marr's hierarchical model of visual information processing, it seemed (and was) reasonable to posit that the anatomical hierarchy acted as a substrate to a hierarchical processing of visual information in the primate brain.

It is important to note that the notion of the *anatomical* hierarchy itself has never been called into question. It is indisputable that, given a set of principled anatomical criteria (see (98)), the visual system can indeed be organized into a hierarchy of visual areas. Rather, the subsequent debate outlined below has been about whether the visual *processing* is also hierarchical, and whether it meaningfully parallels Marr's model.

A second set of findings, primarily neurophysiological, concerned the functional organization of visual areas. Neurophysiological experiments on this topic had started as far back as the 1960s. Until quite recently, a vast majority of these neurophysiological findings came from single electrode neurophysiology, in which a single electrode is inserted transdurally into a given brain area of interest, and extracellular potentials, or spikes, from well-isolated single neurons are painstakingly recorded one neuron at a time, often in an awake animal engaged in a visual task.<sup>5</sup> Needless to say, this necessitated studying one visual area at a time. While these neurophysiological studies provided an understanding of individual neurons with unmatched spatial and temporal resolution, it was nonetheless a keyhole view of the visual system, which made it difficult to gain a holistic understanding of the workings of neuronal ensembles in individual areas, let alone neural networks in the brain at large. These experiments showed that individual neurons in a given visual area tended to be selective to a narrow set of visual features, such as (depending on the area) orientation, color, motion, faces, etc. Thus, it seemed reasonable to posit that the visual system takes apart the various types of visual information in the image, and processes them more or less separately and in parallel in different visual areas that specialize in the processing of various visual features, so that visual processing is distributed across the various areas of the visual cortex.

These empirical neurobiological findings seemed to fit well with Marr's aforementioned notion of visual processing, wherein the visual system initially takes apart different types of visual information, processes them more or less separately, and then puts it back together into a holistic percept (137, 271, 297).<sup>6</sup> Moreover, it was reported that visual sensory information was processed in two distinct, parallel pathways (225) (Fig. 9). One pathway, referred to as the ventral pathway, occipitotemporal pathway, or the "what?" pathway, was thought to specialize in form processing, and thus played a major role in object recognition. The second pathway, referred to as the dorsal pathway, occipitoparietal pathway or the "where?" pathway, was thought to specialize in the processing of spatial information, and thus played a major role in spatial vision.<sup>7</sup> Taken together, these studies appear to support the notion that the visual system took apart the various primitives of the visual image and processed it in different visual areas, just as Marr's model predicted.

We realize, with the benefit of hindsight, that some of these scientific conclusions were colored by our collective inability,



Figure 8 Visual anatomical hierarchy in the macaque monkey described by Felleman and Van Essen, 1991 (98). Each colored rectangle represents a distinct cortical visual area. The open rectangles at top denote higher cortical areas that are not considered primarily visual area. The open rectangles at bottom denote the retinal ganglion cells (RGC) and the lateral geniculate nucleus (LGN). Lines connecting the cortical areas denote interconnections, usually reciprocal, between a given pairs of areas. The various brain regions are represented in a tiered, or hierarchical, fashion based on objective anatomical criteria, most important of which are the laminar patterns of feed-forward and feedback connections (also see (98, 216, 343)). For additional details and abbreviations, see Felleman and Van Essen, 1991 (98). An alternative formulation of the hierarchy, originally formulated by Mishkin, Gross, and their colleagues (130) (also see (78, 129)) is largely similar, but it parcels the cortex into many fewer areas and recognizes fewer interconnections. Also, some the visual area names are different in this scheme. For instance, areas AIT and CIT (anterior and central inferotemporal areas, respectively) in the Felleman and Van Essen scheme are equivalent to area TE (temporal area) in the Gross et al. scheme, and area PIT (posterior inferotemporal area) in the Van Essen scheme is equivalent to area TEO (temporoccipital area) in the Gross et al. scheme. Both schemes are used in this review, based on the scheme used by the study in question. Reproduced, with permission, from (98).



Figure 9 Our evolving understanding of the functional organization of the primate visual system. Panels A and B depict our view of the two main visual processing streams in the macaque brain in the early 1980s. (A) The dorsal and ventral visual pathways as originally formulated by Mishkin and colleagues in 1993 is denoted by *solid arrows* (OB  $\rightarrow$  OA  $\rightarrow$  PG pathway being the anatomically dorsal pathway, and OB  $\rightarrow$  OA  $\rightarrow$  TEO  $\rightarrow$  TE being the ventral pathway] (225). The continuation of these pathways to FD<sub> $\Delta$ </sub> and FDv, respectively (*dashed arrows*) represented the collective outcome of many additional studies. (B) Anatomical locations of the various regions, some renamed according to a more modern naming convention (188). (C) Key changes in the receptive field properties of neurons along the ventral pathway. Areas are color-coded as in panel B. Panel D summarizes our understanding of the same pathways some 30 years later, as summarized by Kravitz and colleagues in 2013 (188). Note that it is clear that the two pathways have turned out to be far more interconnected than previously envisioned. What is lost in terms of pedagogical simplicity is more than made up for by the nuance and granularity of this network picture, presaged decades ago by Kerrigan and Maunsell (221). Self-evidently, far more remains to be learned, including how these networks interact with other networks in the brain and subserve behavior. Adapted, with permission, from (188). Human brain (not shown) is also purported to have two evolutionarily homologous processing pathways, although there is even less empirical information to support this notion.

for technical reasons, to envisage the activity of whole brain networks using our understanding of the activity in individual visual areas—our inability to see the forest for the trees. Thus, it is not so much that these experimental findings in and of themselves were wrong, but that they missed the bigger picture.

Moreover, they unwittingly reinforced the flawed "feedforward thinking" outlined above (57, 315). The myriad of reasons why the feed-forward thinking is flawed will be outlined in context below, but suffice it to make two general observations about it here. First, the sheer prevalence of feedback and lateral connections among brain areas, and the fact that such connections tend to be conserved over the course of evolution indicates they must play functionally crucial roles in information processing (98, 164, 270, 344, 345). Second, even Marr's own field of machine vision has abandoned this purely feed-forward approach in every successful modern application of machine vision to the real world, such as self-driving cars, as alluded to above (for reviews, see (50, 73, 233)).

Both the strength and weakness of models of the real world is that they simplify its complexities. Marr's model had the additional disadvantage of being wrong. But to discard it altogether would be to throw out Marr's baby with the bathwater. The most useful lesson of Marr's unquestionably brilliant work for contemporary practitioners and students of neurophysiology is that anything worth doing in neurophysiology is worth doing within a computational framework, the more explicitly the better. It was just that Marr's model happened to be an unfortunate case where theory got too far ahead of experiments.



Figure 10 What's to infer? Is not it all there in the image? The answer is no. Retinal image is simply a 2D pattern of image intensities. Image intensities of a particular image are represented in panel A as a color-coded surface plot, where the height and color of a point denotes the image intensity at that point. Note that when the image is represented in this fashion, it makes no sense to us. But when the same image is represented as corresponding variations in image intensities (panel B), we readily recognize it as an image of a brook in the woods. However, the information in the two representations is exactly the same. The difference is that, in panel B, the image is in an input "format" that our eyes can process. Beyond that, the inferential processes to "make sense" of the pattern of intensities are exactly the same.

### A Modern Framework for Understanding Vision: Visually Guided Behavior is Statistical Inference Implemented by Brain Networks

In recent years, it has come to be broadly appreciated that the retinal information is inherently ambiguous, and makes little sense without further processing and interpretation by the brain. That is, with the exception of a few, exceedingly simplistic laboratory stimuli that can be safely neglected for our purposes (35,213,283,361), it is impossible to "reverse engineer" the retinal image to faithfully recover actual the visual scene that gave rise to the retinal image. This is because under real-world conditions, for any given retinal image, there are many (oftentimes an infinite number of) real-world scenes that could have generated the image, and it is impossible to identify the unique scene that generated the given image based solely on the retinal image. Therefore, seeing is often referred to an "ill-posed" problem (258). The aforementioned complexities of the image make solving the problem so much harder.

One simple way to intuitively appreciate why the retinal image is not "self-explanatory" is to think about the nature of 3D perception. The retinal image is a completely a flat, 2D image, in which the third dimension (i.e., depth) is missing. If the visual system treated the retinal image as self-explanatory information that needed no further interpretation, we would perceive the world as a flat, depthless entity. Yet, this is not what happens; we indeed perceive the world around us in vivid 3D. This because the visual system implicitly recognizes that the retinal image, while it is itself 2D, contains a variety of implicit cues about depth, and exploits these cues to infer the third dimension. It contains a variety of implicit cues about the third dimension, that is, depth (153). But depth information must still be inferred from the implicit depth cues. Thus, the very fact that we see in 3D at all demonstrates the fact that the brain is an "inference machine" that does its best to make sense of, or interpret, incoming sensory information that is inherently ambiguous. For another demonstration that retinal image makes little sense without proper interpretation, see Figure 10.

In recent years, the aforementioned conventional framework has been largely supplanted by probabilistic inference as the theoretical approach of choice for studying and understanding brain function in general (for overviews, see (80, 83, 178, 365)). This contemporary view, that perception and action are essentially various forms of statistical inference, had its beginnings in the notion of "unconscious inference" that Hermann von Helmholtz (1821-1894) articulated with astonishing prescience (108). The modern probabilistic framework simply mathematically codifies Helmholtz's brilliant intuition. Given its widespread and increasing use in vision research, it is important to gain an intuitive understanding of these approaches.

The quantitative underpinnings of probabilistic ultimately lie in the classical Bayes' law of conditional probability. Since the Bayes' law is the mathematically correct way of describing any given set of conditional relationship among probabilistic variables, the term "probabilistic inference" in general has come to be synonymous with "Bayesian inference" (80, 83, 178, 365). The variables that underlie brain function are inherently probabilistic because of the probabilistic, or "chancy" nature of real-world phenomena, the ambiguities



Figure 11 Visual illusions help illustrate the inferential nature of visual perception. They also demonstrate that the inference need not be a deliberate, volitional, or conscious process. (A) Ames room. Two people stand in opposite corners of the room. One appears to be much taller than the other, even though they are roughly the same height. Instead, it is the room that is distorted to produce this illusion. Picture courtesy of Ian Stannard, Flickr. Reproduced with permission. (B) Hollow-mask illusion. This picture shows the front of the mask (*right*) and the hollow beck of the mask (*left*). Nonetheless, both look like normal, convex faces. For a video demonstration of the hollow mask illusion, see https://www.youtube.com/watch?v=sKa0eaKsdA0. Note that, in this video, the mask appears to flip its direction of rotation at the same time the other side of the mask begins appearing. Theoretical studies show that visual illusions are often perfectly rational inferences given the evidence (see, e.g., (108, 357)).

in the sensory information, and the inherent noisiness of the neural hardware (118). For this reason, it is widely appreciated that the probabilistic approach is the correct quantitative framework for formulating the inferential problems that confront the brain.<sup>8</sup>

The Bayesian formulation of visual perception is that what we think of seeing is really a process by which the brain infers what is "out there" in the external world based on all the information it has, including the information that comes in through the senses such as vision. In other words, visual perception is the process by which the brain comes up with an interpretation of the situation at hand and a course of action that best fits all available information at hand, including the sensory information, its prior knowledge of the nature of the external world,<sup>9</sup> and possible consequences of various actions or lack thereof. Thus, this framework implicitly recognizes that it is neither possible nor, for that matter, necessary for the brain to construct a veridical internal representation of the external world. The brain can and must do its best with the probabilistic and ambiguous information it has. It also implicitly recognizes that visual perception is a probabilistic process, and not a deterministic process where a certain set of conditions, such as a given input image, foreordain a given perceptual outcome (178, 185). Figures 1 through 6 illustrate some of the key visual phenomena that a strictly deterministic feedforward model cannot explain, but the Bayesian framework straightforwardly can.

This is not to say that a given inference is necessarily a deliberative, volitional, or even conscious process, or that the "purpose" of vision is to just make optimal inferences about the visual sensory input. It is important to remind ourselves that biological sensory systems exist in the service of behavior, regardless of whether the behavior itself is consciously planned and executed. For instance, when we open the door by reaching for and turning the knob—a process that is well captured by a Bayesian explanation—we do not consciously infer the shape or position of the knob or calculate the force vector needed to turn the knob. We just turn the knob without being conscious of the underlying calculations and the resulting inferences.

It is also worth noting that the brain's interpretation of the information at hand is not always correct. Visual illusions help illustrate the fact that inferential process can be "gamed" to produce counterfactual percepts (see, e.g., Fig. 11). In fact, the underlying inferences can be so strong that it can be hard to consciously override or "veto" the brain's inference of the data at hand, even when we know that the given percept is illusory or incorrect (see, e.g., Fig. 11). This helps illustrate that the inferential process is neither a primarily conscious process nor is it always swayed by conscious inferences. The Bayesian framework also accounts for how the brain "falls for" visual illusions (178, 185). It turns out that illusions occur because the brain tends to attach more weight to sensory informationwhich, in the case of illusions, is not only ambiguous but also misleading-than to the top-down cognitive information that can clear up the ambiguity. Thus, the brain goes with the illusory percept, because it is the interpretation that best fits the sensory data at hand. Thus, the brain is by no means a hyperrational system.

To get an intuition of these aspects of brain function, it is useful to think of it as akin to the day-to-day functioning of the US government. It is a huge and complex system whose behavior can be understood by a set of more or less sensible operational principles. In most situations, it more or less does the right thing, except when it does not. It does things that may appear smart or stupid, but it usually does not do so consciously, nor is it inclined to change its ways in light of disabusing information, however correct. Note that the Bayes framework, in principle, addresses the computational shortcomings of the aforementioned feedforward framework, in that it explicitly allows for the system to evaluate all available bottom-up and top-down information in a dynamic fashion. Moreover, it does not necessarily limit the system to a single pass through the process, as the feedforward model implicitly does. In fact, it is straightforward to envisage scenarios where the system makes multiple passes through the process, testing and refining hypotheses about the possible outcomes of the neural computation during each pass using an "analysis by synthesis" approach that explicitly allows for feedback and lateral inputs to the process (365).

Finally, it is worth noting that the term "Bayesian framework" by itself does not refer to not a particular hypothesis, model, or theory of how the brain works; it is simply a term of art that describes a set of mathematical methods using which such hypotheses, theories, and models can be constructed and empirically tested. In other words, it is not that brain implements the Bayesian framework, but that what the brain does is best understood and studied in terms of the Bayesian framework. Bayesian inference is not an implementation plan for the brain, but a tool, or "language," to describe brain function.

# Neural computations are implemented by dynamic interactions among brain regions

How, then, does the brain implement this probabilistic inference? Recent findings provide increasing support for a connectomic view, which posits that brain functions are implemented through the dynamic changes in the connectivity within and across brain regions organized as networks. In other words, to understand how the brain works, we must understand not only how individual brain regions process and represent (i.e., carry or convey) information but also how different brain regions "talk" to each other as parts of a larger brain network, and what they are telling each other (for overviews, see (295, 314)). It is increasingly clear that understanding the various brain networks and the patterns of network interaction, or connectivity, is a key to understanding brain function.<sup>10</sup> To help understand how brain interactions are studied, it is helpful to understand the two different types of connectivity that are commonly reported in neurophysiological studies.

# Functional versus effective connectivity among brain regions

In addition to the methods that can be used for characterizing the anatomical (i.e., structural or physical) connections among brain regions, methods also exist to characterize connectivity based on the activity of brain regions. As noted above, two types of activity-based connectivity can be distinguished. *Functional connectivity* is defined as "statistical dependencies among remote neurophysiological events" (104). Neurons or neuronal ensembles can exhibit functional connectivity for many different reasons. For instance, the initial, predominantly stimulus-driven response transients<sup>11</sup> tend to be correlated across nearby neurons, because they represent simultaneous activation elicited by the same underlying stimulus. Response correlations also reflect shared changes in activity mediated by anatomical connections. Intrinsic connectivity inferred by resting state functional magnetic resonance imaging (fMRI) (see below) is a form of functional connectivity (104).

*Effective connectivity* refers "explicitly to the influence that one neural system exerts over another, either at a synaptic or population level" (1, 104, 314). Connectivity revealed by task-related fMRI is a form of effective connectivity. Both functional and effective connectivity imply a level of synchronization of various brain regions.

Different neurophysiological techniques tend to report synchronization at slightly different time scales. fMRI studies tend to report synchronization at fairly slow time scales (hundreds and thousands of milliseconds). Whether and to what extent this synchronization is related to the response synchronization of neurons and neuronal ensembles at much faster time scales (a few milliseconds) reported by microelectrode studies remains unclear (see (41, 297)).

To help make the distinction between the two types of connectivity clearer, it is useful to consider the following analogy. Imagine you are at a large swimming pool. The pool is strewn with many buoys, many of which are tethered together in small subsets. There are also many swimmers swimming in the pool. You are trying to understand how the waves are traveling in the swimming pool by studying the undulations of the buoys and the swimmers. You are analogous to the neurophysiologist trying to understand how information flow causes undulations of neuronal responses in different brain regions. In the swimming pool, you observe that certain subsets of buoys (think brain regions) tend to bob up and down in sync with each other, but out of sync with other subsets of buoys. This correlated undulation occurs because the given subset of buoys happens to be tethered together, or riding the same waves, or both. The buoys did not cause each other's undulations. Rather, their undulation is caused by an unrelated, common source. In other words, their correlated activity does not imply causation. Thus, the buoys that undulate together are functionally connected.

But you also observe that the undulations of specific subset of buoys is correlated with the strokes of a swimmer (think a different combination of brain regions), typically because (in the present example) they happen to be close to the swimmer. The given subset of buoys is effectively connected to the swimmers' strokes, because the activity of the latter is causing the activity of the former. This analogy can also help make explicit two additional aspects of the above definitions. First, the two types of connectivity represent functional definitions that neuroscientists use to characterize the observed patterns of activity, and do not represent biological mechanisms *per se*. Second, effective connectivity is a stricter concept that subsumes the concept of functional connectivity, in the sense that regions that are effectively connected are also functionally connected (but not necessarily vice versa). This also means that the contributions of one or both of types of connectivity to a given set of raw neural responses have to be mathematically teased apart. Fortunately, many user-friendly toolkits that can do this are currently available (24, 99, 267).

### A Word to the Wise: Nothing in Neuroscience Makes Sense Except in Light of Behavior

Given the framework outlined above, "How do we see?" is really a misleading question, both because it implies that vision is a single process with a unitary explanation, and because it implies that seeing is an end unto itself, that vision exists for its own sake. In light of the current data, broadly summarized in the following sections, it makes more sense to ask how vision serves behavior. It also makes more sense to posit that visual perception and the related phenomena of visual awareness or consciousness reflect read-outs of the network states or, in more cognitive terms, introspection about the information being processed.

Theodosius Dobzhansky's famous dictum that "nothing in biology makes sense except in the light of evolution" (81) indisputably applies to neurobiology as well. Behavior is the phenotypical, or outwardly, manifestation of neural systems; it is the substrate upon which evolution operates. Therefore, a useful reformulation of Dobzhansky's principle for the neuroscientist's pocketbook is that nothing in neuroscience makes sense except in light of behavior.

In trying to study, understand, and explain the structure and function of neural systems, it is important to bear in mind that neural systems have evolved to help the organism function in its environment, that they are often optimally adapted to the statistical properties of that environment, and that they were not designed to subserve a computational end, in the way computational machines are designed to serve computational ends. Of course, this is not to say that computational analyses are not useful to neurophysiologists. Quite the contrary. Essentially all of the major advances in our understanding of high-level vision have come from neuroscientists embracing the computational sciences. An illustrative example is the ideal observer analysis, which determines performance of a computationally optimal observer in a given information processing task, say, discriminating between two objects, estimating the direction of illumination of scene, or estimating the optimal trajectory of visually guided hand movement (for reviews, see (12, 106, 108, 178, 186)). It is generally informative to compare the performance of ideal observers with that of actual observers, human or otherwise. It turns out that in many cases, the performance of the human brain hews very close to the computational ideal, but in many other cases it does not. In fact, deviations of actual observers from the computational ideal tend to be even more informative, because

one has to account for why the neural system deviates from the computational ideal. But in doing so, it is always imperative to keep in mind the fact that neural systems are evolved systems, not designed systems (107).

#### Neural systems have evolved primarily to subserve the four F's

To paraphrase the immortal formulation of Paul Maclean (as recounted by Patricia Churchland in (56)), animals who are good at the four fundamental F's—feeding, fleeing, fighting, and, um, reproducing—tend to fare well from an evolutionary point of view.<sup>12</sup> Since better neural systems make for better F's, neural systems have evolved to help deliver better F's. That is, neural systems that better subserve the animal's behavioral goals tend to be more successful over the course of evolution.

The aforementioned evolutionary principles collectively have two main practical imports for the neurophysiologist: First, they highlight the fact that the ultimate "design principle" by which the brain has come to exist and function is evolution, and not necessarily computational performance per se (see, e.g., (349)). We will revisit this principle below in the context of efficient coding. Second, they also highlight the fact that, while the human brain bears an evolutionary relationship with the brains of other animals, there is no sense in which the human brain is more evolved than the brain of any other living being. In other words, evolutionary ladder, that is, the notion that human being represent the pinnacle of evolution and the "lower" species are correspondingly less evolved, is a myth (105, 257). Neither is brain evolution a goal-driven process whose "goal" is to produce brains that are capable of increasingly complex computation (163, 322). Rather, brains capable of more complex computations, such as human brains, sometimes spread in numbers. This caveat is helpful, because it keeps us from mistaking the human brain as somehow an improvement over, say, contemporary monkey brains, and monkey brains as an improvement over contemporary rodent brains. Of course, there are many objective measures by which human brains are more complex than monkey brains, which in turn are more complex than rodent brains, but evolution does not necessarily entail increasing complexity, or vice versa (85, 92, 249, 327). In summary, our focus on the human brain, while eminently principled and justifiable on other grounds, can amount to rank speciesism when carried too far.

#### Another Word to the Wise: Vision is Not a Unitary Process but a Collection of Processes

One of the assumptions in the question "How do we see?" is the assumption that there is a single answer. But there is no reason to believe that vision is a single process in any

principled sense of the term. To cite an admittedly extreme "proof of principle" example, in blindsight, subjects respond to visual stimuli without consciously perceiving them. Indeed, human subjects with blindsight often insist they are blind. When the striate cortex (area V1) is removed or blocked in monkeys, the animals can still accurately perform visual discrimination tasks (4,61,280). This is because the image information from the retina also reaches many other brain regions in the midbrain and thalamus that remain intact even when V1 is removed in its entirety (4, 61, 280, 364). Obviously, these monkeys "see" by any principled definition of seeing, but the process by which they see is radically different from the way monkeys with intact V1 see. Obviously, vision is a many-splendored process that works in more ways than one.

One can cite any number of less extreme examples from various branches of vision research to help illustrate the fact that the nature, and neural mechanisms, of vision differ significantly based on the task, visual stimulus, subject's behavioral state, level of learning, etc. For instance, the information that must be gleaned from the image in Figure 3A will be drastically different based on whether the task is to wade into the water hole in the foreground, drink the water, determine the ages and gender of the animals in the scene, or search for berries in the vegetation in the background. A single-purpose, "one-size-fits-all" process will not only be *inefficient* for all tasks but it is likely to be *insufficient* for any given task. Note also that for each task, our visual system tends to discard (or at least disregard) a whole lot of task-irrelevant information. For instance, if the task is solely to drink the water from the water hole, anything beyond the water hole is unlikely to register very keenly. Since the nature of task-relevant and taskirrelevant information will vary from one task to the next, the system that consistently discards some aspect of scene information will be caught wrong-footed when the discarded information suddenly becomes task-relevant. For reviews of these ideas, see (57, 58, 359).

The notion that vision is many processes, not just one, is important, because it straightforwardly implies that there are multiple underlying corresponding neural mechanisms. Of course, the various visual processes and the underlying mechanisms share many fundamental similarities, some of which will be examined below.

### Rapidly Evolving Repertoire of Neurophysiological Techniques has Changed the Face of Neuroscience

The spatial and temporal resolution afforded by microelectrode recording is matched by no other neurophysiological technique. But until rather recently, microelectrode neurophysiology had a limited spatial reach, and was largely limited to studying a few neurons in a single brain region at a time. Advances in technology now make it possible to simultaneously study the activity of multiple neurons and neuronal ensembles in multiple brain regions (Fig. 12). However, it is still not possible to carry out whole brain studies using these methods, which can make it hard to see the forest for the trees. Moreover, since microelectrode neurophysiology is invasive, it is largely limited to being carried out in animals. Under rare circumstances, they can be carried out in suitably anesthetized human surgical patients who are undergoing unrelated brain surgery for the medical benefit of the patient (for reviews, see (90,231)).

#### Local field potentials carry much information about the responses of, and interactions among, neuronal ensembles

Local field potential (LFP) represents the summed electrical activity of multiple nearby neurons within a relatively small volume of neuronal tissue, typically 50 to 350 µmol/L from the tip of the recording electrode (200). It is important to note that the extracellular potentials, or spiking activity, make only a limited contribution to LFPs. It is thought that LFPs are generated by synchronized synaptic currents arising from cortical neurons (17, 239). LFPs are somewhat similar to EEG (electroencephalography) signals in this regard (269). Because the information carried by spikes and LFPs are mutually quite nonredundant, and because collecting both types of data is often quite straightforward in modern electrophysiological setups, it is generally quite useful to look at both whenever possible.

# Whole brain imaging techniques provide a better view of the forest, but not so much of the trees

One of the main advantages of whole brain imaging techniques is that they can be carried out noninvasively, and afford a view the whole forest. Just as important is the fact that, since they are noninvasive, they can be carried out in human subjects performing complicated tasks and making detailed and nuanced perceptual reports. On the other hand, whole brain imaging techniques have significant limitations of their own. EEG uses the electrical activity measured at the level of the scalp to infer the underlying brain activity. It is one of the least expensive of the whole brain imaging methods, and offers a very high temporal resolution, on the order of milliseconds. However, the spatial resolution of the EEG signals (referred to as evoked response potentials, or ERPs) is quite poor (see (208)). Magnetoencephalography (MEG) is analogous to EEG in the limited sense that it also measures brain activity through sensors mounted on the scalp, but differs from EEG in that it can pick up only a subset of the neural activity detected by EEG (for reviews, see (136, 158, 246, 291)). The temporal resolution of event-related magnetic fields (ERMFs) is comparable to that of ERPs. But MEG has better spatial resolution, especially at greater depths (136, 158). MEG and EEG both suffer from the "inverse problem," which refers inferring the underlying brain activity by working back from the observed, ambiguous signals. This is somewhat like



Figure 12 The spatiotemporal domain of the methods available for the study of the functional organization of nervous system in 2014, compared to the methods available in 1988 (*inset*). Each *colored region* represents a range of spatial and temporal resolutions for a given method. *Open regions* represent measurement techniques; *filled regions*, perturbation techniques; EEG, electroencephalography; MEG, magnetoencephalography; PET, positron emission tomography; VSD, voltage-sensitive dye; TMS, transcranial magnetic stimulation; and 2-DG, 2-deoxyglucose. Redrawn, with permission, from (290).

listening to the din of a crowded room behind closed doors, and trying to figure out who is doing the talking, what they are saying, and where in the room they are. Recent advances in data analysis techniques have now made this problem quite tractable, and MEG and EEG often yield results that can be validated by independent methods (136, 158).

fMRI measures neural activity indirectly, through changes in the blood oxygenation level-dependent (BOLD) activity caused by neural activity. It has a spatial resolution of up to a millimeter, but has a relatively poor temporal resolution of a few seconds. It might seem that this temporal resolution is too low to be useful for measured neurophysiological activity which, after all, occurs at millisecond levels. Nonetheless, as we will see below, much has been learned about brain function using fMRI by itself or in combination with other techniques (see (21, 207, 281)).

In addition to major improvements in the techniques for collecting neuronal data, there have been major advances in the field of *analyzing* neural data as well (for overviews, see (34, 63, 237, 366)). "Big data" approaches are making a huge difference in our understanding of brain function, and the role of informatics is going to broaden going forward. Moreover, the future of neurophysiology is increasingly multidisciplinary and, by necessity, collaborative. For instance,

the neurophysiologist will likely benefit from collaborations with computational neuroscientists to help pose the question and model the data, biostatisticians to help analyze the data, etc. But professional success in neurophysiology will mean a working understanding of all these fields, not in the least to develop and maintain effective collaborations.

# Visual information processing involves two types of computation: Estimation and categorization

From a computational viewpoint, information processing involved in vision can be described in terms of one or both of the following two fundamental types of computation: Estimation and categorization. Estimation is the calculation of some continuous, or analog, property of the image, such as the luminance of an image region, the saturation of a certain color in the image, orientation of a given line segment, the depth of a given point in the scene, or the time of an event. As a general matter, the visual system is quite bad at estimating absolute visual metrics, but generally much better at estimating relative or comparative metrics. For instance, when we see two objects in depth, we perform quite poorly at estimating their absolute distance from each other or from us (see, e.g., (18)). But we are much better at determining which object is closer to us. The reason why our brains perform better as "comparison engines" as opposed to "measurement engines" is presumably because comparisons of space and/or time parameters are more behaviorally useful than absolute measurements. Of course, our brains are not foolproof when it comes to comparisons either, and are better at some type of comparisons than others (18, 153).

The other type of computation involves categorization or classification, that is, making categorical perceptual judgements, such as "predator or prey?" or "up or down?." Note that categorical judgements are not necessarily always two-way, or binary, classifications. Under natural viewing conditions, they often involve multiway classifications: for instance, judging the ethnicity of a person.

Subsequent sections will briefly present some of the highlights of our current fragmentary understanding of the neurophysiological mechanisms of high-level vision.

### Feed-Forward Visual Processing Involves Progressive Refinement of Behaviorally Relevant Information

Along the visual feed-forward pathway, that is, at progressively higher levels of the visual anatomical hierarchy, the response properties of visual neurons change in some important ways. While the response properties vary substantially within any given visual area and overlap substantially with the response properties of neurons in other areas (especially at adjacent levels of the hierarchy), four broad, overall trends are evident (143, 188, 189, 314) (also see Fig. 9C). First, the receptive fields of visual neurons get progressively larger, from about 0.5° in diameter within the foveal representation of V1 to about 10° to 40° in AIT, located at the anterior tip of the temporal lobe. Second, the neurons tend to become progressively more selective for more complex visual properties (see below for details). Third, the latency of visually evoked response tends to get progressively longer. Fourth, the responses tend to be more subject to modulatory influences of the visual context (see below) and of cognitive factors such as attention, behavioral goal and reward; and show greater plasticity (i.e., learning-dependent changes in response).

It is worth repeating the aforementioned caveat that, while these overall trends are apparent across the visual feed-forward pathway considered as a whole, the areas especially at adjacent levels of the hierarchy—overlap substantially enough that no single visual area can be said to have a unique, diagnostic set of response properties that distinguishes it from its neighbors on the hierarchy. That is, no known response property changes in a strictly stepwise or hierarchical fashion.

A large number of studies have examined these aspects of visual information processing in various regions of the visual system. Summarizing all of them is well beyond the purview of this review. In this section, we will focus instead on some representative empirical findings about visual processing that help illustrate the larger, general principles of visual system function.

#### "Early" visual regions play key roles in high-level vision, and are not limited to processing low-level image information

As alluded to above, a large proportion of the cerebral cortex is devoted to the processing of visual information. In the macaque monkey, where most of the invasive neurophysiological studies of the visual system have been carried out to date, up to 50% of the cerebral cortex is responsive to visual stimuli (98,341,344). In humans, about 30% of the cerebral cortex is directly responsive to visual stimuli (127). In macaques, the visual cortex is parcellated into a number of visual areas. In humans, these parcellations are typically referred to as visual *regions*.<sup>13</sup>

Visual processing of low-level properties of the image, or image primitives, such as contrast edges, orientation, and direction of motion, will be addressed by other reviews in this volume (see, e.g., "Lower Visual Pathways" and "Color Vision"), and will not be revisited here. Note that this does not allow us to leapfrog over the early cortical stages of feedforward visual processing, such as the retinotopic visual regions, so called because the neural responses in these regions systematically vary according to the retinal location of the stimulus. It is indeed true that the retinotopic regions, such as V1, V2, V3, V3A (in macaques) or V1, V2, V3, V4v (in humans) play key roles in processing low-level image properties. But to assume that their prominent role in feed-forward processing of sensory information precludes them from playing an important role in high-level vision is to fall prey to the "feedforward thinking" outlined above. It is now abundantly clear that retinotopic visual regions are a key part of dynamic brain networks, and as such play important roles in visual cognition.

Indeed, responses of individual neurons in V1, which represents the earliest stage of cortical processing, can accurately reflect what the viewer is perceiving, which is not what one would expect if low early visual areas were devoted solely to the processing of low-level visual features. In an influential neurophysiological study, Logothetis and colleagues showed that during binocular rivalry, responses of individual neurons in monkey V1 accurately reflect the reported percept of the monkey (33, 201, 206). Briefly, in binocular rivalry, two disparate images are shown to the viewer, one to each eye (Fig. 13A). If the views from the two eyes are similar, as it is during normal viewing, the brain fuses them together into a stereoscopic 3D percept. But if they are too different to be fused to a single, sensible percept (i.e., if the stimulus is dichoptic), only one of the images is perceived at a time, with the percept switching to the other image at random intervals. The investigators presented a grating oriented clockwise at 45° to one of the eyes and concurrently presented a grating



Figure 13 Neuronal responses in monkey visual area V1 during binocular rivalry. See text for details. Adapted, with permission, from (33).

oriented counterclockwise at 45° to the other eye. Neurons in V1 tend to be selective for orientation, so that a given neuron was a priori likely to respond differently to the two gratings when presented separately. But in this case, the gratings were being presented simultaneously and continuously, so that the stimulus itself did not change. The monkey was previously trained to report, by pulling a suitable lever, which way the grating it saw was oriented (201,206). As expected, the monkey reported randomly alternating percepts indicating that, at the perceptual level, the binocular rivalry was working as expected (the bar along the x axis in Fig. 13B, bottom). In many V1 neurons studied, the response of the neuron tracked closely with the reported percept, so that the activity was higher during one of the percepts than the other. Note that, since the stimuli themselves remained unchanged, the stimuli themselves could not have caused the changes in the neuronal response (33, 201, 206).<sup>14</sup> The finding that individual neurons in V1 can reflect a high-level phenomenon such as the viewer's (in this case, the monkey's) percept was novel, in part because it disproved the notion that early visual areas are involved solely in the processing of low-level properties of the image.

Functional MRI studies have found comparable effects in human V1 during rivalry (259). Conversely, studies in both monkeys and humans have found that many visual areas in the higher levels of the visual hierarchy also play role in binocular rivalry, although the network connectivity patterns associated with rivalry are yet to be fully delineated (33, 39). "Early" visual areas are also known to play significant roles in other high-level visual processes. For instance, many "early" visual regions play a role in figure-ground segregation (i.e., the perceptual process whereby an object of behavioral interest is perceptually distinguished from the background) and perceptual grouping (i.e., the perceptual process whereby different image elements are perceived as a larger, more holistic image element) (139, 273, 277). They also play a key role in perceptual grouping (i.e., the complementary of process of mentally grouping related parts of a scene), both of which are important aspects of high-level understanding of visual scenes (Fig. 14) (260, 273).

Many fMRI studies have shown that responses in many retinotopic visual regions are larger when visual recognition is unsuccessful relative to when it is successful. This may represent a form of error-coding (or predictive coding) where the activity in the low-level regions represent the "cognitive residual" or the visual information left unaccounted for by visual perception (304, 324). As also discussed below, the lateral geniculate nucleus of the thalamus plays a key role in reducing the redundancy of information in natural visual scenes. In other words, high-level vision is by no means the exclusive domain of high-level brain regions.

Early visual regions also play a significant role in visual perceptual learning (3,93,113,148,289,335,355). Perceptual learning, or sensory learning, is a distinct form of learning by which the ability of sensory systems to respond to stimuli is improved through experience, typically in adults. It remains



Figure 14 Figure-ground segregation and perceptual organization. (A) How many circles can you see in this image? In this image, referred to as the Coffer Illusion, you should be able to see 16 circles. Figure courtesy of Dr. Anthony Norcia, Stanford University. Reproduced with permission. (B) Camouflage is an extreme case of figure-ground segregation, where the object of interest is hard to recognize even when "in plain view." This image shows two variants of the pepper moth *Biston betularia*, one black and the other with light peppered coloring. The black variant is effectively camouflaged against colored tree bark whereas the light variant is easy to recognize (i.e., it "pops out"). The opposite is true when the same two variants are seen against a background of light bark and lichens. The black variant emerged for the first time in the industrial midlands of Britain in the 19th century, where tree barks were turning black with industrial soot. Soon the black variants because the predators of the moths had much greater success breaking the camouflage themselves against. This was an instance of the prey "gaming" the predators' high-level visual faculties to enhance its own survival (179, 334). Figure from Ford, E.G. (1977) Ecological Genetics. Springer.

unclear precisely how sensory learning in adults (or adult plasticity) differs from learning during development.

In the visual domain, skills ranging from discrimination of low-level visual features (e.g., orientation, contrast, direction of motion of depth) to high-level recognition of objects (e.g., "Labrador retriever or golden retriever?") and scenes (e.g., estimated property values of a neighborhood) can be improved through perceptual learning. As alluded to above, perceptual learning is a prerequisite for perception—the brain needs to learn what an object "looks like" before it can recognize it. There is no perception without perceptual learning. Indeed, the precise effects of perceptual learning in early visual areas continues to be matter of considerable debate and study (for reviews, see (110, 352)). There are some reasons to suspect that perceptual learning affects early sensory areas in the auditory and somatosensory pathways differently than it seems to affect early visual areas.

The neural mechanisms of perceptual learning are not fully understood. Neurophysiological studies in macaques have shown that perceptual learning occurs even at earliest stages of cortical processing, that is, in area V1 (111,113,284,335). Human neuroimaging studies using fMRI and EEG have found broadly comparable results have found in early visual areas, including V1 (262,285,352). In general, neural responses in higher visual areas tend to show greater adult plasticity (101,151,266,306,329,363).

A detailed examination of the neural mechanisms of perceptual plasticity is beyond the purview of this review. For additional information on this topic, readers should consult the relevant reviews cited above and the review "Network Supervision of Adult Experience and Learning Dependent Sensory Cortical Plasticity" in this series (32).

# Brain regions in monkeys and humans do not necessarily have a one-to-one correspondence

As noted earlier, much of what we know about the neural mechanisms of high-level vision comes from studies in monkeys and humans. As noted in the previous section, many of the original monkey neurophysiology findings have been repeated in humans, typically using noninvasive neuroimaging methods. The functional similarity between monkey and human visual processing can be particularly striking in early visual areas, where the responses are retinotopic. Many early visual regions, for example, V1 and V2, have come to be referred to by identical names between the two species.

It is therefore tempting to assume that monkey brain regions have exact counterparts in humans and vice versa, especially when they have identical or similar names. But it is important to remember that such direct comparisons can be unwise. Human and monkey brains are related to each other by virtue of having evolved from a common evolutionary ancestor. It is clear enough that, in evolutionary terminology, monkey and human brains are homologous organs, just the way human hands are homologous to bird's wings (149, 210, 330). But it would be a mistake to assume that every substructure in the hand, every muscle or bone, has an exact counterpart in a bird's wing, or vice versa. Some structures, such as hairs and feathers, have no counterpart in the other species at all. In other words, individual brain regions in human versus monkey brains do not necessarily have a one-to-one correspondence, just like the fingers in human versus panda forelimbs do not.<sup>15</sup>

Strictly speaking, it is impossible to be certain as to whether *any* single region in the human brain has an exact evolutionary homolog in the monkey brain or vice versa. This is because homology is defined solely based on common evolutionary origins. Thus, ascertaining the homology of, say, rhesus macaque visual area V2 with human V2 would necessitate examining the V2 of their common ancestor that lived 25 to 28 million years ago (274) but, of course, is now extinct. Besides, the human brain has many more distinct brain regions than the monkey brain, so that that is *a priori* evident that there is no one-to-correspondence at the level of individual brain areas of the two species.

There are indeed some cases where individual brain regions seem to have selectivity for similar visual properties (e.g., selectivity for faces) and seem to occupy similar places in the information processing pathways. But it is important to bear in mind that such properties do not amount to being evolutionary homologies, or evolutionary correspondence, any more than the wings of a fly correspond to the wings of a bird. It is also important to remember that, generally speaking, the criteria that have been used to define regions in monkey brains and human brains are quite different. Human brain regions have been defined mostly based on differential activation to stimuli or task conditions (e.g., face vs. house) or patterns of correlated activity (also see the section on resting state fMRI below). By contrast, visual areas in the monkey brain have been defined using a large array of principled criteria (6, 98, 293, 341), including responses of individual neurons; laminar patterns of feedforward, lateral, and feedback connectivity; colossal projections (i.e., interhemispheric connections carried through the corpus callosum), and shared representations of vertical meridians.<sup>16</sup> Conversely, even in cases where the case for evolutionary homology seems most compelling, such as monkey V1 versus human V1, we know they differ substantially in their functional organization (165, 166, 205).

The import of the foregoing discussion for the neurophysiologist is of course not that one should avoid the implications of evolutionary homology altogether, but that one must proceed with caution when making cross-species comparisons of response properties of individual areas (167).

# Neurons in the temporal lobe are highly selective to object shape

#### Face processing in the macaque IT

Regions along the aforementioned "what?" pathway represent progressively more complex information about object shape (Fig. 9C). One of the more intensely studied response properties in the temporal lobe is the selectivity for faces. It was first reported in the superior temporal polysensory area (STP) by Gross and colleagues (40) (Fig. 15; see legend for



Figure 15 Selectivity for faces in the macaque superior temporal polysensory area (STP). (A) Anatomical location of area STP (yellow highlight). (B) Responses of a single STP neuron to various visual objects, including variations in face stimuli. Note that, among the stimuli tested, the neuron responds best to a face with all the key facial features (left column, second stimulus from top). Cutting this stimulus into 16 pieces and showing the pieces in a shuffled order essentially eliminated the response (right column, third stimulus form top). The icon at bottom right denotes the size and visual field location of the receptive field. C, contralateral visual field (i.e., contralateral to the recording location). I, ipsilateral. Such neurons with large receptive fields that span both the visual hemifields are common in the visually responsive areas of the central and anterior temporal cortex. Adapted, with permission, from (40).

details). Face selectivity has since been demonstrated in the human temporal lobe, especially the fusiform face area (FFA) (14, 242).

As noted above, our perception of the visual scene is not instantaneous (see, e.g., Figs. 2 and 3A). Instead, it evolves over time, but rapidly enough that its temporal evolution is not readily apparent. The neuronal dynamics that underlies the perceptual dynamics has been studied in some detail over the last decades or so.

The temporal dynamics of face-selectivity of macaque IT neurons show an intriguing pattern: The response selectivity changes in a coarse-to-fine fashion in time, so that spikes fired shortly after the stimulus onset tend to convey coarse category information, that is, information relevant to distinguishing among broad categories of objects (see (140) for details). An important study by Sugase and colleagues (323) examined the time course of face representation by individual IT cells in awake, fixating macaques (Fig. 16). The stimulus set consisted of 38 stimuli, made up of three different human faces with four different expressions each, four different monkey faces with four different expressions each, and ten different geometric shapes (see Fig. 16 (inset) for a representative subset of the stimuli). The three stimulus types (human faces, monkey faces, and geometric stimuli) constituted the global or coarse categories, and the four fine categories consisted of the identity of the human faces, expression on the human faces, identity of the monkey faces, and expression on the monkey faces. The stimuli were presented for 350 ms each, and the information transmission rate for the global and fine categories was measured during each given 50 ms sliding window between 50 and 500 ms after the stimulus onset. The authors found that the responses during the initial transients carried significant information about the global categories, but much less information about fine categories. Fine-category information emerged much later, during the sustained response after the initial transients. In other words, the posttransient responses conveyed significant information about both global and fine categories, whereas the initial transient conveyed significant information only about the global categories.

Note that in the above study, the image itself was static, and the animal, its head, and the eyes were stationary, for all practical purposes. As alluded to above, real-world scenes are far more dynamic, among other reasons because the objects in the scene may move, conditions of lighting and shadows may change, and viewer, viewer's head and the viewer's eyes



Figure 16 Coarse-to-fine tuning of shape categories in the macaque IT. The stimulus set consisted of 38 stimuli, a subset of which is shown in the inset. The global shape categories (*inset, vertical axis*) consisted of human faces, monkey faces, and geometric shapes. The fine categories (*inset, horizontal axis*) consisted of the various facial identities and expressions. The plots show the cumulative information transmission rate of a sample of IT neurons about both the global and fine categories (*red and blue lines,* respectively). The thick horizontal line along the x axis denotes the stimulus duration. Adapted, with permission, from (140, 323).

all tend to move. This greatly increases the dynamicity of the retinal image, and a variety of top-down factors help the information processing even more dynamic (also see below).

# Regions in the "dorsal pathway" represent spatial information about the external world

As noted above, it is increasingly clear that, in view of our current understanding of the brain as a dynamic network, the notion of dorsal and ventral pathways is simplistic and antiquated. But it is not uncommon to find references to these pathways in textbooks and research articles. It is important to bear in mind that these references are simply terms of convenience.

Operating in our complex visual world entails not only recognizing the various objects in the scene, but also the spatial relationship among the objects in the scene and between the viewer and the objects. Thus, understanding how the brain computes the spatial parameters required for operating in the real world is important to understanding high-level vision.

Regions in the parietal lobe play a crucial role in processing the spatial information required for such behaviors (for reviews, see (9, 10, 47, 168, 294, 325, 350); also see the other relevant articles in this series). Various regions in the parietal lobe specialize in the processing of various types of spatial information, but no single region represents all the information that is required to carry out all spatial tasks. In other words, spatial information processing is fairly widely distributed, just as form processing is. Another similarity to form processing is the progressive, although not strictly hierarchical, refinement of spatial information: Parietal regions at higher levels of the visual anatomical hierarchy tend to carry more refined spatial information. Moreover, representations of form and space overlap considerably in many brain regions (292). For instance, individual cells in area MT are selective for motion, stereopsis and many other 3D spatial cues, as well as for shape cues (74-76, 79, 102).

#### Output of the spatial processing in the parietal lobe serve as inputs to three cortical systems involved in different types of visually guided behaviors

In using visual information to inform behavior, the brain needs to map the visual information to some type of real-world spatial coordinates. It turns out that the brain maps this information to not just one, but to three different types of real-world coordinate systems, each of which is useful for a particular set of visually guided behaviors (261, 310). The initial stages of this sensorimotor processing occur in various regions of the parietal lobe (9, 10), and the resulting information is relayed to three cortical systems of visually guided behavior: First, the output to the premotor cortex contains information necessary to navigating the extrapersonal space. This includes information necessary for moving one or more body parts, including moving the eyes, head, or the limbs from one position to the next (see the reviews "Eye Movements" and "Sensory Systems in the Control of Movement" (264) in this series). The second type of parietal output is relayed to the prefrontal cortex, and carries information critical for executive functions, including spatial working memory (see the review "Prefrontal Cortex in Motor Control" in this series). Third, the pathway to the medial temporal lobe (MTL) plays an important role in spatial declarative memory and navigation in the extrapersonal space. We will examine the processing in MTL in some detail below, because it provides a compelling illustration of how vision ultimately subserves behavior.

How exactly is the spatial information combined with visual information to produce the requisite real-world maps? The answer remains spotty for each of the aforementioned three systems of visually guided behavior, but it is clearest for one of them: the conversion of the visuospatial information to navigational information in various regions of the MTL. We will examine this pathway further below, especially to appreciate the simple, elegant neurocomputational logic of it.

# Brain regions in the medial temporal lobe (MTL) represent navigational space

To get around in the world, an organism needs information about itself in relation to the external world. That is, it needs an internal GPS map of its surrounding with a "You are here" (YAH) sign on it. Of course, the YAH sign must be constantly updated as the organism moves in its environment. How does the brain go about doing this?

Moser and Moser found part of the answer to this in rats, when they discovered grid cells in the entorhinal cortex that represent the external world as a grid (134). Individual grid cells fire preferentially when the animal is at specific spatial locations along its navigational path, and these locations of preferential response are organized as a spatial grid (Fig. 17A). it is as though each grid cell is a type of YAH sign, and each location has its own YAH sign, and the YAH sign corresponding to a given location "beeps" when you get to that location. By looking up which cell is "beeping" at any given time, the rat can figure out where it is in the navigational space. If individual grids cells correspond to individual YAH signs, where is the map?

It turns out that the YAH signs collectively help make up the map: Grids of different grid cells are slightly staggered, or phase-shifted from each other, so that grid cells as a population help bracket the navigational space and sample it densely in a combinatorial fashion (228-230).

Grid cells have been found in many other species, including monkeys (Fig. 17B). In monkeys, the grid cells are evident when the animal itself is stationary, as is its head, but it is freely moving its eyes while examining a real-world picture (180), indicating the grid cells can represent the extrapersonal space in many different ways, not all of which involve self-motion in physical space.



Figure 17 Grid cells help represent the external visual space. Grid cells were originally reported in rats, but have since been found in many species, including monkeys. (A) Responses of a single grid cell in the entorhinal cortex of the rat. *Left, black lines* denote the trajectory of a rat freely moving in a box. The cell fired spikes (*red dots*) when the rat was at specific locations within the box. These locations were organized in a grid-like fashion that spanned the box. *Right,* the firing rates of the cell represented as a heatmap, where "warmer" colors denote higher firing rates. Adapted, with permission, from (134). (B) Responses of a single grid cell in the entorhinal cortex of the macaque monkey. Unlike the rat referred to in panel A, the monkey was stationary. It sat in a primate chair with its head held steady, but freely moved its eyes as it looked at real-world pictures (not shown). *Left, red dots* denote the locations in a picture (not shown) that the monkey fixated, or gazed steadily at for a brief period. *Center,* firing rate of the grid cell shown in heatmap format. The *scale bars at bottom* each denote 6° of visual angle. Adapted, with permission, from (180).

O'Keefe and colleagues have shown a comparable phenomenon elsewhere in the MTL of freely moving rats, in the hippocampus (46, 230). Individual "place cells" respond preferentially when the animal is in a particular spatial location. Thus, place cells convey the YAH information explicitly, whereas grid cells only convey it implicitly. Models have proposed that place cells represent the input from multiple grid cells with differing grid periodicities but similar phase (46, 230).

The above survey provided only the briefest of outlines of how the feed-forward processing of sensory information in the retinal image gives rise to "actionable" information for visually guided behaviors. Later in this review, we will develop a fuller view of how information flow in other directions within the brain network, especially feedback and lateral processing, augments visually guided behavior.

### Select Urban Myths About Response Selectivity in Visual Areas

From late 1970s to about 2010s, there was an explosion of microelectrode neurophysiological studies that characterized in detail how individual neurons and neuronal ensembles in different extrastriate visual areas (i.e., all areas other than the striate cortex, or  $V1^{17}$ ) represent information about the underlying visual image. For an overview of these response properties, see (358).

Understandably enough, neurophysiological studies of visual processing tend to highlight the novel or unexpected aspects of the findings, which tend to be that neural responses in a given region represent information that no other brain region is known to represent, Naturally, the uniqueness of

**High-Level Vision** 

each area tends to get more noticeable than the similarities among the regions in the scientific community at large, especially if the novel findings lend themselves to pithy slogans. One unfortunate consequence of this has been the spawning of several undying urban myths about functional hyperspecialization in the visual system. The aforementioned tale of two pathways is one such example. It has been clear since the early 1990s that this model is flawed at best (for an overview of the early anatomical evidence (221)). Mounting evidence from the subsequent years indicates that, beyond pedagogical contexts, the notion of separate pathways is fundamentally misleading, in that it represents an oversimplification of the underlying complex network, the precise architecture of which varies dynamically (358).

Many variations of the two-pathway idea have also gained some currency. One such notion is that the dorsal and ventral pathways are magnocellular and parvocellular pathways, respectively, and that their input derives from the respective cell types in the lateral geniculate nucleus. Tracing the source of this myth is as appealing as tracing the source of the odor in the elevator. Suffice it to say that this is a demonstrable myth (48, 220, 308). A related notion, rooted in the magnocellularparvocelluar bifurcation but more directly relevant to highlevel vision, is that the dorsal pathway is geared toward fast but coarse-grained processing of visual information, and ventral pathway is more geared toward finer-grained, but slower processing (21). This model is flawed at best, not the least because it is based on discredited notions of magno-parvo bifurcation (140). Another version of the two-pathway idea, primarily based on neuropsychological studies of human brain lesions, is that the ventral and dorsal pathways are specialized for perception and action, respectively (120). For an appraisal of this model, see (219).

The notion that "V4 is the color area" (272) is another urban myth. Three main reasons why this is untrue bear stating here, because they apply to functional specialization of any visual area: First, while V4 cells do indeed convey information about color, they also convey information about many other visual features, including shape, disparity, motion, etc. (343). Second, conversely, color is also represented in many other areas of the extrastriate cortex and in the frontal lobe (232). Third, it is now clear that color is represented in a distributed fashion across many regions of the primate cortex, so that no single visual area can claim to be a color area.

In fact, a large body of neurophysiological studies, as well as many microstimulation and lesion studies have shown that distributed representation or coding is a broadly applicable principle of visual processing. No single visual feature is exclusively or primarily represented in a single area, nor is any visual area specialized exclusively for the processing of any given visual feature. Thus, another myth, that MT is exclusively the motion area, has been debunked (74-76, 102).

This is not to say, however, that visual representation is so distributed that no single area represents sufficient information to support visual perception by itself. On the contrary, many studies have shown that individual brain areas, neuronal ensembles, or in some cases a relatively small number of individual neurons, carry sufficient information to support, in principle, the perceptual outcome (also see the section about Multivoxel Pattern Analysis below). In this specific sense, coding can be quite local (as opposed to distributed) in the visual cortex, and a given visual parameter can be precisely represented by multiple cells representing different values of the parameter in a combinatorial fashion (185, 244). Indeed, in macaque areas MT and MST (middle temporal area and medial superior temporal area, respectively), responses of relatively small number of neurons can, in principle, reliably represent the perceptual outcome on a trial-to-trial basis (60, 153).

### Top-Down Processes Significantly Affect Visual Processing

Recent neurophysiological studies, especially whole-brain neuroimaging studies that have examined connectivity patterns across the entire brain, have made it abundantly clear that the visual anatomical hierarchy by itself cannot be taken as a blueprint for visual information processing. The four main lines of evidence that have led to this revelation are briefly summarized below.

# Recurrent and lateral inputs can bring top-down cognitive information to bear on visual processing

Recent studies have made it increasingly clear that visual information also flows through the lateral and feedback connections within and across areas. Such interactions, collectively referred to as recurrent or reentrant processing in some contexts, play a crucial role in visual information processing (45, 148). For instance, it is known that prior knowledge of the visual world is critical to visual perception. That is, to recognize something, we must know what it is. Such prior knowledge is acquired through perceptual learning during development and adulthood, and are brought to bear on visual perception through recurrent processing (e.g., Figs. 2 and 3A). It is now abundantly clear that recurrent processing itself does not necessarily follow a hierarchical pattern either (88, 147, 188, 340), although some earlier theories had posited that perception followed a reverse hierarchical pattern (3, 148).

It is also clear that information flow through lateral connections affects visual information. For instance, the responses of a given neuron to a stimulus within its classical receptive field, or CRF, primarily reflect the stimulusdriven, or feed-forward component of the response. But the stimuli outside the CRF, that is, in the nonclassical surround, can greatly modulate the neuron's response to the stimulus within the CRF (see, e.g., Fig. 18; see legend for details). In sum, all nonretinal factors that influence the processing of



Figure 18 Coding of visual context. The classical receptive field (CRF) of a given neuron is the portion of the visual field in which the neuron is most responsive to visual stimuli. The surrounding region in which visual stimulation modulates the responses to CRF stimulation is referred to as the nonclassical receptive field (nCRF) or nonclassical surround. This figure shows one of the earliest demonstrations of the modulatory effect of nCRF on CRF, in this case the response of a single neuron in macaque MT. Individual neurons in MT often response best when the stimulus moves in a particular direction, often referred to as its preferred or optimal direction. The neuron shown responded best when the dots in the CRF moved horizontally from left to right. CRF is denoted by the *dashed rectangle* in the icons at top. (Left panel) Modulatory effect of stationary surround on motion stimuli in the CRF. The direction of the movement of dots in the CRF was systematically varied, while the dots in the nCRF were held stationary. (Right panel) Modulatory effect of surround motion on motion stimuli in the CRF. The responses of the same neuron when the dots moved in its optimal direction, while the direction of the dots in the nCRF was systematically varied. Note that the neuron's responses vary systematically (i.e., the responses are "tuned"<sup>21</sup>) with respect to both CRF motion, and motion in both the center and surround. Thus, neuron can convey information of the motion "context," or motion of a given moving object relative to nearby stationary or moving objects. Adapted, with permission, from (5).

the retinal image information are brought to bear by recurrent connections.

It is important to note that there do seem to be visual tasks that require little or no recurrent processing. For instance, some detection tasks can be successfully performed within 150 ms, a time frame that affords little time for reentrant processing (see (348) and the references therein). However, such tasks that require little or no recurrent processing appear to be special-case scenarios or laboratory curiosities, since most natural scenes tend to be complex enough to warrant recurrent processing. Indeed, there is no evidence that vision can function without recurrent processing under natural viewing conditions.

### Extracortical regions play crucial roles in visual information processing

The cortical anatomical hierarchy relies solely on the interconnections among the various cortical areas. But it is clear that, in addition to these corticocortical pathways (Fig. 7), information also flows from the cortex to subcortical striatal regions, and back to the cortex, in a set of four mutually distinguishable but overlapping pathways referred to as the corticostriatal loops, sometimes referred to as corticothalamic or transthalamic pathways (Fig. 19; also see below and the review "Functioning of Circuits Connecting Thalamus and Cortex" in this series). Only one of these pathways, the so-called visual loop (Fig. 19A), is predominantly visual, although visual information can influence the neural responses in each of the other three loops. In case of all four loops, information travels in the cortex  $\rightarrow$  striatum  $\rightarrow$  thalamus  $\rightarrow$  cortex direction (Fig. 19B). The various brain regions in the corticostriatal loops do not conform to a strict hierarchical pattern, much less the same hierarchical pattern as the corticocortical connections (176, 301, 303).

While much remains to be discovered about information processing in corticostriatal loops, it is already clear that they play a vital role in brain function. The aforementioned visual loop, for instance, is known to play key roles in visual perception, visual-guidance of movements,



Figure 19 Corticostriatal loops in human brain. (A) Four distinguishable but mutually overlapping loops are usually recognized (colored labeled arrows), based primarily on the types of tasks in which they play prominent roles. Cortical inputs arrive largely via the striatum and ultimately are directed back into the cortex via the thalamus. The (ultimate) cortical output of the basal ganglia reaches largely to the same cortical areas that give rise to the initial inputs to the basal ganglia. The *visual loop* is known to play a prominent role in the learning of visual object categories, but during object categorization tasks using learned categories, the executive loop also plays a prominent role. Corticostriatal loops in the nonhuman primate brain are largely similar (not shown). Adapted, with permission, from (287); also see (222,250). (B) A more detailed circuit map of the visual loop shows the flow of information within the loop. GPe: Globus pallidus, external portion. GPi: Globus pallidus, internal portion. SNr: Substantia nigra pars reticulata. SNc: Substantia nigra pars compacta. STN: Subthalamic nucleus. VTA: Ventral tegmental area. Adapted, with permission, from (286).

perceptual learning, and multisensory processing (for reviews, see (197, 287, 288, 299)).

It should be noted in this context that much of visual neurophysiological research so far has focused on the visual cortex. The role of subcortical and cerebellar regions in the processing of visual information is largely unclear and remains to be explored (see Table 2). This is arguably because these regions are smaller in size and less neurophysiologically accessible. However, we know enough about the importance of these regions to know that we cannot fully understand vision without fully understanding the role of these regions in vision.

#### There is no prespecified flow-chart for visual information; flow of visual information can and does change dynamically

A naive functional interpretation of the anatomical hierarchy would be that the visual information flow follows an obligatory, preset pathway. But recent studies of functional connectivity have shown that the pathways of visual information flow are task-dependent (57). Moreover, only a subset of the available anatomical connections, that is, only a subset of the functional connectivity patterns, are activated at any given time. In no known instance are the patterns of functional connectivity strictly hierarchical (see, e.g., (143, 314)).

Another dynamic aspect of brain function is that the functional properties of visual neurons change over time. Moreover, the responses often change in an adaptive manner, so that 
 Table 2
 Some Major Unanswered, Interrelated Questions in Neurophysiology of High-Level Vision

- 1. How does vision work under real-world conditions?
- 2. How do brain networks interact dynamically with each other to produce behavior, and how does this influence the manner visual information is processed?
- 3. How does vision change over time and through various stages of life?
- 4. What roles do brain regions outside the cerebral cortex, such as the various subcortical regions and the cerebellum, play in vision?
- 5. What are the neuronal mechanisms of perceptual learning? How do mechanisms of adult plasticity differ from mechanisms of plasticity during development?
- 6. How does the brain integrate visual information with information from other senses?
- How does the activity of nonneuronal cells in the brain, such as the glia affect the processing of visual information?
   How do the various mental and metabolic states, such as
- 8. How do the various mental and metabolic states, such as various states of sleep, wakefulness, arousal affect visual information processing?
- 9. What are the neural bases of individual differences? What are the neuronal bases of visual expertise, creativity, and visually mediated affect (e.g., sensation of beauty, love, humor, or horror)?
- 10. What are the neural mechanisms of brain impairment and dysfunction?
- Which aspects of biological vision can be reproduced, for practical intents and purposes, in machines, and why?
   Which visual deficits and which visual shortcomings in
- 12. Which visual deficits and which visual shortcomings in healthy subjects can be compensated for using machines, and how?

they help enhance the information carried by the responses about certain feature/s of the image. Such adaptive changes can occur at different time frames, ranging from milliseconds to years, or even decades (187). The hierarchical view of visual processing cannot explain these dynamic properties of the visual system.

Yet another dynamic dimension of visual perception is that it is profoundly influenced by the cognitive context. You are more likely to recognize the checkout person of your neighborhood grocery store at the checkout counter than waiting for a taxi at an overseas airport. That is, our expectation of what we are likely to see influences what we do see (for reviews, see (19, 20, 145, 146); also see (19, 20, 30, 146)).

# Dynamic interplay of feed-forward and feedback information shapes visual perception

Traditional methods of measuring perceptual outcomes do not adequately characterize changes in high-level visual percepts. This is in part because natural scenes tend to be very complex, but also because our understanding of visual scenes, or at least the introspective or reported aspect of it (see below), tends to have a significant semantic or subjective component (see (140) for details). In a study notable for its novel design and somewhat unexpected results, Li and colleagues (97) have shown how the reported percepts of natural scenes change with varying viewing durations (also see (146, 265)).

Advances in neurophysiological and neuroimaging methods and in methods of data analysis now make it possible to monitor the activity of multiple brain regions concurrently, and determine how various regions interact with each other (see (175, 254, 313, 314) for recent reviews). These studies collectively show that feedback information continuously influences the processing of feed-forward information, and vice versa. In fact, in most brain regions, with the possible exception of the retina itself, neuronal responses at any given moment have both feed-forward and feedback components, so that it is generally not possible to classify the response as purely feed-forward or purely feedback even under controlled laboratory conditions, let alone under natural viewing conditions. This intermingling of information becomes progressively more complex as one moves up the levels of the visual anatomical hierarchy.

However, it is straightforward enough to experimentally demonstrate the aforementioned principle of dynamic interplay of feedback versus feed-forward information. In a 2006 fMRI study (324), Summerfield and colleagues showed subjects low-contrast, ambiguous pictures of faces, houses, and cars in random order during each given block of the scan. Subjects were required to categorize each given picture in one of two ways depending on the block. During the "face block," subjects had to report whether a given picture was that of a face or not. Similarly, during the "house block," subjects had to report whether a given picture was that of a house or not. Car stimuli, which did not have a block of their own, served as control stimuli. As expected, responses in some regions were more directly influenced by the stimuli (i.e., the feed-forward or bottom-up information) than the task (i.e., feedback or topdown information) and vice versa. Responses in the temporal cortex, specifically in inferior occipital gyrus (IOG), fusiform face area (FFA), temporoparietal junction (TP), and amygdala were more stimulus-dependent than task-dependent. Responses in two regions of the frontal cortex, the dorsal and ventral medial frontal cortex (dMFC and vMFC, respectively) showed the opposite pattern.

A straightforward hypothesis about how these brain regions help carry out the task hand is that the frontal cortical regions "instruct" the temporal regions what to expect during a given block. To test this hypothesis, Summerfield and colleagues used a type of hypothesis-dependent technique of connectivity analysis called dynamic causal modeling (DCM), a technique that can be used to test specific hypotheses about functional connectivity among brain regions. DCM results showed that presentation of face stimuli selectively strengthened feed-forward connections from IOG to FFA and amygdala. On the other hand, feedback connection from vMFC to FFA and amygdala were selectively enhanced during face blocks, that is, when subjects had to perform a face detection task. This is consistent with the above hypothesis that top-down information from the frontal regions help modulate the processing of the bottom-up information in the temporal regions. Note that, without being able to analyze effective connectivity patterns in this fashion, one would have deemed the selfsame results eminently consistent with the conventional hierarchical processing model, where the temporal regions do their part in analyzing the image information on their own, and send off the processed information higher up the hierarchy.

While the above study used DCM to analyze effective connectivity, analytical methods quite distinct from DCM can be also used in connectivity analyses. After all, DCM is not suited for exploratory analyses. It can only be used for testing specific hypotheses formulated independently. Fortunately, many methods such as various types of autoregression or cointegration methods are suitable for exploratory analyses as well hypothesis testing. Granger causality (GC), for instance, can be a powerful and principled alternative or complement to DCM (13, 103). When GC is used for testing causal relations between, say, two time series A and B (e.g., responses of two neurons or two brain regions, etc.), time series B is said to be "Granger-caused" by A if successive values of B are better predicted by taking into account both A and B, rather than A alone. In other words, when time series A Granger-causes time series B, the temporal patterns in A are approximately repeated in B after some delay (Fig. 20).

These techniques can be applied to any type of neurophysiological data, including microelectrode recording, LFPs, or neuroimaging data from EEG, MEG, or fMRI. As such, familiarity with such techniques of determining effective connectivity is essential to the contemporary neurophysiologist. User-friendly toolkits from implementing such analyses are easily available (24, 99).



Figure 20 An illustration of Granger causality. This figure shows cause and effect relationship between two hypothetical neural responses (*top* and *bottom* rows, respectively). The causeand-effect relationship exists throughout the responses, but is most readily apparent by visual inspection of those portions of the responses where the response is prominently modulated (*arrows*). Granger causality uses the entire length of both responses to quantitatively measure the cause-and-effect relationship even when the relationship may be too subtle or complex to be visually evident. In the present case, the response shown in the top row is said to "Grangercause" the response in the bottom row. Note that the term "Granger causation" denotes an *inferred* cause-and-effect relationship, which may or not include direct causation. Thus, the concept of Granger causality is in some respects narrower, and in some other respects broader, than the concept of direct causation (13, 29, 224). But in either case, the cause must necessarily precede the effect.

### Attentional Selection and Eye Movements Help Make Natural Scene Perception Even More Dynamic

The brain continuously receives a vast amount of sensory information from the external world, only a small fraction of which tends to be relevant to the behavioral goal at hand at a given moment. To optimally allocate its finite computational resources to help process this information, the brain needs a mechanism of selecting behaviorally relevant aspects of the incoming information. Attention is one such mechanism of perceptually selecting a particular stimuli or tasks relevant to behavior. In an oft-used definition, William James (160) formulated attention as "... the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought."

Unsurprisingly, all sensory modalities have attentional mechanisms. Attentional mechanisms can also operate across two or more sensory modalities (84). The neural mechanisms and behavioral manifestations of attention are discussed in detail in the review "Attention: Behavior and neural mechanisms" elsewhere in this series (also see (96, 238)). In this review, we will examine the ways in which visual attention can dynamically change neural responses to visual stimuli even when the underlying visual stimulus remains unchanged (223).

# Attention has four different types of modulatory effects on neuronal responses

In neurophysiological experiments, attentional effects are usually studied as the responses to a pair of stimuli, both, either, or none of which is attended, depending on the condition. In general, effects of attention tend to be progressively larger in higher visual areas. Neurophysiological studies in awake, behaving monkeys have elucidated four distinct types of effects of attention on the responses of individual neurons (59, 217, 218). It is important to bear in mind that two or more of these effects may be simultaneously operational at any given time, especially under real-world conditions.

First, attention can enhance the individual neuron's response to a given stimulus in its CRF. That is, visual neurons tend to respond better to a given stimulus presented in its CRF when attention is directed to either a particular aspect of the stimulus—such as a particular object or feature of that object (*object attention* or *feature attention*, *see below*) or to the spatial location of the CRF (*spatial attention*).

Second, attention can increase, in the signal theoretic sense, the sensitivity and discriminability of the neuron's responses to different stimuli. Increased sensitivity means that the neuron is more likely to respond to its preferred stimulus than to its nonpreferred stimuli. Increased discriminability results from a reduction in the variance of responses to one or both of the stimuli, larger spread in the mean responses to the stimuli, or both. Third, attention can suppress the response to the unattended stimulus. That is, when two stimuli are presented simultaneously within the cell's CRF, the response of the cell is usually different than when either stimulus is presented alone within the CRF. But when both stimuli are presented within the CRF and only one of them is attended, the cell tends to respond as though only the attended stimulus is presented; the response to the unattended stimulus is suppressed.

Fourth, by far the biggest effect of attention is that attention decreases the correlation among the responses of neurons, so that the neurons become mutually less redundant. Finally, whether or not attention alters the selectivity of the neuron (e.g., change its preferred stimulus) is largely unclear. Attention and eye movements are closely linked in many animals, including humans and other primates. We typically shift visual attention by moving our eyes to the portion of the visual scene that we need to attend to, that is, select for closer visual scrutiny. This is because in many animals including primates, not all parts of the retina are equally sensitive to visual information. Instead, photoreceptors are densely packed in a relatively small portion of the retina (subtending 0.5°-1°, depending on the criteria used) near the back of the eye, called the fovea. The visual image has the highest spatial resolution at the fovea. Spatial resolution decreases, and spatial blurring increases, rather steeply away from the fovea. Therefore, animals with fovea generally deploy attention by foveating the location of interest, that is, moving the eyes so as to focus the image region of interest on the fovea. However, eye movements and attention can be mutually dissociated. For instance, one can shift attention without moving the eyes. Moreover, many animals such as rodents lack fovea and eye movements, but can nonetheless have sophisticated attentional mechanisms (223). Eye movements are examined in greater detail in the accompanying review "Eye Movements."

#### Insights from resting state fMRI: Brain contains many different networks dynamically recruited according to behavioral task

Resting state fMRI (rsfMRI or R-fMRI) is a method of functional brain imaging that can be used to evaluate interactions among brain regions that occur when a subject is simply resting (i.e., not engaged in a behavioral task) (31, 43). While there is no universally accepted definition of resting state, rsfMRI are typically carried on alert subjects lying still with their eyes closed. Even in this state, the activity of various brain regions is correlated in time, so that the response levels in these regions tend to rise and fall together.<sup>18</sup> The level of this correlation provides a metric of how strong the intrinsic connections between a pair of brain regions are (see Figs. 21 and 22). Thus, rsfMRI can be thought of as a roadmap that shows the network of roadways that connect various geographic locations.

Five insights into brain function provided by rsfMRI are particularly noteworthy: First, it reveals that not all visual regions are equally well-connected to others. Some "hub" regions, including many regions in the striate and extrastriate visual cortex, are much more heavily interconnected with other brain regions than others (316, 339). Moreover, hub regions tend to be well-connected to other hubs, forming a "rich club" of brain regions (339). Such well-connectedness among the well-connected is obviously unique to the brain.

Second, the networks revealed by rsfMRI tend to be similar, but not identical to, the underlying anatomical networks. Indeed, the anatomical connections help account for the pattern of connectivity in rsfMRI because, after all, the underlying anatomical connections provide the "communication infrastructure" that the various brain regions use for communicating with each other.

Third, brain networks identified by rsfMRI tend to overlap a great deal with those revealed by the fMRI during active tasks (309) (also see Fig. 21). Resting state networks (RSNs) often form the core of network activity during active tasks. It is as though the intrinsic connectivity networks during rest represent an operational repertoire, and the network connectivity during various cognitive tasks represents elaborations or variations of the themes in this repertoire.

A fourth, related, insight is that the RSNs are remarkably similar across subjects, bolstering the notion that the RSNs do represent "themes" of brain function (see, e.g., Fig. 21B). The fact that these networks are similar across subjects raises the intriguing possibility that the *differences* of network connectivity between subjects may help explain individual differences. The extent to which individual differences can be accounted for in this fashion remains an intriguing area of future research.

Finally, the RSNs are largely homologous among many mammalian species, including humans, nonhuman primates, and cats (316, 339). This is consistent with the fact that all these brains share a well-established evolutionary homology of various degrees. But more importantly, it points to how the brain came to implement a particular processing pathway.

### Relating Neuronal Activity to the Behavioral Outcome, and Vice Versa

# Microstimulation in MT and MST can influence perceptual decisions

In a landmark set of studies, Newsome and colleagues compellingly demonstrated that responses of individual neurons can account for the perceptual outcome in a quantitative fashion (37,236,278). Specifically, they studied how the neuronal responses in monkey middle temporal area (MT) reflect perceptual decisions in awake, behaving monkeys.

MT cells tend to respond strongly to motion stimuli, and individual cells respond selectively to a given direction of motion, referred to as the cell's "preferred direction." The studies by Newsome and colleagues used dots moving either in the preferred direction of the cell under study, or in the



Figure 21 Resting state brain networks are very similar to brain networks active during tasks. (A) Smith and colleagues (309) extracted 20 mutually independent patterns of activation in the resting state networks (RSNs) from a database of 36 adult human subjects using the independent components analysis (ICA) (319). Brain regions identified by each of these ICAs can be thought of as an independent network. Smith and colleagues then compared RSN ICAs with ICAs of taskactivated networks from nearly 30,000 subjects from the BrainMap (BM) database (brainmap.org). Comparisons for ten most informative ICAs are shown side by side in this panel. For each ICA, activation is shown in a color-coded format for a coronal, sagittal, and horizontal section (top, middle, bottom row, respectively). (B) Smith and colleagues (309) then analyzed the extent to which the top ten of the RSN ICA's play a role in various types of behavioral tasks (or "behavioral task domains" defined by the BrainMap database). Higher color values denote a correspondingly larger role by the given network in a given task. Note that each given type of task recruits different RSNs to different extents. Conversely, each RSN is active during multiple, different task paradigms, with the degree of participation varying according to the task. Thus, the RSNs represent a repertoire that the brain recruits and employs to various degrees depending on the task at hand. For details and some important caveats, see (309). Adapted, with permission, from (309).

opposite ("null") direction, depending on the trial. Depending on the trial, a certain proportion of the dots remained stationary (Fig. 23A). The stimuli were presented, one per trial, within the CRF of the neuron under study while the animal indicated, by pressing a lever, whether the dots moved in the preferred or the null direction, and was rewarded with a drop of juice. Since the neuronal responses also depend on the overall strength of motion information, or motion "energy," in the stimulus, the investigators "tuned" the strength of the motion information by changing the proportion of dots that moved in the same direction, that is, percentage correlation of motion. The tuning of the motion strength in the stimuli elicited an



Figure 22 Networks in human brain at rest. Nodes denote the center of mass of the corresponding brain regions, and the edges (i.e., *colored lines*) represent intrinsic connections between a pair of regions identifiable in the resting brain. *Blue nodes* represent the "rich club" brain regions, which are well-connected brain regions that are wellconnected with other well-connected brain regions. *Gray dots* denote nonrich club regions. *Red lines* denote connections between rich club regions. *Orange lines* denote "feeder" connections that connect a rich club region with a nonrich club region. *Yellow lines* denote "feeder" connections that connect a nonrich club region with another nonrich club region. Adapted, with permission, from (339). Note that this figure does not show subcortical or cerebellar networks, which decidedly play crucial roles in brain function.

expected sigmoidal tuning in response from the neurons (solid dots and solid fitted curve in Fig. 23B; see (37,236) for analytical details). This "neurometric function" of the neuron's response closely paralleled the "psychometric function" of the animal's behavioral responses (open dots and dashed fitted curve in Fig. 23B), suggesting that the visual sensory information represented by this single neuron can, in principle, support the animal's perceptual decision. The extent to which the responses of a given neuron reflect the animal's perceptual decision-or perceptual choice between the available alternatives-can be captured by a metric called the choice probability (CP). CP is a numerical measure of the probability of a certain perceptual or behavioral outcome (62, 68). Each neuron can be assigned a CP. The higher the CP, the higher the given neuron's ability to reflect the animal's perceptual decisions.

The close parallels between the neurometric versus psychometric functions held across different animal subjects, and was not idiosyncratic to a particular subject tested (Fig. 23C). Note, however, that the linear trend indicated by the *dashed line* notwithstanding, there was considerable individual variation among the subjects (see (37) for details), as there often are in neurophysiology. As outlined below, accounting for these individual differences is likely to be a major future trend in brain research.

#### Microstimulation demonstrated a causal relationship between neuronal and perceptual responses

When a brief electrical pulse was delivered near the neuron under study, the animal's performance actually improved (Fig. 23D), in the sense that less motion energy was needed to support the same level of behavioral performance (see (278) for details). This panel shows the psychometric function of an animal with or without microstimulation (*solid dots and fitted line and open and dashed fitted line*, respectively) when two different neurons were stimulated. Note that the parallels between the psychometric versus neurometric functions imply local coding, that is, that enough motion information is conveyed by the responses of individual neurons to support perceptual outcomes. The results of the microstimulation further strengthen the evidence for local coding<sup>19</sup> of motion information in MT.

#### Multivoxel pattern analysis can accurately predict perceptual outcome based on BOLD activity

MVPA is essentially the fMRI equivalent of the aforementioned CP analyses of single neuron data. Since the application of MVPA to various forms of neurophysiological data, especially microelectrode data, is likely to be a major data analysis trend in the near term, we will consider them in some detail here.

MVPA takes advantage of the fact that the spatial pattern of BOLD activation can often, but not always, be diagnostic of the perceptual outcome. Thus, it simply finds, using advanced machine learning methods beyond the purview of this review, a classifier that can reliably distinguish voxel activation patterns that are associated with various perceptual outcomes. The choice of the particular classifier can depend on the particular implementation of MVPA (i.e., the particular toolkit used), the user, or the data. In this sense, MVPA is a collection of classification tools, rather than a single classification technique. There is no a priori optimal classifier that is suitable for all data but, in general, techniques that transform the data into a dataspace that makes them easier to classify (or the so-called kernel techniques) make for good classifiers for neural data. Of the kernel techniques, support vector machines (SVMs) are the most versatile and powerful classifiers, and hence make for principled initial choice. In some cases, the classifier has to be trained using the data at



Figure 23 Neuronal responses in monkey middle temporal area (MT) reflect perceptual decisions in awake, behaving monkeys. (A) Stimuli and task paradigm. The studies by Newsome and colleagues (37, 236) (278) used dots moving either in the preferred direction of the cell under study, or in the opposite ("null") direction, depending on the trial. The investigators "tuned" the strength of the motion information by changing the proportion of dots that moved in the same direction (i.e., percentage correlation of motion). (B) The "neurometric function" (solid dots and solid fitted curve) an individual neuron that closely paralleled the "psychometric function" of the animal's behavioral responses (open dots and dashed fitted curve). (C) The close parallels between the neurometric versus psychometric functions held across different animal subjects. The dashed line notwithstanding indicates the best fitting linear trend. (D) Demonstration of a causal relationship between the responses of individual neurons and the animal's percepts using microstimulation. The psychometric function of an animal with or without microstimulation (solid dots and fitted line, respectively). See text and (37, 236) for details. Adapted, with permission, from (37).

hand before it performs reliably at predicting the perceptual outcome. In these cases, the training data must be independent of the testing data, or the performance of the classifier will be subject to, among other things, the statistical cardinal sin of sampling bias. Often, however, the classifier can perform reliably without prior training, which is the preferred scenario.

A large number of studies have used MVPA to show that responses in various individual brain regions or a set of regions can reliably predict the perceptual outcome. Sometimes such exercises are referred to as "mind reading," in the sense that one can read the brain responses to judge what the mind is perceiving (see, e.g., (174, 234, 331, 362)).

#### Consciousness

There is no rigorous neuroscientific definition of consciousness. Crick and Koch (66) famously, and plausibly, argued that it is unwise to define it prematurely, before we better understand what it is and how it works. However, they pointed out that we all have "a rough idea of what is meant by being conscious" (also see (67, 77)). It is easy enough to dismiss some of the public interest in the topic of consciousness a peculiar affliction of those with metaphysical proclivities. But this topic is of interest in the narrow context of the present review, because visual consciousness, or visual awareness, is often thought of as a stand-in for visual perception (see, e.g., (64)). But neuroscientific treatments of consciousness (65-67) often ignore the fact that visual awareness, at least as it is (not) defined, involves some introspection that is not commonplace during visually guided real-world behavior. Mainly for this reason, this review adopts the position that consciousness is a side-effect of brain function, and not its *raison d'etre*.

### Visual Perception Changes Dynamically Across Time Scales Ranging from Milliseconds to Decades

# Short-term dynamics of the adult brain: Multiple factors influence the temporal dynamics of neuronal responses

To understand the temporal dynamics of the responses of visual neurons, it is useful to bear in mind a few basic facts. To begin with, there is an intrinsic dynamicity to the responses that is not attributable to external factors (for a review, see (140)). The responses of visual cortical neurons can and do change quite rapidly over time, even as the stimulus itself remains unchanged. Moreover, neurons in the visual cortex typically fire spikes at low, "background" levels even in the absence of overt visual stimulation. When a static visual stimulus is presented in the neuron's CRF, there is usually a brief delay before the response rises above background levels (282). This delay—or in more overtly neurophysiological lingo, latency-is usually on the order of a few tens of milliseconds in early visual cortical areas such as V1 or V2, and tends to become progressively longer, lasting several tens of milliseconds, in higher visual areas, such as those in the parietal, temporal, and frontal lobes (144, 282). In the macaque, where the latencies have been measured in detail, at least some neurons in all visually responsive areas, including frontal and motor cortices, will have been activated by about 190 ms following the stimulus onset (for reviews, see (133, 140, 245)). The response latencies in the human brain are generally longer by a few tens of milliseconds, depending on the brain region (328). This is thought to be primarily because the human brain is physically larger, so that the neuronal signals must travel farther.

The feed-forward and feedback connections conduct information at a velocity of about 2 to 3.5 m/s (see (11, 44, 115, 195) and the references therein). The lateral connections, or connections between neurons within a given area, including connections that subserve the aforementioned center-surround interactions, conduct information about ten times slower, at about 0.33 m/s (11). Across the visual cortex,

the synaptic delays are negligible for electrical synapses, but tend to be about 5 to 20 ms per synapse for chemical synapses (54, 353). Since a vast majority of cortical connections are believed to use chemical synapses, synaptic delay is a major factor that influences information processing and transfer in the brain (54, 353).

Under controlled laboratory conditions, the responses of visual neurons show, to a first approximation, the following intrinsic temporal dynamic pattern in the absence of changes in stimulation. When a static visual stimulus is presented within the CRF of a typical visual neuron, its firing rate continues to be at background levels for a characteristic latency period, following which it rises rapidly, peaking after a few tens of milliseconds, depending on the neuron and the cortical area (282). After this initial transient (or onset transient) response, the firing rate decays, but typically more slowly than during the rising phase, before largely stabilizing at a lower response level over the next few hundred milliseconds, depending on the stimulus, the individual neuron, and given cortical area (140). It is thought that, under controlled laboratory conditions where one can ensure that the response from the previous trial decays to background levels before the next trial is started, feed-forward inputs fully account for the response transients, whereas recurrent processing plays a major role in shaping the posttransient response (196). Obviously, this is unlikely to be true under natural viewing conditions, where the visual system receives a continuous stream of overlapping stimulations.

It is important to emphasize that subcortical visual neurons differ considerably from cortical neurons in terms of the temporal dynamics of the response. The most notable difference is that, in subcortical neurons, the response decays to a much lesser degree after the initial transient.

# Redundancy reduction and adaptive filtering in early visual processing

Obviously, the fact that neuronal responses vary over time even in the absence of any changes in the stimulus itself has computational consequences because it means that information conveyed by the neurons changes over time even when the stimulus itself does not. But is this response change a necessary evil that simply reflects that fact that neurons, being biological entities, can keep up a given response only for so long? Or does it also have an adaptive value which serves a useful computational purpose? It is clear enough that these two possibilities are not mutually exclusive, but is there a reason to believe that the temporal variations in response serve a computational end?

A study by Dan and colleagues (71) provided empirical evidence that it might. They studied the responses of neurons in the lateral geniculate nucleus (LGN) in anesthetized cats. They stimulated the neurons with relatively long (20-60 min) stretches of grayscale (or "black and white") natural movies.<sup>20</sup> The control stimuli were white noise (e.g.,



Figure 24 A schematic illustration of correlation, decorrelation, and sparsening of the responses at the population level during the initial rapid transient responses (panel A), or at later stages (panels B-E). Each panel shows a highly idealized "population" consisting of four neurons (*circles*). Each *quadrant* of a given circle denotes the response of the neuron to a given stimulus, color-coded according to the color scale at *bottom left*. See text for details. Adapted, with permission, from (140).

TV "snow") movies of similar length. An analysis of the frequency composition of the various movies showed, as expected, that there was substantial temporal correlation, or redundancy, among the natural movies. The movies had more information (or, technically, more power) at the lower temporal frequencies (or "red" frequencies), so that the power spectrum of the movies was, on average, "pink." One would expect that the information in the stimulus would be reflected in the corresponding neuronal responses. However, the responses of LGN neurons to the movies were much less "pink" than expected. Instead the responses were largely "whitened," so that contributions from the lower temporal frequencies did not dominate the neuronal responses as they did in case of the stimuli.

More recent studies have shown that cells in the cat primary visual area (striate cortex or area 17, homologous to monkey V1) dynamically adapt to visual stimuli, and over even shorter time courses. Neurons in the striate cortex of the anesthetized cat show stimulus-dependent adaptive changes over the course of a few tens to several thousand milliseconds (298). These temporal changes appear to make the cells ultimately more sensitive to underrepresented spatial frequencies so that, after the adaptive change, these cells act more like a set of filters optimally suited to process the visual input.

# Neuronal responses of individual neurons and neuronal populations sparsen over time

In many visual areas, neuronal responses during the onset transient tends to be similar, or in technical terms, correlated, not only across cells, but across stimuli as well (Fig. 24A). Following the transients, the responses tend to become

dissimilar from one neuron to the next, so that population response becomes decorrelated. But how exactly do the neuronal responses become dissimilar across the population? One scenario is that the responses decrease differentially from one cell to the next after the transient. Another, nonexclusive, possibility is that the responses become dissimilar from one stimulus to the next.

Figures 24B to E schematically illustrate four such scenarios. It turns out that all these scenarios are applicable to some degree to a given neuronal population, depending on the exact set of circumstances. In some instances, responses decorrelate in a restricted subpopulation of cells (Fig. 24B, *bottom row*), whereas the responses of the remaining cells remain largely unchanged (Fig. 24B, *top row*). In other cases, all cells decorrelate after the transient, but different subpopulations of cells decorrelate similarly, so that the responses remain correlated within the given subpopulation (Fig. 24C; also see (198)).

#### There is no "one-size-fits-all" coding strategy

It is useful, from the point of view of a neurophysiologist, to emphasize the following three interrelated observations about efficient coding, in the event that they are not evident from the first principles alone. First, the best efficient-coding strategy is likely to vary depending on the precise nature of the underlying computation. There is no single efficient-coding strategy that is likely to be successful for every single information processing task (241, 307). For instance, the optimal strategy for efficiently coding an indoor scene, with its geometric regularities and unique statistical redundancies, is unlikely to be the same as the optimal strategy for coding an outdoor scene.



Figure 25 Visual cliff demonstrates development of depth perception. (A) The visual cliff is a laboratory apparatus that helps test depth perception in human infants and animals. It consists of an actual cliff covered with a sturdy but transparent plexiglass (112). The cliff is textured with a high-contrast checkerboard pattern, so that the cliff is clearly visible through the plexiglass. An infant called by his mother from the opaque side of the apparatus readily crawls to her (112). On the other hand, he is reluctant to venture over the perceived cliff (panel B). Even when the infants know by patting the glass that it is solid, they still tend to be reluctant to cross. Infants' decision as to whether or not to cross the visual cliff are also influenced by whether the gestures of the parent are encouraging, neutral, or discouraging (268). Such behaviors show sophisticated inferences based on a joint evaluation of various depth cues, risks, and rewards. Studies show that healthy human infants have such depth perception even before they are able to crawl (49). The visual cliff effect has been reported in many mammalian species (94). For a video of visual cliff effect, see https://www.youtube.com/watch?v=p6cqNhHrMJA.

Needless to say, the optimal strategy will also depend on what is being coded, and to what end.

Second, *computational* efficiency by itself may not determine how the brain performs information coding. There is mounting evidence that *metabolic* efficiency may be a crucial determinant in brain function. For instance, various types of sparsening (52, 162, 241, 311) can be computationally inefficient under certain circumstances, because most neurons are "idle" in a sparsened network (Fig. 24; also see (52, 162, 241, 311)). However, such idling is self-evidently efficient from a metabolic viewpoint. Indeed, this may be the reason why sparsening at every level of neuronal activity seems to be so widespread in the brain (131).

Finally, the functional organization of the brain is, in the ultimate analysis, shaped by a large number of evolutionary factors. No single factor, efficiency of whatever ilk, is likely to be determinative. The brain is not an organ *designed* to do anything—it simply is a part of organisms that have evolved. It is important to remember that, in the ruthless calculus of life, no single evolutionary trait, including a great brain, will a successful organism make.

### Dynamics on a Slower Time Scale: Visual Cognition Changes in Complex Ways from Cradle to Grave

#### Changes in early development

Visual development in nonhuman primates is relatively well understood, in part because invasive lesion and neurophysiological studies are possible in this case (for reviews, see (7,182,235)). One of the most intriguing insights gained from these studies is that, for the most part, receptive field properties of individual neurons in the infant brain are substantially more mature than infant visual function. In most cases, the properties of single neurons alone are not sufficient to account for visual development. This discrepancy is attributable to many factors, including weak signaling by infant neurons, correlated firing among neurons, and comparatively delayed development of neuronal mechanisms for pooling and reading out the responses of populations of neurons (182, 183, 305).

Another important insight, gained from behavioral studies, is that, surprisingly sophisticated high-level visual faculties are apparent at very early stages of development (see, e.g., Fig. 25; see legend for details). These faculties become rapidly become even more sophisticated, in a process referred to as "perceptual narrowing," over the course of early development (235). For instance, 6-month old human infants discriminate individual faces of monkeys as well as they discriminate the faces of individual humans. Discrimination performance of 9month old infants was comparable to that of the adults, in that they discriminated individual human faces well, but monkey faces much less well (248). The neural mechanisms of such perceptual narrowing remain largely unclear.

Compared to our understanding of the adult visual system, comparatively less is known about the functional organization of the visual system in very young human infants, mainly because of the difficulties in assuring the safety and comfort of infants inside the loud, dark, and scary high-field magnetic environments of MR or MEG scanners. However, this situation is slowly changing (for reviews, see (124, 263)). The few neuroimaging studies that exist suggest that the functional organization of the visual system undergoes fairly



Figure 26 fMRI in young human infants. (A) Renderings of what infants at various ages are likely to see when they view a teddy bear. (B) Visual responses in the neonate. Visually responsive regions are located in the anterior aspect of the calcarine sulcus in either hemisphere. Moreover, the visually evoked responses are lower compared to the periods of rest. (C) Visual responses in a different 5-month old infant. Visual stimulation activates a much posterior aspect of the calcarine sulcus. Also, the visually evoked responses are higher than the responses during rest. Panels B and C are courtesy of Dr. Ernst Martin (215) and reproduced with permission.

large-scale changes during the first few weeks after birth (Fig. 26).

#### Healthy and pathological aging are accompanied by generic and specific deficits in high-level vision

As is well known, healthy aging in human and nonhuman primates results in the general diminishment of the senses and visual cognitive faculties (91, 117, 150, 177). Older adults also show significant deficits in social cognition (226), although social cognition is a complex process in which vision plays only a part. It is also a matter of common experience that healthy, age-related changes show considerable individual variation, variously attributable to genetic differences, differences in lifestyle, etc. While aging results in a general decrease in performance in visual tasks (e.g., increases in reaction times not accounted for by slowing of motor responses), some aspects of vision diminish faster than others. By the eighth decade of life, most people have significantly reduced ability to discriminate colors and luminance. One prevalent notion, termed the "sensory deficit hypothesis of aging," posits that such deficits in visual cognitive abilities are largely attributable to the aging of the sensory organs *per se*. For instance, when subjects are asked to match digits and symbols written in low-contrast font, the performance of older adults is comparable to those of much younger adults (114).

However, healthy aging is also known to be accompanied by a number of changes in the functional organization of the visual system. For instance, in a PET study by Grady and colleagues, younger and older adults performed a location- or face-matching task using the same set of stimuli (122, 123). Both young and old subjects showed occipitotemporal rCBF (regional cerebral blood flow) activation during face matching and occipitoparietal activation during location matching when these conditions were compared to the control task. However, in both tasks, young subjects showed greater activation of the prestriate cortex (Brodmann's area 18), and old subjects had larger rCBF increases in the occipitotemporal cortex (area 37). Areas in the prefrontal cortex, as well as in the inferior and medial parietal cortices, were more activated in the older subjects during location matching. These findings suggest that: (i) Younger adults make more efficient use of occipital visual areas than older subjects. (ii) In older subjects, visual processing, especially of spatial information, is much more distributed across the visual cortex and less dependent on the occipital cortex, presumably because of the reduced processing efficiency of the occipital cortex. (iii) Spatial vision may be affected to a greater degree by aging than is object vision.

While Grady and colleagues did not specifically seek to test the aforementioned sensory deficit hypothesis, their results nonetheless illustrate the difficulty of testing (and, if applicable, falsifying) it. It is possible that larger effects on spatial vision in older adults may be due to differential aging of the relevant parts of the brain, or differential sensitivity of various parts of the brain due to age-related changes in the eye or, to varying degrees depending on the visual task, both.

Changes in cognitive faculties in various age-related diseases are better understood, but the underlying neurophysiological causes still remain poorly understood, in part because the underlying causes vary a great deal depending on the disease and clinical treatment or management of the disease (177).

Altogether, life cycle changes in the neurophysiology of vision, especially during early development and aging, are a trending area of research. The neural factors that contribute to the differential development of vision-related faculties (e.g., reading), and the diminishment of visual cognitive faculties in healthy aging, age-related diseases and various clinical interventions (e.g., chemotherapy) remain poorly understood and also constitute a major research trend.

### Subcortical and Cerebellar Structures Play Crucial, But Still Poorly Understood, Roles in Visually Guided Behavior

# Pulvinar and basal ganglia are important to visual function

Subcortical structures such as the thalamus (of which the pulvinar is a part) and the basal ganglia have been unintended victims of the persistent corticocentric bias in our understanding of visual cognition and sidelining of the subcortical structures. In all fairness, it should be noted that modern connectomic treatments of brain function have not rushed to correct this oversight, either (for an overview, see (314); also see the other relevant articles in this series, (300)). This neglect of subcortical structures is in no small part due to the technical difficulties of studying them: The individual nuclei tend to be relatively small and deep in the brain, so that they are, compared to areas in the cerebral cortex, substantially harder to locate, access and study. Basal ganglia do receive significant input from extrastriate visual areas, and indirectly project back to the cortex, indirectly through the pulvinar. This corticostriatal loop plays a major role in goal-directed and reward-based behaviors, visual and otherwise (for overviews, see (38, 152, 181)). There are multiple corticostriatal loops with overlapping circuitries (see Fig. 19). Studies in monkeys and humans have shown that one of the corticostriatal loops, the visual loop, is involved in the perceptual learning of visual object categories (38, 152, 181). However, much remains to be learned about the precise role of subcortical structures in high-level visual processes.

# Cerebellum plays key modulatory role in visually guided behavior

Historically, the cerebellum has been thought of as a structure related to motor functions, especially motor control, and its role in cognitive process had been long ignored. The fact that people born with a complete lack of a cerebellum, a condition known as cerebellar agenesis, can live to ripe old age helped further the notion that the cerebellum is not indispensable to brain function.

However, it is now clearer than ever that the cerebellum plays a major role in visual cognition, especially when adaptive processing, sensory-motor learning, and timing control are involved (for reviews, see (320, 351); also see the other relevant articles in this series). Strikingly, recent anatomical studies have shown that, in humans, the majority of the human cerebellum maps to the association cortex in the cerebrum (for a review, see (42)). Even more remarkably, the lateralization (i.e., interhemispheric asymmetry) evident in the cerebral substrates of language and attention are mirrored in this mapping (42, 321, 354).

In one of the earliest neuroimaging studies that demonstrated the role of the cerebellum in high-level cognitive functions, Petersen and colleagues measured brain function using PET while human subjects viewed written words and engaged in progressively more elaborate tasks based on the words (255, 256). At the most basic task level, participants passively viewed the words (e.g., nouns like cake, dog, and tree). At the next level, the subjects read the words aloud. At the most demanding level, the subjects provided action verbs that were semantically related to the written words (e.g., eat, walk, and climb, respectively). When subjects generated words in this fashion, right lateral cerebellum was preferentially activated (Fig. 27). Since these preferential responses were not attributable to motor activity per se (255, 256), these results indicate that the cerebellum was involved in intricate computations that involve vision, language, and action.

The precise neurophysiological mechanisms that underlie the role of cerebellum in higher level cognitive functions are almost entirely unclear. This aspect of cerebellar function is likely to be a major "growth area" in neurophysiological research in the future (Table 2).



Figure 27 Preferential responses in the human cerebellum during high-level cognitive task. Panel A shows the differential PET responses in a heatmap format, where brighter colors represent greater response. Panel B schematically summarizes the regions (*red squares*) that showed the task-dependent preferential response. Note that the responses are highly lateralized. Adapted, with permission, from (256).

### Dysfunction of High-Level Vision Helps Elucidate Mechanisms of High-Level Vision

As noted in Figure 12 above, lesion studies represent a highly useful, perturbative approach to understanding brain function (27, 89, 212, 246). But when it comes to high-level vision, lesions present a conundrum: On the one hand, targeted lesions can be only made in animal systems, but it is hard to measure nuanced high-level percepts in animals, including monkeys. For this reason, although lesion studies in animals have shed much light on visual processing, they have had limited success in elucidating high-level visual processes (7, 128, 169, 220, 356). On the other hand, while one can measure nuanced, high-level percepts in human subjects, human lesions, by definition, are uncontrolled, and vary greatly from patient to the next. The fact that each patient is unique makes it very hard to quantitatively relate data across patients. Such methodological difficulties have hampered progress in this field for decades.

However, situation is gradually improving, thanks in part to the development of modern multivariate methods of quantitatively studying, individual subjects (see below). Moreover, since understanding brain lesions is crucial to treating them, studies of brain dysfunction in general, and of lesions in particular, continue to be of great clinical significance. For all these reasons, neuropsychology is likely to be another promising "growth area" for research (Table 2). We will briefly summarize one specific impairment, hemineglect, as an exemplar of current limitations and future directions of lesion studies.

# Hemineglect has a complex, multivariate symptomology

Hemineglect, also referred to as hemispatial neglect, unilateral neglect, or spatial neglect, is a neurological impairment in which patients fail to be aware of one side of the visual world or extrapersonal space (161, 172, 312) (Fig. 28). While, for a variety of familiar reasons, hemineglect is most often clinically evident as a visual deficit, other sensory faculties are



Figure 28 Hemineglect. This figure shows the results from a drawing test (119, 135) from a single patient with left hemineglect, resulting from a localized lesion in the right temporal lobe. The patient was asked the draw the dial of a clock. In most clinical cases, lesions tend to be less circumscribed and more widespread than in the patient whose drawing is shown here. Hence, drawings by most hemineglect patients tend to be much more complex, and less clear-cut, than the "text book" case shown in this figure (see (119, 135) for reviews). Figure courtesy of Scholarpedia.

often also impaired in hemineglect. In other words, the symptomology of hemineglect tends to be complex in any given patient, with deficits and impairments not fully captured by the label "hemineglect" (170,204,347). For instance, hemineglect patient may have varying degrees of lateralized weakness of limb or eye muscles; varying degrees of deficits of spatial reference frames, motor abilities, motor control, and memory deficits. They may have varying degrees of visual field deficits or no visual field deficits at all. It is also important to note parenthetically that, for a variety of reasons, hemineglect is nearly not as well-recorded in nonhuman animals, including apes and monkeys.

# Left hemineglect is much more common than right hemineglect

Localized temporal lesions to a given side of the hemisphere generally result in the hemineglect of the opposite hemifield. Hemineglect is most prominent and long-lasting after damage to the right hemisphere, as opposed to after damage to the left hemisphere.

In other words, neglect of the left hemifield is much more common than the neglect of the right hemifield, even though damage to the two corresponding contralateral hemispheres is about equally common. This is a clear enough manifestation that the two hemispheres process the visual information in markedly different fashions. The exact neurophysiological reasons for this are not yet clear (for an overview of possible mechanisms, see (172, 312)).

# Hemineglect can result from lesioning of any of many different brain regions

The conventional wisdom has been that neglect is associated with lesions of the right posterior parietal cortex, especially lesions of the inferior parietal lobe (IPL) or of temporoparietal junction (TPJ) (338). Other studies have suggested other candidate locations, including subcortical ones (171,227). But a striking insight to emerge from a fairly large body of work is that there is no single brain location which, if lesioned, will lead to hemineglect.

How does one go about quantitatively relating multivariate structural deficit data in hemineglect (and brain lesions in general) to the resulting multivariate clinical symptoms? When the question is posed this way, it is straightforwardly clear that this a field where major advances in the future will come from informatic, "big data" approaches that quantitatively relate structural versus functional data. Quantitative databases of are already available for hemineglect (171,227), and it should be relatively straightforward to create similar databases of structural deficits of other brain lesion syndromes. What is currently lacking are detailed, multivariate quantifications of the functional deficits.

#### Vision Acts in Concert with Other Senses: A Brief Overview of Cross-Modal Perception

For eminently understandable practical reasons, we tend to study each of the senses in isolation. But behavioral studies have long made it clear that, especially under real-world conditions, the brain concurrently uses information from all senses, and often combines them so as to compensate for what is missing in the information from one of the senses is compensated for, or augmented, from the corresponding information from another sense (fore reviews, see (317, 333)). Thus, we recognize speech both from the sound and lip movements. We test whether a fruit is ripe by jointly using the senses of vision, touch and smell in a mutually interactive fashion. When we eat, the external senses of touch, smell, and taste, as well as the visceral senses of hunger and satiety, come into play. In other words, multisensory processing is not a novelty act for the brain, it is the brain's native mode of operation. Needless to say, combining information across senses is even more important when one of the senses is impaired. A classic example is braille, wherein visually impaired people use fingertips as "seeing eyes."

The visual system has extensive, reciprocal anatomic and functional connections with other sensory systems at multiple levels of processing. The superior colliculus, a midbrain center that plays a well-known role in eye movements, is famously multimodal (318). In the cortex, many of the areas along the superior temporal sulcus that are well-studied for their role in high-level visual perception, including STP (superior temporal polysensory area) that has long been known for its selectivity for faces (Fig. 15), have long been known to be polysensory. Multimodal connectivity is evident at an early stage during development, and the functional properties of multimodal neurons undergo narrowing and specialization during development, with multimodal sensory experience being an important factor in this specialization (318).

But recent anatomical, neurophysiological, and neuropsychological studies have shown that cortical areas conventionally thought to be exclusively involved in visual processing, such as areas V1 and V2, have extensive connections with the auditory system, especially in the peripheral representation of the visual field (Fig. 29). Multimodal functional interactions have also been found in the human brain (Fig. 30; see legend for details). That is, visual processing affects, and is affected by, other sensorimotor modalities (for overviews, see (194, 279, 317, 333)).

Collectively, these discoveries of multimodal interactions have raised the possibility that much of the neocortex is essentially multisensory (109). This notion would once have seemed radical. But in light of the available and mounting evidence, it is simply an acknowledgement of the fact that vision does not act alone, and that the brain is set up to combine, and jointly evaluate, the available information from all the senses.



Figure 29 Our evolving understanding of multimodal anatomical connections with the visual system. (A) Traditional view of the cortical anatomy of the primate brain recognized very few areas with multimodal anatomical connections (*colored areas*). (B) A more modern scheme of the cortical anatomy of multisensory areas. *Colored areas* represent regions where anatomical and/or electrophysiological studies have demonstrated multisensory interactions. *Dashed gray outlines* represent opened sulci. See (109) for details, including the criteria used for determining multimodal connectivity at the anatomical level. Adapted, with permission, from (109).

It also helps further illustrate the fact that "vision for vision's sake" is a misleading, antiquated notion.

# Synesthesia is an unusual consequence of sensory "cross-talk"

Synesthesia is a cognitive phenomenon in which stimulation of one sensory modality or pathway leads to a perceptual experience in another sensory modality or pathway (25,69). It is relatively rare, with a reported incidence of about 4% (70). It is fairly clear that at least some forms of synesthesia have a genetic basis, although synesthesia can also be induced pharmacologically (70).

Synesthesia is of interest in the study of high-level vision, essentially for the same reason as binocular rivalry and other types of perceptual instability are: They represent perceptual phenomena in which the actual percept is dissociated from the stimulus, that is, the same stimulus can produce more than one percept. The fact that they are relatively easy to induce reproducibly in the laboratory is also important. After all, high-level phenomena such as dreams or sensations of déjà vu are also intriguing high-level perceptual phenomena, but they are all but impossible to induce reproducibly. People who experience synesthesia, or synesthetes, show considerable individual differences (25, 69).

There are many forms of synesthesia. One comparatively common form of synesthesia is grapheme-color synesthesia

or color-graphemic synesthesia, wherein letters or numbers are perceived as inherently colored (Fig. 31A). Cross-modal forms of synesthesia include instances where sounds are perceived as colors (a phenomenon referred to as chromesthesia), or where tasting food elicits a sensation of tactile shapes (25,69).

Neural mechanisms have not been well understood for any form of synesthesia. In one of the most notable studies of the neural mechanisms of synesthesia, Hubbard and colleagues used fMRI to compare the neural responses of graphemecolor synesthetes versus control subjects to graphemic and nongraphic stimuli. They found that color graphemic stimuli activate some brain regions similarly in both synesthetes and controls subjects. But synesthetes show additional activation of the color-selective portion of the retinotopic region V4 (hV4), whereas nonsynesthetes do not show hV4 activation (Fig. 31B).

There is some neuroimaging evidence that suggests that synesthesia may be caused by an increased cross-talk between brain regions that specialize in the processing of information from different sensory modalities (154, 155).

### Individual Differences in High-Level Perception

It has long been appreciated that individuals differ considerably in every known facet of visual perception (see, e.g., (72)).



Figure 30 Visual-haptic object processing activates lateral occipital complex (LOC) in the occipitotemporal pathway. Ahmedi and colleagues (8) compared BOLD responses to four conditions: visual objects, somatosensory (or haptic) objects, visual textures, and haptic textures. Statistical map of the contralateral hemisphere from a single subject are shown in panel A (3D folded view), panel B (inflated view of the same hemisphere), and panel C (flattened view of the same hemisphere). Bottom, BOLD responses to the four conditions are shown in the somatosensory cortex (bottom left) and LOC in the occipitotemporal junction (bottom right). Col S, collateral sulcus; Cal S, calcarine sulcus; CS, central sulcus; IPS, intraparietal sulcus; lateral S, lateral sulcus; STS, superior temporal sulcus. Adapted, with permission, from (8).

Strictly speaking, it remains possible that our subjective experiences (sometimes referred to as qualia) are different, even when the underlying stimuli are the same (for discussions of this intriguing topic, see (77, 87, 190)).

Vision research has traditionally ignored individual differences, partly because of practical necessity. Most of the available data analysis techniques use a so-called "frequentist" approach, which estimate the underlying variables from a sufficiently large sample of individuals. Individual variations are necessarily "averaged out" using such an approach. This is not to say that statistical analyses of individual differences are impossible in frequentist statistics. For instance, *predictions* can be made for an individual subject based on *estimations* of the underlying variable based on a sample. Nonetheless, a frequentist approach makes it inherently harder to analyze individual variations. But many new approaches to analyzing



Figure 31 Color-graphemic synesthesia. (A) (*left*) A stimulus that can elicit synesthesia in color-graphemic synesthetes. (A) (*right*) A rendition of what the synesthete likely to have perceived. Note that, since color-graphemic synesthetes tend to perceive different numbers as different colors, the triangle made up of 2's stands out, or "pops out," perceptually for them. By contrast, nonsynesthetes perceive all the numbers to be of the same color, so that for them, the triangle is not readily distinguishable from the background. (B) Neural responses during color-graphemic synesthetes (*left*) and control subjects and results are rendered on inflated, bottom-up views of brains of representative subjects. Both control subjects and synesthetes but not in nonsynesthetes. See (156) for details. Adapted, with permission, from (156).

and characterizing individual differences have been developed (see, e.g., (55, 192)).

It should be noted that characterizing individual variations is more than just an analytical pastime. Nothing makes us more unique as humans than our brains (or sometimes the figurative lack thereof). Thus, understanding individual variations in brain activity is critical to explaining what makes each of us unique. Understanding individual differences, and helping other fields exploit and cater to individual differences (e.g., in evidence-based medicine), is a major growth area of neuroscientific research.

### Major Challenges Lie Ahead in Porting the Laboratory Studies of Vision to the Real World

# Laboratory studies differ from real-world situations in many ways

### Introspection and active reporting can have confounding brain activations

On the one hand, under real-world conditions, vision contends with daunting complexities and ambiguities that are hard to capture under controlled conditions. On the other hand, typical laboratory visual tasks evidently require explicit, active reporting that requires or elicits introspection, memorization, etc. That is, these tasks require the subject to think introspectively about what they saw and actively report what it is that they saw. Obviously, this is not what happens under realworld conditions. In the real-world, instead, visual perception is a continuous stream, and we rarely dip into it to actively introspect, or actively report, what we just saw.

But does such introspection or reporting make a difference? In an intriguing recent study, Frässle and colleagues (100) monitored brain responses during binocular rivalry with or without active reports by the subjects. They took advantage of the fact that the subjects' eye movements and pupil sizes during rivalry were diagnostic of the subjects' percepts, so that these objective eye tracking measures could be reliably used as markers for what the subjects did perceive when they were not required to actively report their percepts. The investigators then simply compared the brain responses with or without active reporting. They found that many occipital and parietal areas responded similarly under both conditions, indicating that the responses in these areas reflected the visual percept, rather than the introspection or reporting of the percept. However, the activity in many areas of the frontal lobe was active only under the active reporting conditions. Thus, these frontal areas, previously reported to be involved in binocular rivalry in studies using active reporting, appear to be involved in introspection and active reporting. In a larger sense, these results constitute a compelling "proof of principle" that brain activity under laboratory task conditions is unlikely to be the same as brain activity under real-world conditions.

A variety of rsfMRI studies have shown that the learning of new tasks, including task-related motor responses to new stimuli, is known produce major changes in the intrinsic connectivity among brain regions including, but not limited to, retinotopic visual areas that have been previously known to be activated during various experimental visual tasks (125, 132, 203, 209). This suggests that laboratory studies produce brain activity related to the study paradigm *per se* that may not occur under real-world conditions.

Rewards (and in some cases, punishments) under laboratory conditions are unlike those in the real world in two main respects. First, visual perception is rarely associated with rewards as periodic, immediate, or explicit in the real world as they typically are under laboratory conditions. Second, rewards are intricately associated with reporting under laboratory conditions, but the two do not regularly cooccur in the real world. In the real world, rewards, when they come, typically do so even when the percept is not reported. Altogether, rewarding regimes typically used in laboratory studies create conditions that are quite different from those in the real world.

#### Is the connectomic view just another fad?

The connectomic view essentially asserts that, to fully understand how the brain works, understanding the responses of individual regions is not enough. We also need to understand how the various parts work as a whole, that is, how the various regions talk to each other when the brain "does" something (e.g., when it produces a given behavior). Thus, it simply seeks a broader understanding that subsumes all our previous understanding-a larger, holistic explanation that accounts for narrower ones. An argument for considering the whole instead of just parts is unlikely to turn out to be just a fad. Besides, as noted throughout this review and within many of the references cited, there is compelling empirical evidence that brain function is mediated by dynamic networks. However, it is not only possible, but also likely, that our view of brain networks are likely to evolve and refine as we learn about how they work. Throughout this review, I have noted various areas of "low-hanging fruit" for future research, where research is likely to make a significant impact.

### Conclusion: Connectome Cometh!

Over the last few decades we have learned a lot of about how various parts of the visual system work. What is less clear is how various parts of the visual system function as a whole, especially under real-world conditions. As noted above, we know enough to know that the older view that visual processing proceeds in a feed-forward, hierarchical manner is quite wrong (57, 315), but we do not yet have a new framework of understanding to replace it. However, as to the question of how vision works, while we do not yet know the answer itself, we do know the *nature* of the answer, and it is fundamentally connectomic: Vision is a process in which visual sensory information from the eyes is mapped to various effector organs of behavior by dynamic networks of the brain. Visual perception and visual awareness are the results of the readout of this process. While much remains to be discovered to elucidate exactly how this happens, it is clear that the connectomic sea change in our understanding of brain function in general, and of vision in particular, is already underway, and has been underway for some time now (28, 116, 211, 296, 346). The truly exciting prospect is that we will have the outlines of an answer fairly soon, probably within the lifetimes of most of us.

### Acknowledgments

Work in my laboratory is supported by continuing and past grants from the U.S. Army (grants #W911NF-11-1-0105 and

W911NF-15-1-0311), National Institutes of Health (grant #R21 NS086356), and National Science Foundation (grant #IOS-1147097). I thank Ms. Jennevieve Sevilla for invaluable assistance during the preparation of the manuscript. I am grateful to two anonymous reviewers for helpful comments about the manuscript, and to Dr. Norman Johnson for useful discussions.

### Endnotes

- 1. In majority of cases, the macaque monkey (*Macaca* spp.), especially *Macaca mulatta*, or the Rhesus macaque.
- 2. Crowding and clutter refer to essentially the same aspect of the image. To a first approximation, the former is a term of art common in the psychophysical literature, and the latter is used in neurophysiology and machine vision.
- 3. For a more detailed and engaging recounting of this history by three great scientists who helped make it, see Gross CG. Brain, Vision, Memory. Tales in the History of Neuroscience. Cambridge: MIT Press, 1998; and Hubel DH and Wiesel TN, Brain and Visual Perception: The Story of a 25-Year Collaboration. New York: Oxford University Press, 2005.
- 4. There are two main systems of defining and naming macaque visual areas. One is the widely adopted system delineated by Van Essen and his colleagues (see, e.g., Felleman DJ and Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1: 1-47, 1991, and the references therein). The other is the system delineated by Mishkin, Gross, and their colleagues (see, e.g., Gross CG, Bruce CJ, Desimone R, Fleming R, and Gattass R. Cortical visual areas of the temporal lobe: Three areas in the macaque. In: Cortical Sensory Organization, edited by Woolsey CN. New York: Humana Press, 1981, pp. 187-216; Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, and Mishkin M. The ventral visual pathway: An expanded neural framework for the processing of object quality. Trends Cogn Sci 17: 26-49, 2013; and the references therein). The two systems of cortical parcellation/nomenclature overlap considerably, especially at the early and intermediate levels of the visual hierarchy where the defining criteria of the various areas are objectively clearer. The present review mostly adopts the nomenclature of Van Essen et al., except in case of Figure 9, which follows the Mishkin et al. nomenclature. Which nomenclature system to choose is mostly a matter of personal preference; neurophysiologists tend to adopt the system of nomenclature using which they were originally trained.
- 5. During a typical microelectrode recording session involving a visual task, the animal sits comfortably in a chair with its head fixed and performs, for a reward, a behavioral task in which it is trained. The animal's eye positions are monitored and are often controlled (e.g., brief fixations are required), the responses of isolated neurons or groups of neurons are measured using one or more microelectrodes inserted through the dura. A great majority of previous studies recorded the activity of one neuron at a time. While it is now technically possible to concurrently record from a large number of neurons from a given area, or even from multiple different areas across the brain, it is still not possible to carry out whole-brain neurophysiological recordings. Therefore, how the responses of individual neurons add up to the activity of brain networks is largely unclear.

- 6. This process of putting together the individual features that are processed apart into a unified, holistic percept is often referred to as "binding" (for reviews, see Hardcastle VG. Consciousness and the neurobiology of perceptual binding. Semin Neurol 17: 163-170, 1997, 234; Robertson LC. Binding, spatial attention and perceptual awareness. Nat Rev Neurosci 4: 93-102, 2003, 253; Shadlen MN and Movshon JA. Synchrony unbound: A critical evaluation of the temporal binding hypothesis. Neuron 24: 67-77, 111-125, 1999). Whether binding is a real process is a matter of debate. It has been proposed that synchronization of neuronal firing is a mechanism for, and is diagnostic of, binding. It should be noted that, while the synchronization of neuronal firing is an empirically demonstrated fact, this by itself does not prove that binding is a real process, any more than the undisputed existence of Loch Ness proves the existence of Loch Ness Monster. The main reason for suspecting that binding is not real process is that, as outlined elsewhere in this review, it increasingly clear that the visual system does not really process individual visual features apart from each other in the first place, thus obviating the need to putting them back together. Thus, at best, it remains to be seen whether or not binding is epiphenomenal as a brain process (see September 1999 issue of Neuron). Alternatively, binding may turn out to be just a digestive phenomenon not discussed in polite company.
- 7. It is worth noting parenthetically in this context that one of the persistent myths about these pathways is that dorsal pathway is primarily a magnocellular pathway and the ventral pathway is primarily a parvocellular pathway. This is simply untrue. See, e.g., Callaway EM. Neural substrates within primary visual cortex for interactions between parallel visual pathways. Prog Brain Res 149: 59-64, 2005; and Sincich LC and Horton JC. The circuitry of V1 and V2: Integration of color, form, and motion. Annu Rev Neurosci 28: 303-326, 2005.) for additional info.
- 8. In the context of visual perception, the Bayes' law of conditional probability can be expressed as  $p(S|I) = [p(S) \times p(I|S)]/p(I)$ , where *I* is the retinal image of the given visual scene *S* in the real world, and  $p(I) \neq 0$ . This reformulation of Bayes law represents the best possible way for a computational system to combine the information about S and I to infer which real-world scene S is "out there" in front of the eyes given the image I that the scene forms on the retina. The quantity p(I) is simply a normalizing constant, which means that it can be safely left out the equation, which then reduces the equation to  $p(S|I) = p(S) \times p(I|S)$ . The quantities p(S|I), p(S), and p(I|S) are referred to as the posterior distribution, prior distribution (or "prior"), and likelihood distribution (or likelihood function), respectively. The prior essentially represents the a priori belief about the chances of encountering various scenes. Note that when the prior probability of encountering all scenes is the same, or when there is no reason to believe that they are not the same, the prior becomes uniform or "flat," and conveys no information. In this case, the above equation further reduces p(S|I) = p(I|S), but the Bayes' law still holds. This illustrates that priors, while useful, are dispensable in the Bayesian framework. One of the common misconceptions about the Bayes framework is that it fundamentally consists of updating one's prior belief based on new sensory evidence I. The above thought exercise demonstrates that this is not strictly true, and illustrates why, that is, when the prior is flat, there is no prior belief to update, and the likelihood function solely determines the posterior distribution. On the other hand, if the likelihood function is flat and the prior is not, the posterior distribution is

given solely by the prior distribution. Thus, no single quantity on the right hand of the equation is indispensable, nor is the computation necessarily limited to using information about only S and/or I. Bayesian framework may reduce or expand depending on the given case to include all probabilistic variables that are relevant to inferring S given I. As noted in the next, the brain's inferences are often influenced by the possible risks and benefits of a given behavior. The Bayes framework provides a way of expanding (or contracting) the above Bayes' equation to accommodate such additional factors (or lack thereof). This flexibility of the Bayes' equation is a big reason why it is so useful in realworld computations. For details, see Knill D and Richards W (Eds.), Perception as Bayesian Inference. New York: Cambridge University Press, 1996; and Kersten D, Mamassian P, and Yuille A. Object perception as Bayesian inference. Annu Rev Psychol 55: 271-304, 2004.

- 9. It is important to distinguish the concept of "prior knowledge of the visual world" from the aforementioned concept of priors. The brain "knows" more about the visual world than just the *a priori* chances of encountering a particular scene. The brain has a vast amount of information about the visual world (including information about what the various visual objects and scenes "look like") that it acquires during development and through perceptual learning. This prior knowledge is ultimately incorporated in the Bayes' framework in the form of the likelihood function, and the prior belief is incorporated though the prior distribution. So "prior knowledge" and "priors" have very distinct meanings, to appreciate which is an important aspect of Bayes literacy. For details, see Knill D and Richards W (Eds.), Perception as Bayesian Inference. New York: Cambridge University Press, 1996; and Kersten D, Mamassian P, and Yuille A. Object perception as Bayesian inference. Annu Rev Psychol 55: 271-304, 2004.
- 10. For an exemplar graphical representation of a brain network in action, see https://sms.cam.ac.uk/media/1587539.
- "Transient" or "response transient" refers to the burst of spikes fired by the neuron at the start of its stimulus-evoked response. That is, the response transient refers to that portion of the response in which the firing rate rapidly rises and peaks.
- 12. Strictly speaking, there are some very narrow, special cases of evolution, such as drift and extinction, where this principle of natural selection does not necessarily hold, and the adaptive value of the four F's does not matter. Thus, natural selection, drift, and extinction constitute the neo-Darwinian evolutionary framework that accounts for the evolution of all neural systems.
- 13. Studies of the human brain tend to refer to parcellations of the brain as "regions," as opposed to "areas." This both because "area" is a rigorous term that refers to a specific portion of the cerebral cortex defined using a set of objective criteria that include architectonics, connectivity, and neurophysiology. These criteria typically require invasive studies, and have been carried out in animal systems, including monkeys. In case of human brain, where invasive studies highly difficult, if not impossible, brain regions are parcellated solely using functional criteria, that is, on the basis of their neural responses. The locations and boundaries of these parcellations can vary substantially, depending on the exact functional test used for defining them. Therefore, it would be inappropriate to refer to these parcellations as areas. Besides, the term "region" has the incidental advantage that it can be used for referring to noncortical parcellations of the brain, e.g., subthalamic nucleus or globus

pallidus. The term "area" is not suitable for referring to these or any other noncortical structures, in humans or animals; it is typically reserved for cortical structures defined using the aforementioned criteria.

- 14. Indeed, such stimuli that induce bistable or multistable percepts in the complete absence of physical changes in the stimulus are a highly useful experimental tool for dissociating stimulusdriven factors from other factors that affect perception. See Kim CY and Blake R. Psychophysical magic: Rendering the visible "invisible." Trends Cogn Sci 9: 381-388, 2005.
- 15. Indeed, while panda's thumb may appear to be very similar to the human thumb, it is actually an extension of a completely different bone, as explained by Stephen Jay Gould in his elegant book *Panda's Thumb* (New York, W.W. Norton and Co. 1980).
- 16. Vertical meridians are generally given greater weight than horizontal meridians in determining the boundaries between visual areas, because vertical meridian representations are related to connectivity between the two hemispheres (and therefore the two visual hemifields), mainly through the corpus callosum.
- 17. In the cat, the striate cortex is referred to as Area 17, because it is equivalent to Area 17 in the human brain described by Brodmann (see Finger S. Origins of Neuroscience: A History of Explorations into Brain Function. New York: Oxford University Press, 1994).
- For a video demonstration of this phenomenon, see http:// journals.plos.org/plosone/article/asset?unique&id=info:doi/ 10.1371/journal.pone.0158504.s002
- 19. Note that the term "local" here does not mean that the relevant information is *spatially* localized. Instead, it means that the information is limited to a small, local region of the *coding space*. Thus, when all or most of the information necessary to support a given percept is represented by just one neuron or a small set of neurons, we say that the information is locally coded. The proverbial "grandmother neuron" or the neuron that responds only when your grandmother comes into view, is a prototypical example of local coding, because this neuron by itself can signal when your grandmother comes into view. But when individual neurons code only part of the information, so that no individual neuron carries all of the relevant information, but a large number of neurons collectively do, the information is said to be distributed across such neurons.
- 20. One of the movies they showed happened to be Casablanca.
- 21. Tuning is somewhat of a touchstone for visual feature selectivity in neurophysiological investigations, because it denotes a systematic dose-response relationship, that is, a systematic relationship with a given stimulus variable and a neuronal response. For an overview of this principle, see Parker AJ and Newsome WT. Sense and the single neuron: Probing the physiology of perception. Annu Rev Neurosci 21: 227-277, 1998.

### References

- Aertsen A, Preißl H. Dynamics of activity and connectivity in physiological neuronal networks. In: HG S, editor. *Nonlinear Dynamics and Neuronal Networks*. New York: Wiley-VCH, 1991, pp. 281-302.
- 2. Aghdam HH, Heravi EJ. *Guide to Convolutional Neural Networks*. New York: Springer, Berlin, Heidelberg, 2017.
- Ahissar M, Hochstein S. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci* 8: 457-464, 2004.
- Ajina S, Bridge H. Blindsight and unconscious vision: What they teach us about the human visual system. *Neuroscientist* 23: 529-541, 2016.

- Allman J, Miezin F, McGuinness E. Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci* 8: 407-430, 1985.
- 6. Allman JM, Kaas JH. A representation of the visual field in the caudal third of the middle tempral gyrus of the owl monkey (*Aotus trivirgatus*). *Brain Res* 31: 85-105, 1971.
- Alvarado MC, Bachevalier J. Revisiting the maturation of medial temporal lobe memory functions in primates. *Learn Mem* 7: 244-256, 2000.
   Amedi A, Malach R, Hendler T, Peled S, Zohary E. Visuo-haptic object-
- Amedi A, Malach R, Hendler T, Peled S, Zohary E. Visuo-haptic objectrelated activation in the ventral visual pathway. *Nat Neurosci* 4: 324-330, 2001.
- Andersen RA, Cui H. Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* 63: 568-583, 2009.
   Andersen RA, Snyder LH, Bradley DC, Xing J. Multimodal represen-
- Andersen RA, Snyder LH, Bradley DC, Xing J. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu Rev Neurosci* 20: 303-330, 1997.
- Angelucci A, Bressloff PC. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extraclassical receptive field surround of primate V1 neurons. *Prog Brain Res* 154: 93-120, 2006.
- Archambault PS, Ferrari-Toniolo S, Caminiti R, Battaglia-Mayer A. Visually-guided correction of hand reaching movements: The neurophysiological bases in the cerebral cortex. *Vision Res* 110: 244-256, 2015.
- Ashby FG. Statistical Analysis of fMRI Data. Cambridge, MA: MIT Press, 2011.
- Atkinson AP, Adolphs R. The neuropsychology of face perception: Beyond simple dissociations and functional selectivity. *Philos Trans R* Soc Lond B Biol Sci 366: 1726-1738, 2011.
- 15. Attneave F. Some informational aspects of visual perception. *Psychol Rev* 61: 183-193, 1954.
- 16. Bagloee SA, Tavana M, Asadi M, Oliver T. Autonomous vehicles: Challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation* 24: 284-303, 2016.
- Banerjee A, Dean HL, Pesaran B. A likelihood method for computing selection times in spiking and local field potential activity. *J Neurophysiol* 104: 3705-3720, 2010.
- Banks MS, Gepshtein S, Landy MS. Why is spatial stereoresolution so low? J Neurosci 24: 2077-2089, 2004.
- 19. Bar M. Visual objects in context. Nat Rev Neurosci 5: 617-629, 2004.
- Bar M, Aminoff E. Cortical analysis of visual context. *Neuron* 38: 347-358, 2003.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hamalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E. Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103: 449-454, 2006.
- Barlow H. The mechanical mind. *Annu Rev Neurosci* 13: 15-24, 1990.
   Barlow HB. Pattern recognition and the responses of sensory neurons.
- Ann N Y Acad Sci 156: 872-881, 1969.
- Barnett L, Seth AK. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *J Neurosci Methods* 223: 50-68, 2014.
- Baron-Cohen S, Harrison JE. Synaesthesia: Classic and Contemporary Readings. Cambridge, MA: Blackwell, 1997.
- Bart E, Hegdé J. Invariant Recognition of Visual Objects. Lausanne, Switzerland: Frontiers Media SA, 2013.
- 27. Barton JJ. Higher cortical visual function. *Curr Opin Ophthalmol* 9: 40-45, 1998.
- Bassett DS, Sporns O. Network neuroscience. Nat Neurosci 20: 353-364, 2017.
- Berzuini C, Dawid P, Bernardinelli L. Causality: Statistical Perspectives and Applications. Hoboken, NJ: Wiley, 2012.
- Biederman I. Perceiving real-world scenes. *Science* 177: 77-80, 1972.
   Biswal BB. Resting state fMRI: A personal history. *Neuroimage* 62:
- BISWAI BB. Resting state IMRI: A personal history. *Neuroimage* 52: 938-944, 2012.
   Blake DT Network supervision of adult experience and learning depen-
- Blake DT. Network supervision of adult experience and learning dependent sensory cortical plasticity. *Compr Physiol* 7: 977-1008, 2017.
- Blake R, Logothetis N. Visual competition. Nat Rev Neurosci 3: 13-21, 2002.
- Bloom FE, Morrison JH, Young WG. Neuroinformatics: A new tool for studying the brain. J Affect Disord 92: 133-138, 2006.
- Braddick O. Encyclopedia of Perception. Thousand Oaks, CA: SAGE Publications, Inc., 2010.
- Bremmer F, Krekelberg B. Seeing and acting at the same time: Challenges for brain (and) research. *Neuron* 38: 367-370, 2003.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745-4765, 1992.
- Brown LL, Schneider JS, Lidsky TI. Sensory and cognitive functions of the basal ganglia. *Curr Opin Neurobiol* 7: 157-163, 1997.
- Brown RJ, Norcia AM. A method for investigating binocular rivalry in real-time with the steady-state VEP. *Vision Res* 37: 2401-2408, 1997.

- Bruce C, Desimone R, Gross CG. Visual properties of neurons in a 40 polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46: 369-384, 1981.
- 41 Bruno RM. Synchrony in sensation. Curr Opin Neurobiol 21: 701-708, 2011
- Buckner RL. The cerebellum and cognitive function: 25 years of insight 42. from anatomy and neuroimaging. Neuron 80: 807-815, 2013. Buckner RL, Krienen FM, Yeo BT. Opportunities and limitations of
- 43 intrinsic functional connectivity MRI. Nat Neurosci 16: 832-837, 2013.
- Bullier J. Integrated model of visual processing. Brain Res Brain Res 44. Rev 36: 96-107, 2001.
- 45. Bullier J, Hupe JM, James AC, Girard P. The role of feedback connections in shaping the responses of visual cortical neurons. *Prog Brain* Res 134: 193-204, 2001.
- Burgess N, O'Keefe J. Models of place and grid cell firing and theta rhythmicity. *Curr Opin Neurobiol* 21: 734-744, 2011. 46.
- Cabeza R, Ciaramelli E, Moscovitch M. Cognitive contributions of the 47. ventral parietal cortex: An integrative theoretical account. Trends Cogn Sci 16: 338-352, 2012.
- 48. Callaway EM. Neural substrates within primary visual cortex for interactions between parallel visual pathways. Prog Brain Res 149: 59-64, 2005
- Campos JJ, Langer A, Krowitz A. Cardiac responses on the visual cliff in prelocomotor human infants. *Science* 170: 196-197, 1970. 49.
- 50
- Cao JW, Lin ZP. Extreme learning machines on high dimensional and large data applications: A survey. *Math Probl Eng* 2015: 1-13, 2015. Cavanagh P. What's up in top-down processing? In: Gorea A, edi-tor. *Representations of Vision: Trends and Tacit Assumptions in Vision* 51. Research. New York: Cambridge University Press, 1991, pp. 295-304.
- Chacron MJ, Longtin A, Maler L. Efficient computation via sparse 52. coding in electrosensory neural networks. Curr Opin Neurobiol 21: 752-760, 2011.
- Chen Y, Elenee Argentinis JD, Weber G. IBM Watson: How cogni-53. tive computing can be applied to big data challenges in life sciences research. *Clin Ther* 38: 688-701, 2016. Chentanez T, Keatisuwan W, Akaraphan A, Chaunchaiyakul R,
- 54 Lechanavanich C, Hiranrat S, Chaiwatcharaporn C, Glinsukon T. Reaction time, impulse speed, overall synaptic delay and number of synapses in tactile reaction neuronal circuits of normal subjects and thinner sniffers. Physiol Behav 42: 423-431, 1988.
- Christensen R. Bayesian Ideas and Data Analysis: An Introduction for 55. Scientists and Statisticians. Boca Raton, FL: CRC Press, 2011.
- 56. Churchland PS. A neurophilosophical slant on consciousness research. Progress in Brain Research 149: 285-293, 2005.
- Churchland PS, Ramachandran VS, Sejnowski TJ. A critique of pure 57. vision. In: Koch C, Davis J, editors. *Large-Scale Neuronal Theories of the Brain*. Cambridge, MA: MIT Press, 1994, pp. 23-60. Coates P, Coleman S. *Phenomenal Qualities: Sense, Perception, and Scale Neuronal Complexity*.
- 58. Consciousness. New York: Oxford University Press, 2015.
- 59 Cohen MR, Maunsell JH. Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12: 1594-1600, 2009
- Cohen MR, Newsome WT. Estimates of the contribution of single neu-60. rons to perception depend on timescale and noise correlation. JNeurosci 29: 6635-6648, 2009
- Cowey A, Stoerig P. Blindsight in monkeys. Nature 373: 247-249, 61. 1995.
- Crapse TB, Basso MA. Insights into decision making using choice probability. *J Neurophysiol* 114: 3039-3049, 2015. Crasto CJ. *Neuroinformatics*. Totowa, NJ: Humana, 2007. 62
- 63
- Crick F. Visual perception: Rivalry and consciousness. Nature 379: 64. 485-486, 1996.
- Crick F, Koch C. Are we aware of neural activity in primary visual 65. cortex? Nature 375: 121-123, 1995.
- Crick F, Koch C. Consciousness and neuroscience. Cereb Cortex 8: 66. 97-107, 1998
- 67. Crick F, Koch C. A framework for consciousness. Nat Neurosci 6: 119-126, 2003.
- Cumming BG, Nienborg H. Feedforward and feedback sources of 68 choice probability in neural population responses. Curr Opin Neurobiol 37: 126-132, 2016.
- Cytowic RE. Synesthesia: A Union of the Senses. Cambridge, MA: MIT 69. Press. 2002
- 70. Cytowic RE, Eagleman D. Wednesday is Indigo Blue: Discovering the Brain of Synesthesia. Cambridge, MA: MIT Press, 2009.
- Dan Y, Atick JJ, Reid RC. Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. J Neurosci 16: 3351-3362, 1996.
- Davidoff JB. Differences in Visual Perception: The Individual Eye. London: Crosby Lockwood Staples, 1975. 72.
- Davies ER. Computer and Machine Vision: Theory, Algorithms, Prac-73 ticalities. Amsterdam; Boston: Elsevier, 2012.

- DeAngelis GC, Cumming BG, Newsome WT. Cortical area MT 74. and the perception of stereoscopic depth. Nature 394: 677-680, 1998.
- DeAngelis GC, Newsome WT. Organization of disparity-selective neu-75. rons in macaque area MT. J Neurosci 19: 1398-1415, 1999. DeAngelis GC, Newsome WT. Perceptual "read-out" of conjoined
- 76. direction and disparity maps in extrastriate area MT. PLoS Biol 2: E77, 2004.
- Dennett DC. Consciousness Explained. Boston: Little, Brown and Co., 77. 1991.
- Desimone R, Ungerleider LG. Neural mechanisms of visual processing 78. in monkeys. In: Boller F, Grafman J, editor. Handbook of Neuropsvchology. Amsterdam: Elsevier, 1989.
- DeYoe EA, Van Essen DC. Concurrent processing streams in monkey visual cortex. *Trends Neurosci* 11: 219-226, 1988. 79.
- 80 Dienes Z. Understanding Psychology As A Science: An Introduction to Scientific and Statistical Inference. New York: Palgrave Macmillan, 2008
- 81. Dobzhansky T. Nothing in biology makes sense except in the light of evolution. Am Biol Teach 35: 125-129, 1973.
- Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RS, Friston KJ. How the brain learns to see objects and faces in an impoverished context. Nature 389: 596-599, 1997.
- Doya K. Bayesian Brain: Probabilistic Approaches to Neural Coding. 83.
- Cambridge, MA: MIT Press, 2007. Driver J, Spence C. Cross-modal links in spatial attention. *Philos Trans R Soc Lond B Biol Sci* 353: 1319-1331, 1998. Einhauser W, Kruse W, Hoffmann KP, Konig P. Differences of monkey 84
- 85. and human overt attention under natural conditions. Vision Res 46: 1194-1209, 2006.
- Einhauser W, Mundhenk TN, Baldi P, Koch C, Itti L. A bottom-up 86. model of spatial attention predicts human error patterns in rapid scene recognition. J Vis 7: 6.1-6.13, 2007.
- Eliasmith C, Mandik P. Qualia. In: Hochstein E, editors. Dictionary of 87. Philosophy of Mind, Waterloo, Canada: Sites.google.com, 2004.
- Elston GN. Cortical heterogeneity: Implications for visual processing and polysensory integration. *J Neurocytol* 31: 317-335, 2002. 88
- Emmans D, Laihinen A, Halsband U. Comparative Neuropsychology and Brain Imaging. Zürich: LIT Verlag, 2015. 89.
- 90. Engel AK, Moll CK, Fried I, Ojemann GA. Invasive recordings from the human brain: Clinical insights and beyond. Nat Rev Neurosci 6: 35-47, 2005.
- 91. Erwin J, Hof PR. Aging in Nonhuman Primates. New York: Karger, 2002.
- 92 Fabre-Thorpe M, Richard G, Thorpe SJ. Rapid categorization of natural images by rhesus monkeys. Neuroreport 9: 303-308, 1998.
- 93. Fahle M, Poggio T. Perceptual Learning. Cambridge, MA: MIT Press, 2002
- 94. Fantz RL. The origin of form perception. Sci Am 204: 66-72, 1961.
- Farah MJ, Ratcliff G. The Neuropsychology of High-Level Vision: Col-lected Tutorial Essays. Hillsdale, NJ: Lawrence Erlbaum Associates, 95. 1994
- Fawcett JM, Risko EF, Kingstone A. The Handbook of Attention. Cam-96. bridge, MA: MIT Press, 2015.
- 97. Fei-Fei L, Iyer A, Koch C, Perona P. What do we perceive in a glance of a real-world scene? J Vis 7: 10, 2007.
- Felleman DJ, Van Essen DC. Distributed hierarchical processing in the 98.
- Frinate cerebral cortex. Cereb Cortex 1: 1-47, 1991. Frassle S, Lomakina EI, Razi A, Friston KJ, Buhmann JM, Stephan KE. Regression DCM for fMRI. Neuroimage 155: 406-421, 2017. 99
- 100. Frassle S, Sommer J, Jansen A, Naber M, Einhauser W. Binocular rivalry: Frontal activity relates to introspection and action but not to perception. J Neurosci 34: 1738-1747, 2014.
- Freedman DJ. Familiarity breeds plasticity: Distinct effects of experi-101. ence on putative excitatory and inhibitory neurons in inferior temporal cortex. *Neuron* 74: 8-11, 2012.
- 102. Freedman DJ, Assad JA. Experience-dependent representation of visual categories in parietal cortex. Nature 443: 85-88, 2006.
- Friston K, Moran R, Seth AK. Analysing connectivity with Granger 103. causality and dynamic causal modelling. Curr Opin Neurobiol 23: 172-178, 2013
- 104. Friston KJ. Functional and effective connectivity: A review. Brain Connect 1: 13-36, 2011.
- 105. Futuyma DJ. Evolution. Sunderland, MA: Sinauer Associates, 2009.
- Geisler WS. Contributions of ideal observer theory to vision research. *Vision Res* 51: 771-781, 2011. 106.
- 107. Geisler WS, Diehl RL. Bayesian natural selection and the evolution of perceptual systems. Philos Trans R Soc Lond B Biol Sci 357: 419-448, 2002.
- 108. Geisler WS, Kersten D. Illusions, perception and Bayes. Nat Neurosci 5: 508-510, 2002.
- Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? 109 Trends Cogn Sci 10: 278-285, 2006.

- 110. Ghose GM. Learning in mammalian sensory cortex. Curr Opin Neuro-biol 14: 513-518, 2004.
- Ghose GM, Yang T, Maunsell JH. Physiological correlates of perceptual learning in monkey V1 and V2. *J Neurophysiol* 87: 1867-1888, 2002. 111.
- Gibson EJ, Walk RD. The "visual cliff." Sci Am 202: 64-71, 1960. 112. 113. Gilbert CD, Sigman M, Crist RE. The neural basis of perceptual learn-
- ing. Neuron 31: 681-697, 2001.
- 114. Gilmore GC, Spinks RA, Thomas CW. Age effects in coding tasks: Componential analysis and test of the sensory deficit hypothesis. Psychol Aging 21: 7-18, 2006.
- 115. Girard P, Hupe JM, Bullier J. Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction
- velocities. J Neurophysiol 85: 1328-1331, 2001.
   116. Glasser MF, Smith SM, Marcus DS, Andersson JL, Auerbach EJ, Behrens TE, Coalson TS, Harms MP, Jenkinson M, Moeller S, Robin-Berge SC, 2018, 2018, 2018. son EC, Sotiropoulos SN, Xu J, Yacoub E, Ugurbil K, Van Essen DC. The Human Connectome Project's neuroimaging approach. Nat Neurosci 19: 1175-1187, 2016.
- 117. Gliga T. Handbook of developmental social neuroscience. Neuropsychol Rehabil 20: 637-638, 2010.
- Glimcher PW. Indeterminacy in brain and behavior. Annu Rev Psychol 118. 56: 25-56, 2005.
- Goldstein LH, McNeil JE. Clinical Neuropsychology: A Practical Guide to Assessment and Management for Clinicians. Chichester, West
- Sussex; Malden, MA: Wiley-Blackwell, 2013.
  Goodale MA, Milner AD. Separate visual pathways for perception and action. *Trends Neurosci* 15: 20-25, 1992.
- 121. Goswami G, Powell BM, Vatsa M, Singh R, Noore A. FR-CAPTCHA: CAPTCHA based on recognizing human faces. PLoS One 9: e91708, 2014.
- 122. Grady CL. Age-related changes in cortical blood flow activation during perception and memory. Ann N Y Acad Sci 777: 14-21, 1996
- 123. Grady CL, Maisog JM, Horwitz B, Ungerleider LG, Mentis MJ, Salerno JA, Pietrini P, Wagner E, Haxby JV. Age-related changes in cortical blood flow activation during visual processing of faces and location. *J Neurosci* 14: 1450-1462, 1994.
   Graham AM, Pfeifer JH, Fisher PA, Lin W, Gao W, Fair DA. The
- potential of infant fMRI research and the study of early life stress as a
- promising exemplar. *Dev Cogn Neurosci* 12: 12-39, 2015. Gregory MD, Agam Y, Selvadurai C, Nagy A, Vangel M, Tucker M, Robertson EM, Stickgold R, Manoach DS. Resting state connectiv-125. ity immediately following learning correlates with subsequent sleepdependent enhancement of motor task performance. Neuroimage 102 (Pt 2): 666-673, 2014.
- 126. Grill-Spector K, Kanwisher N. Visual recognition: As soon as you know it is there, you know what it is. Psychol Sci 16: 152-160, 2005.
- Grill-Spector K, Malach R. The human visual cortex. Annu Rev Neurosci 27: 649-677, 2004. 127.
- Gross CG. Processing the facial image: A brief history. Am Psychol 60: 128. 755-763, 2005.
- 129. Gross CG, Bruce CJ, Desimone R, Fleming R, Gattass R. Cortical visual areas of the temporal lobe: Three areas in the macaque. In: Woolsey CN, editor. Cortical Sensory Organization. New York: Humana Press, 1981, pp. 187-216.
- 130. Gross CG, Rodman HR, Gochin PM, Colombo MW. Inferior temporal cortex as a pattern recognition device. In: Baum E., editor. Computational Learning and Cognition. Philadelphia: Society for Industrial and Applied Mathematics, 1993, pp. 44-73. 131. Grossberg S. Linking visual development and learning to information
- processing: Preattentive and attentive brain dynamics. In: DeWeerd P, Pinaud R, Tremere L, editors. Plasticity in the Visual System: From Genes to Circuits. New York: Springer/Kluwer Academic Press, 2005, pp. 323-346.
- 132. Ĝuidotti R, Del Gratta C, Baldassarre A, Romani GL, Corbetta M. Visual learning induces changes in resting-state fMRI multivariate pattern of information. J Neurosci 35: 9786-9798, 2015.
- 133. Guyonneau R, Vanrullen R, Thorpe SJ. Temporal codes and sparse representations: A key to understanding rapid processing in the visual system. J Physiol Paris 98: 487-497, 2004.
- 134. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. Microstructure
- of a spatial map in the entorhinal cortex. *Nature* 436: 801-806, 2005.
  135. Halligan PW, Kischka U, Marshall JC. *Handbook of Clinical Neuropsychology*. Oxford; New York: Oxford University Press, 2010.
  136. Hansen PC, Kringelbach ML, Salmelin R. *MEG: An Introduction to*
- Methods. New York: Oxford University Press, 2010.
- 137. Hardcastle VG. Consciousness and the neurobiology of perceptual binding. Semin Neurol 17: 163-170, 1997.
- Harnad S. To cognize is to categorize: Cognition is categorization. In: 138. Cohen H, Lefebyre C, editors. *Handbook of Categorization in Cognitive Science*. San Diego, CA: Elsevier, 2005, pp. 19-43.
   Hasson U, Hendler T, Ben Bashat D, Malach R. Vase or face? A neural
- correlate of shape-selective grouping processes in the human brain. *J Cogn Neurosci* 13: 744-753, 2001.

- Hegdé J. Time course of visual perception: Coarse-to-fine processing and beyond. *Prog Neurobiol* 84: 405-439, 2008.
- Hegdé J, Felleman DJ. Reappraising the functional implications of the primate visual anatomical hierarchy. Neuroscientist 13: 416-421, 2007.
- 142. Hegde J, Kersten D. A link between visual disambiguation and visual memory. *J Neurosci* 30: 15124-15133, 2010. 143. Hegdé J, Van Essen DC. A comparative study of shape representation in
- macaque visual areas V2 and V4. Cereb Cortex 17: 1100-1116, 2007.
- 144. Hegdé J, Van Essen DC. Temporal dynamics of shape analysis in macaque visual area V2. *J Neurophysiol* 92: 3030-3042, 2004.
- 145. Henderson JM, Hollingworth A. Eye movements and visual memory: Detecting changes to saccade targets in scenes. Percept Psychophys 65: 58-71, 2003.
- Henderson JM, Hollingworth A. High-level scene perception. Annu Rev Psychol 50: 243-271, 1999.
- 147. Henson RN. Repetition suppression to faces in the fusiform face area: A personal and dynamic journey. Cortex 80: 174-184, 2016.
- Hochstein S, Ahissar M. View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron* 36: 791-804, 2002. 148
- 149. Hodge MJS, Radick G. The Cambridge Companion to Darwin. Cambridge; New York: Cambridge University Press, 2009.
- 150. Hof PR, Mobbs CV. Functional Neurobiology of Aging. San Diego, CA: Academic Press, 2001.
- 151. Hoffman KL, Logothetis NK. Cortical mechanisms of sensory learning and object recognition. Philos Trans R Soc Lond B Biol Sci 364: 321-329, 2009.
- 152. Hoshi E. Cortico-basal ganglia networks subserving goal-directed behavior mediated by conditional visuo-goal association. Front Neural Circuits 7: 158, 2013.
- 153. Howard IP, Rogers BJ. Perceiving in Depth. New York: Oxford University Press, 2012.
- 154. Hubbard EM, Arman AC, Ramachandran VS, Boynton GM. Individual differences among grapheme-color synesthetes: Brain-behavior corre-lations. *Neuron* 45: 975-985, 2005.
- 155. Hubbard EM, Piazza M, Pinel P, Dehaene S. Interactions between number and space in parietal cortex. Nat Rev Neurosci 6: 435-448, 2005
- Hubbard EM, Ramachandran VS. Neurocognitive mechanisms of synesthesia. *Neuron* 48: 509-520, 2005.
- 157. Huurneman B, Boonstra FN, Cox RF, Cillessen AH, van Rens G. A systematic review on 'Foveal Crowding' in visually impaired children and perceptual learning as a method to reduce crowding. BMC Ophthalmol 12: 27, 2012.
- 158. Hyder F. Dynamic Brain Imaging: Multi-Modal Methods and In Vivo Applications. New York: Humana, 2009.
- 159. Isik L, Meyers EM, Leibo JZ, Poggio T. The dynamics of invariant object recognition in the human visual system. J Neurophysiol 111: 91-102, 2014.
- James W, Rouben Mamoulian Collection (Library of Congress). *The Principles of Psychology*. New York: H. Holt and Company, 1890. 160.
- Jeannerod M. Neurophysiological and Neuropsychological Aspects of 161. Spatial Neglect. Amsterdam; New York; North-Holland: Elsevier Science Pub. Co., 1987.
- Jefferis GS, Livet J. Sparse and combinatorial neuron labelling. Curr 162. Opin Neurobiol 22: 101-110, 2012.
- 163. Johnson NA. Darwinian Detectives: Revealing the Natural History of Genes and Genomes. Oxford; New York: Oxford University Press, 2007
- 164. Kaas JH. The evolution of brains from early mammals to humans. Wiley Interdiscip Rev Cogn Sci 4: 33-45, 2013.
- 165. Kaas JH. Evolution of columns, modules, and domains in the neocortex of primates. Proc Natl Acad Sci U S A 109 (Suppl 1): 10655-10660, 2012
- 166. Kaas JH. The evolution of neocortex in primates. Prog Brain Res 195: 91-102, 2012
- Kaas JH. Why does the brain have so many visual areas? Journal of 167. Cognitive Neuroscience 1: 121-135, 1989.
- 168. Kaas JH, Stepniewska I. Evolution of posterior parietal cortex and parietal-frontal networks for specific actions in primates. *J Comp Neurol* 524: 595-608, 2016.
- 169. Kaldy Z, Sigala N. The neural mechanisms of object working memory: What is where in the infant brain? *Neurosci Biobehav Rev* 28: 113-121, 2004
- 170. Karnath HO. Spatial attention systems in spatial neglect. Neuropsychologia 75: 61-73, 2015.
- 171. Karnath HO, Fruhmann Berger M, Kuker W, Rorden C. The anatomy of spatial neglect based on voxelwise statistical analysis: A study of 140 patients. Cereb Cortex 14: 1164-1172, 2004.
- 172. Karnath HO, Milner AD, Vallar G. The Cognitive and Neural Bases of Spatial Neglect. Oxford; New York: Oxford University Press, 2002. 173. Karthik S, Paul A, Karthikeyan N. Deep Learning Innovations and
- their Convergence with Big Data. Hershey, PA: Information Science Reference, 2018.

- Kay KN, Naselaris T, Prenger RJ, Gallant JL. Identifying natural images from human brain activity. *Nature* 452: 352-355, 2008.
   Kelly C, Biswal BB, Craddock RC, Castellanos FX, Milham MP. Char-
- acterizing variation in the functional connectome: Promise and pitfalls. Trends Cogn Sci 16: 181-188, 2012.
- 176. Kendall AL, Hantraye P, Palfi S. Striatal tissue transplantation in nonhuman primates. Prog Brain Res 127: 381-404, 2000.
- Kensinger EA. Cognition in ageing and age-related disease. In: Handbook of the Neuroscience of Aging. New York: Academic Press, 2009
- 178. Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. *Annu Rev Psychol* 55: 271-304, 2004.
- 179. Kettlewell HB. Insect survival and selection for pattern: Most camouflage and survival mechanisms, though highly perfected, can be adapted to changing environments. *Science* 148: 1290-1296, 1965.
- 180. Killian NJ, Jutras MJ, Buffalo EA. A map of visual space in the primate entorhinal cortex. Nature 491: 761-764, 2012.
- Kim HF, Hikosaka O. Parallel basal ganglia circuits for voluntary and 181. automatic behaviour to reach rewards. Brain 138: 1776-1800, 2015.
- 182. Kiorpes L. Visual development in primates: Neural mechanisms and critical periods. Dev Neurobiol 75: 1080-1090, 2015.
- 183. Kiorpes L, Price T, Hall-Haro C, Movshon JA. Development of sensitivity to global form and motion in macaque monkeys (Macaca nemest-
- 184. Knierim JJ, Van Essen DC. Visual cortex: Cartography, connectivity, and concurrent processing. *Curr Opin Neurobiol* 2: 150-155, 1992.
  185. Knill DC, Pouget A. The Bayesian brain: The role of uncertainty in a concurrent processing. *Curr Opin Neuropiol* 2: 1710-710, 2004.
- neural coding and computation. Trends Neurosci 27: 712-719, 2004.
- 186. Kording KP, Wolpert DM. Bayesian integration in sensorimotor learning. Nature 427: 244-247, 2004.
- Kourtzi Z, Connor CE. Neural representations for object perception: 187. Structure, category, and adaptive coding. Annu Rev Neurosci 34: 45-67, 2011.
- 188. Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway: An expanded neural framework for the process-ing of object quality. *Trends Cogn Sci* 17: 26-49, 2013.
   Krekelberg B, Lappe M. Neuronal latencies and the position of moving
- objects. Trends Neurosci 24: 335-339, 2001.
- 190. Kriegel U. Current Controversies in Philosophy of Mind. New York: Routledge, 2013.
- 191. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: NIPS (Neural Information Processing Systems). San Francisco, CA: Morgan Kaufmann Publishers, 2012.
- Kruschke JK. Doing Bayesian Data Analysis: A Tutorial with R and BUGS. Burlington, MA: Academic Press, 2011.
   Kumar P, Tiwari A. Ubiquitous Machine Learning and Its Applications. Hershey, PA: IGI Global, 2017.
   Lacey S, Sathian K. Multisensory object representation: Insights from User Statistics and there here Pairs Pairs 101 (5) 176 (2011)
- studies of vision and touch. Prog Brain Res 191: 165-176, 2011.
- Lamme VA, Roelfsema PR. The distinct modes of vision offered by 195. feedforward and recurrent processing. Trends Neurosci 23: 571-579, 2000.
- 196. Lamme VA, Super H, Spekreijse H. Feedforward, horizontal, and feedback processing in the visual cortex. Curr Opin Neurobiol 8: 529-535, 1998
- 197. Laubach M. Who's on first? What's on second? The time course of learning in corticostriatal systems. Trends Neurosci 28: 509-511, 2005.
- Laurent G. Olfactory network dynamics and the coding of multidimen-sional signals. *Nat Rev Neurosci* 3: 884-895, 2002.
- 199. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 521: 436-444, 2015
- 200. Legatt AD, Arezzo J, Vaughan HG, Jr. Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: Effects of volume-conducted potentials. J Neurosci Methods 2: 203-217, 1980
- 201. Leopold DA, Logothetis NK. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. Nature 379: 549-553, 1996.
- 202. Levi DM. Crowding-An essential bottleneck for object recognition: A mini-review. Vision Res 48: 635-654, 2008.
- 203. Lewis CM, Baldassarre A, Committeri G, Romani GL, Corbetta M. Learning sculpts the spontaneous activity of the resting human brain. *Proc Natl Acad Sci U S A* 106: 17558-17563, 2009.
- 204. Li K, Malhotra PA. Spatial neglect. Pract Neurol 15: 333-339, 2015.
- 205. Lindenfors P. Neocortex evolution in primates: The "social brain" is for females. Biol Lett 1: 407-410, 2005.
- Logothetis NK, Leopold DA, Sheinberg DL. What is rivalling during binocular rivalry? *Nature* 380: 621-624, 1996.
- 207. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412: 150-157, 2001.
- 208. Luck SJ, Kappenman ES. Oxford Handbook of Event-Related Potential Components. Oxford: Oxford University Press, 2012.

- 209. Ma L, Narayana S, Robin DA, Fox PT, Xiong J. Changes occur in resting state network of motor system during 4 weeks of motor skill learning. Neuroimage 58: 226-233, 2011.
- 210. Malashichev YB, Deckel AW. Behavioral and Morphological Asymmetries in Vertebrates. Georgetown, TX: Landes Bioscience: Eureka.com, 2006.
- 211. Marcus GF, Freeman JA. The Future of the Brain: Essays by the World's
- Leading Neuroscientists. Princeton: Princeton University Press, 2015. Mariën P, Abutalebi J. Neuropsychological Research: A Review. Hove 212.
- England; New York: Psychology Press, 2008. 213. Marmarelis PZ, Marmarelis VZ. The White-Noise Method in System identification. In: Analysis of Physiological Systems: The White-Noise Approach. Boston, MA: Springer US, 1978, pp. 131-180.
- 214. Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco: W.H. Freeman, 1982.
- 215. Martin E. Imaging of brain function during early human development. In: Rutherford MA, editor. MRI of the Neonatal Brain. London: Mary A. Rutherford, 2016.
- 216. Maunsell JH, van Essen DC. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3: 2563-2586, 1983.
- McAdams CJ, Maunsell JH. Effects of attention on the reliability of 217. individual neurons in monkey visual cortex. Neuron 23: 765-773, 1999.
- McAdams CJ, Maunsell JH. Attention to both space and feature mod-ulates neuronal responses in macaque area V4. J Neurophysiol 83: 218. 1751-1755, 2000
- 219. McIntosh RD, Schenk T. Two visual streams for perception and action: Current trends. Neuropsychologia 47: 1391-1396, 2009.
- 220. Merigan WH, Maunsell JH. Macaque vision after magnocellular lateral geniculate lesions. Vis Neurosci 5: 347-352, 1990.
- 221. Merigan WH, Maunsell JH. How parallel are the primate visual pathways? Annu Rev Neurosci 16: 369-402, 1993.
- Middleton FA, Strick PL. Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cereb Cortex* 12: 926-935, 2002. 222.
- 223. Mill J. Rodent models: Utility for candidate gene studies in human attention-deficit hyperactivity disorder (ADHD). J Neurosci Methods 166: 294-305, 2007. 224. Mills TC. Time Series Econometrics: A Concise Introduction. Hound-
- mills, Basingstoke, Hampshire; New York: Palgrave Macmillan, 2015.
- 225. Mishkin M, Ungerleider LG, Macko KA. Object vision and spatial vision: Two cortical pathways. Trends in Neurosciences 6: 414-417, 1983
- 226. Moran JM, Jolly E, Mitchell JP. Social-cognitive deficits in normal aging. J Neurosci 32: 5553-5561, 2012. Mort DJ, Malhotra P, Mannan SK, Rorden C, Pambakian A, Kennard
- 227. C, Husain M. The anatomy of visual neglect. Brain 126: 1986-1997, 2003.
- Moser EI, Moser MB, Roudi Y. Network mechanisms of grid cells. *Philos Trans R Soc Lond B Biol Sci* 369: 20120511, 2014.
   Moser EI, Roudi Y, Witter MP, Kentros C, Bonhoeffer T, Moser MB.
- Grid cells and cortical representation. Nat Rev Neurosci 15: 466-481, 2014
- Moser MB, Rowland DC, Moser EI. Place cells, grid cells, and memory. 230. Cold Spring Harb Perspect Biol 7: a021808, 2015.
- Mukamel R, Fried I. Human intracranial recordings and cognitive neu-231. roscience. Annu Rev Psychol 63: 511-537, 2012.
- Murphey DK, Yoshor D, Beauchamp MS. Perception matches selec-tivity in the human anterior color center. *Curr Biol* 18: 216-220, 2008. Mutz F, Veronese LP, Oliveira-Santos T, de Aguiar E, Cheein FAA, 232
- 233. de Souza AF. Large-scale mapping in complex field scenarios using an autonomous car. Expert Syst Appl 46: 439-462, 2016
- 234. Naselaris T, Kay KN, Nishimoto S, Gallant JL. Encoding and decoding in fMRI. Neuroimage 56: 400-410, 2011.
- Nelson CA, Collins ML. Handbook of Developmental Cognitive Neu-235. roscience. Cambridge, MA: MIT Press, 2008.
- 236. Newsome WT, Britten KH, Salzman CD, Movshon JA. Neural mechanisms of motion perception. In: Cold Spring Harbor Symposia in Quantitative Biology. Cold Spring Harbor, NY: Cold Spring Harbor Lab Press, 1990, p. 697-705.
- Nielsen FA, Christensen MS, Madsen KH, Lund TE, Hansen LK. fMRI neuroinformatics. *IEEE Eng Med Biol Mag* 25: 112-119, 2006.
- Nobre K, Kastner S. The Oxford Handbook of Attention. Oxford; New 238 York: Oxford University Press, 2014.
- 239. Nunez PL, Srinivasan R. Electric Fields of the Brain: The Neurophysics
- of EEG. Oxford; New York: Oxford University Press, 2006. Okuyama-Uchimura F, Komai S. Mouse ability to perceive subjective 240. contours. Perception 45: 315-327, 2016.
- Olshausen BA, Field DJ. Sparse coding with an overcomplete basis set: A strategy employed by V1? Vision Research 37: 3311-3325, 1997.
- 242. Op de Beeck HP, Dicarlo JJ, Goense JB, Grill-Spector K, Papanastassiou A, Tanifuji M, Tsao DY. Fine-scale spatial organization of face and object selectivity in the temporal lobe: Do functional magnetic

resonance imaging, optical imaging, and electrophysiology agree? J. Neurosci 28: 11796-11801, 2008.

- 243. Orger MB, Smear MC, Anstis SM, Baier H. Perception of Fourier and non-Fourier motion by larval zebrafish. Nat Neurosci 3: 1128-1133, 2000
- 244. Osborne LC, Palmer SE, Lisberger SG, Bialek W. The neural basis for combinatorial coding in a cortical population response. J Neurosci 28: 13522-13531, 2008.
- 245. Panzeri S, Brunel N, Logothetis NK, Kayser C. Sensory neural codes using multiplexed temporal scales. Trends Neurosci 33: 111-120, 2010.
- 246. Papanicolaou AC. The Oxford Handbook of Functional Brain Imaging in Neuropsychology and Cognitive Neurosciences. New York: Oxford University Press, 2017.
- Parodi S, Riccardi G, Castagnino N, Tortolina L, Maffei M, Zoppoli G, Nencioni A, Ballestrero A, Patrone F. Systems medicine in oncology: Signaling network modeling and new-generation decision-support systems. Methods Mol Biol 1386: 181-219, 2016.
- 248. Pascalis O, de Haan M, Nelson CA. Is face processing species-specific during the first year of life? Science 296: 1321-1323, 2002.
- 249. Passingham R. How good is the macaque monkey model of the human brain? Curr Opin Neurobiol 19: 6-11, 2009.
- 250. Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433: 873-876, 2005.
- 251. Pelli DG. Crowding: A cortical constraint on object recognition. *Curr Opin Neurobiol* 18: 445-451, 2008.
- Pelli DG, Tillman KA, Freeman J, Su M, Berger TD, Majaj NJ. Crowd-ing and eccentricity determine reading rate. J Vis 7: 20 21-36, 2007.
- 253. Perry CJ, Baciadonna L, Chittka L. Unexpected rewards induce dopamine-dependent positive emotion-like state changes in bumble-bees. *Science* 353: 1529-1531, 2016.
- 254. Pessoa L, McMenamin B. Dynamic networks in the emotional brain. Neuroscientist 23: 383-396, 2016.
- Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron 255. emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331: 585-589, 1988.
   Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron
- emission tomographic studies of the processing of single words. J Cogn Neurosci 1: 153-170, 1989.
- 257. Platek SM, Shackelford TK. Foundations in Evolutionary Cognitive Neuroscience. Cambridge; New York: Cambridge University Press, 2009
- Poggio T, Torre V, Koch C. Computational vision and regularization theory. *Nature* 317: 314-319, 1985. 258.
- 259. Polonsky A, Blake R, Braun J, Heeger DJ. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat Neurosci* 3: 1153-1159, 2000. 260. Poort J, Raudies F, Wannig A, Lamme VA, Neumann H, Roelfsema
- PR. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. Neuron 75: 143-156, 2012.
- 261. Pouget A, Snyder LH. Computational approaches to sensorimotor transformations. Nat Neurosci 3 (Suppl): 1192-1198, 2000.
- 262. Pourtois G, Rauss KS, Vuilleumier P, Schwartz S. Effects of perceptual learning on primary visual cortex activity in humans. Vision Res 48: 55-62, 2008.
- 263. Prior J, Van Herwegen J. Practical Research with Children. New York: Routledge, 2016.
- 264.
- Prochazka A, Ellaway P. Sensory systems in the control of movement. *Compr Physiol* 2: 2615-2627, 2012. Pylyshyn ZW. Visual indexes, preconceptual objects, and situated vision. *Cognition* 80: 127-158, 2001. 265.
- 266. Raiguel S, Vogels R, Mysore SG, Orban GA. Learning to see the difference specifically alters the most informative V4 neurons. J Neurosci 26: 6589-6602, 2006.
- 267. Ribeiro AS, Lacerda LM, Ferreira HA. Multimodal Imaging Brain Connectivity Analysis (MIBCA) toolbox. PeerJ 3: e1078, 2015.
- 268. Richards JE, Rader N. Affective, behavioral, and avoidance responses
- on the visual cliff: Effects of crawling onset age, crawling experience, and testing age. *Psychophysiology* 20: 633-641, 1983. Riera JJ, Ogawa T, Goto T, Sumiyoshi A, Nonaka H, Evans A, Miyakawa H, Kawashima R. Pitfalls in the dipolar model for the neo-cortical EEG sources. *J Neurophysiol* 108: 956-975, 2012. 269.
- Rilling JK. Comparative primate neuroimaging: Insights into human brain evolution. *Trends Cogn Sci* 18: 46-55, 2014. 270.
- Robertson LC. Binding, spatial attention and perceptual awareness. *Nat Rev Neurosci* 4: 93-102, 2003. 271.
- 272. Roe AW, Chelazzi L, Connor CE, Conway BR, Fujita I, Gallant JL, Lu H, Vanduffel W. Toward a unified theory of visual area V4. Neuron 74: 12-29, 2012
- 273. Roelfsema PR. Cortical algorithms for perceptual grouping. Annu Rev Neurosci 29: 203-227, 2006.
- 274. Rogers J, Gibbs RA. Comparative primate genomics: Emerging patterns of genome content and dynamics. Nat Rev Genet 15: 347-359, 2014.

- putational bases, and a model, VisNet. *Front Comput Neurosci* 6: 35, 2012. 275. Rolls ET. Invariant visual object and face recognition: Neural and com-
- 276. Rosch E. Natural categories. Cognitive Psychology 4: 328-350, 1973.
- 277. Rubin N. Figure and ground in the brain. Nat Neurosci 4: 857-858, 2001.
- 278. Salzman CD, Britten KH, Newsome WT. Cortical microstimulation influences perceptual judgements of motion direction. Nature 346: 174-177, 1990.
- 279. Sathian K, Buxbaum LJ, Cohen LG, Krakauer JW, Lang CE, Corbetta M, Fitzpatrick SM. Neurological principles and rehabilitation of action disorders: Common clinical deficits. Neurorehabil Neural Repair 25: 21S-32S, 2011.
- 280. Schmid MC, Maier A. To see or not to see-Thalamo-cortical networks during blindsight and perceptual suppression. Prog Neurobiol 126: 36-48, 2015.
- Schmidt JT. Activity-driven sharpening of the retinotectal projection: 281. The search for retrograde synaptic signaling pathways. J Neurobiol 59: 114-133, 2004.
- 282. Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG. Signal timing across the macaque visual system. J Neurophysiol 79: 3272-3278, 1998.
- 283. Schooler LJ, Shiffrin RM, Raaijmakers JG. A Bayesian model for implicit effects in perceptual identification. Psychol Rev 108: 257-272,  $20\bar{0}1.$
- 284. Schoups A, Vogels R, Qian N, Orban G. Practising orientation identification improves orientation coding in V1 neurons. Nature 412: 549-553, 2001
- 285. Schwartz S, Maquet P, Frith C. Neural correlates of perceptual learning: A functional MRI study of visual texture discrimination. Proc Natl Acad Sci U S A 99: 17137-17142, 2002.
- 286. Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neuroscience and Biobehavioral Reviews 32: 265-278, 2008
- Seger CA, Miller EK. Category learning in the brain. Annu Rev Neurosci 287. 33: 203-219, 2010.
- 288. Seger CA, Peterson EJ. Categorization = decision making + generalization. *Neurosci Biobehav Rev* 37: 1187-1200, 2013.
- 289 Seitz AR, Dinse HR. A common framework for perceptual learning. *Curr Opin Neurobiol* 17: 148-153, 2007. Sejnowski TJ, Churchland PS, Movshon JA. Putting big data to good
- 290. use in neuroscience. Nature Neuroscience 17: 1440-1441, 2014.
- 291. Sekihara K, Nagarajan SS. Electromagnetic Brain Imaging. New York: Springer, Berlin, Heidelberg, 2015.
- 292 Sereno AB, Maunsell JH. Shape selectivity in primate lateral intraparietal cortex. Nature 395: 500-503, 1998.
- 293 Sereno MI, Allman JM. Cortical visual areas in mammals. In: Leventhal AG, editor. The Neural Basis of Visual Function. London: Macmillan,
- 1991, pp. 160-172. Sereno MI, Huang RS. Multisensory maps in parietal cortex. *Curr Opin* 294. Neurobiol 24: 39-46, 2014.
- 295 Seung S. Connectome: How the Brain's Wiring Makes us Who We Are. Boston: Houghton Mifflin Harcourt, 2012.
- 296. Seung S. Connectome: How the Brain's Wiring Makes us Who We Are. Boston: Mariner Books, Houghton Mifflin Harcourt, 2013.
- Shadlen MN, Movshon JA. Synchrony unbound: A critical evaluation 297 of the temporal binding hypothesis. Neuron 24: 67-77, 111-125, 1999.
- Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller 298. KD. Adaptive filtering enhances information transmission in visual cortex. *Nature* 439: 936-942, 2006.
- 299. Shepherd GM. Corticostriatal connectivity and its role in disease. Nat Rev Neurosci 14: 278-291, 2013.
- Sherman SM. Functioning of circuits connecting thalamus and cortex. Compr Physiol 7: 713-739, 2017.
- 301. Sherman SM, Guillery RW. Exploring the Thalamus and Its Role in Cortical Function. Cambridge, MA: MIT Press, 2006. Shettleworth SJ. Cognition, Evolution, and Behavior. Oxford; New
- 302. York: Oxford University Press, 2010.
- 303. Shipp S. The functional logic of corticostriatal connections. Brain Struct Funct 222: 669-706, 2017
- Shmuel A, Augath M, Oeltermann A, Logothetis NK. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nat Neurosci* 9: 569-577, 2006.
- 305. Shooner C, Hallum LE, Kumbhani RD, Ziemba CM, Garcia-Marin V, Kelly JG, Majaj NJ, Movshon JA, Kiorpes L. Population representation of visual information in areas V1 and V2 of amblyopic macaques. *Vision* Res 114: 56-67, 2015.
- 306. Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature 415: 318-320, 2002.
- 307. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annu Rev Neurosci 24: 1193-1216, 2001.
- Sincich LC, Horton JC. The circuitry of V1 and V2: Integration of color, form, and motion. *Annu Rev Neurosci* 28: 303-326, 2005. 308

- Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filip-pini N, Watkins KE, Toro R, Laird AR, Beckmann CF. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106: 13040-13045, 2009.
- 310. Snyder LH. Coordinate transformations for eye and arm movements in the brain. Curr Opin Neurobiol 10: 747-754, 2000.
- 311. Spanne A, Jorntell H. Questioning the role of sparse coding in the brain. Trends Neurosci 38: 417-427, 2015.
- 312. Spiteri Y, Galea EM. Psychology of Neglect. Hauppauge, NY: Nova Science Publishers, 2012
- 313. Sporns O. The human connectome: A complex network. Ann NYAcad Sci 1224: 109-125, 2011.
- Sporns O. Discovering the Human Connectome. Cambridge, MA: MIT 314. Press, 2012.
- 315. Sporns O. Cerebral cartography and connectomics. Philos Trans R Soc Lond B Biol Sci 370: 1-12, 2015.
- 316. Sporns O, Honey CJ, Kotter R. Identification and classification of hubs in brain networks. PLoS One 2: e1049, 2007.
- Stein BE. The New Handbook of Multisensory Processing. Cambridge, MA: MIT Press, 2012.
- Stein BE, Jiang W, Wallace MT, Stanford TR. Nonvisual influences 318. on visual-information processing in the superior colliculus. Prog Brain Res 134: 143-156, 2001.
- 319. Stone JV. Independent Component Analysis: A Tutorial Introduction.
- Cambridge, MA: MIT Press, 2004. Stoodley CJ. The cerebellum and cognition: Evidence from functional imaging studies. *Cerebellum* 11: 352-365, 2012. 320.
- 321. Stoodley CJ, Valera EM, Schmahmann JD. Functional topography of the cerebellum for motor and cognitive tasks: An fMRI study. Neuroimage 59: 1560-1570, 2012.
- Striedter GF, Avise JC, Ayala FJ. In the Light of Evolution: Volume 322. VI: Brain and Behavior. Washington (DC): National Academies Press, 2013.
- 323. Sugase Y, Yamane S, Ueno S, Kawano K. Global and fine information coded by single neurons in the temporal visual cortex. Nature 400: 869-873, 1999
- 324. Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J. Predictive codes for forthcoming perception in the frontal cortex. Science 314: 1311-1314, 2006.
- 325. Teixeira S, Machado S, Velasques B, Sanfim A, Minc D, Peressutti C, Bittencourt J, Budde H, Cagy M, Anghinah R, Basile LF, Piedade R, Ribeiro P, Diniz C, Cartier C, Gongora M, Silva F, Manaia F, Silva JG. Integrative parietal cortex processes: Neurological and psychiatric aspects. *J Neurol Sci* 338: 12-22, 2014.
- Thilagam PS. Advanced Computing, Networking and Security: Interna-tional Conference, ADCONS 2011, Surathkal, India, December 16-18,
- 2011, Revised selected papers. Berlin; New York: Springer, 2012.
  327. Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature* 381: 520-522, 1996.
- Thorpe SJ, Fabre-Thorpe M. Neuroscience. Seeking categories in the 328. brain. Science 291: 260-263, 2001.
- 329. Tolias AS, Keliris GA, Smirnakis SM, Logothetis NK. Neurons in macaque area V4 acquire directional tuning after adaptation to motion stimuli. Nat Neurosci 8: 591-593, 2005.
- 330. Tommasi L, Peterson MA, Nadel L. Cognitive Biology: Evolutionary and Developmental Perspectives on Mind, Brain, and Behavior. Cambridge, MA: MIT Press, 2009.
- Tong F, Pratte MS. Decoding patterns of human brain activity. Annu Rev Psychol 63: 483-509, 2012.
- Treue S. Visual attention: The where, what, how and why of saliency. *Curr Opin Neurobiol* 13: 428-432, 2003. 332.
- Trommershauser J, Kording K, Landy MS. Sensory Cue Integration. Oxford; New York: Oxford University Press, 2011. 333.
- 334. Troscianko T, Benton CP, Lovell PG, Tolhurst DJ, Pizlo Z. Camouflage and visual perception. Philos Trans R Soc Lond B Biol Sci 364: 449-461, 2009
- 335 Tsodyks M, Gilbert C. Neural networks and perceptual learning. Nature 431: 775-781, 2004.
- 336. Tyler CW, Likova LT. Crowding: A neuroanalytic approach. J Vis 7: 16.11-16.19, 2007.
- Ullman S. High-level Vision: Object Recognition and Visual Cognition. 337. Cambridge, MA: MIT Press, 2000.
- Vallar G, Perani D. The anatomy of unilateral neglect after right-338. hemisphere stroke lesions. A clinical/CT-scan correlation study in man. Neuropsychologia 24: 609-622, 1986.

- 339. van den Heuvel MP, Sporns O. An anatomical substrate for integration among functional networks in human cortex. J Neurosci 33: 14489-14500, 2013.
- 340. van der Helm PA. Cognitive architecture of perceptual organization: From neurons to gnosons. *Cogn Process* 13: 13-40, 2012. Van Essen DC. Functional organization of primate visual cortex. In:
- 341. Jones EG, Peters A, editors. Cerebral Cortex. New York: Plenum Press, 1975, pp. 259-329.
- 342. Van Essen DC. Visual areas of the mammalian cerebral cortex. *Annu Rev Neurosci* 2: 227-263, 1979.
  343. Van Essen DC. Organization of visual areas in macaque and human
- cerebral cortex. In: Chalupa L, Werner JS, editors. *The Visual Neuro-*sciences. Cambridge, MA: MIT Press, 2004, pp. 507-521.
- 344. Van Essen DC, Anderson CH, Felleman DJ. Information processing in the primate visual system: An integrated systems perspective. Science 255: 419-423, 1992
- Van Essen DC, Felleman DJ. On hierarchies: Response to Hilgetag et al. *Science* 271: 777, 1996.
   Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz
- R, Chang A, Chen L, Corbetta M, Curtiss SW, Della Penna S, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E, Consortium WU-MH. The Human Connectome Project: A data acquisition perspective. Neuroimage 62: 2222-2231, 2012
- 347. Van Vleet TM, DeGutis JM. The nonspatial side of spatial neglect and related approaches to treatment. *Prog Brain Res* 207: 327-349, 2013.
- 348. VanRullen R, Thorpe SJ. Surfing a spike wave down the ventral stream. Vision Res 42: 2593-2615, 2002.
- Verrelli BC, Tishkoff SA. Signatures of selection and gene conversion 349. associated with human color vision variation. Am J Hum Genet 75: 363-375, 2004.
- 350. Vesia M, Crawford JD. Specialization of reach function in human posterior parietal cortex. *Exp Brain Res* 221: 1-18, 2012. 351. Voogd J, Schraa-Tam CK, van der Geest JN, De Zeeuw CI. Visuomotor
- cerebellum in human and nonhuman primates. Cerebellum 11: 392-410, 2012
- Wandell BA, Smirnakis SM. Plasticity and stability of visual field maps in adult primary visual cortex. *Nat Rev Neurosci* 10: 873-884, 2009. Wang C, Cleland BG, Burke W. Synaptic delay in the lateral geniculate 352.
- 353. nucleus of the cat. Brain Res 343: 236-245, 1985.
- 354. Wang D, Buckner RL, Liu H. Cerebellar asymmetry and its relation to cerebral asymmetry estimated by intrinsic functional connectivity. J Neurophysiol 109: 46-57, 2013.
- 355. Watanabe T, Sasaki Y. Perceptual learning: Toward a comprehensive theory. *Annu Rev Psychol* 66: 197-221, 2015.
- 356 Webster MJ, Ungerleider LG, Bachevalier J. Development and plasticity of the neural circuitry underlying visual recognition memory. Can *J Physiol Pharmacol* 73: 1364-1371, 1995. 357. Weiss Y, Simoncelli EP, Edelson EH, Motion illusions as optimal
- percepts. Nat Neurosci 5: 598-604, 2002. Werner JS, Chalupa LM. The New Visual Neurosciences. Cambridge,
- 358 MA: The MIT Press, 2014.
- 359. Westwood DA, Goodale MA. Converging evidence for diverging pathways: Neuropsychology and psychophysics tell the same story. Vision Res 51: 804-811, 2011
- 360. Whitney D, Levi DM. Visual crowding: A fundamental limit on conscious perception and object recognition. Trends Cogn Sci 15: 160-168, 2011
- 361. Wu MC, David SV, Gallant JL. Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci 29: 477-505, 2006
- 362. Xiao Y, Rao R, Cecchi G, Kaplan E. Improved mapping of information distribution across the cortical surface with the support vector machine. Neural Netw 21: 341-348, 2008.
- 363. Yang T, Maunsell JH. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24: 1617-1626, 2004.
- 364. Yoshida M, Itti L, Berg DJ, Ikeda T, Kato R, Takaura K, White BJ, Munoz DP, Isa T. Residual attention guidance in blindsight monkeys watching complex natural scenes. *Curr Biol* 22: 1429-1434, 2012.
- Yuille A, Kersten D. Vision as Bayesian inference: Analysis by synthe-365. sis? Trends Cogn Sci 10: 301-308, 2006.
- Zanzotto FM. Brain Informatics: International Conference, BI 2012, 366. Macau, China, December 4-7, 2012: Proceedings. Berlin; New York: Springer, 2012.