IX

CONSCIOUSNESS

Introduction

GIULIO TONONI

SCIENTISTS ARE finally approaching the relationship between the brain and consciousness in several complementary ways. Over the past 20 years, many have attempted to identify the neural correlates of consciousness (NCC)—the minimal neuronal mechanisms jointly sufficient for any one specific conscious percept. For example, one can examine how brain activity changes when a sensory stimulus is experienced or not, everything else being held as similar as possible. In his chapter, Kreiman illustrates what can be learned through such an approach using binocular rivalry as a case study. He then moves on to consider what we know about the NCC of voluntary actions, which have now been investigated using electroencephalography (EEG), functional MRI (fMRI), and intracranial recordings of neuronal activity.

A complementary strategy is to consider conditions in which consciousness is globally diminished and ask what has changed in the brain. The chapter by Massimini considers the loss of consciousness that occurs in dreamless sleep, in general anesthesia, and after massive brain lesions. Recent work using transcranial magnetic stimulation and high-density EEG shows that level of consciousness can be measured objectively, without requiring behavioral reports, by considering to what extent brain activity is both integrated and differentiated across these diverse conditions and in individual subjects. Demertzi and Laureys examine in detail how a combination of careful clinical tests, passive paradigms that use EEG, positron emission tomography, fMRI, and a combination of transcranial magnetic stimulation and EEG to evaluate brain responses to sensory stimuli, and active paradigms that examine the neural responses to commands in behaviorally unresponsive

patients, can shed light on level of consciousness. Blumenfeld summarizes what we have learned about how and when consciousness is lost in epileptic seizures such as absences, generalized tonic-clonic seizures, and complex partial seizures. This is an area of study that was nearly nonexistent even 10 years ago, but has by now revealed a remarkable diversity of individual differences in the way consciousness changes during a seizure and in underlying mechanisms.

A further strategy is to try and distinguish consciousness and its neural correlates from other aspects of cognition, both conceptually and experimentally. Tsuchiya and Koch make a strong case for a double dissociation between consciousness and attention. They also point out that some putative NCC may actually be related to stimulus processing or response preparation rather than to consciousness itself. This is a conceptual distinction that will certainly promote new experimental paradigms and lead to much-needed refinements. Kouider and Sackur address another classic distinction in consciousness research—that between phenomenal and access consciousness. They argue instead for a unified framework in which consciousness is pragmatically identified with access to cognitive content, but access itself can be more or less partial, and discuss the underlying neural mechanisms. In Blanke's survey, one sees clearly how a combination of neuroimaging with clinical and behavioral studies has already refined our understanding of what awareness of the body contributes to consciousness itself, and of how one can distinguish the awareness of one's body (self-identification), of where it is in space (self-location), and from where one perceives the world (first-person perspective). The chapter by Haggard examines instead the way awareness of agency—the intentional initiation and execution of actions—relates to the underlying brain mechanisms, both in experimental settings and in pathological conditions. He also evaluates what our growing understanding of the control of intention, decision making, action, and movement says about free will and social responsibility.

A final strategy that should complement experimental and clinical studies is the development of a theoretical framework that clarifies what consciousness is, how it can be generated by a physical system, and how it can be measured. This brings us back to the chapter by Massimini, which proposes a theoretically motivated measure of level of consciousness and tests theoretical predictions against data from clinical and experimental neuroscience. Altogether, the section on consciousness shows clearly how much we have learned since the days when research on consciousness was nearly barred from psychology and neuroscience, or banished to the afterthoughts of a textbook, pointing to a future in which consciousness may instead regain its central seat in the cognitive neurosciences.

AUTHOR QUERY FORM

Dear Author,

During the preparation of your manuscript for publication, the questions listedbelow have arisen. Please attend to these matters and return this form with yourproof.

Many thanks for your assistance.

Query	References Query	Remarks
1	AU: Unable to verify citation. Is this the paper "Neural 'ignition': Enhanced activation linked to perceptual awareness in human ventral stream visual cortex," pp. 562–574 in *Neuron* vol. 64 issue 4?	

68 Neural Correlates of Consciousness: Perception and Volition

GABRIEL KREIMAN

ABSTRACT   Consciousness is the result of interactions among neuronal networks in our brains. Although our mechanistic understanding of consciousness remains tentative, the last two decades have seen intense and increasing efforts aimed toward elucidating the neural circuits and spatiotemporal dynamics underlying certain aspects of conscious experience. Experimental paradigms such as those involving bistable perception (wherein a constant stimulus can give rise to different percepts), combined with neurophysiological recordings and computational theories, have provided hints at neuronal signals that correlate with subjective perception. It is also revealing to determine which firing patterns do not correlate with consciousness or which ones correlate with preconscious sensations. Both nonconscious and preconscious signals have been described in perceptual studies as well as in studies of volitional decisions. Discriminating neuronal activity linked to internal perceptual changes in the absence of external changes provides empirical constraints and initial glimpses into how signaling cascades in the brain can give rise to consciousness.

The phenomenological feeling of *consciousness* is central to our moment-to-moment experiences. What we see, hear, feel, or reminisce forms the content of conscious sensations. Most scientists would agree that consciousness is ultimately encoded and orchestrated by the activity of neurons in the brain, but the elucidation of where, when, how, and why neural ensemble dynamics lead to consciousness remains a deep and fundamental mystery.

Our brains are physical entities. In many ways, the chemical components of brains are similar to those encountered in plants and trees and distinct from those in chairs and tables. An important distinguishing feature of the organization of the chemical components in our brains is the existence of interconnected neurons—yet interconnected neural circuits are also present in *Caenorhabditis elegans* and related worms. Whether the neuronal circuits in worms lead to conscious sensations somewhat akin to the ones experienced by humans is not clear. Out of the bewildering complexity (Tononi, 2004; Tononi & Edelman, 1998) that results from vast numbers of interconnected neurons arises the feelings of love and pain and the capacity to plan our future and prove new mathematical theorems. An elegant theoretical framework to define when, how, and why a given circuit of neurons can give rise to consciousness, whereas an apparently similar circuit or ensemble firing pattern does not has been proposed by Tononi. The theory proposes that "the level of consciousness of a physical system is related to the repertoire of causal states (*information*) available to the system as a whole (*integration*)" (2004, p. 253, emphasis in original). This theoretical framework can help distinguish why some physical systems may experience consciousness while others do not, which species can experience consciousness, why different neuronal circuits show distinct contributions to consciousness, and even why different activity patterns in the same neuronal circuit may show different correlations with conscious experience.

To formulate a principled strategy to begin to investigate the problem of which aspects of brain function correlate with specific contents of conscious sensations and which do not, Crick and Koch, 1990, defined the neural correlates of consciousness (NCC). The NCC represents a "minimal set of neuronal events and mechanisms sufficient for a specific conscious percept" (Koch, 2005). Over the last two decades, growing enthusiasm and blossoming efforts have capitalized on tools to interrogate brain activity at the circuit level in order to take initial steps toward investigating the NCC.

There are multiple fascinating aspects of conscious experience that need to be explained. We focus the discussion here on two of them, sensory perception and volition. With some degree of approximation, we can think of these as representative of brain input and output, respectively. We aim to provide an overview of the advances, difficulties, experiments, and theories that have shaped the discussions around the relationship between neural activity and conscious perception in a few specific instances, which we cover more thoroughly at the expense of other themes in the field. In doing so, we are not doing justice to a large body of heroic efforts. We refer the reader to several reviews (see Baars, 1989; Blanke, 2012; Blumenfeld, 2011; Cotterill, 2001; Crick, Koch, Kreiman, & Fried, 2004;

Dehaene & Changeux, 2011; Jackendoff, 1987; Kim & Blake, 2005; Koch, 2005; Kouider & Dehaene, 2007; Logothetis, 1998; Metzinger, 2000; Rees, Kreiman, & Koch, 2002; Posner, 1994; Searle, 1998, 2005; Singer, 1998; Tononi, 2005; Tononi & Koch, 2011).

Articulating a set of requisites to guide the search for neural correlates of consciousness

It is helpful to ponder what it would take to *understand* the neural circuits that orchestrate a given aspect of consciousness. For the sake of argument, let us consider a given putative NCC for a certain percept *P*. The putative NCC could take the form of activity from a particular subset of neurons, in a specific brain area, at a certain time point and with a structured firing pattern (Koch, 2005). We discuss below what we do and do not know about those neurons, areas, times, and patterns; for the moment, we refer to these as putative NCC for short. We are also not specifying what the percept *P* is. As an example and to frame the discussion, consider the perception of a given face. What sort of evidence would reinforce our belief in a certain activity pattern to be part of the NCC? To a reasonable first-order approximation, we can borrow from the articulation of requisites to correlate neurophysiological responses with behavior (Parker & Newsome, 1998):

(1) Activity in the putative *NCC* should correlate with the percept *P*. The putative NCC should be elicited by any situation that leads to percept *P* (e.g., presentation of stimulus *P*, visual imagery of *P*, dreaming about *P*, perception of *P* during a bistable perception or a masking experiment, etc.).

(2) Conversely, the putative NCC should *not* be elicited by situations that do *not* lead to percept P regardless of the presence or absence of *P* (e.g., presentation of a different percept *Q*, perception of *Q* during a bistable perception test even when the stimulus that would otherwise lead to *P* may be present on the screen, etc.). Note that the same neurons can still be involved in representing *P* and *not P* as long as the NCC as a whole is correlated only with *P*.

(3) No NCC, no consciousness. If we somehow abolish the activity related to the NCC, we expect to observe a concomitant impairment in consciousness. Abolishing activity could take the form of a deliberate lesion in animal models, a certain neurological condition in patients, pharmacological interventions, optogenetic interventions, and so on.

(4) Activation of the NCC should lead to the percept *P*. This activation could be caused by presentation of an external stimulus, electrical stimulation, optogenetic techniques, or any other means. According to this proposal, how the NCC is activated is not relevant; detection of the NCC implies eliciting percept *P*.

This set of requirements implies that we should be able to "read out" percepts *P* by interrogating neural activity patterns. We should only be able to detect NCC whenever the subject experiences *P*. Detecting the NCC in the absence of percept *P* or vice versa would violate (1–4). Violations of (1–4) suggest that we have not fully captured the NCC. Admittedly, this is a tall order.

Consciousness and perception

We focus here on the quest to characterize the neuronal circuits involved in eliciting visual percepts because there has been more work in this domain. Several aspects of the discussion here will also be applicable to other sensory modalities, and perhaps even to other aspects of conscious experience. Following on the definition outlined above for the NCC, we seek to describe the correlates of a specific visual percept (say, seeing a particular face) in terms of specific brain areas, neuronal types, ensemble activity patterns, and timing.

Visual inputs are rapidly transformed into conscious perception. Visual signals impinging on the retinae go through a cascade of processes that lead to perception in a fraction of a second (Blumberg & Kreiman, 2010; Hung, Kreiman, Poggio, & DiCarlo, 2005; Lamme & Roelfsema, 2000; Liu, Agam, Madsen, & Kreiman, 2009; Logothetis & Sheinberg, 1996; Schmolesky et al., 1998; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2002). When exactly perception arises along this rapid and approximately hierarchical sequence of steps remains unclear (Baars, 1989; Fisch et al., 2009; Koch, 2005). A rough outline of the brain areas and pathways involved in this cascade has been documented through anatomical and physiological studies, particularly in monkeys (Bullier, 2001; Felleman & Van Essen, 1991). Significant efforts are currently being directed toward describing these cascades in rodents (Wang & Burkhalter, 2007). Due to the difficulties inherent in anatomical mapping and neurophysiological recordings, much less is known about the architecture of the human visual system.

Information from the retina is needed to trigger a visual percept, yet several arguments suggest that we are not *directly* aware of the activity of neurons in our retinae. Investigators have argued that one of the requirements for the NCC is that of an *explicit* representation. A set of neurons is said to explicitly represent a certain aspect of information *y* if an accurate estimate of *y* can be obtained by a one-layer network acting on

the output of those neurons (Koch, 2005). Information about local contrast changes is explicitly encoded in the retina. Information about the presence or absence of a face is *not* explicitly encoded in the retina (even though the information can be decoded from retinal activity through a cascade of multiple operations). Several aspects of the properties of retinal photoreceptors argue against their explicit representation of conscious information: (1) retinal photoreceptors can follow rapid spatiotemporal changes that we are not aware of (e.g., a monitor's refresh rate); (2) there are no photoreceptors at the blind spot, but we do not see a hole in the corresponding location in an image; (3) there is no feedback to the retina from other parts of the brain; (4) visual percepts can be elicited through imagery or dreaming in the absence of retinal activation; (5) the type of visual information that we are typically conscious of cannot be decoded in a single step from the retina; (vi) retinal photoreceptors show major activity changes in response to multiple small eye movements that we constantly make and that we are completely unaware of. Similar comments can be made about the output cells in the retina (retinal ganglion cells) as well as their target cells in the thalamus (within the lateral geniculate nucleus, or LGN). Several, but not all, of these arguments are also pertinent in the case of primary visual cortex (V1; Crick & Koch, 1995; Leopold, 2012).

EXPERIMENTAL PARADIGMS   As we ascend through the visual hierarchy, neuronal responses become more sophisticated, and the neurons' firing preferences gradually begin to acquire some of the properties that we associate with our subjective visual percepts (Connor, Brincat, & Pasupathy, 2007; Logothetis & Sheinberg, 1996; Tanaka, 1996). The arguments used above to rule out a role for the retina or LGN in conscious perception lose their weight, and there is a need to use more sophisticated experimental paradigms in an attempt to dissociate perception and sensory inputs. Common to many of these empirical approaches is to consider two conditions in which the external inputs are identical, where other internal variables are as similar as possible and the percepts are different (Kim & Blake, 2005). Consider the famous Necker cube (figure 68.1), which can be perceived in two different configurations. Let us assume that fixation, attention, arousal, and other variables are identical in two trials in which the viewer perceives the two different possible configurations (figure 68.1B, C). By assumption (identical fixation), retinal activity should be similar in both trials (except that neurons are capricious and may not show the exact same spiking pattern in two seemingly identical trials, not even in the retina; see, e.g., Van Stevenick, Lewen, Strong, Koberle, & Bialek, 1997). Yet, for the percept to be distinct in the two conditions, something *must* be different in the brain.
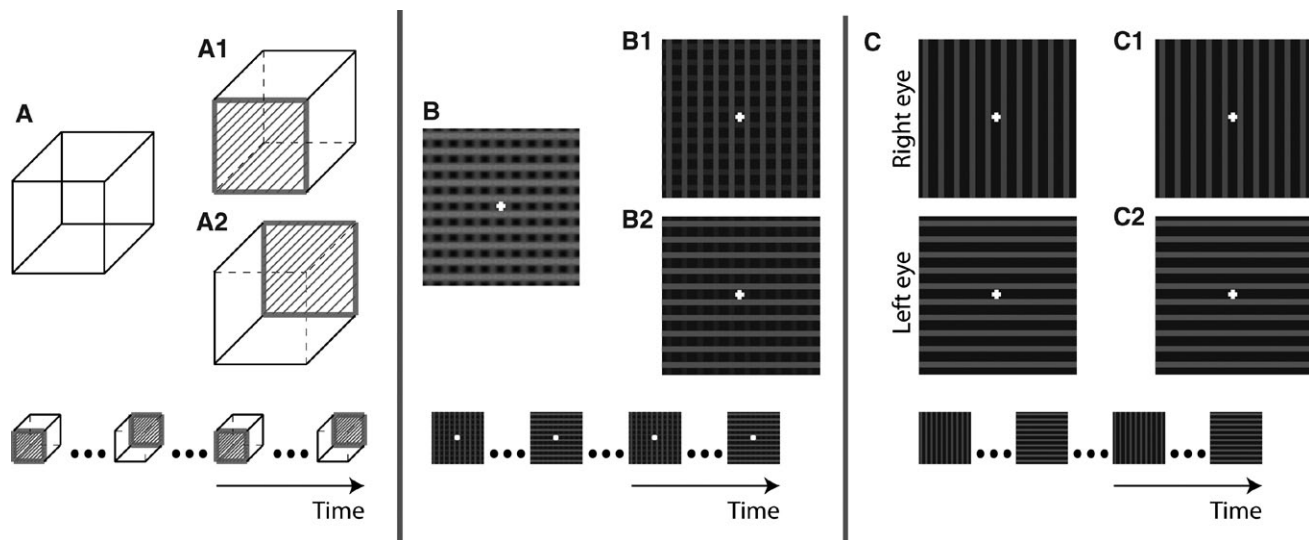

FIGURE 68.1   Example of bistable percepts. A. The Necker cube can be interpreted in two possible configurations (A1 and A2). Upon viewing the Necker cube, subjects can voluntarily switch from one configuration to the other (bottom). B. Monocular rivalry. Upon fixating on the grid of horizontal red lines and vertical blue lines, subjects report a percept that alternates between predominantly blue lines (B1) and predominantly red lines (B2). C. Binocular rivalry. When presenting different stimuli to each eye, subjects report a percept that alternates between the two inputs (C1, C2) in a seemingly random fashion. During binocular rivalry, it is difficult to voluntarily switch between the two percepts. (See color plate 59.)

We know almost nothing about what changes in cortex when perception switches while viewing the Necker cube. A few other paradigms have proven more fruitful for investigation. We focus below on discussing bistable percepts, including binocular rivalry (Blake & Logothetis, 2002), flash suppression (Wilke, Logothetis, & Leopold, 2003; Wolfe, 1984), and structure from motion (Bradley, Chang, & Andersen, 1998). Several other experimental paradigms have been examined, including near-threshold or degraded stimulation, visual crowding, visual masking (Kouider & Dehaene, 2007; Macknik, 2006), and different forms of transiently induced "blindness" such as motion-induced blindness (Bonneh, Cooperman, & Sagi, 2001), inattentional blindness (Rees, Russell, Frith, & Driver, 1999), change blindness (O'Regan, Rensink, & Clark, 1999; Simons & Rensink, 2005), and the attentional blink (Raymond, Shapiro, & Arnell, 1992; Sergent, Baillet, & Dehaene, 2005).

Neurophysiological Changes Correlated with Bistable Percepts   Several experimental paradigms exploit the observation that our brains impose a single interpretation on the outside world at any given time (figure 68.1). In cases where the visual input is ambiguous and is compatible with two (or more) different possible interpretations, our brains settle on one of them. Under many circumstances, this decision is only transient, and our brains may toggle between one interpretation and the other. The Necker cube discussed above constitutes one such case. In ambiguous depth from motion stimuli, a set of dots is set to rotate in such a way that the image can be interpreted to rotate in either of two possible directions (Bradley et al., 1998; Siegel & Andersen, 1988). Another example is the phenomenon of monocular rivalry (Campbell & Howell, 1972; figure 68.1B). Consider a grid of horizontal red bars and vertical blue bars; perception mostly alternates between horizontal red bars or vertical blue bars at the expense of an interpretation containing both horizontal and vertical bars. Rivalry is much stronger when elicited in a binocular fashion. Binocular rivalry arises when a stimulus R is shown to the right eye and a different stimulus L is shown to the left eye (figure 68.1C). Typically, perception alternates between R and L in a seemingly random fashion. The rate of alternation depends on the characteristics of R and L. In a trivial case, if L is blank, R completely dominates perception (e.g., if you cover the left eye). If R and L are very large (say more than 5–6 degrees of visual angle), in addition to perceptual alterations, parts of the image are perceived as a mixture of R and L (known as piecemeal rivalry). Stimuli with high contrast typically dominate

perception. The psychophysics of binocular rivalry and the variables that govern dominance and alterations have been carefully studied (Alais & Blake, 2005).

Which neuronal changes accompany the perceptual changes evident upon viewing bistable stimuli? Access to neuronal responses in the human brain is rather limited (Engel, Moll, Fried, & Ojemann, 2005; Kreiman, 2007; Mukamel & Fried, 2012); most neurophysiological studies of binocular rivalry to date have focused on examining responses in macaque monkeys. Before delving into the neurophysiological responses in macaque monkeys, it is important to ask whether monkeys perceive bistable stimuli in the same way that humans do. Because it is difficult to access subjective experiences in monkeys, investigators have used ingenious controls and behavioral measurements to evaluate perceptual alternations while monkeys view bistable stimuli. For example, embedded within a binocular rivalry experiment are catch trials consisting of monocular stimulus presentation where there is no ambiguity about what the behavioral responses should be. Throughout these studies, the answer is consistently that monkeys seem to perceive changes in the interpretation of bistable stimuli in the same way that humans do (Libedinsky, Savage, & Livingstone, 2009; Macknik & Livingstone, 1998; Myerson, Miezin, & Allman, 1981; Sheinberg & Logothetis, 1997; Siegel & Andersen, 1988).

In binocular rivalry experiments, ascending through the visual hierarchy, there is a progressive increase in the proportion of neurons that show a correlation with perception (Leopold & Logothetis, 1999). While binocularly presenting orthogonal drifting gratings to fixating monkeys, none of the neurons in the LGN showed any evidence of alternations indicative of binocular rivalry (but it should be noted that monkeys were passively fixating and not reporting their perceptual alterations; Lehky & Maunsell, 1996). Similar conclusions were reached in a motion-induced perceptual suppression experiment wherein the activity of LGN neurons did not correlate with subjective perception, but pulvinar neurons did (Wilke, Mueller, & Leopold, 2009).

Logothetis and colleagues conducted an elegant series of experiments in monkeys that were trained to report their alternating percepts during binocular rivalry by pressing and holding levers. In primary visual cortex, only about 18% of the neurons showed perceptual modulation (Leopold, 2012; Leopold & Logothetis, 1996). This small percentage of neurons showing perceptual modulation in primary visual cortex is consistent with the idea that perceptual alternations arise from competitive mechanisms beyond V1 (Leopold & Logothetis, 1996, 1999; Logothetis, Leopold, &

Sheinberg, 1996; Wilke, Leopold, & Logothetis, 2002; see, however, Blake, 1989; Blake & Logothetis, 2002). In areas V4 and MT, 38% and 43% of the neurons showed perceptual modulation (Leopold & Logothetis, 1996; Logothetis & Schall, 1989). Intriguingly, in area MT, half of the neurons increased their firing rate when their preferred stimulus was perceived, while the other half showed the reverse pattern; that is, they increased their activity when their preferred stimulus was suppressed. The correlation between subjective perception and the activity of some, but not all, MT neurons was also apparent in a structure-from-motion task during conditions in which the input was ambiguous (Bradley et al., 1998). The bewildering variety of neurons that show different degrees of correlation with perception is perhaps a reflection of the intricacy of connectivity patterns in neural circuits. The advent of techniques based on optogenetics may enable the possibility of interrogating (and manipulating) different layers and different types of neurons. Analysis of circuitry at the level of "brain areas" may prove insufficient to uncover the NCC. The neurons in V1, V4, and MT that show stronger correlation with subjective perception may have distinct properties (in terms of their location, inputs, and outputs, and how they interact with other neurons) from their neighboring neurons whose firing is oblivious or anticorrelated with the percepts.

In the highest echelons of the ventral visual stream, 90% of the neurons in the inferior temporal cortex and the superior temporal sulcus showed a correlation between their firing rate and the subjective report of the neuron's preferred stimulus (Sheinberg & Logothetis, 1997). In a variant of binocular rivalry–denominated flash suppression, a stimulus is flashed monocularly followed by presentation of another stimulus to the other eye (Wolfe, 1984). The flashed stimulus dominates perception, even though the initial stimulus remains present. In parallel to the binocular rivalry results, neurons in the macaque inferior temporal cortex and superior temporal sulcus show a strong correlation with the percept (Sheinberg & Logothetis, 1997). Similarly, most neurons in the human medial temporal lobe also show a correlation with subjective perception during flash suppression (Kreiman, Fried, & Koch, 2002).

ATTENTION, IMAGING, AND STIMULUS CONFIGURATION DEPENDENCIES In contrast to neurophysiological recordings in macaque V1, several human functional neuroimaging studies have suggested that activity in primary visual cortex also correlates with subjective perception (Haynes, Deichmann, & Rees, 2005; Polonsky, Blake, Braun, & Heeger, 2000; Tong & Engel, 2001; see also similar claims in the LGN: Wunderlich, Schneider,

& Kastner, 2005). These discrepancies between blood flow measurements and neurophysiological signals have also been observed in other experiments (Logothetis, 2002; Posner & Gilbert, 1999; Sirotin & Das, 2009; Watanabe et al., 2011). An elegant study by Maier and colleagues examined blood oxygen level–dependent (BOLD) functional MRI signals, local field potentials (LFPs), and spiking responses in primary visual cortex during binocular rivalry (Maier et al., 2008). During physical removal of a stimulus, BOLD signals, LFPs, and spiking responses agreed with each other. However, during perceptual suppression, there were small changes in BOLD signals and LFP responses that were not observed at the spiking level. These observations highlight the notion that blood flow and spiking signals measure different aspects of brain function, and the nature of the relationship between these signals may prove to be important to gain further mechanistic insights into the circuitry involved in conscious perception (Leopold, 2012). A potential explanation of these results is that blood flow signals in the LGN and in primary visual cortex reflect feedback modulation from higher visual areas after perceptual rivalry has been resolved, consistent with the notion that V1 does not have a direct role in visual awareness (Crick & Koch, 1995; Leopold, 2012).

Strong modulation in blood flow signals in early visual areas in the absence of concomitant strong modulation at the spike level has also been observed in studies of attentional modulation (Kastner & Ungerleider, 2000; Reynolds & Chelazzi, 2004). It has been argued that modulatory changes observed during binocular rivalry experiments could reflect attentional fluctuations (Macknik & Martinez-Conde, 2009). The extent to which attention and awareness can be dissociated has been a matter of significant debate (Dehaene & Changeux, 2011; Koch & Tsuchiya, 2012; Posner, 1994). Under most everyday circumstances, attention and awareness go hand in hand, yet psychophysical experiments suggest that it is possible to attend to a stimulus even in the absence of awareness (e.g., Koch & Tsuchiya, 2012; van Boxtel, Tsuchiya, & Koch, 2010). Furthermore, a functional imaging study showed that blood flow signals in V1 are modulated by attention but not by changes in awareness (Watanabe et al., 2011). These results are consistent with recent neurophysiological recordings that lend further support to distinct signals giving rise to attentional modulation and awareness (Maier, 2012).

Other factors beyond attention can also influence the relationship between neuronal activity and awareness. The discussion in the previous section described differences in the correlations between subjective perception

and neuronal responses across areas. In a particularly intriguing study, Maier and colleagues asked whether, for a given individual neuron, this correlation depended on the details of the stimulus configuration (Maier, Logothetis, & Leopold, 2007). The authors recorded neurons in area MT in the macaque monkey during binocular rivalry flash suppression. For a given stimulus, the results were consistent with earlier recordings (Logothetis & Schall, 1989). However, when the authors changed the stimulus configuration (e.g., different motion directions or drifting gratings instead of random dots), the extent to which the neuron signaled subjective percepts was significantly altered. For example, a neuron may show changes that correlate with perception when the right eye sees a left moving grating and the left eye sees a downward moving grating, but not when the stimulus in the left eye is an upward-moving grating. These idiosyncratic correlations force us to revisit the notion that the NCC for a particular percept may invoke a fixed set of neurons. These puzzling observations further suggest the urgent need to relate theoretical ideas of how consciousness arises to neurophysiological recordings (Tononi, 2004).

Consciousness and volition

Consciousness is clearly not restricted to awareness of sensory events. At the other end of the sensory/motor spectrum, volitional actions carry a strong sensation of ownership (e.g., "I want to raise my hand"). Before discussing the neuronal manifestations that correlate with our awareness of intention to act, we need to discuss the controversial notion that our volition and intentions are dictated by neurons.

VOLITION AND FREE WILL   Studying the neural signatures that correlate with consciousness about volitional actions may have important implications in settling the age-old questions about "free will." Intuitively defined, the word "free" associated with will implies that, for most of our actions, we experience the strong subjective feeling that we could have opted to act otherwise (see also Haggard, chapter 75 in this volume). If we are asked to pick between a blue pen and a black pen and we pick the black one, it seems that we might as well have picked the blue one.

The extent to which free will is truly free or merely an illusion has been a matter of debate for millennia, with strong advocates on both sides (for an overview, see Haggard, 2008; Heisenberg, 2009). At least two main cautionary notes should be discussed here. The first involves the distinction between "determinism" and "chaos." A system is said to be deterministic if its future state is entirely defined by the initial conditions (and any external forces). A system is said to be chaotic if it displays extreme sensitivity to initial conditions (minuscule differences in the initial state can lead to widely different future states, as in the famous parable of the butterfly effect; Devaney, 2003). These two words are *not* antagonistic. A system can be both deterministic and chaotic. Consider the act of flipping a coin: obtaining heads or tails can depend on a lot of factors (exact initial angle, torque, speed, wind, properties of the surface where the coin lands, etc.) but the physics underlying the problem are well defined and purely deterministic. Nobody would claim that the coin "wanted" to land on heads. The other cautionary note has to do with *computability*. There are many reasons why certain functions may be difficult (or impossible) to compute: there are problems that are not computable (Garey & Johnson, 1979), and there are computable problems that require unrealistic computational resources or data that we do not have access to. Chaos and questions about computational resources may make it very difficult and perhaps impractical to make predictions in certain systems but neither speaks against determinism or in favor of free will.

NEUROPHYSIOLOGICAL CORRELATES OF VOLITIONAL DECISIONS   The majority of studies about decision making have focused on situations in which a cue indicates the target behavior. The cue may be noisy, interpreting the cue may require training and memory, and the relationship between the cue and reward could be a probabilistic one. Yet common to many of these experiments are a temporal trigger and an incentive to choose one action versus another. In contrast, the study of volitional decisions requires situations where different actions are equally likely and attractive.

A few studies in macaque monkeys have examined neurophysiological responses in the parietal and frontal cortex while monkeys performed volitional decisions (Maimon & Assad, 2006a, 2006b; Okano & Tanji, 1987; Romo & Schultz, 1992). Single neurons in the lateral intraparietal area, cortical area 5, the basal ganglia, and frontomotor areas exhibited gradual increases in firing rate during execution of volitional arm movements. Some of the neurons showed activation during both visually triggered movements and proactive movements. However, the slow ramp in firing rates was characteristic of internally generated movements only.

This slow increase in activity is reminiscent of gradual changes in scalp electroencephalographic signals in the human brain during execution of volitional movements (Brass & Haggard, 2008; Deecke et al., 1987; Haggard, 2008). In a variant of this type of experiment,

Libet asked subjects to tap their index finger at will and also report the time of their intention to act based on an analog clock present on the screen during the experiment (Libet, 1985; Libet, Gleason, Wright, & Pearl, 1983). These experiments revealed that the averaged scalp electroencephalographic signals preceded the conscious intention of the urge to move by several hundreds of milliseconds. The interpretation of these experiments has been the subject of much debate in the field (e.g., see Libet, 1985, 2002, and discussions in the same issue).

In a recent study, Fried and colleagues took advantage of a rare clinical opportunity to record from >1,000 neurons in the human frontal and temporal lobes when subjects performed willed action and reported the time of volition onset, as in the Libet experiment. Consistent with earlier studies using noninvasive methods, they found evidence at the single neuron level in humans for an anatomically localized early frontal cortex signal that preceded conscious will. Over a time period of ~1,500 msec prior to the awareness of will, an increasing number of neurons in two specific brain regions, the supplementary motor area and anterior cingulate cortex, were progressively recruited. The subjectively reported onset of volition could be accurately predicted on a single trial basis based on neural activity in the supplementary motor area well before the subject's awareness. Based on these findings, the authors proposed a computational model and a biophysically plausible mechanism for the emergence of conscious will in humans based on progressive recruitment of neuronal ensembles in frontal cortex until a threshold is crossed. The model is consistent with the notion that the all-or-none nature of consciousness is the result of gradual accumulation reaching a threshold (Crick & Koch, 2003). Furthermore, in another study, Fried showed that electrical stimulation in the human supplementary motor area triggered an "urge" to perform motor actions (Fried et al., 1991).

Outlook

The fundamental problem of understanding how neuronal circuits give rise to conscious sensations has risen from nebulous beginnings and debates to become a major effort in cognitive neuroscience. Progress in the field should be interpreted with cautious optimism. Caution is important because this is undoubtedly a difficult problem, and current theories are as diverse as they could be. And yet there is optimism and steady progress. The last two decades have seen the blossoming of a young generation of energetic and heroic investigators who have dared to ask difficult questions and

approach them with a new arsenal of tools that is making rapid strides in elucidating other aspects of cognition, including multielectrode arrays, computational modeling, microstimulation, optogenetics, and so on.

Advances and controversies in trying to correlate neural signals and conscious perception have led to a theoretical framework that provides quantitative definitions of how neuronal interactions could lead to consciousness (Tononi, 2004). Additionally, initial but significant steps have been made toward better defining questions about conscious processing (e.g., dissociating pure attentional effects from conscious perception; Koch and Tsuchiya, 2012), toward sharpening experimental tools (e.g., noticing that blood flow signals may not reveal underlying spiking; Maier, 2008), and toward a richness of experimental paradigms and approaches (Dehaene & Changeux, 2011).

There is currently significant excitement in cognitive neuroscience with the advent of tools that enable the manipulation of circuits at unprecedented resolution (e.g., Han et al., 2009). These tools open the doors to exciting and promising opportunities to attempt to transiently inactivate and also directly stimulate local circuits, and thus bias subjective decisions in ways that have not been possible before. While it is anyone's guess whether the quest for the NCC will be resolved in the near future, there is no question that we should expect fascinating surprises and novel insights in the community's efforts to elucidate how physical systems lead to consciousness.

REFERENCES

Alais, D., & Blake, R. (2005). *Binocular rivalry.* Cambridge, MA: MIT Press.

Baars, B. (1989). *A cognitive theory of consciousness.* Cambridge, UK: Cambridge University Press.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychol Rev, 96,* 145–167.

Blake, R., & Logothetis, N. (2002). Visual competition. *Nat Rev Neurosci, 3,* 13–21.

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nat Rev Neurosci, 13,* 556–571.

Blumberg, J., & Kreiman, G. (2010). How cortical neurons help us see: Visual recognition in the human brain. *J Clin Invest, 120,* 3054–3063.

Blumenfeld, H. (2011). Epilepsy and the consciousness system: Transient vegetative state? *Neurol Clin, 29,* 801–823.

Bonneh, Y., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature, 411,* 798–801.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of 3D structure from motion by primate area MT neurons. *Nature, 392,* 714–717.

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *Neuroscientist, 14,* 319–325.

Bullier, J. (2001). Integrated model of visual processing. *Brain Res Brain Res Rev, 36,* 96–107.

Campbell, F. W., & Howell, E. R. (1972). Monocular alternation: A method for the investigation of pattern vision. *J Physiol, 225,* 19–21.

Connor, C. E., Brincat, S. L., & Pasupathy, A. (2007). Transformation of shape information in the ventral pathway. *Curr Opin Neurobiol, 17,* 140–147.

Cotterill, R. (2001). Evolution, cognition and consciousness. *J Consciousness Stud, 8,* 3–17.

Crick, F., & Koch, C. (1990). Some reflections on visual awareness. *Cold Spring Harbor Symposia on Quantitative Biology, 55,* 953–962.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature, 375,* 121–123.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci, 6,* 119–126.

Crick, F., Koch, C., Kreiman, G., & Fried, I. (2004). Consciousness and neurosurgery. *Neurosurgery, 55,* 273–282.

Deecke, I., Lang, W., Heller, H., Hufnagl, M., & Kornhuber, H. (1987). Bereitschaftspotential in patients with unilateral lesions of the supplementary motor area. *J Neurol, 50,* 4.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70,* 200–227.

Devaney, R. L. (2003). *An introduction to chaotic dynamical systems.* Boulder, CO: Westview.

Engel, A. K., Moll, C. K., Fried, I., & Ojemann, G. A. (2005). Invasive recordings from the human brain: Clinical insights and beyond. *Nat Rev Neurosci, 6,* 35–47.

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

Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., … Malach, R. (2009). Neuronal "ignition": Non-linear activation at the threshold of perceptual awareness in human ventral stream visual cortex. *Neuron.*

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., & Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *J Neurosci, 11,* 3656–3666.

Garey, M., & Johnson, D. (1979). *Computers and intractability: A guide to the theory of NP-completeness.* New York, NY: Freeman.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nat Rev Neurosci, 9,* 934–946.

Han, X., Qian, X., Bernstein, J. G., Zhou, H. H., Franzesi, G. T., Stern, P., … Boyden, E. S. (2009). Millisecond-timescale optical control of neural dynamics in the nonhuman primate brain. *Neuron, 62,* 191–198.

Haynes, J. D., Deichmann, R., & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature, 438,* 496–499.

Heisenberg, M. (2009). Is free will an illusion? *Nature, 459,* 2.

Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast read-out of object identity from macaque inferior temporal cortex. *Science, 310,* 863–866.

Jackendoff, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci, 23,* 315–341.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible "invisible." *Trends Cogn Sci, 9,* 381–388.

Koch, C. (2005). *The quest for consciousness.* Los Angeles, CA: Roberts.

Koch, C., & Tsuchiya, N. (2012). Attention and consciousness: Related yet different. *Trends Cogn Sci, 16,* 103–105.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philos Trans R Soc Lond B Biol Sci, 362,* 857–875.

Kreiman, G. (2007). Single neuron approaches to human vision and memories. *Curr Opin Neurobiol, 17,* 471–475.

Kreiman, G., Fried, I., & Koch, C. (2002). Single neuron correlates of subjective vision in the human medial temporal lobe. *Proc Natl Acad Sci USA, 99,* 8378–8383.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci, 23,* 571–579.

Lehky, S. R., & Maunsell, J. H. R. (1996). No binocular rivalry in the LGN of alert monkeys. *Vis Res, 36,* 1225–1234.

Leopold, D. A. (2012). Primary visual cortex: Awareness and blindsight. *Annu Rev Neurosci, 35,* 91–109.

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature, 379,* 549–553.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends Cogn Sci, 3,* 254–264.

Libedinsky, C., Savage, T., & Livingstone, M. (2009). Perceptual and physiological evidence for a role for early visual areas in motion-induced blindness. *J Vis, 9*(1), 1–10.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav Brain Sci, 8,* 529–566.

Libet, B. (2002). The timing of mental events: Libet's experimental findings and their implications. *Conscious Cogn, 11,* 291–299.

Libet, B., Gleason, C., Wright, E., & Pearl, D. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain, 106,* 623–642.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron, 62,* 281–290.

Logothetis, N. K. (1998). Single units and conscious vision. *Philos Trans R Soc Lond B Biol Sci, 353,* 1801–1818.

Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci, 357,* 1003–1037.

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivaling during binocular rivalry? *Nature, 380,* 621–624.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245,* 761–763.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annu Rev Neurosci, 19,* 577–621.

Macknik, S. (2006). Visual masking approaches to visual awareness. *Prog Brain Res, 155,* 177–215.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci*, *1*, 144–149.

Macknik, S. L., & Martinez-Conde, S. (2009). The role of feedback in visual attention and awareness. *Consciousness*, 1163–1177.

Maier, A. (2012). The cortical microcircuitry of conscious perception and selective attention. Paper presented the Annual Meeting of the Society for Neuroscience, New Orleans, LA.

Maier, A., Logothetis, N., & Leopold, D. (2007). Context-dependent perceptual modulation of single neurons in primate visual cortex. *Proc Natl Acad Sci USA*, *104*, 5620–5625.

Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F. Q., & Leopold, D. A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci*, *11*, 1193–1200.

Maimon, G., & Assad, J. A. (2006a). Parietal area 5 and the initiation of self-timed movements versus simple reactions. *J Neurosci*, *26*, 2487–2498.

Maimon, G., & Assad, J. A. (2006b). A cognitive signal for the proactive timing of action in macaque LIP. *Nat Neurosci*, *9*, 948–955.

Metzinger, T. (2000). *Neural correlates of consciousness: Empirical and conceptual questions*. Cambridge, MA: MIT Press.

Mukamel, R., & Fried, I. (2012). Human intracranial recordings and cognitive neuroscience. *Annu Rev Psychol*, *63*, 511–537.

Myerson, J., Miezin, F., & Allman, J. (1981). Binocular rivalry in macaque monkeys and humans: A comparative study in perception. *Behav Anal Lett*, *1*, 149–159.

Okano, K., & Tanji, J. (1987). Neuronal activities in the primate motor fields of the agranular frontal cortex preceding visually triggered and self-paced movement. *Exp Brain Res*, *66*, 155–166.

O'Regan, J., Rensink, R., & Clark, J. (1999). Change-blindness as a result of mudsplashes. *Nature*, *398*, 34.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annu Rev Neurosci*, *21*, 227–277.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat Neurosci*, *3*, 1153–1159.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proc Natl Acad Sci USA*, *91*, 7398–7403.

Posner, M. I., & Gilbert, C. D. (1999). Attention and primary visual cortex. *Proc Natl Acad Sci USA*, *96*, 2585–2587.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *J Exp Psychol Hum Percept Perform*, *18*, 849–860.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nat Rev Neurosci*, *3*, 261–270.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science*, *286*, 2504–2507.

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

Romo, R., & Schultz, W. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements. III. Neuronal activity in the supplementary motor area. *Exp Brain Res*, *91*(3), 396–407.

Schmolesky, M., Wang, Y., Hanes, D., Thompson, K., Leutgeb, S., Schall, J., & Leventhal, A. (1998). Signal timing across the macaque visual system. *J Neurophysiol*, *79*, 3272–3278.

Searle, J. (1998). How to study consciousness scientifically. *Philos Trans R Soc Lond B Biol Sci*, *353*, 1935–1942.

Searle, J. (2005). Consciousness: What we still don't know [Review of the book *The Quest for Consciousness* by C. Koch]. *The New York Review of Books*. Available from http://www.nybooks.com/articles/archives/2005/jan/13/consciousness-what-we-still-dont-know/.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat Neurosci*, *8*, 1391–1400.

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal areas in perceptual organization. *Proc Natl Acad Sci USA*, *94*, 3408–3413.

Siegel, R. M., & Andersen, R. A. (1988). Perception of three-dimensional structure from motion in monkey and man. *Nature*, *331*, 259–261.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends Cogn Sci*, *9*, 16–20.

Singer, W. (1998). Consciousness and the structure of neuronal representations. *Philos Trans R Soc Lond B Biol Sci*, *353*, 1829–1840.

Sirotin, Y. B., & Das, A. (2009). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature*, *457*, 475–479.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annu Rev Neurosci*, *19*, 109–139.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

Tong, F., & Engel, S. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, *411*, 195–199.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci*, *5*, 42.

Tononi, G. (2005). Consciousness, information integration, and the brain. *Prog Brain Res*, *150*, 109–126.

Tononi, G., & Edelman, G. (1998). Consciousness and complexity. *Science*, *282*(5395), 1846–1851.

Tononi, G., & Koch, C. (2011). The neural correlates of consciousness: An update. *Ann NY Acad Sci*, *1124*, 239–261.

Van Boxtel, J. J., Tsuchiya, N., & Koch, C. (2010). Opposing effects of attention and consciousness on afterimages. *Proc Natl Acad Sci USA*, *107*, 8883–8888.

Van Steveninck, R., Lewen, G. D., Strong, S. P., Koberle, R., & Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science*, *275*, 1805–1808.

VanRullen, R., & Thorpe, S. (2002). Surfing a spike wave down the ventral stream. *Vis Res*, *42*, 2593–2615.

Wang, Q., & Burkhalter, A. (2007). Area map of mouse visual cortex. *J Comp Neurol*, *502*, 339–357.

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothesis, N. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science*, *334*(6057), 829–831.

WILKE, M., LEOPOLD, D., & LOGOTHETIS, N. (2002). Flash suppression without interocular conflict. *Soc Neurosci Abstracts, 161.15.*

WILKE, M., LOGOTHETIS, N. K., & LEOPOLD, D. A. (2003). Generalized flash suppression of salient visual targets. *Neuron, 39,* 1043–1052.

WILKE, M., MUELLER, K. M., & LEOPOLD, D. A. (2009). Neural activity in the visual thalamus reflects perceptual suppression. *Proc Natl Acad Sci USA, 106,* 9465–9470.

WOLFE, J. (1984). Reversing ocular dominance and suppression in a single flash. *Vis Res, 24,* 471–478.

WUNDERLICH, K., SCHNEIDER, K. A., & KASTNER, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nat Neurosci, 8,* 1595–1602.