Neuron
Perspective

On the Perception of Probable Things: Neural Substrates of Associative Memory, Imagery, and Perception

Thomas D. Albright[1,*]
[1]Center for the Neurobiology of Vision, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
*Correspondence: tom@salk.edu
DOI 10.1016/j.neuron.2012.04.001

Perception is influenced both by the immediate pattern of sensory inputs and by memories acquired through prior experiences with the world. Throughout much of its illustrious history, however, study of the cellular basis of perception has focused on neuronal structures and events that underlie the detection and discrimination of sensory stimuli. Relatively little attention has been paid to the means by which memories interact with incoming sensory signals. Building upon recent neurophysiological/behavioral studies of the cortical substrates of visual associative memory, I propose a specific functional process by which stored information about the world supplements sensory inputs to yield neuronal signals that can account for visual perceptual experience. This perspective represents a significant shift in the way we think about the cellular bases of perception.

> You cannot count the number of bats in an inkblot because there are none. And yet a man—if he be "bat-minded"—may "see" several. (Gregory Bateson, 1972)

It should come as no surprise that what you see is not determined solely by the patterns of light that fall upon your retinae. Indeed, that visual perception is more than meets the eye has been understood for centuries, and there are several *extraretinal* factors known to interact with the incoming sensory data to yield perceptual experience. Perhaps foremost among these factors is information learned from our prior encounters with the visual world—our memories—which enables us to infer the cause, category, meaning, utility, and value of retinal images. By this process, the inherent ambiguity and incompleteness of information in the image—what is out there? Have I seen it before? What does it mean? How is it used?—is overcome, nearly instantaneously and generally without awareness, to yield unequivocal and behaviorally informative percepts.

How does this transformation occur, and what are the underlying neuronal structures and events? Viewed in the context of a hierarchy of visual processing stages, prior knowledge of the world is believed to be manifested as "top-down" neuronal signals that influence the processing of "bottom-up" sensory information arising from the retina. Although the primate visual system has been a subject of intense study in neurobiological experiments for a half-century now, the primary focus of this research has been on the processing of visual signals as they ascend bottom-up through various levels of the hierarchy. Thus, with the notable exception of work on visual attention (for review, see Reynolds and Chelazzi, 2004), the neuronal substrates of top-down influences on visual processing have only recently come under investigation. Several of these recent experiments specifically address the interactions between top-down signals that reflect visual memories and bottom-up signals that convey retinal image content. The results of these experiments call for a significant shift in the way we think about the neuronal processing of visual information, and they are the subject of this review.

The first part of this review explores neuronal changes that parallel the acquisition of long-term memories of associations between visual stimuli, such as between a knife and fork or a train and its track. The second part considers neuronal events that correspond to memories recalled via such learned associations and the relationship of this recall to the phenomenon of visual imagery. Finally, evidence is presented for a specific functional process by which—in the prescient words of 19th century perceptual psychologist James Sully (1888)—the mind "supplements a sense impression by an accompaniment or escort of revived sensations, the whole aggregate of actual and revived sensations being solidified or 'integrated' into the form of a percept."

Visual Associative Learning and Memory

The concept of association is fundamental to learning and memory. Although this point was appreciated by the Ancient Greeks, it was by way of John Locke (1690) and the emergent Associationist philosophy that the content of the human mind became viewed as progressively accumulating and diversifying throughout one's lifetime via the "associations of ideas." Locke defined "ideas" broadly, but the simplest form of idea consists of sensation itself. Indeed, the learning of associations between sensory stimuli is a pervasive feature of human cognition.

Formally speaking, learned associations between sensory stimuli constitute acquired information about statistical regularities in the observer's environment, which may be highly beneficial for predicting and interpreting future sensory inputs. Learned associations also help define the semantic properties of stimuli, as the meaning of a stimulus can be found, in large part, in the other stimuli with which it is associated.

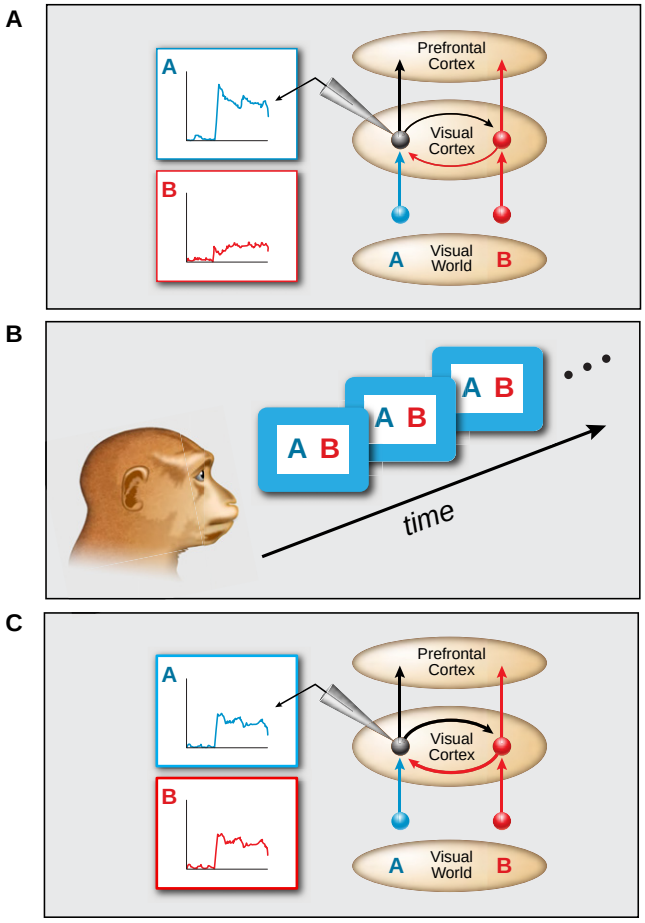

Figure 1. Schematic Depiction of Change in Local Cortical Connectivity and Neuronal Signaling Predicted to Underlie Acquisition of Visual Associative Memories

(A) Nervous system consists of two parallel information processing channels, which independently detect and represent visual stimuli "A" and "B." The flow of information is largely feed-forward from the sensory periphery, but there exist weak lateral connections that provide the potential for crosstalk between channels. The stimulus selectivity of each channel can be revealed by monitoring neuronal responses in visual cortex. (Small plots at left indicate spike rate as function of time.) The cortical neuron in the A channel responds strongly to stimulus A and weakly or not at all to stimulus B. The B channel neuron does the converse (not shown).

(B) Subject learns association between stimuli A and B by repeated temporal pairing with reinforcement. Following sufficient training, the sight of one stimulus comes to elicit pictorial recall of its pair.

(C) Associative learning is believed to be mediated by the strengthening of connections—the lateral projections in this schematic—between the independent representations of the paired stimuli. Each channel now receives inputs from both stimuli, though via different routes. The neurophysiological signature of this anatomical change is thus a convergence of responses to the paired stimuli. This signature has been observed for neurons in the inferior temporal (IT) cortex of rhesus monkeys (Sakai and Miyashita, 1991; Messinger et al., 2001).

Associative learning can take place with or without an observer's awareness. It may be the product of simple temporal coincidence of stimuli—your grandmother (stimulus 1) is always seated in her favorite chair (stimulus 2)—or it may be facilitated by conditional reinforcement—emotional rewards may strengthen, for example, an association between the face of your lover (stimulus 1) and the song that the jukebox played on your first date (stimulus 2).

A Neuronal Foundation for Associative Learning

The neuronal bases of associative learning have been the subject of speculations and detailed theoretical accounts for well over 100 years. Many of these proposals have at their core an idea first advanced concretely by William James (1890): the behavioral learning of an association between two stimuli is accomplished by the establishment or strengthening of a functional connection between the neuronal representations of the associated stimuli.

At some level, James' hypothesis must be correct, and it is useful to consider the implications of this idea for the neuronal representation of visual information. This can be done using a simple example based on a nervous system composed of two parallel visual information processing channels (Figure 1A). These channels extend from the retina up through visual cortex and beyond. One channel is dedicated to the processing of stimulus A and the other stimulus B. The flow of information through these channels is largely feed-forward, but there exist weak lateral connections that provide limited opportunities for crosstalk between the two channels. Recordings of activity from the A neuron in visual cortex should reveal a high degree of selectivity for stimulus A, relative to B, simply attributable to the different routes by which the signals reach the recorded neuron.

Now, suppose the subject in whose brain these two channels exist is trained to associate stimuli A and B, by repeated temporal pairing of the stimuli in the presence of reinforcement (Figure 1B). By the end of training, stimuli A and B are highly predictive of one another—in some sense A means B, and vice versa. The Jamesian hypothesis predicts that the neuronal correlate of this associative learning is the strengthening of crosstalk between the two channels (Figure 1C). Now recordings from the A neuron should reveal similar responses to stimuli A and B, because both channels now have comparable access (albeit via different routes) to the recorded neuron. Thus, according to this simple model, the predicted neuronal signature of associative learning in visual cortex is a convergence of response magnitudes—as A and B become associated, neurons initially responding selectively to one or the other of these stimuli will generalize to the associated stimulus.

Neural Correlates of Visual Associative Learning

An explicit test of the Jamesian hypothesis was first conducted by Miyashita and colleagues (Sakai and Miyashita, 1991). These investigators trained monkeys to associate a large number of pairs of visual stimuli: A with B, C with D, etc. Following behavioral acquisition of the associations, recordings were made from isolated neurons in the inferior temporal (IT) cortex (Figure 2), a region known to be critical for visual object recognition and memory (see below). Sakai and Miyashita (1991) found that paired stimuli (e.g., A&B) elicited responses of similar magnitude, whereas stimuli that were not paired (e.g., A&C) elicited uncorrelated responses. This finding of "pair-coding" neurons provided seminal support for the Jamesian view, as the similar

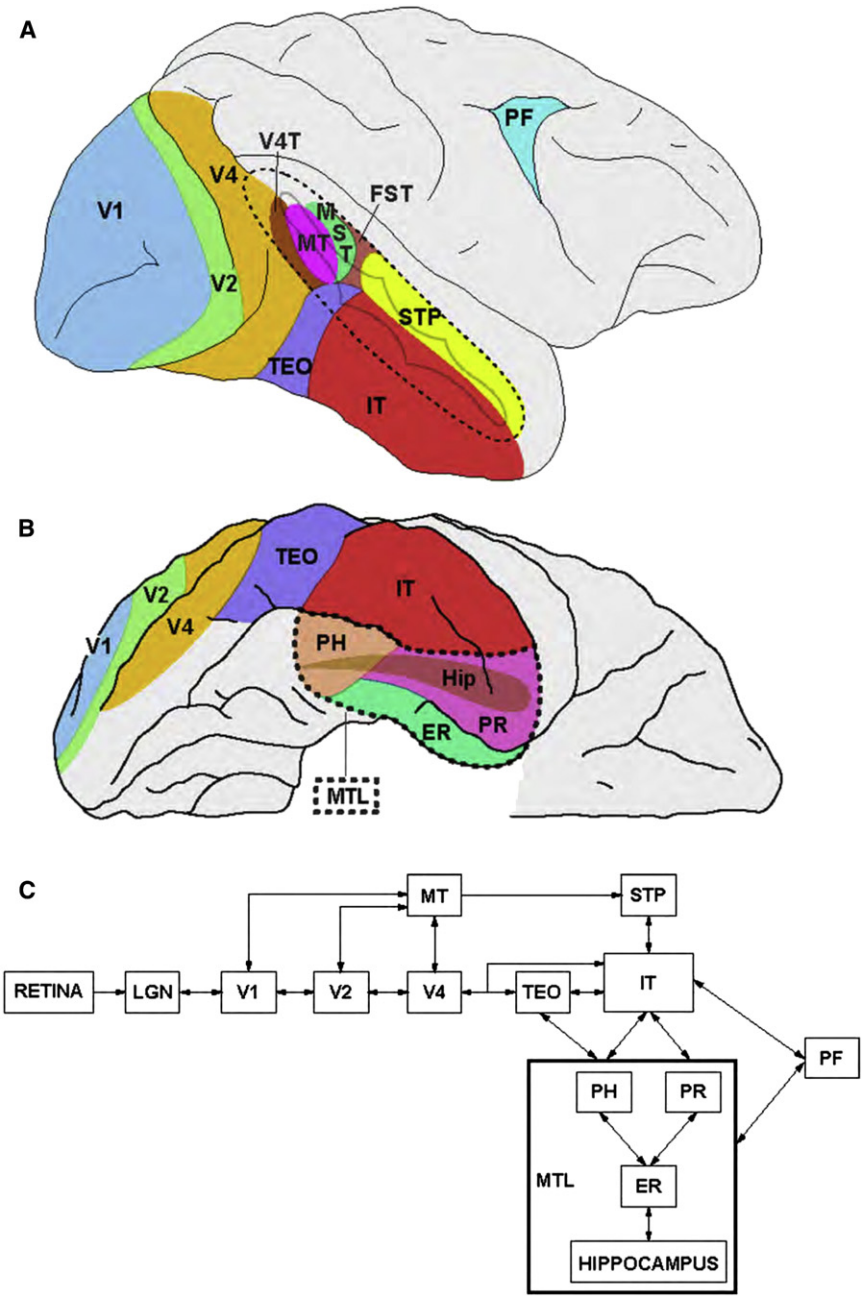

Figure 2. Locations and Connectivity of Cerebral Cortical Areas of Rhesus Monkey (*Macaca mulatta*) Involved in Associative Memory, Visual Imagery, and Visual Perception

(A) Lateral view of cortex. Superior temporal sulcus (STS) is partially unfolded to show relevant cortical areas that lie within. Distinctly colored regions identify a subset (visual areas V1, V2, V4, V4t, MT, MST, FST TEO, IT) of the nearly three dozen cortical areas involved in the processing of visual information.

(B) Ventral view of cortex. Distinctly colored regions identify inferior temporal cortex (IT) and a collection of medial temporal lobe (MTL) areas critical for learning and memory (ER, entorhinal cortex; PH, parahippocampal cortex; PR, perirhinal cortex; H, hippocampal formation, lies in the interior of the temporal lobe).

(C) Connectivity diagram illustrating known anatomical projections from primary visual cortex (V1) up through the inferior temporal (IT) cortex and on to MTL areas. Most projections are bidirectional.

responses to paired stimuli were taken to be a consequence of the learning-dependent connections formed between the neuronal representations of these stimuli.

To directly explore the emergence of pair-coding responses, Messinger et al. (2001) recorded from IT neurons while monkeys learned new stimulus pairings. For many neurons, the pattern of stimulus selectivity changed incrementally as pair learning progressed: responses to paired stimuli became more similar and responses to stimuli that had not been paired became less similar. The time course of this "associative neuronal plasticity" matched the time course of learning and the presence of neuronal changes depended upon whether learning actually occurred (i.e., if the monkey failed to learn new pairings, neuronal selectivity did not change). A snapshot of the Messinger et al. (2001) results taken at the end of training reveals a pattern of neuronal selectivity that closely matches the findings of Sakai and Miyashita (1991).

The emergence of pair-coding responses in IT cortex supports the conclusion that learning strengthens connectivity between the relevant neuronal representations. That enhancement of connectivity may be regarded as the process of associative memory formation, the product of which is a neuronal state that captures the memory, i.e., the memory trace. This is precisely the interpretation that Miyashita and colleagues (e.g., Miyashita, 1993), and subsequently Messinger et al. (2001), have applied to the finding of pair-coding neurons in IT cortex, and it is consistent with neuropsychological data that identifies IT cortex as a long-term repository of visual memories (see below).

Mechanisms of Associative Neuronal Plasticity in IT Cortex

Visual paired association learning is dependent upon the integrity of the hippocampus and cortical areas of the medial temporal lobe (MTL) (Murray et al., 1993). These areas, which include the entorhinal, perirhinal and parahippocampal cortices, receive inputs from and are a source of feedback to IT cortex (see Figure 2; Webster et al., 1991). The learning impairment following

MTL lesions appears to be one of memory formation and the MTL areas are thus, under normal conditions, believed to exert their influence by enabling structural reorganization of local circuits in the presumed site of storage, i.e., IT cortex (Miyashita, 1993; Squire et al., 2004; Squire and Zola-Morgan, 1991). This hypothesis is supported by the finding that MTL lesions also eliminate the formation of pair-coding responses in IT cortex (Higuchi and Miyashita, 1996).

Exactly how MTL regions contribute to the strengthening of connections between the neuronal representations of paired stimuli—with the attendant associative learning and neuronal response changes—is unknown. There are, nonetheless, good reasons to suspect the involvement of a Hebbian mechanism for enhancement of synaptic efficacy. Specifically, the temporal coincidence of stimuli during learning may cause coincident patterns of neuronal activity, which may lead, in turn, to a strengthening of synaptic connections between the neuronal representations of the paired stimuli (e.g., Yakovlev et al., 1998). This conclusion is supported by the finding that associative plasticity in IT cortex is correlated with the appearance of molecular-genetic markers for synaptic plasticity: mRNAs encoding for brain-derived neurotrophic factor (BDNF) and for the transcription factor zif268 (Miyashita et al., 1998; Tokuyama et al., 2000). BDNF is known to play a role in activity-dependent synaptic plasticity (Lu, 2003). zif268 is a transcriptional regulator that leads to gene products necessary for structural changes that underlie plasticity (Knapska and Kaczmarek, 2004).

Is Associative Neuronal Plasticity Unique to IT Cortex?

The inferior temporal cortex was chosen as the initial target for study of associative neuronal plasticity for a number of reasons. This region of visual cortex was, for many years, termed "association cortex." Although this designation originally reflected the belief that the temporal lobe represents a point at which information from different sensory modalities is associated (Flechsig, 1876), the term was later used to refer, more generally, to the presumed site of Locke's "association of ideas."

This view received early support from neuropsychological studies demonstrating that temporal lobe lesions in both humans and monkeys selectively impair the ability to recognize visual objects, while leaving basic visual sensitivities intact (Alexander and Albert, 1983; Brown and Schafer, 1888; Kluver and Bucy, 1939; Lissauer, 1988). Along the same lines, the classic explorations of the neurosurgeon Wilder Penfield (Penfield and Perot, 1963) revealed that electrical stimulation of the human temporal lobe commonly elicits reports of visual memories.

The anatomical connections of IT cortex also support a role in object recognition and visual memory (Figure 2). IT cortex lies at the pinnacle of the ventral cortical visual processing stream and its neurons receive convergent projections from many visual areas at lower ranks, thus affording integration of information from a variety of visual submodalities (Desimone et al., 1980; Ungerleider, 1984). As noted above, IT cortex is also reciprocally connected with MTL structures that are critical for acquisition of declarative memories (Milner, 1972; Mishkin, 1982; Murray et al., 1993; Squire and Zola-Morgan, 1991).

Finally, the visual response properties of IT neurons, which have been explored in much detail over the past 40 years, also exhibit features that suggest a role in object recognition and visual memory (for review see Gross et al., 1985; Miyashita, 1993). Most importantly, IT neurons are known to respond selectively to complex objects—often those with some behavioral significance to the observer, such as faces (Desimone et al., 1984; Gross et al., 1969).

Based on this collective body of evidence, it would seem that IT cortex is unique among visual areas and strongly implicated as a storage site for long-term associative memories. Yet, there are reasons to suspect that associative neuronal plasticity may be a general property of sensory cortices. Evidence for this comes in part from functional brain imaging studies that have found learning-dependent activity changes in early cortical visual areas (e.g., Shulman et al., 1999; Wheeler et al., 2000). Motivated by these findings, Schlack and Albright (2007) explored the possibility that associative learning might influence response properties in the middle temporal visual area (area MT), which occupies a relatively early position in the cortical visual processing hierarchy (Ungerleider and Mishkin, 1979).

MT Neurons Exhibit Associative Plasticity

In an experiment that represents a simple analog to the paired-association learning studies of Sakai and Miyashita (1991) and Messinger et al. (2001), Schlack and Albright (2007) trained monkeys to associate directions of stimulus motion with stationary arrows. Thus, for example, monkeys learned that an upward-pointing arrow was associated with a pattern of dots moving in an upward direction, a downward arrow was associated with downward motion, etc. (Figures 3A and 3B).

Moving stimuli were used for this training because it is well known that such stimuli elicit robust responses from the vast majority of neurons in cortical visual area MT (Albright, 1984). In macaque monkeys, where it has been most intensively studied, area MT is a small cortical region (Figure 2) that lies posteriorly along the lower bank of the superior temporal sulcus (Gattass and Gross, 1981) and which receives direct input from primary visual cortex (Ungerleider and Mishkin, 1979). MT neurons are highly selective for the direction of stimulus motion, and the area is believed to be a key component of the neural substrates of visual motion perception (for review, see Albright, 1993).

If MT neurons have potential for associative plasticity similar to that seen in IT cortex, the behavioral pairing of motion directions with arrow directions should lead to a convergence of responses to the paired stimuli, overtly detectable in MT as emergent responses to the arrows. Moreover, those responses should be tuned for arrow direction, and the form of that tuning should depend on the specific associations learned. Schlack and Albright (2007) tested these hypotheses by recording from MT neurons after the motion-arrow associations were learned. Many MT neurons exhibited selectivity for the direction of the static arrow—a property not seen prior to learning, and seemingly heretical to the accepted view that MT neurons are primarily selective for visual motion. Moreover, for individual neurons, the arrow-direction tuning curve was a close match to the motion-direction tuning curve (Figures 3C and 3D).

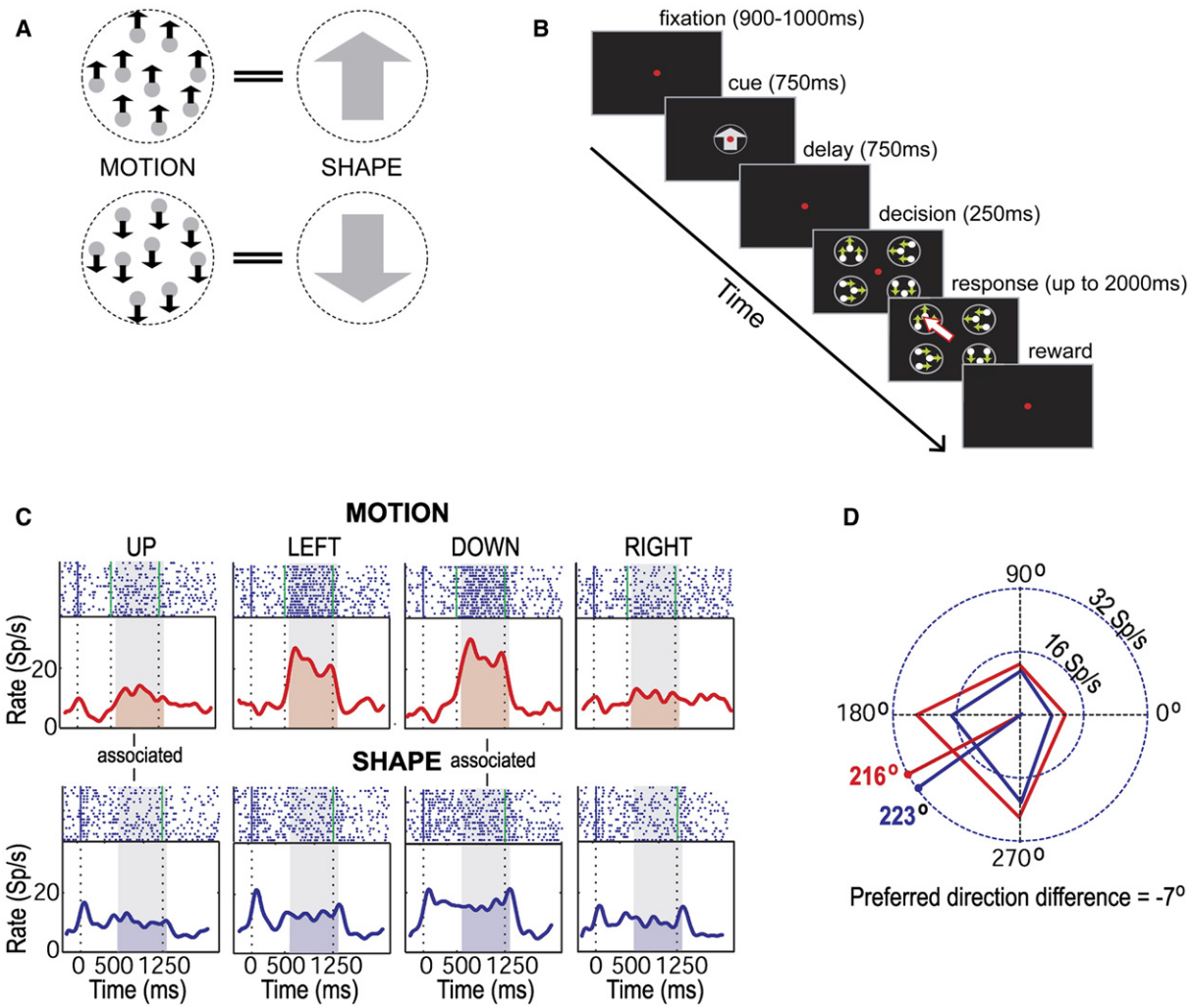

Figure 3. Emergent Stimulus Selectivity of Neurons in Cortical Visual Area MT following Paired Association Learning

(A) Rhesus monkeys learned to associate up and down motions with up and down arrows.

(B) Schematic depiction of task used to train motion-arrow pairings. Trial sequence is portrayed as a series of temporal frames. Each frame represents the video display and operant response (eye movement to chosen stimulus). All neuronal data were collected following extensive training on this task, and during behavioral trials in which monkeys were simply required to fixate a central target.

(C) Data from representative MT neuron. Top row illustrates responses to four motion directions. Spike raster displays of individual trial responses are plotted above cumulative spike-density functions. Vertical dashed lines correspond from left to right to stimulus onset, motion onset, and stimulus offset. Gray rectangle indicates analysis window. The cell was highly directionally selective. Bottom row illustrates responses to four static arrows. The animal previously learned to associate arrow direction with motion direction. Plotting conventions are same as in upper row. The cell was highly selective for arrow direction.

(D) Mean responses of neuron shown in (C) to motion directions (red curve) and corresponding static arrow directions (blue curve), indicated in polar format. Preferred directions for the two stimulus types (red and blue vectors) are nearly identical.

From Schlack and Albright (2007).

To confirm that the emergent responses to arrows reflected the learned association with motions rather than specific physical attributes of the arrow stimulus, Schlack and Albright (2007) trained a second monkey on the opposite associations (e.g., upward motion associated with downward arrow). As expected from the learning hypothesis, the emergent tuning again reflected the association (e.g., if the preferred direction for motion was upward, the preferred direction for the arrow was downward) rather than the specific properties of the associated stimulus.

What Is Represented by Learning-Dependent Neuronal Selectivity in Area MT?

On the surface of things, the plasticity seen in area MT appears identical to that previously observed in IT cortex: the neuronal response change is learning-dependent and can be characterized as a convergence of responses to the paired stimuli. One might suppose, therefore, that the phenomenon in MT also reflects mechanisms for long-term memory storage. There are, however, several reasons to believe that the plasticity observed in MT reflects rather different functions and mechanisms.

To begin with, IT and MT cortices are distinguished from one another by the availability of substrates for long-term memory storage. In the IT experiments described above the paired stimuli (arbitrary complex objects) are in all cases plausibly represented by separate groups of IT neurons, which means that connections between those representations could be forged locally within IT cortex. The same is not true for area MT, as there exists no native selectivity for stationary arrows (or for most other nonmoving stimuli).

IT and MT are also distinguished from one another by the presence versus absence of feedback from cortical areas of the medial temporal lobe (see Figure 2). As noted above, these MTL areas are essential for learning of visual paired-associates (presumably also including those between arrows and motions), and they are believed to enable memory formation via selective modification of local circuits at the targets of their feedback projections. IT cortex is one of those targets, but area MT is not (Suzuki and Amaral, 1994). Although it remains to be seen whether MTL lesions block the emergence of pair-coding responses in area MT, as they do in IT cortex, the evident connectional dissimilarities between MT and IT suggest that the associative neuronal plasticity in MT is not the basis of memory storage.

If not memory storage, what then is represented by the observed learning-dependent responses in MT? One possibility is that they simply represent the properties of the retinal stimulus, i.e., the direction of the arrow. Alternatively, the learning-dependent responses may have nothing directly to do with the retinal stimulus but, rather, represent the motion that is recalled in the presence of the arrow. The distinction between these two possibilities—a response that represents the bottom-up stimulus versus a response that represents top-down associative recall—is fundamental to this discussion.

According to the bottom-up argument, the cortical circuitry in area MT has been co-opted, as a result of extensive training on the motion-arrow association task, for the purpose of representing a novel stimulus type. This argument maintains that motion processing is the default operation in MT, but the inherent plasticity of cortex allows these neurons to take on other functional roles as dictated by the statistics of the observer's environment. Although the evidence to date cannot rule out this possibility, it defies the not unreasonable assumption that properties of early visual neurons must remain stable in order to yield a stable interpretation of the world (van Wezel and Britten, 2002). By contrast with the bottom-up argument, there is considerable parsimony in the view that the emergent responses to arrow stimuli are manifestations of a top-down signaling process, the purpose of which is to achieve associative recall. Importantly, this view asserts that area MT remains stably committed to motion processing, with recognition that the same motion-sensitive neurons may become activated by either bottom-up or top-down signals.

Visual Associative Recall

The storage of information in memory and the subsequent retrieval of that information are generally viewed as interdependent processes rooted in overlapping neuronal substrates (e.g., Anderson and Bower, 1973). Evidence reviewed above suggests that the associative neuronal plasticity—the emergence of pair-coding responses—seen in IT cortex is a manifestation of memory storage. At the same time, the response to a paired stimulus is a demonstration of retrieval, and thus can also be viewed as "recall-related" activity.

By contrast with IT cortex, evidence indicates that the learning-dependent responses to arrows in area MT are solely a manifestation of retrieval. They are, in a literal sense, a cued top-down reproduction of the activity pattern that would be elicited in MT by a moving stimulus projected upon the retina. In other words, the recall-related activity seen in area MT is a neural correlate of visual imagery of motion. This provocative proposal naturally raises two important questions: (1) what is the source of the top-down recall-related activity, and (2) what is it for? These questions will be addressed in detail after a brief consideration of other evidence for neural correlates of visual imagery.

A Common Neuronal Substrate for Visual Imagery and Perception

> Why don't you just go ahead and imagine what you want? You don't need my permission. How can I know what's in your head? (Haruki Murakami, 2005, *Kafka on the Shore*)

The arguments summarized above maintain that the selective pattern of activity in MT to static arrows reflects the recalled pictorial memory—imagery—of motion, which is represented in the same cortical region and by the same neuronal code as the original motion stimulus. Although the evidence is striking in this case, the concept of common substrates for imagery and perception is not new. This idea can be traced to 1644, when Rene Descartes (1972), argued that visual signals originating in the eye and those originating from memory are both experienced via the "impression" of an image onto a common brain structure. (Descartes incorrectly believed that structure to be the pineal gland.) The same argument—known as the "principle of perceptual equivalence" (Finke, 1989)—has been developed repeatedly and explicitly over the past century by psychologists, neuroscientists, and cognitive scientists alike (e.g., Behrmann, 2000; Damasio, 1989; Farah, 1985; Finke, 1989; Hebb, 1949; James, 1890; Kosslyn, 1994; Merzenich and Kaas, 1980; Nyberg et al., 2000; Shepard and Cooper, 1982).

Modern-day enthusiasm for the belief that imagery and perception are mediated by common neuronal substrates and events grew initially from the commonplace observation that the subjective experiences associated with imagery and sensory stimulation are similar in many respects (e.g., Finke, 1980; Podgorny and Shepard, 1978). Empirical support for the hypothesis followed with studies demonstrating that perception reflects interactions between imagery and sensory stimulation (e.g., Farah, 1985; Ishai and Sagi, 1995; Peterson and Graham, 1974): for example, imagery of the letter "T" selectively facilitates detection of a "T" stimulus projected on the retina (Farah, 1985).

More recently, the common substrates hypothesis has received backing in abundance from human functional brain imaging studies. These studies, in which subjects are either asked to image specific stimuli, or studies in which imagery is "forced" by cued associative recall, have documented patterns

of activity during imagery in a variety of early- and midlevel cortical visual areas (e.g., D'Esposito et al., 1997; Ishai et al., 2000; Knauff et al., 2000; Kosslyn et al., 1995; O'Craven and Kanwisher, 2000; Reddy et al., 2010; Slotnick et al., 2005; Stokes et al., 2009, 2011; Vaidya et al., 2002; Wheeler et al., 2000), including area MT (Goebel et al., 1998; Kourtzi and Kanwisher, 2000; Shulman et al., 1999)—patterns that appear similar in many respects to those elicited by a corresponding retinal stimulus. Along the same lines, electrophysiological recordings from deep electrodes in the temporal cortex of human subjects have revealed responses that were highly selective for the pictorial content of volitional visual imagery (Kreiman et al., 2000).

Neurophysiological studies that have addressed this issue in animals are rare, in part because visual imagery is fundamentally subjective and thus not directly accessible to anyone but the imager. A solution to this problem involves inducing imagery through the force of association. This is, of course, the approach used in the aforementioned studies of association learning in visual areas IT (Messinger et al., 2001; Sakai and Miyashita, 1991) and MT (Schlack and Albright, 2007). Although these stand as the only explicit studies of visual imagery at the cellular level, there are several other indications of support in the neurophysiological literature.

For example, Assad and Maunsell (1995) presented monkeys with a moving spot that followed a predictable path from the visual periphery to the center of gaze. Recordings were made from motion-sensitive neurons in cortical visual area MST. Receptive fields were selected to lie along the motion trajectory, and the passing of the spot elicited the expected response. On some trials, however, the spot disappeared and reappeared along its trajectory, as if passing behind an occluding surface. Although the stimulus never crossed the receptive field on occlusion trials, its inferred trajectory did, and many MST neurons responded in a manner indistinguishable from the response to real receptive field motion. A plausible interpretation of these findings is that the neuronal response on occlusion trials reflects pictorial recall of motion, elicited by the presence of associative cues, such as the visible beginning and end points of the trajectory (see Albright, 1995).

Such effects are not limited to the visual domain. Haenny, Maunsell and Schiller (1988) trained monkeys on a tactile-visual orientation match-to-sample task (cross-modal match-to-sample is a special case of paired-association learning), in an effort to explore the effect of attentional cuing on visual responses. Recordings in area V4 of visual cortex revealed, among other things, orientation-tuned responses to the tactile cue stimulus, prior to the appearance of the visual target (see Figure 4 in Haenny et al., 1988). The authors refer to this response as "an abstract representation of cued orientation," which may be true in some sense, but in light of the findings of Schlack and Albright (2007), one can interpret the V4 response to a tactile stimulus as a neural correlate of the visually recalled orientation.

Early experiments by Frank Morrell might also be interpreted in this vein (for review, see Morrell, 1961). In one set of studies, Morrell reported auditory responses in primary visual cortex of animals that had been trained to associate auditory and visual stimuli (Morrell et al., 1957). While highly controversial at the time, these results now seem consistent with the common substrates hypothesis. Similarly, using cross-modal associative learning, Joaquin Fuster and colleagues (e.g., Zhou and Fuster, 2000) have provided several electrophysiological demonstrations of recall-related activity in the auditory and somatosensory cortices.

What Is the Source of Recall-Related Signals in Visual Cortex?

As summarized above, the neuronal plasticity in IT cortex that accompanies paired-association learning is likely to be mediated via local circuit changes within this visual area (Figure 4A), which in turn provide the foundation for associative recall. Evidence indicates that this retrieval process takes two basic forms: automatic and active (Miyashita, 2004). In the automatic case, a bottom-up cue stimulus directly activates the neuronal representation of an associated stimulus, via the pre-established links in IT cortex. In the active case, retrieval is presumed to occur under executive control mediated by the prefrontal cortex. In this scenario, prefrontal cortex maintains stimulus and task-relevant information in working memory. Top-down signals from prefrontal cortex reactivate associative memory circuits in IT cortex as dictated by the behavioral context at hand (Tomita et al., 1999).

The situation in MT differs primarily in that the paired stimuli are unlikely to be associated via changes in local connections within this visual area. One possibility is that the visual associations learned in the experiment of Schlack and Albright (2007) are stored via circuit changes in IT cortex, in a manner no different from that seen in earlier studies of pair-coding responses in IT (Messinger et al., 2001; Sakai and Miyashita, 1991). According to this hypothesis, the recall-related activity observed in MT reflects a backward spread of feature-specific activation, originating with the memory trace in IT (via automatic or active processes) and descending through visual cortex (Figure 4B).

Whatever the source of the feedback, there are several provocative features of the recall event that may inform an understanding of the underlying mechanism. To begin with, the neurophysiological data indicate that recall-related signals are highly specific. Indeed, in area MT the selectivity for stimuli associated with directions of motion is nearly indistinguishable from the selectivity for the motions themselves (Schlack and Albright, 2007). This selectivity suggests a high degree of anatomical specificity in the feedback signals that activate MT neurons under these conditions.

Second, the feedback signals would seem to possess enormous content flexibility, given that the number of learnable associations for a given stimulus is vast (if not infinite). One can, for example, learn associations between directions of motion and many arbitrary visual stimuli (in addition to the arrows used by Schlack and Albright [2007]), such as colors, shapes, faces, or alphanumeric characters, as well as with non-visual stimuli, such as tones (A. Schlack et al., 2008, Soc. Neurosci., abstract) or tactile movements. The obvious implications are that the source of top-down signaling has access to a wide range of types of sensory information, and that this range may be manifested in the recall-related responses in visual cortex.

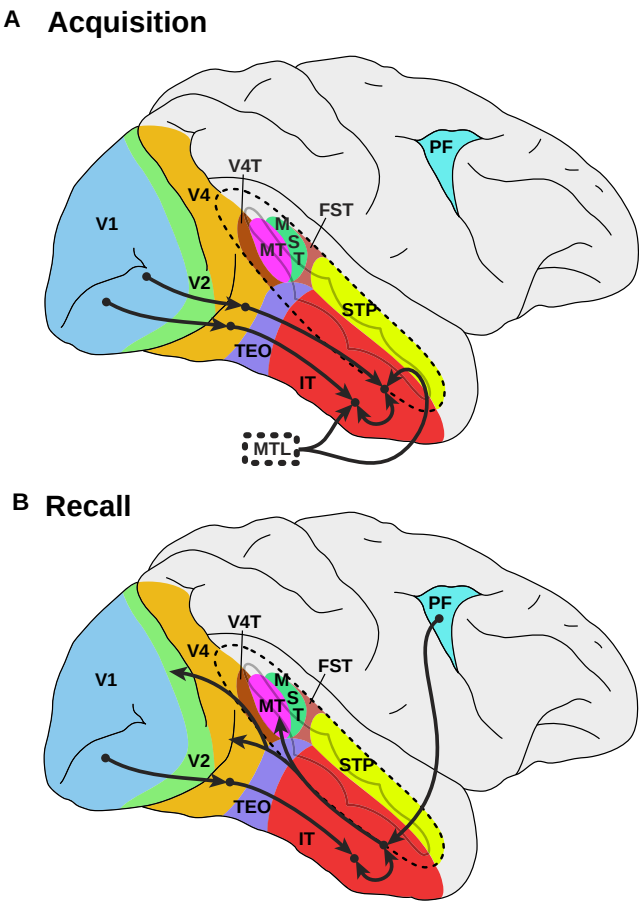

Figure 4. Stylized Depiction of Hypothesized Neuronal Circuits for Acquisition of Visual Associative Memories and Pictorial Recall of Those Memories

See Figure 2 for areal abbreviations.

(A) Acquisition of visual associative memory. Black arrows indicate flow of information from primary visual cortex (V1) up to inferior (IT) cortex. The two arrows so ascending indicate generic connections that underlie representation of two different visual stimuli (e.g., A and B). Learning of an association between the two stimuli is mediated by the formation of reciprocal connections between the corresponding neuronal representations in IT cortex. This associative learning and circuit reorganization are dependent on feedback from the medial temporal lobe (MTL).

(B) Pictorial recall of visual associative memory. If *object B* is viewed, a selective pattern of activation ascends through visual cortex, ultimately activating the neuronal representation of *object B* in area IT. This neuronal representation of *object B* may also be activated indirectly by either of two means when *object B* is not visible. In "automatic" recall mode, the neuronal representation of *object A* is activated (ascending arrow from V1 to IT) by viewing that stimulus. The neuronal representation of the paired stimulus (*object B*) becomes activated in turn via local connections within IT. In "active" recall mode, the neuronal representation of *object B* is activated in IT cortex when that stimulus is held in working memory (descending arrow from prefrontal cortex to IT). In both cases, a visual image of the stimulus so recalled results from a descending cascade of selective activation in visual cortex, which matches the pattern that would normally be elicited by viewing the stimulus. Under most conditions, active and automatic modes correspond, respectively, to the processes underlying what we have termed explicit and implicit imagery.

Third, the feedback signals would appear to be temporally flexible, inasmuch as cued associative recall is context-dependent. The visual images recalled by the sight of a shovel, for example, may depend upon whether the shovel is viewed in the garden or the cemetery. Although it remains to be seen whether recall-related neuronal responses in areas MT and IT are context dependent (but see Naya et al., 1996), the context dependence of imagery itself implies that the relevant top-down signals are dynamically engaged rather than hard-wired. The task of identifying feedback mechanisms and circuits that satisfy these multiple constraints is daunting, to say the least, but their recognition casts new light on cortical visual processing.

What Is the Function of Visual Imagery?

Additional insights into top-down signaling and its contribution to perceptual experience may come from consideration of what purpose it serves. Much has been written about the functions of visual imagery (e.g., Farah, 1985; Hebb, 1968; James, 1890; Kosslyn, 1994; Neisser, 1976; Paivio, 1965; Shepard and Cooper, 1982). To understand these functions, it is useful to consider two types of imagery: explicit and implicit.

Explicit Visual Imagery

Scientific and colloquial discussions of visual imagery have most commonly focused on a class of operations that enable an individual to evaluate the properties of objects or scenes that are not currently visible. This type of imagery is typically both explicit and volitional—corresponding to the "active" retrieval process described above (see Miyashita, 2004)—and is conjured on demand to serve specific cognitive or behavioral goals. Explicit imagery may be retrospective or prospective. The retrospective variety involves scrutiny via imagery of material previously seen and remembered, such as the examination in one's mind's eye of the kitchen counter in order to determine whether the car keys are there. Prospective imagery—what Schacter et al. (2007) call "imagining the future"—includes the evaluation of visual object or scene transformations, or wholesale fabrication of objects and scenes based on information from other sources, such as language. For example, one might imagine the placement of the new couch in the sitting room, without the trouble of actually moving the couch. (Watson [1968] famously used this form of visual imagery to transpose base pairs—"I happily lay awake with pairs of adenine residues whirling in front of my closed eyes"—as he narrowed in on the structure of DNA.) Similarly, any reader of the *Harry Potter* series has surely manufactured rich pictorial representations of the fictional Hogwarts Castle.

For the present discussion, it is noteworthy that explicit imagery often occurs in the presence of retinal stimuli to which the conjured image has no perceptual bearing—physical, semantic, or otherwise. For example, I can readily and richly picture the high-stepping march of Robert Preston's Music Man (trailed of course by the River City Boys' Band), but that dynamic image is (thankfully) perceptually distinct from the world in front of me (though perhaps causing interference; see Segal and Fusella [1970], for example).

Evidence for neural correlates of explicit visual imagery is plentiful. In particular, the numerous functional brain imaging studies cited above (as evidence localizing visual imagery to visual cortex) were conducted primarily under conditions of

Figure 5. Demonstration of the Influence of Associative Pictorial Recall (Top-Down Signaling) on the Interpretation of a Retinal Stimulus (Bottom-Up Signaling)
To most observers, this figure initially appears as a random pattern with no clear figural interpretation. The perceptual experience elicited by this stimulus is radically (and perhaps permanently) different after viewing the pattern shown in Figure 8.

explicit imagery, in which human subjects were simply asked to generate images of specific stimuli.

Implicit Visual Imagery

There exists a second functional role for visual imagery, which is, by contrast, implicit ("automatic") and externally driven, and which plays a fundamental and ubiquitous, albeit less commonly recognized, role in normal visual perception. This function follows from the proposition that perceptual experience falls at varying positions along a continuum between the extremes of pure stimulus and pure imagery (e.g., Thomas, 2011), with the position at any point in time determined primarily by stimulus quality and knowledge of the environment (James, 1890). Under most circumstances, implicit visual images are elicited by learned associative cues and serve to augment sensory data with "likely" interpretations, in order to overcome the ever-present noise, ambiguity, and incompleteness of the retinal image. For example, with little scrutiny, I regularly perceive the blurry and partially occluded stimulus that passes my office window to be my colleague Chuck Stevens, simply because experience tells me that Chuck is a common property of my environment. Similarly, the pattern in Figure 5 may be ambiguous and uninterpretable upon first viewing, but perceived clearly after experience with Figure 8. According to this view, imagery is not simply a thing apart, an internal representation distinct from the scene before our eyes, but rather it is part-and-parcel of perception.

This take on visual imagery is not new. The 19th century Associationist philosopher John Stuart Mill (1865) viewed perception as an internal representation of the "permanent possibilities of sensation." Accordingly, perception derives from inferences about the environment in the absence of complete sensory cues. Similarly, David Hume (1967) noted a "universal tendency among mankind… to transfer to every object, those qualities with which they are familiarly acquainted." William James (1890) expanded upon this theme by noting that "perception is of probable things" and that visual experience is completed by "farther facts associated with the object of sensation." Helmholtz (1924) developed a similar idea in his concept of unconscious inference, according to which perception is based on both sensory data and inferences about probabilities based upon experience.

More recently, these arguments have been echoed in the concept of "amodal completion" (Kanizsa, 1979)—the imaginal restoration of occluded image features, whose "perceptual existence is not verifiable by any sensory modality." Bruner and Postman (1949) spoke of "directive" factors, which reflect an observer's inferences about the environment and operate to maximize percepts consistent with those inferences ("one smitten by love does rather poorly in perceiving the linear characteristics of his beloved"). Finally, this view has acquired the weight of logical formalism through Bayesian approaches to visual processing (e.g., Kersten et al., 2004; Knill and Richards, 1996): learned associations constitute information about the statistics of the observer's environment, which come into play lawfully as the visual system attempts to identify the environmental causes of retinal stimulation (see also Brunswik, 1956). More generally, this line of thinking incorporates a key feature of associative recall—completion of a remembered whole from a sensory part—while assigning a vital functional role to visual imagery in this process.

Empirical support for the implicit imagery hypothesis derives from a long-standing literature addressing the influence of associative experience on perception (e.g., Ball and Sekuler, 1980; Bartleson, 1960; Bruner et al., 1951; Farah, 1985; Hansen et al., 2006; Hurlbert and Ling, 2005; Ishai and Sagi, 1995, 1997a, 1997b; Mast et al., 2001; Siple and Springer, 1983), which dates at least to Ewald Hering's (1878) concept of "memory colors"—e.g., perceived color should be biased toward yellow if the color originates from a banana. In one of the most provocative experiments of this genre (made famous for its use by Thomas Kuhn [1962] as a metaphor for scientific discovery), Bruner and Postman (1949) used "trick" playing cards to demonstrate an influence of top-down imaginal influences on perception. The trick cards were created simply by altering the color of a given suit—a red six of spades, for example. Human subjects were shown a series of cards with brief presentations; some cards were trick and the remainder normal. With startling frequency, subjects failed to identify the trick cards and instead reported them as normal. Upon questioning, these subjects often defended their perceptual reports, even after being allowed to scrutinize the trick cards, thus demonstrating that strongly learned associations between color and pattern are capable of sharply biasing perceptual judgments toward the imagery end of the of the stimulus-imagery continuum.

A Neuronal Representation of Probable Things

The two forms of imagery identified above are phenomenologically and functionally distinct, but they may well rely upon common substrates for selective top-down activation of visual cortex, i.e., recall-related activity (Figure 4B). It is instructive to consider how that neuronal activity relates to perceptual state under different imagery conditions. The studies of recall-related neuronal activity in areas IT and MT summarized above were conducted under conditions deemed likely to elicit explicit imagery. For example, from the study of Schlack and Albright

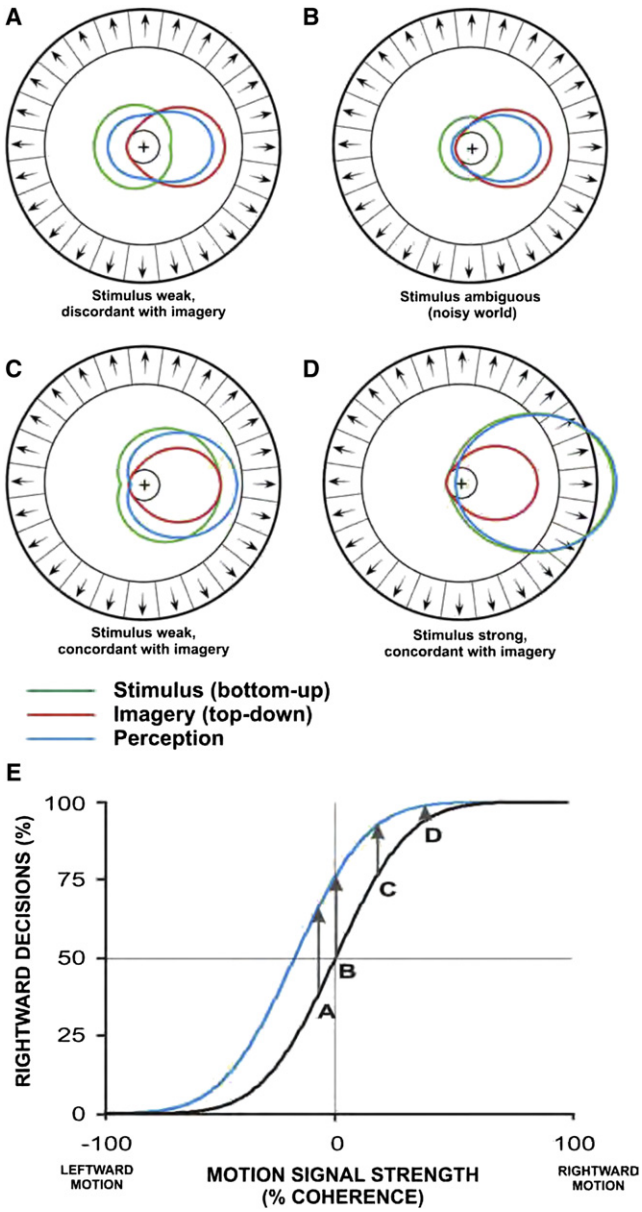

Figure 6. Conceptual Model to Account for Perceptual Consequences of Interactions between Stimulus and Imagery Signals in Visual Cortex

(A–D) Represent hypothesized patterns of activity elicited in area MT by bottom-up signals of different direction and magnitude and a top-down signal of fixed direction and magnitude. Arrowed segments symbolize cortical direction columns (plotted in circle for graphical convenience). Green and red polar plots indicate hypothesized activations of each directional column elicited, respectively, by bottom-up stimulus and top-down imagery signals. Blue curve indicates weighted sum of the two signals (stronger signals have disproportionately large weights). Black circle represents baseline activity of each column. (A) Stimulus signal (green) corresponds to leftward motion and the activity pattern is modeled as low coherence, high directional variance. Imagery signal (red) corresponds to rightward motion and the activity pattern is modeled as mid-level coherence, low variance. The weighted sum of these discordant activity patterns (blue) exhibits a bias toward the imagery direction (rightward). The ratio of rightward to leftward perceptual reports is predicted to be proportional to the ratio of activities (blue curve) for the corresponding neurons, favoring rightward in this case, despite a leftward stimulus. (B) Stimulus signal (green) corresponds to directional noise and the activity pattern is modeled as 0% coherence. Imagery signal (red) is same as (A). The weighted sum of these discordant activity patterns (blue) exhibits a bias toward the imagery direction (rightward), despite an incoherent stimulus. The ratio of perceptual reports is predicted to favor rightward in this case, despite an ambiguous stimulus. (C) Stimulus signal (green) corresponds to rightward motion and the activity pattern is modeled as low coherence, high directional variance. Imagery signal (red) is same as (A). The weighted sum of these activity patterns (blue) reflects the synergy between stimulus and imagery signals. The ratio of perceptual reports in this case is predicted to exhibit a moderate rightward bias above that resulting from stimulus signal alone. (D) Stimulus signal (green) corresponds to rightward motion and the activity pattern is modeled as high coherence, low directional variance. Imagery signal (red) is same as (A). The weighted sum of these activity patterns (blue) reflects the synergy between stimulus and imagery signals. Because the stimulus is strong and unambiguous, the imagery signal yields an insignificant rightward bias above that resulting from stimulus signal alone.

(E) Plot of expected psychometric functions for right-left direction discrimination. Direction discrimination performance is predicted to be proportional to the relative strengths of activation of neurons in opposing (rightward versus leftward) direction columns. Stimulus-only condition is indicated in black. Imagery condition, for which rightward motion has been associatively paired with the color red, is indicated in blue. The upward shift of the psychometric function reflects the perceived directional bias toward rightward motion in the red condition. The four arrows correspond to the imagery-induced directional biases elicited for conditions (A)–(D) above. The bias is large for conditions below threshold (when the stimulus is ambiguous), but the imagery-induced bias is small when the stimulus signal is robust and ambiguous.

(2007) one might suppose that the thing recalled (a patch of moving dots) appears in the form it has been previously seen and serves as an explicit template for an expected target. Under these conditions, the image may have no direct or meaningful influence over the percept of the retinal stimulus that elicited it. Correspondingly, the observed recall-related activity in area MT may have no bearing on the percept of the arrow stimulus that was simultaneously visible.

It seems likely, however, that the retrieval substrate that affords explicit imagery is more commonly—indeed ubiquitously—employed for implicit imagery, which is notable for its functional interactions with the retinal stimulus. Indeed, one mechanistic interpretation of the claim that perceptual experience falls routinely at varying positions along a stimulus-imagery continuum is that bottom-up stimulus and top-down recall-related signals are not simply coexistent in visual cortex, but perpetually interact to yield percepts of "probable things."

This mechanistic proposal can be conveniently fleshed-out and employed to make testable predictions following the logic that Newsome and colleagues (e.g., Nichols and Newsome, 2002) have used to address the interaction between bottom-up motion signals and electrical microstimulation of MT neurons. (This analogy works because microstimulation can be considered a crude form of top-down signal.) As illustrated schematically in Figure 6, bottom-up (stimulus) and top-down (imaginal) inputs to area MT should yield distinct activity patterns across the spectrum of direction columns (Albright et al., 1984). According to this simple model, perceptual experience is determined as a weighted average of these activity distributions (an assumption consistent with perceived motion in the presence of two real moving components [Adelson and Bergen, 1985; Qian et al., 1994; Stromeyer et al., 1984; van Santen and Sperling, 1985]). Under normal circumstances, the imaginal component—elicited by cued associative recall—would be expected to reinforce the stimulus component, which has obvious

functional benefits (noted above) when the stimulus is weak (e.g., Figure 6C).

Potentially more revealing predictions occur for the unlikely case in which stimulus and imaginal components are diametrically opposed (Figure 6A). The resulting activity distribution naturally depends upon the relative strengths of the stimulus and imaginal components. It follows that if the imaginal component is constant, its sway over perceived direction of motion will depend dramatically upon the strength of the retinal stimulus (Figures 6B–6D and 6E). In the extreme, this model predicts that a stimulus that is directionally ambiguous or composed of dynamic noise will yield a percept of directional motion when the imaginal component is directionally strong (Figure 6B).

Support for this mechanistic interpretation comes in part from an experiment by Backus and colleagues (Haijiang et al., 2006). These investigators used classical conditioning to train associations between two directions of motion and two values of a covert second cue (e.g., stimulus position). Following learning, human subjects were presented with directionally ambiguous (bistable) motion stimuli along with one or the other cue value. Subjects exhibited marked biases in the direction of perceived motion, which were dictated by the associated cue, even though subjects professed no awareness of the cue or its meaning. The discovery of recall-related activity in area MT (Schlack and Albright, 2007) suggests that these effects of association-based recall on perception are mediated through integration of bottom-up (ambiguous stimulus) and top-down (reliable implicit imagery) signals at the level of individual cortical neurons.

One important prediction of this mechanistic hypothesis is that the influence of top-down associative recall on perception should, under normal circumstances, be inversely proportional to the "strength" of the bottom-up sensory signal (Figure 6). To test this prediction, A. Schlack et al. (2008, Soc. Neurosci., abstract) designed an experiment in which the influence of associative recall on reports of perceived direction of motion could be systematically quantified over a range of input strengths. The visual stimuli used for this experiment consisted of dynamic dot displays, in which the fraction of dots moving in the same direction (i.e., "coherently") could be varied from 0% to 100%, while the remaining (noncoherent) dots moved randomly. By varying the motion coherence strength, the relative influence of bottom-up and top-down signals could be evaluated over a range of input conditions. These stimuli lend the additional advantage that there is an extensive literature in which they have been used to quantify perceptual and neuronal sensitivity to visual motion (e.g., Britten et al., 1992; Croner and Albright, 1997, 1999; Newsome et al., 1989).

The experiment conducted by Schlack et al. (2008, Soc. Neurosci., abstract) consisted of three phases. In the first ("pretrain") phase, human subjects performed an up-down direction discrimination task using stimuli of varying motion signal strength. The observed psychometric functions confirmed previous reports: the point of subjective equality (equal frequency of responses in the two opposite directions) occurred where the motion signal was at or near 0%. In the second ("training") phase, subjects were exposed to repeated pairings of the directions and colors of moving dot patterns, e.g., upward-green, downward-red. This classical associative conditioning continued 1 hr/day for 20 days and was followed by the third ("posttrain") phase of the experiment, in which direction discrimination performance was reassessed using dot patterns of the two colors employed in phase two (red and green).

Schlack et al. (2008, Soc. Neurosci., abstract) argued that the associative training of phase two would result in cue-dependent recall-related activity in area MT. Reports of perceived direction of motion in phase three should thus reflect a combination of top-down (imaginal) and bottom-up (stimulus) motion signals. Furthermore, the influence of the imaginal component should depend inversely upon the strength of the stimulus component. This is precisely what was observed: the psychometric functions for direction discrimination obtained for red and for green moving dot patterns were displaced relative to one another in a manner consistent with perceptual biases introduced by the associated color cue. These psychophysical findings, in conjunction with the previous discovery of recall-related activity in area MT (Schlack and Albright, 2007), lead to the strong prediction that functions for neuronal discriminability (neurometric functions) of motion direction will exhibit biases that mirror the psychophysical bias, reflect cued associative recall, and are accountable by the simple model outlined in Figure 6.

Distinguishing Stimulus from Imagery

Considerations of the balance between stimulus and imagery naturally raise the larger question of whether (and how) an observer can distinguish between the two if they are both manifested as activation of visual cortex. And, if so, under what conditions does it make a difference? These questions are not new, of course, having been raised repeatedly since the 19th century in discussions of the clinical phenomenon of hallucination (e.g., James, 1890; Richardson, 1969; Sully, 1888). The studies reviewed herein allow these questions to be addressed in a modern neurobiological context.

Most modern neurobiological approaches to these questions skirt the "perceptual equivalence" problem and begin instead with the premise that the perceptual states elicited independently by stimulus versus explicit imagery are, in fact, quite distinct. While visual cortex may provide a common substrate for representation, the perceptual distinction implies that there are different neuronal states associated with stimulus versus imagery. Human neuropsychological (see Behrmann, 2000, and Bartolomeo, 2002, for review) and fMRI studies (e.g., Lee et al., 2012) support this view. Broadly speaking, lesions of more anterior regions along the ventral visual cortical stream—particularly visual areas of the temporal lobe—may impair the capacity to generate explicit visual images while leaving intact the ability to perceive retinal stimuli (Farah et al., 1988; Moro et al., 2008). Conversely, lesions of more posterior regions of visual cortex—low- and mid-level visual processing areas—may disrupt the perception of retinal stimuli without affecting the ability to generate visual images (Bartolomeo et al., 1998; Behrmann et al., 1992; Bridge et al., 2011; Chatterjee and Southwood, 1995). Similarly, although fMRI studies reveal that retinal

stimulation and explicit visual imagery yield largely overlapping patterns of activity in visual cortex (Kosslyn et al., 1997), there are readily detectable differences between these patterns (e.g., Amedi et al., 2005; Lee et al., 2012; Ishai et al., 2000; Roland and Gulyas, 1994), which corroborate the neuropsychological evidence for a stimulus-imagery dissociation and are presumed to account for the differences in perceptual state.

These findings help to resolve a paradox posed by the findings of Schlack and Albright (2007), in which bottom-up and top-down activity patterns in area MT are seemingly equivalent (see Figure 3), but the perceptual states associated with these neuronal activities are not likely to be so. Simply put, isolated recordings from area MT do not tell the full story; MT may be part of the common neuronal substrate for representing stimulus and imagery, but the perceptual states elicited in these experiments are presumably distinguished by differential activation of other cortical regions, such as those identified in the neuropsychological and fMRI studies cited above.

While the presumption that stimulus and imagery elicit different perceptual and neuronal states may generally hold for explicit imagery, a more nuanced view emerges from implicit imagery. Here, the stimulus-imagery distinction is largely moot, as this view posits that perception reflects an ongoing integration of stimulus and imagery signals in visual cortex—observers are simply unaware of the source of the signals. In most cases, imagery corroborates the retinal stimulus by filling in detail based on prior experience. The possibility exists, however, that the imagery signal reflects an incorrect association or flawed premises about the environment, and perceptual experience is none the wiser. If the imaginal component dominates, as it often does in such cases, the result is a commonplace illusion: the coat rack may look like an intruder in the hall, or the shrubbery is mistaken for a police car. The Bruner and Postman (1949) "trick card" study, cited above, is a prime example of such conditions, in which "imagination has all the force of fact" (James, 1890).

There also exists a genre of magical performance art that capitalizes upon illusions derived from flawed inferences—it is the observer's failure to distinguish stimulus from imagery that makes this art possible. Consider, for example, the "vanishing ball illusion": in this simple yet compelling trick, the magician repeatedly tosses a ball into the air. On the final toss, the ball vanishes in mid flight (for video demonstration, see Kuhn and Land [2006], http://www.cell.com/current-biology/supplemental/S0960-9822(06)02331-1). In reality, the ball never leaves the hand. The illusion is effected by the use of learned cues that are visible to the observer, including the magician's hand and arm movements previously associated with a ball toss, and the magician's gaze directed along the usual path of the ball. The observer's inferences about environmental properties and events are probabilistically determined (from the associated cues) but the inferences are incorrect. According to the implicit imagery hypothesis, these flawed inferences are nonetheless manifested as imagery of motion along the expected path. Moreover, this imaginal contribution to perceptual experience is likely to be mediated by top-down activation of directionally selective MT neurons, in a manner analogous to the effects reported by Schlack and Albright (2007).

In other cases of implicit imagery, however, such as a cloud that looks like a poodle or a toast that resembles the Virgin Mary, the imagined component may be robust but it is scarcely confusable with the stimulus. A well-documented and experimentally tractable form of this perceptual phenomenon is variously termed "representational momentum" (Freyd, 1987; Kourtzi, 2004; Senior et al., 2000), "implied motion" (Kourtzi and Kanwisher, 2000; Krekelberg et al., 2003; Lorteije et al., 2006), or "illusions of locomotion" (Arnheim, 1951), in which a static image drawn from a moving sequence (such as an animal in a predatory pounce) elicits an "impression" of the motion sequence. This phenomenon is the basis of a common technique in painting, well-described since Leonardo (da Vinci, 1989), in which static visual features are employed to bring a vibrant impression to canvas. Such impressions are ubiquitous, perceptually robust, and nonvolitional (unlike explicit imagery), but they are not confusable with stimulus motion. Evidence nonetheless suggests that they also reflect top-down pictorial recall of motion—the product of associative experience, in which static elements of a motion sequence have been naturally linked with the movement itself (Freyd, 1987). In support of this view, static implied motion stimuli have been shown to elicit fMRI signals selectively in human areas MT and MST (Kourtzi and Kanwisher, 2000; Lorteije et al., 2006; Senior et al., 2000). Krekelberg et al. (2003) have discovered similar effects for single neurons in cortical areas MT and MST.

What then differentiates cases in which imagery and stimulus are inseparable from cases in which they are distinct? We have already seen that the distinct experiences associated with *explicit* imagery versus retinal stimulation are linked to activation of anterior versus posterior regions of visual cortex. We hypothesize that the same cortical dissociation can hold for implicit imagery. Moreover, for both explicit and implicit forms of imagery this cortical dissociation will only occur under conditions in which the perceptual consequences of stimulus and imagery are dissociable based on "content."

One content factor that is correlated with the stimulus-imagery distinction is the strength and quality of evidence for sensation (see James, 1890). When the stimulus is robust and unambiguous, the stimulus is distinctly perceived. Imagery is inconsequential (as in Schlack et al. [2008, Soc. Neurosci., abstract], reviewed above) or irrelevant (drastically improbable, as in clouds that look like things, or contrived, as in explicit imagery). When the stimulus is weak, by contrast, stimulus-imagery confusion may result (as in phantoms). Empirical support for this view comes originally from a widely cited experiment of the early 20th century (Perky, 1910) in which human observers were instructed to imagine specific objects (e.g., a banana) while viewing a "blank" screen. Unbeknownst to the observers, very low-contrast (but suprathreshold) images of the same object were projected on the screen during imagery. Under these conditions, the perceptual experience was consistently attributed to imagery—a phenomenon known as the "Perky effect"—observers evinced no awareness of the projected stimuli, although the properties of those stimuli (e.g., the orientation of the projected banana) could readily influence the experience. If the contrast of the projected stimuli were made sufficiently large, or if subjects were told that projected stimuli

would appear, by contrast, the perceptual experience was consistently attributed to the stimulus.

Neurobiological support for the possibility that the stimulus-imagery distinction is based, in part, on the strength and quality of evidence for sensation comes from studies of the effects of electrical microstimulation of cortical visual area MT (Salzman et al., 1990). This type of stimulation can be thought of as an artificial form of top-down activation, and the stimulus-imagery problem applies here as well. Newsome and colleagues have shown that this activation is confused with sensation, in that it is added (as revealed by perceptual reports) to the simultaneously present retinal stimulus (Salzman et al., 1990). But this is only true when the stimulus is weak. When the stimulus is strong, microstimulation has little measurable effect on behavior.

A related content factor that differentiates cases in which imagery and stimulus are inseparable from cases in which they are distinct is the a priori probability of the imagined component. If the retinal stimulus is weak or ambiguous, some images come to mind because they are statistically probable features of the environment, and the stimulus and imaginal contributions are inseparable. But other images come to mind on a lark or by a physical resemblance to something seen before (such as the Rorschach ink blot that looks like a bat). Images of the latter variety are commonly indifferent to known statistics of the observer's environment and they are rarely confused with properties of that environment received as sensory stimuli. (As with the old military adage, "When the terrain differs from the map, trust the terrain.")

Although little is known of the neuronal mechanisms by which probability influences this process (but see Girshick, Landy and Simoncelli, 2011), there are well known psychopathologies and drug-induced alterations of sensory processing in which the imaginal component dominates regardless of its likelihood or the quality of stimulation, and perceptual experience becomes hallucination. By this view, visual hallucinations are a pathological product of the same top-down system for pictorial recall that serves perceptual inference—a view supported by the finding of activity patterns in visual cortex that are correlated with visual hallucinations in cases of severe psychosis (Oertel et al., 2007). Moreover, evidence indicates that sensory cortex is less sensitive to exogenous stimulation during hallucinations (Kompus et al., 2011), suggesting that the imaginal component is given a competitive advantage.

A particularly striking pathological case of overreaching imaginal influences on perception is Charles Bonnet Syndrome (CBS)—a bizarre disorder characterized by richly detailed visual imagery in individuals who have recently lost sight from pathology to the retina (e.g., macular degeneration) or optic nerve (Gold and Rabins, 1989). The images perceived are commonly elicited by associative cues. For example, upon hearing an account of the revolutionary war, one patient with CBS reported a vivid percept of a winking sailor: "He had on a cap, a blue cap with a polished black beak and he had a pipe in his mouth" (Krulwich, 2008). Similar imagery-dominated perceptual experiences have been reported for normal human subjects artificially deprived of vision for extended periods (Merabet et al., 2004).

In all of these cases in which stimulus properties and probabilities, or myriad pathological and pharmacological states, influence the perceptual distinction between stimulus and imagery, we can assume that there are patterns of neuronal signaling correlated with that distinction. Likely candidates are those brain regions found to be differentially engaged in the neuropsychological and fMRI studies of explicit imagery cited above. Much additional work is needed, however, to identify the specific mechanisms and neuronal events that underlie these effects.

Intermodal Associations and Perceptual Experience

This review has focused on vision because it is the sensory system for which there exists the greatest understanding of perceptual experience as well as relevant neuronal organization and function. There are nonetheless good reasons to believe that the same principles for associative recall and perception pertain to all senses. Moreover, these principles apply well to interactions between sensory modalities. Perceptual phenomena reflecting such interactions can be robust and dramatic. To illustrate the point, William James (1890) offered the phrase "Pas de lieu Rhône que nous," which any Frenchman will tell you makes no sense at all. If, however, the listener is informed that the spoken phrase is English, the very same sounds are perceived as "paddle your own canoe." James noted further that "as we seize the English meaning *the sound itself appears to change*" (my italics).

Along the same lines, Sumby and Pollack (1954) showed that visibility of a speaker's lips improves auditory word recognition, particularly when spoken words are embedded in auditory noise. The McGurk Effect (McGurk and MacDonald, 1976) demonstrates, furthermore, that moving lips can markedly bias the interpretation of clearly spoken phonemes. Just as argued for vision, the visual cue stimulus in such cases elicits associative auditory recall, which interacts with the bottom-up auditory stimulus. The product is a percept fleshed out by auditory imagery derived from probabilistic rules. These conclusions are supported by neurobiological evidence for intermodal associative recall, which comes from both human brain-imaging studies (e.g., Calvert et al., 1997; Sathian and Zangaladze, 2002; Zangaladze et al., 1999) and single-cell electrophysiology (e.g., Haenny et al., 1988; Zhou and Fuster, 2000).

A special case of intermodal interactions, termed "synesthesia," occurs when a stimulus arising in one sensory modality or submodality (the "inducer") elicits a consistent perceptual experience (the "concurrent") in another modality. For example, grapheme-color synesthesia is characterized by the perception of specific colors upon viewing specific graphical characters (e.g., the number "2" may elicit a percept of the color blue). Owing to its intriguing nature, synesthesia has been a subject of study in psychology and neuroscience for well over 100 years (Galton, 1880), yet there remains much debate about its etiology. Evidence suggests a heritable contribution in some cases (Baron-Cohen et al., 1996), but in other cases the condition appears dependent upon prior experience (Howells, 1944; Mills et al., 2002; Ward and Simner, 2003; Witthoft and Winawer, 2006). These experience-based cases argue that synesthetes have learned associations between stimuli representing the

Figure 7. *Hoarfrost at Ennery (Gilee Blanche)*, Camille Pissarro, Oil on Canvas, 1873, Musée d'Orsay, Paris
Pissarro's impressionist depiction of frost on a plowed field was the target of a satirical review by the Parisian art critic Louis Leroy (1874), which questioned the legitimacy, value, and aesthetics of this new form of art. The impressionists maintained that a few simple and often crudely rendered features were sufficient to trigger a perceptual experience richly completed by the observer's own preposessions. Neuroscientific evidence reviewed herein suggests that this perceptual completion occurs via the projection of highly-specific top-down signals into visual cortex. Image used with permission of Art Resource.

inducer and concurrent and that subsequent presentation of the inducer elicits recall of the concurrent. We add to this argument the hypothesis that the recall event constitutes implicit imagery of the concurrent, which is mediated by top-down activation of visual cortex. This appears to be a case in which a learned association is so idiosyncratic that the resulting imaginal contribution to perception, albeit highly significant, has no inherent value or adaptive influence over behavior.

Imagery, Categorical Perception, and Perceptual Learning

Top-down signaling in visual cortex benefits perception by enabling stimuli to be seen as they are likely to be. One might easily imagine how this same system could facilitate discrimination of unfamiliar stimuli by inclining them to be perceived as familiar stereotypes or caricatures. In his discussion of perceptual learning—the improved discriminative capacity that comes with practice—William James (1890) raised this possibility:

> "I went out the other day and found that the snow just fallen had a very odd look, different from the common appearance of snow. I presently called it a "micaceous" look; and it seemed to me as if, the moment I did so, the difference grew more distinct and fixed than it was before. The other connotations of the word "micaceous" dragged the snow farther away from ordinary snow and seemed even to aggravate the peculiar look in question."

What James speaks of is a form of categorical perception, in which a sensory stimulus (snow, in this example) becomes bound by association with a large category of stimuli (things that look like mica) that share unique sensory characteristics. This phenomenon is a common feature of human perceptual learning: category concepts or labels can predictably bias judgments of visual similarity (e.g., Goldstone, 1994; Goldstone et al., 2001; Gauthier et al., 2003; Yu et al., 2008). All else being equal, stimuli that are members of the same category are commonly less discriminable from one another than are members of different categories. Gauthier et al. (2003) have argued that the key element is semantic association, as it is meaning that defines category. While the emphasis on semantic assignment may be valid, it is arguably true that any sensory-sensory association is semantic, as the meaning of a sensory stimulus is given in part by the sensory stimuli with which it is associated.

Ryu and Albright (2010) explored this sensory association hypothesis more fully in an attempt to link the perceptual consequences of category learning to existing evidence for top-down signaling in sensory cortex. These investigators assessed performance of human observers on a difficult orientation discrimination task before and after learning of specific visual-auditory associations. After the initial orientation discrimination assessment, observers were trained to associate the orientations individually with one of two very distinct tones: for example, an orientation of 10° was paired with a tone frequency of 200 Hz and an orientation of 16° was paired with 1,000 Hz. Orientation discrimination performance improved markedly following orientation-tone pairing. As for James' varieties of snow, one can interpret these findings as resulting from differential category assignment of the two orientations. The category labels (auditory tones) in this case are simply symbols that represent the paired visual orientations.

These effects can be understood mechanistically using the stimulus-imagery framework described above. This interpretation begins with the indubitable assumption that the discriminability of two stimuli is determined, in part, by the degree of overlap between the patterns of neuronal activity that they elicit (e.g., Gilbert et al., 2001). The orientation discriminanda used in these experiments (6° difference) would be expected to activate highly overlapping distributions of neurons in primary visual cortex, yielding a difficult discrimination. The findings of Schlack and Albright (2007) and others (e.g., Zhou and Fuster, 2000), however, imply that orientation-tone associative learning should lead to selective top-down activation of cortical neurons representing the stimuli recalled by association. By this logic, viewing of each of the orientation discriminanda will not only drive orientation-selective neurons in visual cortex but should also activate the corresponding frequency-selective neurons in auditory cortex. If the distributions of recall-related neuronal activity in auditory cortex are sufficiently distinct (as would be expected for 200 Hz versus 1,000 Hz tones) those activations may be the basis for improved discrimination of the visual orientations (relative to the untrained state). In other words, the improved discriminability of visual orientations is made possible through the use of neuronal proxies, which are established by the learned category labels (tones). This is recognizably the same process that I have termed implicit imagery, but in this case it serves perceptual learning.

Figure 8. Demonstration of the Influence of Associative Pictorial Recall (Top-Down Signaling) on the Interpretation of a Retinal Stimulus (Bottom-Up Signaling)
Most observers will experience a clear meaningful percept upon viewing this pattern. After achieving this percept, refer back to Figure 5. The perceptual interpretation of the pattern should now be markedly different, with a figural interpretation that is driven largely by imaginal influences drawn from memory.

It Always Comes Out of Our Own Head

"You see… a hoarfrost on deeply plowed furrows."

"Those furrows? That frost? But they are palette-scrapings placed uniformly on a dirty canvas. It has neither head nor tail, top nor bottom, front nor back."

"Perhaps… but the impression is there."

This fictional exchange between two 19th century painters was penned by the Parisian critic Louis Leroy (1874) after viewing Camille Pissarro's painting titled *Hoarfrost at Ennery (Gilee Blanche)* (Figure 7) at the first major exhibition of impressionist art (in Paris, 1874). Leroy was not a fan and his goal was satire, but his critic's assertion, "but the impression is there," nonetheless captures the essence of the art (and Leroy's term "impressionism" was, ironically, adopted as the name of the movement). Indeed, it is precisely what the artist intended, and the art form's legitimacy—and ultimately its brilliance—rests on the conviction that the "impression" (the retinal stimulus) is merely a spark for associative pictorial recall. The impressionist painter does not attempt to provide pictorial detail, but rather creates conditions that enable the viewer to charge the percept, to complete the picture, based on his/her unique prior experiences. ("The beholder's share" is what Gombrich [1961] famously and evocatively termed this memory-based contribution to the perception of art.)

Naturally, both the beauty and the fragility of the method stem from the fact that different viewers bring different preconceptions and imagery to bear. Leroy's critic saw only "palette-scrapings on a dirty canvas." Legend has it that, upon viewing a particularly untamed (by the standards of the day) sunset by the pre-impressionist J.M.W. Turner, a young woman remarked, "I never saw a sunset like that, Mr. Turner." To which Turner replied, "Don't you wish you could, madam?" The undeniable pleasure that many viewers take in this art form is an example of what James (1890) termed "the victorious assimilation of the new," the coherent perceptual experience of the unknown, something we have never quite seen before, by its association with things familiar. The alternative is perceptual rejection of the new—it bears and elicits no meaning—leaving the observer's (e.g., Leroy's critic and Turner's companion) experience mired in the literal and commonplace world of retinal stimuli.

These knotty concepts of perception, memory, and individual human experience stand amid a myriad of cognitive factors long thought to lie beyond the reach of one's microelectrode. The recent work reviewed here suggests otherwise, and it identifies a novel perspective that can now guide the neuroscientific study of perception forward—ever bearing in mind James' "general law of perception": "Whilst part of what we perceive comes through our senses from the object before us, another part (and it may be the larger part) always comes out of our own head" (James, 1890).

ACKNOWLEDGMENTS

I am indebted to many colleagues and collaborators—particularly Gene Stoner, Larry Squire, Sergei Gepshtein, Charlie Gross, and Terry Sejnowski—for insights and provocative discussions of these topics in recent years. I also owe much to the late Margaret Mitchell for unparalleled administrative assistance delivered with pride and an unforgettable spark of wit.

REFERENCES

Adelson, E.H., and Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A *2*, 284–299.

Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. *52*, 1106–1130.

Albright, T.D. (1993). Cortical processing of visual motion. In Visual Motion and its Use in the Stabilization of Gaze, J. Wallman and F.A. Miles, eds. (Amsterdam: Elsevier), pp. 177–201.

Albright, T.D. (1995). 'My most true mind thus makes mine eye untrue.' Trends Neurosci. *18*, 331–333.

Albright, T.D., Desimone, R., and Gross, C.G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. J. Neurophysiol. *51*, 16–31.

Alexander, M.P., and Albert, M.I. (1983). The anatomical basis of visual agnosia. In Localization in Neuropsychology, A. Kertesz, ed. (New York: Academic Press).

Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., and Naumer, M.J. (2005). Functional imaging of human crossmodal identification and object recognition. Exp. Brain Res. *166*, 559–571.

Anderson, J.R., and Bower, G.H. (1973). Human Associative Memory (Washington, D.C.: Winston).

Arnheim, R. (1951). Perceptual and aesthetic aspects of the movement response. J. Pers. *19*, 265–281.

Assad, J.A., and Maunsell, J.H. (1995). Neuronal correlates of inferred motion in primate posterior parietal cortex. Nature *373*, 518–521.

Ball, K., and Sekuler, R. (1980). Models of stimulus uncertainty in motion perception. Psychol. Rev. *87*, 435–469.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., and Bolton, P. (1996). Synaesthesia: prevalence and familiality. Perception *25*, 1073–1079.

Bartleson, C.J. (1960). Memory colors of familiar objects. J. Opt. Soc. Am. *50*, 73–77.

Bartolomeo, P. (2002). The relationship between visual perception and visual mental imagery: a reappraisal of the neuropsychological evidence. Cortex *38*, 357–378.

Bartolomeo, P., Bachoud-Lévi, A.C., De Gelder, B., Denes, G., Dalla Barba, G., Brugières, P., and Degos, J.D. (1998). Multiple-domain dissociation between impaired visual perception and preserved mental imagery in a patient with bilateral extrastriate lesions. Neuropsychologia *36*, 239–249.

Bateson, G. (1972). Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology (Chicago: University of Chicago Press).

Behrmann, M. (2000). The mind's eye mapped onto the brain's matter. Curr. Psychol. Sci. *9*, 50–54.

Behrmann, M., Winocur, G., and Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. Nature *359*, 636–637.

Bridge, H., Harrold, S., Holmes, E.A., Stokes, M., and Kennard, C. (2011). Vivid visual mental imagery in the absence of the primary visual cortex. J. Neurol., in press. Published online November 8, 2011. 10.1007/s00415-011-6299-z.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J. Neurosci. *12*, 4745–4765.

Brown, S., and Schafer, E.S. (1888). An investigation into the functions of the occipital and temporal lobes of the monkey's brain. Philos. Trans. R. Soc. Lond. B Biol. Sci. *179*, 303–327.

Bruner, J.S., and Postman, L. (1949). On the perception of incongruity; a paradigm. J. Pers. *18*, 206–223.

Bruner, J.S., Postman, L., and Rodrigues, J. (1951). Expectation and the perception of color. Am. J. Psychol. *64*, 216–227.

Brunswik, E. (1956). Perception and the Representative Design of Psychological Experiments (Berkeley, CA: University of California Press).

Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, R., Williams, S.C., McGuire, P.K., Woodruff, P.W., Iversen, S.D., and David, A.S. (1997). Activation of auditory cortex during silent lipreading. Science *276*, 593–596.

Chatterjee, A., and Southwood, M.H. (1995). Cortical blindness and visual imagery. Neurology *45*, 2189–2195.

Croner, L.J., and Albright, T.D. (1997). Image segmentation enhances discrimination of motion in visual noise. Vision Res. *37*, 1415–1427.

Croner, L.J., and Albright, T.D. (1999). Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. J. Neurosci. *19*, 3935–3951.

da Vinci, Leonardo (1989). Leonardo on Painting. M. Kemp, ed., M. Kemp and M. Walker, trans. (New Haven, CT: Yale University Press).

D'Esposito, M., Detre, J.A., Aguirre, G.K., Stallcup, M., Alsop, D.C., Tippet, L.J., and Farah, M.J. (1997). A functional MRI study of mental image generation. Neuropsychologia *35*, 725–730.

Damasio, A.R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. Cognition *33*, 25–62.

Descartes, R. (1972). Treatise on Man, Translated by T.S. Hall (Cambridge, MA: Harvard University Press).

Desimone, R., Fleming, J., and Gross, C.G. (1980). Prestriate afferents to inferior temporal cortex: an HRP study. Brain Res. *184*, 41–55.

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

Farah, M.J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. J. Exp. Psychol. Gen. *114*, 91–103.

Farah, M.J., Péronnet, F., Gonon, M.A., and Giard, M.H. (1988). Electrophysiological evidence for a shared representational medium for visual images and visual percepts. J. Exp. Psychol. Gen. *117*, 248–257.

Finke, R.A. (1980). Levels of equivalence in imagery and perception. Psychol. Rev. *87*, 113–132.

Finke, R.A. (1989). Principles of Mental Imagery (Cambridge, MA: MIT Press).

Flechsig, P. (1876). Die Leitungsbahnen im Gehirn and Ruckenmard des Menschen auf grund entwicklungsgeschichtlicher untersuchungen (Leipzig, Germany: Englemann).

Freyd, J.J. (1987). Dynamic mental representations. Psychol. Rev. *94*, 427–438.

Galton, F. (1880). Visualised numerals. Nature *21*, 494–495.

Gattass, R., and Gross, C.G. (1981). Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. J. Neurophysiol. *46*, 621–638.

Gauthier, I., James, T.W., Curby, K.M., and Tarr, M.J. (2003). The influence of conceptual knowledge on visual discrimination. Cogn. Neuropsychol. *20*, 507–523.

Gilbert, C.D., Sigman, M., and Crist, R.E. (2001). The neural basis of perceptual learning. Neuron *31*, 681–697.

Girshick, A.R., Landy, M.S., and Simoncelli, E.P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. Nat. Neurosci. *14*, 926–932.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., and Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. Eur. J. Neurosci. *10*, 1563–1573.

Gold, K., and Rabins, P.V. (1989). Isolated visual hallucinations and the Charles Bonnet syndrome: a review of the literature and presentation of six cases. Compr. Psychiatry *30*, 90–98.

Goldstone, R. (1994). Influences of categorization on perceptual discrimination. J. Exp. Psychol. Gen. *123*, 178–200.

Goldstone, R.L., Lippa, Y., and Shiffrin, R.M. (2001). Altering object representations through category learning. Cognition *78*, 27–43.

Gombrich, E.H. (1961). Art and Illusion: A Study in the Psychology of Pictorial Representation (Princeton, NJ: Princeton Univ Press).

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

Gross, C.G., Desimone, R., Albright, T.D., and Schwartz, E.L. (1985). Inferior temporal cortex and pattern recognition. In Study Group on Pattern Recognition Mechanisms, C. Chagas, ed. (Vatican City: Pontifica Academia Scientiarum), pp. 179–200.

Haenny, P.E., Maunsell, J.H., and Schiller, P.H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. Exp. Brain Res. *69*, 245–259.

Haijiang, Q., Saunders, J.A., Stone, R.W., and Backus, B.T. (2006). Demonstration of cue recruitment: change in visual appearance by means of Pavlovian conditioning. Proc. Natl. Acad. Sci. USA *103*, 483–488.

Hansen, T., Olkkonen, M., Walter, S., and Gegenfurtner, K.R. (2006). Memory modulates color appearance. Nat. Neurosci. *9*, 1367–1368.

Hebb, D.O. (1949). The Organization of Behavior: A Neuropsychological Theory (New York: Wiley).

Hebb, D.O. (1968). Concerning imagery. Psychol. Rev. *75*, 466–477.

Helmholtz, H.V. (1924). Treatise on Physiological Optics, J.P.C. Southall, trans (New York: Optical Society of America).

Hering, E. (1878). Zur Lehre vom Lichtsinne (Principles of a new theory of the color sense). In Color Vision, Selected Readings, R.C. Teevan and R.C. Birney, eds., K. Butler, trans. (New York: Van Nostrand Reinhold).

Higuchi, S., and Miyashita, Y. (1996). Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. Proc. Natl. Acad. Sci. USA *93*, 739–743.

Howells, T.H. (1944). The experimental development of color-tone synesthesia. J. Exp. Psychol. *34*, 87–103.

Hume, D. (1967). The Natural History of Religion (Stanford: Stanford University Press).

Hurlbert, A.C., and Ling, Y. (2005). If it's a banana, it must be yellow: the role of memory colors in color constancy. J. Vis. *5*, 787a.

Ishai, A., and Sagi, D. (1995). Common mechanisms of visual imagery and perception. Science *268*, 1772–1774.

Ishai, A., and Sagi, D. (1997a). Visual imagery: effects of short- and long-term memory. J. Cogn. Neurosci. *9*, 734–742.

Ishai, A., and Sagi, D. (1997b). Visual imagery facilitates visual perception: psychophysical evidence. J. Cogn. Neurosci. *9*, 476–489.

Ishai, A., Ungerleider, L.G., and Haxby, J.V. (2000). Distributed neural systems for the generation of visual images. Neuron *28*, 979–990.

James, W. (1890). Principles of Psychology (New York: Henry Holt).

Kanizsa, G. (1979). Organization in Vision: Essays on Gestalt Perception (New York: Praeger).

Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. Annu. Rev. Psychol. *55*, 271–304.

Kluver, H., and Bucy, P.C. (1939). Preliminary analysis of functions of the temporal lobes in monkeys. Arch. Neurol. Psychiatry *42*, 979–1000.

Knapska, E., and Kaczmarek, L. (2004). A gene for neuronal plasticity in the mammalian brain: Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK? Prog. Neurobiol. *74*, 183–211.

Knauff, M., Kassubek, J., Mulack, T., and Greenlee, M.W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. Neuroreport *11*, 3957–3962.

Knill, D.C., and Richards, W. (1996). Perception as Bayesian Inference (Cambridge: Cambridge University Press).

Kompus, K., Westerhausen, R., and Hugdahl, K. (2011). The "paradoxical" engagement of the primary auditory cortex in patients with auditory verbal hallucinations: a meta-analysis of functional neuroimaging studies. Neuropsychologia *49*, 3361–3369.

Kosslyn, S.M. (1994). Image and Brain (Cambridge, MA: The MIT Press).

Kosslyn, S.M., Thompson, W.L., Kim, I.J., and Alpert, N.M. (1995). Topographical representations of mental images in primary visual cortex. Nature *378*, 496–498.

Kosslyn, S.M., Thompson, W.L., and Alpert, N.M. (1997). Neural systems shared by visual imagery and visual perception: a positron emission tomography study. Neuroimage *6*, 320–334.

Kourtzi, Z. (2004). "But still, it moves.". Trends Cogn. Sci. (Regul. Ed.) *8*, 47–49.

Kourtzi, Z., and Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. J. Cogn. Neurosci. *12*, 48–55.

Kreiman, G., Koch, C., and Fried, I. (2000). Imagery neurons in the human brain. Nature *408*, 357–361.

Krekelberg, B., Dannenberg, S., Hoffmann, K.P., Bremmer, F., and Ross, J. (2003). Neural correlates of implied motion. Nature *424*, 674–677.

Krulwich, R. (2008). Blind Man 'Sees.' National Public Radio, All Things Considered. January 30, 2008. http://www.npr.org/templates/story/story.php?storyId=18487511.

Kuhn, T.S. (1962). The Structure of Scientific Revolutions (Chicago: University of Chicago Press).

Kuhn, G., and Land, M.F. (2006). There's more to magic than meets the eye. Curr. Biol. *16*, R950–R951.

Lee, S.H., Kravitz, D.J., and Baker, C.I. (2012). Disentangling visual imagery and perception of real-world objects. Neuroimage *59*, 4064–4073.

Leroy, L. (1874). Exposition des Impressioinnistes (Exhibition of the Impressionists). Le Charivari, April 25, 1874.

Lissauer, H. (1988). A case of visual agnosia with a contribution to theory. M. Jackson, trans. Cogn. Neuropsychol. *5*, 157–192.

Locke, J. (1690). An Essay Concerning Human Understanding (London: Basset).

Lorteije, J.A., Kenemans, J.L., Jellema, T., van der Lubbe, R.H., de Heer, F., and van Wezel, R.J. (2006). Delayed response to animate implied motion in human motion processing areas. J. Cogn. Neurosci. *18*, 158–168.

Lu, B. (2003). BDNF and activity-dependent synaptic modulation. Learn. Mem. *10*, 86–98.

Mast, F.W., Berthoz, A., and Kosslyn, S.M. (2001). Mental imagery of visual motion modifies the perception of roll-vection stimulation. Perception *30*, 945–957.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. Nature *264*, 746–748.

Merabet, L.B., Maguire, D., Warde, A., Alterescu, K., Stickgold, R., and Pascual-Leone, A. (2004). Visual hallucinations during prolonged blindfolding in sighted subjects. J. Neuroophthalmol. *24*, 109–113.

Merzenich, M.M., and Kaas, J.H. (1980). Principles of organization of sensory-perceptual systems in mammals. Prog. Psychobiol. Physiol. Psychol. *9*, 1–42.

Messinger, A., Squire, L.R., Zola, S.M., and Albright, T.D. (2001). Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc. Natl. Acad. Sci. USA *98*, 12239–12244.

Mill, J.S. (1865). Examination of Sir Willilam Hamilton's Philosophy and of the Principal Philosophical Questions Discussed in His Writings (London: Longmans, Green, Reader and Dyer).

Mills, C.B., Viguers, M.L., Edelson, S.K., Thomas, A.T., Simon-Dack, S.L., and Innis, J.A. (2002). The color of two alphabets for a multilingual synesthete. Perception *31*, 1371–1394.

Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. Clin. Neurosurg. *19*, 421–446.

Mishkin, M. (1982). A memory system in the monkey. Philos. Trans. R. Soc. Lond. B Biol. Sci. *298*, 83–95.

Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. Annu. Rev. Neurosci. *16*, 245–263.

Miyashita, Y. (2004). Cognitive memory: cellular and network machineries and their top-down control. Science *306*, 435–440.

Miyashita, Y., Kameyama, M., Hasegawa, I., and Fukushima, T. (1998). Consolidation of visual associative long-term memory in the temporal cortex of primates. Neurobiol. Learn. Mem. *70*, 197–211.

Moro, V., Berlucchi, G., Lerch, J., Tomaiuolo, F., and Aglioti, S.M. (2008). Selective deficit of mental visual imagery with intact primary visual cortex and visual perception. Cortex *44*, 109–118.

Morrell, F. (1961). Electrophysiological contributions to the neural basis of learning. Physiol. Rev. *41*, 443–494.

Morrell, F., Naquet, R., and Gastaut, H. (1957). Evolution of some electrical signs of conditioning. I. Normal cat and rabbit. J. Neurophysiol. *20*, 574–587.

Murakami, H. (2005). Kafka on the Shore (New York: Knof).

Murray, E.A., Gaffan, D., and Mishkin, M. (1993). Neural substrates of visual stimulus-stimulus association in rhesus monkeys. J. Neurosci. *13*, 4549–4561.

Naya, Y., Sakai, K., and Miyashita, Y. (1996). Activity of primate inferotemporal neurons related to a sought target in pair-association task. Proc. Natl. Acad. Sci. USA *93*, 2664–2669.

Neisser, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology (San Francisco: W.H. Freeman).

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. Nature *341*, 52–54.

Nichols, M.J., and Newsome, W.T. (2002). Middle temporal visual area microstimulation influences veridical judgments of motion direction. J. Neurosci. *22*, 9530–9540.

Nyberg, L., Habib, R., McIntosh, A.R., and Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. Proc. Natl. Acad. Sci. USA *97*, 11120–11124.

O'Craven, K.M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. J. Cogn. Neurosci. *12*, 1013–1023.

Oertel, V., Rotarska-Jagiela, A., van de Ven, V.G., Haenschel, C., Maurer, K., and Linden, D.E. (2007). Visual hallucinations in schizophrenia investigated with functional magnetic resonance imaging. Psychiatry Res. *156*, 269–273.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-association learning. J. Verbal Learn. Verbal Behav. *4*, 32–38.

Penfield, W., and Perot, P. (1963). The brain's record of auditory and visual experience. a final summary and discussion. Brain *86*, 595–696.

Perky, C.W. (1910). An experimental study of imagination. Am. J. Psychol. *21*, 422–452.

Peterson, M.J., and Graham, S.E. (1974). Visual detection and visual imagery. J. Exp. Psychol. *103*, 509–514.

Podgorny, P., and Shepard, R.N. (1978). Functional representations common to visual perception and imagination. J. Exp. Psychol. Hum. Percept. Perform. *4*, 21–35.

Qian, N., Andersen, R.A., and Adelson, E.H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. J. Neurosci. *14*, 7357–7366.

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

Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. Annu. Rev. Neurosci. *27*, 611–647.

Richardson, A. (1969). Mental Imagery (New York: Springer).

Roland, P.E., and Gulyas, B. (1994). Visual imagery and visual representation. Trends Neurosci. *17*, 281–287.

Ryu, J.-J., and Albright, T.D. (2010). The context of stimulus association influences the perception of visual similarity. Proceedings of the Association for the Scientific Study of Consciousness *14*, 77.

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

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. Nature *346*, 174–177.

Sathian, K., and Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. Behav. Brain Res. *135*, 127–132.

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. *8*, 657–661.

Schlack, A., and Albright, T.D. (2007). Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron *53*, 881–890.

Segal, S.J., and Fusella, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. J. Exp. Psychol. *83*, 458–464.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E.T., Brammer, M., and David, A.S. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. Curr. Biol. *10*, 16–22.

Shepard, R.N., and Cooper, L.A. (1982). Mental Images and Their Transformations (Cambridge, MA: MIT Press/Bradford Books).

Shulman, G.L., Ollinger, J.M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Petersen, S.E., and Corbetta, M. (1999). Areas involved in encoding and applying directional expectations to moving objects. J. Neurosci. *19*, 9480–9496.

Siple, P., and Springer, R.M. (1983). Memory and preference for the colors of objects. Percept. Psychophys. *34*, 363–370.

Slotnick, S.D., Thompson, W.L., and Kosslyn, S.M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. Cereb. Cortex *15*, 1570–1583.

Squire, L.R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. Science *253*, 1380–1386.

Squire, L.R., Stark, C.E., and Clark, R.E. (2004). The medial temporal lobe. Annu. Rev. Neurosci. *27*, 279–306.

Stokes, M., Thompson, R., Cusack, R., and Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. J. Neurosci. *29*, 1565–1572.

Stokes, M., Saraiva, A., Rohenkohl, G., and Nobre, A.C. (2011). Imagery for shapes activates position-invariant representations in human visual cortex. Neuroimage *56*, 1540–1545.

Stromeyer, C.F., 3rd, Kronauer, R.E., Madsen, J.C., and Klein, S.A. (1984). Opponent-movement mechanisms in human vision. J. Opt. Soc. Am. A *1*, 876–884.

Sully, J. (1888). Outlines of Psychology, with Special References to the Theory of Education (New York: Appleton).

Sumby, W.H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. *26*, 212–215.

Suzuki, W.A., and Amaral, D.G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J. Comp. Neurol. *350*, 497–533.

Thomas, N.J.T. (2011). Mental Imagery. Stanford Encyclopedia of Philosophy, Winter 2011 Edition. E.N. Zalta, ed. http://plato.stanford.edu/archives/win2011/entries/mental-imagery/.

Tokuyama, W., Okuno, H., Hashimoto, T., Xin Li, Y., and Miyashita, Y. (2000). BDNF upregulation during declarative memory formation in monkey inferior temporal cortex. Nat. Neurosci. *3*, 1134–1142.

Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., and Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature *401*, 699–703.

Ungerleider, L.G. (1984). Constrasts between the cortiocortical pathways for pattern and spatial vision. In Study Group on Pattern Recognition Mechanisms, C. Chagas, ed. (Vatican City: Pontifical Academy of Sciences).

Ungerleider, L.G., and Mishkin, M. (1979). The striate projection zone in the superior temporal sulcus of Macaca mulatta: location and topographic organization. J. Comp. Neurol. *188*, 347–366.

Vaidya, C.J., Zhao, M., Desmond, J.E., and Gabrieli, J.D. (2002). Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. Neuropsychologia *40*, 2136–2143.

van Santen, P.H., and Sperling, G. (1985). Elaborated reichardt detectors. J. Opt. Soc. Am. A. *2*, 300–321.

van Wezel, R.J., and Britten, K.H. (2002). Multiple uses of visual motion. The case for stability in sensory cortex. Neuroscience *111*, 739–759.

Ward, J., and Simner, J. (2003). Lexical-gustatory synaesthesia: linguistic and conceptual factors. Cognition *89*, 237–261.

Watson, J.D. (1968). The Double Helix: A Personal Account of the Discovery of the Structure of DNA (New York: Atheneum).

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

Wheeler, M.E., Petersen, S.E., and Buckner, R.L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. USA *97*, 11125–11129.

Witthoft, N., and Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. Cortex *42*, 175–183.

Yakovlev, V., Fusi, S., Berman, E., and Zohary, E. (1998). Inter-trial neuronal activity in inferior temporal cortex: a putative vehicle to generate long-term visual associations. Nat. Neurosci. *1*, 310–317.

Yu, N.Y., Yamauchi, T., and Schumacher, J. (2008). Rediscovering symbols: the role of category labels in similarity judgment. J. Cogn. Sci. *9*, 89–100.

Zangaladze, A., Epstein, C.M., Grafton, S.T., and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. Nature *401*, 587–590.

Zhou, Y.D., and Fuster, J.M. (2000). Visuo-tactile cross-modal associations in cortical somatosensory cells. Proc. Natl. Acad. Sci. USA *97*, 9777–9782.