Consciousness Regained: Disentangling Mechanisms, Brain Systems, and Behavioral Responses

Storm, J.F.; Boly, M.; Casali, A.G.; Massimini, M.; Olcese, U.; Pennartz, C.M.A.; Wilke, M.

Published in:
The Journal of Neuroscience

DOI:
10.1523/JNEUROSCI.1838-17.2017

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Symposium

Consciousness Regained: Disentangling Mechanisms, Brain Systems, and Behavioral Responses

Johan F. Storm, Melanie Boly, Adenauer G. Casali, Marcello Massimini, Umberto Olcense, Cyriel M.A. Pennartz, Melanie Wilke

How consciousness (experience) arises from and relates to material brain processes (the “mind-body problem”) has been pondered by thinkers for centuries, and is regarded as among the deepest unsolved problems in science, with wide-ranging theoretical, clinical, and ethical implications. Until the last few decades, this was largely seen as a philosophical topic, but not widely accepted in mainstream neuroscience. Since the 1980s, however, novel methods and theoretical advances have yielded remarkable results, opening up the field for scientific and clinical progress. Since a seminal paper by Crick and Koch (1998) claimed that a science of consciousness should first search for its neural correlates (NCC), a variety of correlates have been suggested, including both content-specific NCCs, determining particular phenomenal components within an experience, and the full NCC, the neural substrates supporting entire conscious experiences. In this review, we present recent progress on theoretical, experimental, and clinical issues. Specifically, we (1) review methodological advances that are important for dissociating conscious experience from related enabling and executive functions, (2) suggest how critically reconsidering the role of the frontal cortex may further delineate NCCs, (3) advocate the need for general, objective, brain-based measures of the capacity for consciousness that are independent of sensory processing and executive functions, and (4) show how animal studies can reveal population and network phenomena of relevance for understanding mechanisms of consciousness.

Key words: brain complexity; consciousness; content-specific NCC; functional connectivity; neural correlates of consciousness; no-report paradigm; Perturbational Complexity Index; unresponsive wakefulness; vegetative state

Introduction

What is consciousness? How do our inner, subjective experiences fit into our scientific world view? Why does consciousness exist at all? Why are we not simply “zombies,” responding to sensory input and producing behaviors according to the laws of nature, but with no subjective experiences?

Questions like these (the mind-body problem, mind versus matter, consciousness versus brain) have been pondered by thinkers for centuries, and are widely regarded as among the deepest unsolved problems in science, called “the ultimate intellectual challenge of this new millennium” (Dehaene and Changeux, 2004) and “the major unsolved problem in biology” (Crick, 2004) with wide-ranging theoretical, clinical, and ethical implications.

Until the last few decades, the fundamental problems of consciousness were largely regarded as philosophical, but not widely accepted as scientific topics in mainstream neuroscience. Thus, influential textbooks from that period contained little about these issues, and in the 46 previous annual meetings of the Society for Neuroscience (1969–2016) there have been few symposia explicitly focused on consciousness: a Society for Neuroscience...
Satellite symposium on experimental approaches to consciousness chaired by Christof Koch (1998); and a minisymposium on the neural basis of consciousness chaired by Tsuchiya and Maier (2012). This is a paradox because the widespread interest in the brain surely stems much from its unique ability to generate our conscious experiences, including thoughts, feelings, and dreams. During much of the last century, the influential behaviorist school in psychology regarded “introspective” reports of subjective experiences as unscientific. Since the 1980s, however, pioneers have helped making consciousness a researchable, scientific topic; and novel methods and theoretical advances (e.g., the global workspace theory (GWT)/global neuronal workspace hypothesis, and the integrated information theory of consciousness (IIT)) have yielded remarkable results, opening up the field for scientific and clinical progress (Leopold and Logothetis, 2003; Owen et al., 2006; Seth et al., 2008; Dehaene and Changeux, 2011; Aru et al., 2012a; Casali et al., 2013; Hobson et al., 2014; Oizumi et al., 2014; Pennartz, 2015; Tsuchiya et al., 2015; Koch et al., 2016b; Tononi et al., 2016).

Nevertheless, some uncertainty and skepticism about the scientific status of this topic still remain, which may be seen as natural, given its exceptional methodological and philosophical challenges. Because the concept of consciousness is fundamentally derived from our subjective perspective, it seems hard to reconcile with science in general, which is based on objectivity, although certain aspects of experience are verifiable on an interpersonal basis. In addition, the many different concepts and meanings of the term “consciousness” remain a source of confusion. In this review, we focus on the most basic concept of consciousness: the existence of subjective experiences, (i.e., perceptions, thoughts, dreams, imagery) that disappear in general anesthesia, coma, or certain stages of sleep (Dehaene and Changeux, 2011; Oizumi et al., 2014; Tononi et al., 2016). It is widely accepted that consciousness in this sense encompasses two main dimensions: (1) the “level” of consciousness (arousal/wakefulness), largely controlled by the ascending activation systems of the brainstem and basal forebrain; and (2) the “content” of consciousness, thought to be largely represented by the thalamocortical system (Boly et al., 2013; Sanders et al., 2012; but for a critical review of “levels” of consciousness, see Bayne et al., 2016). Whereas arousal appears to be necessary for consciousness, it is not sufficient. Thus, brain-damaged patients can remain in a “vegetative” state of unresponsive wakefulness for years, with eyes open but no apparent signs of conscious content, leading to severe diagnostic and ethical problems. Thus, consciousness research spans from the deepest theoretical and philosophical questions to practical, experimental, clinical, and ethical issues.

Since the seminal paper by Crick and Koch (1998), which claimed that a science of consciousness should first search for its neural correlates (NCCs), a range of NCCs have been suggested in humans and animals. Recently, two varieties of NCC have been defined: (1) content-specific NCC determines a particular phenomenal component within an experience (e.g., a specific experience of a face), whereas (2) the full NCCs are the “neural substrates supporting conscious experiences in their entirety, irrespective of their specific contents” (Koch et al., 2016b). And as an alternative to the classic distinction between “content” and “level” of consciousness, the “level” may be regarded as reflecting “background conditions for being conscious,” such as appropriate neuromodulatory input to “ensure adequate cortical excitability”, whereas the content-specific and full NCCs are both seen as reflecting content of consciousness (Koch et al., 2016b). And now, after maturing for almost 30 years, the field reflects on its earlier findings, asking which of the reported NCCs truly represent consciousness and which are merely enabling factors or consequences (Miller, 2007; Dehaene and Changeux, 2011; Tsuchiya et al., 2015; Koch et al., 2016b; Tononi et al., 2016).

In the four sections in this review, we present recent progress on theoretical, experimental, and clinical issues. Specifically, we (1) review methodological advancements that are important for dissociating conscious experience from related enabling and executive functions, (2) suggest how critically reviewing the role of the frontal cortex may further delineate the NCCs, (3) advocate the need for general, objective, brain-based measures of the capacity for consciousness that are independent of sensory processing and executive functions, and (4) show how animal studies can reveal population and network phenomena of relevance for understanding mechanisms of conscious states.

How to disentangle content-specific NCC from decision making and visuomotor processes

A fruitful approach for finding correlates of visual consciousness has been the use of bistable visual stimuli, where the physical stimulus gives rise to different percepts (Leopold and Logothetis, 1999) and perceptual suppression/masking paradigms where visual stimuli are erased from consciousness (Leopold and Logothetis, 1999; Kim and Blake, 2005; Dehaene and Changeux, 2011; Tsuchiya et al., 2015) (Fig. 1). In such paradigms, the difference between the trial-averaged neural activities from one perceptual state versus the other (e.g., “visible” vs “invisible”) is typically considered to be a “neural correlate of visual consciousness” (NCC).

This approach identified NCCs in a wide range of brain regions and at different spatiotemporal scales, using diverse methods such as single cell, local field potential and EEG/MEG recordings as well as fMRI in monkeys and humans (Blake and Logothetis, 2002; Sterzer et al., 2009; Tsuchiya et al., 2015; Koch et al., 2016b). From these studies, the view has emerged that NCCs, measured as number of neurons and signal amplitudes that differentiate between two percepts, gradually increase from early stages of visual processing (i.e., LGN/V1) toward “higher-order” association areas and subcortical nuclei (i.e., V4/IT/frontoparietal cortices/thalamic pulvinar) (Boly et al., 2013). Electrophysiological studies in monkeys and humans revealed several signatures of visual consciousness, including a latetime broad electric ERP component, called P3b and oscillations in α/β (9–30 Hz) and γ (> 40 Hz) bands within and between visual and frontoparietal cortices (Dehaene and Changeux, 2011). The multitude of brain regions and neural signals reported as “NCC” already suggests that they fulfill different functions in the process, leading to the critical question, which is now at the center of current debates in consciousness research (Aru et al., 2012b; Frässle et al., 2014; Tsuchiya et al., 2015; Koch et al., 2016b; Brascamp et al., 2018): which brain regions and signals previously reported as NCC constitute the content of visual consciousness, and which reflect either prerequisites (pre-NCC) or consequent decision and motor processes (post-NCC) that enable the perceptual report? To answer this question, researchers adopted several strategies, including the following: (1) refinement of behavioral paradigms allowing the dissociation between task demands and conscious stimulus processing; (2) development of no-report paradigms; (3) local inactivation and lesion studies; and (4) noninvasive brain stimulation techniques, such as transcranial magnetic stimulation (TMS) or direct/alternating current stimulation. Below we will summarize the latest develop-
ments for points 1–3 (for a review of noninvasive brain stimulation studies, see, e.g., de Graaf and Sack, 2014).

Task refinements and No-report paradigms
To study the neural basis of mental states, such as conscious contents, behavioral reports, such as button presses or eye movements, seem necessary. However, relying on overt responses confounds the putative NCC with other cognitive variables, such as selective attention, working memory, motor planning, and performance monitoring (Koch and Tsuchiya, 2012; Tsuchiya et al., 2015) (Fig. 2). Recent intracranial studies in the thalamic pulvinar and visual cortices in monkeys (Wilke et al., 2009) and an MEG study across visual cortices in humans (Kloosterman et al., 2015) have shown that the α/β power decrease typically associated with perceptual suppression vanishes when no behavioral report is required. By systematically varying task relevance and visual perception in a masking paradigm, a recent EEG study also came to the conclusion that the P3b ERP component might reflect a report-related post-NCC process (Pitts et al., 2014b). Similarly, recent fMRI studies in humans have shown that most frontoparietal BOLD activity changes related to perceptual switches during binocular rivalry vanish when task difficulty is held constant (Knapen et al., 2011) or without report when perceptual switches are inferred from involuntary physiological measures, such as pupil diameter and optokinetic nystagmus (Frässel et al., 2014). It needs to be noted that an apparent absence of percept-related prefrontal activation does not exclude its participation in sustaining visual consciousness since fMRI analysis relies on somewhat arbitrary thresholding. Indeed, there is evidence from single-cell recordings in monkeys showing image-specific binocular-rivalry modulation in dorsolateral prefrontal cortex without a report (Panagiotaropoulos et al., 2012). Also, current no-report paradigms might be problematic because, for example, pupil size is also modulated by cognitive factors, such as decision uncertainty, reward prediction, and motor preparation; (Hupé et al., 2009; Laeng et al., 2012; Kloosterman et al., 2015; Overgaard and Fazekas, 2016) and seemingly “perception-related” optokinetic nystagmus during binocular rivalry is also observed in monkeys undergoing ketamine anesthesia (Leopold et al., 2002). Thus, albeit promising, the perfect no-report paradigm to study visual consciousness has not been developed yet, and evidence must be derived from orthogonal research approaches.

Lesion and inactivation studies
Human patients and monkeys with lesions in primary visual cortex (V1) are blind at the corresponding visual field location; therefore, the role of V1 as a prerequisite for visual consciousness seems unequivocal (Tong, 2003; Maier et al., 2008). However, the presence of “blindsight” (Stoerig and Cowey, 1997) (i.e., visually guided actions without accompanying visual consciousness) underlines that correct stimulus detection and localization do not necessarily implicate visual consciousness (Goodale et al., 1991; Cowey and Stoerig, 1995; Schmid et al., 2010; Sahraie et al., 2013; Ajina et al., 2015). On the other hand, the causal contribution of neural populations in area V1 (or LGN) as a possible substrate of conscious perception is difficult to evaluate in lesion studies because they also constitute prerequisites that might be subtracted out in the typical contrastive approach. Due to intact primary visual pathways, conscious perception is easier to evaluate in patients with higher-order thalamic, superior temporal, parietal, and frontal lesions, which often lead to impaired responses and exploration of the contralateral space called “spatial neglect” (Karnath, 2001; Kerkhoff, 2001; Karnath et al., 2002) (Fig. 3). However, biased spatial behavior can occur for different reasons, including impaired visual perception and/or motor decision/planning deficits. Indeed, using a spatial decision paradigm with directed eye movements versus button presses showed that inferior parietal lesions in humans interfere with directed saccade choices but not with visual perception reported with button presses (Ro et al., 2001).

Similarly, by using local inactivation in specific parietal regions in monkeys (lateral intraparietal and medial intraparietal/parietal reach region; Fig. 3), several studies showed effector-specific visuospatial impairments (e.g., for hand but not eye movement reports and vice versa), which are hard to reconcile with the notion that the entire parietal cortex serves a critical function in visual consciousness (Hwang et al., 2012; Christopoulos et al., 2015; Katz et al., 2016). A similar case can be made for the dorsal pulvinar, which strongly interconnects with frontoparietal cortices. Spiking rates of pulvinar neurons in monkeys strongly correlate with reported perception (Wilke et al., 2009; Komura et al., 2013), but local inactivation and microstimulation there indicate visuomotor impairments with intact visual detection performance (Wilke et al., 2010, 2013; Dominguez-Vargas et al., 2017). However, pulvinar as well as parietal cortices might
contain subregions that are more closely related to visual consciousness as opposed to visuomotor behavior (Driver and Mattingley, 1998; Ward et al., 2002; Harvey and Rossit, 2012; Komura et al., 2013; Zhou et al., 2016). Thus, deciphering regional specificity and the further development and application of behavioral paradigms that allow the distinction between perceptual and action planning deficits will remain important. From the beginnings of NCC studies, it has been widely assumed that the prefrontal cortex plays a critical role in enabling both contents and levels of consciousness (Crick and Koch, 1998; Leopold and Logothetis, 1999; Cruse et al., 2011; Dehaene and Changeux, 2011). Because the role of the prefrontal cortex has recently emerged as one of the most controversial issues in NCC research (Pennartz, 2015; Koch et al., 2016a), the section below will review the recent correlative and lesion studies on the prefrontal cortex in more detail.

Are the neural correlates of consciousness mainly in the front or in the back of the cerebral cortex?

It has been widely assumed that prefrontal cortex is critical for consciousness as opposed to a frontoparietal network (Del Cul et al., 2009; Bor and Seth, 2012; Laureys and Schiff, 2012). However, recent evidence from lesion, stimulation, and neuroimaging studies challenges this assumption.

Lesion studies have provided several reports of patients with a normal level of consciousness after extensive frontal damage. For example, patient A (Brickner, 1952), after surgical removal of bilateral prefrontal cortex, once (sic) “toured the Neurological Institute” with “two [..] distinguished neurologists, [..] none of them noticing anything unusual until their attention was especially called to A” after more than an hour. By contrast, traumatic lesions of the posterior corpus callosum, connecting large parts of the posterior cortex, carry a 214-fold increased risk of permanent vegetative state (Kampfl et al., 1998). With regard to the contents of consciousness, there is no clear evidence for loss of specific experiential contents after frontal damage (Penfield and Jasper, 1954). Similarly, while lesions in frontal cortices have been reported to lead to overt visually guided orienting deficits, they hardly lead to impairments of conscious contents per se (Na et al., 1998; Mataró et al., 2001; Ro et al., 2001). In contrast, abundant neurological evidence points to lesions in the posterior cortex causing a loss of specific contents of experience (Farah, 2004). For example, lesions of the right fusiform face area may cause unconsciousness of faces, and lesions of inferolateral occipital cortex unconsciousness of colors (Barton, 2011).

Electrical stimulation of most of the frontal cortex fails to elicit content-specific changes in experience (Penfield and Jasper, 1954), although it can induce involuntary movements or interfere with
In humans, lesions in frontal (Brodmann area 44), inferior parietal cortex (Brodmann area 40), superior temporal gyrus (STG), basal ganglia, and pulvinar have been reported to lead to spatially biased demands; in monkeys, visual perception might be preserved. Ventral pulvinar (vPULV, purple shading) inactivation leads to change detection deficits resembling visual neglect. Dorsal pulvinar (dPULV, orange shading) inactivation leads to spatial orienting bias for both eye and hand movements which can be compensated by visual reward cues, suggesting that visual perception might be preserved. Ventral pulvinar (vPULV, purple shading) inactivation leads to change detection deficits resembling visual neglect. Summarized from local inactivation studies in monkeys (Wardak et al., 2002; Wilke et al., 2010; Hwang et al., 2012; Wilke et al., 2012; Wilke et al., 2013; Hwang et al., 2014; Christopoulos et al., 2015; Katz et al., 2016; Zhou et al., 2016).

The advent of large-scale neuroimaging databases can now help to quantify systematic associations between specific conscious contents and the activation of specific cortical areas by using meta-analytic reverse inference (Poldrack, 2006; Yarkoni et al., 2010; Moran and Zaki, 2013; Poldrack and Yarkoni, 2016). In agreement with lesion and stimulation studies, reverse inference analyses performed, for example, with Neurosynth (www.neurosynth.org) locate the best predictors for specific conscious contents within specific areas of posterior cortex, such as the right fusiform face area for face percepts, occipital areas for visual words or motion, temporal areas for speech sounds, or parietal areas for touch percepts. In all these cases, reverse inference fails to highlight frontal areas as predictive for the presence of specific contents of consciousness (Boly et al., 2017).

Multivariate decoding techniques can also help to identify the true NCC as the brain activity patterns that have the highest predictive value for specific conscious percepts (Haynes, 2009; Sandberg et al., 2014). Using such approaches, the best predictor for the presence versus absence of consciousness has been located in temporo-parieto-occipital cortices both during non-rapid eye movement (NREM) sleep (Siclari et al., 2017) and in patients with brain damage (Demertzi et al., 2015; Stender et al., 2016). While studies comparing brain function during anesthesia versus wakefulness consistently highlighted differences in both frontal and parietal cortices and in thalamocortical connectivity (Alkire et al., 2008), activity in frontal regions was shown to be poorly predictive for the presence versus absence of connected consciousness (Gaskell et al., 2017). On the other hand, the predictive value of posterior cortex activity for consciousness during anesthesia still needs to be assessed (Boly et al., 2017). With regard to the contents of consciousness, numerous studies in both awake and dreaming subjects could decode the presence of specific conscious contents from specific activity patterns in posterior cortex (Nishimoto et al., 2011; Horikawa et al., 2013; Siclari et al., 2017), whereas the evidence for the frontal cortex is much less clear.
Working memory contents can also be more reliably decoded from the back than from the front of the cortex (Emrich et al., 2013).

Together, evidence from lesion, stimulation, and neuroimaging studies consistently points to regions of the posterior cortex, including temporal, parietal, and occipital areas, as a “hot zone” for consciousness, in which different brain activity patterns may specify different experiential contents. By contrast, evidence for a content-specific contribution of the prefrontal cortex is scarce or indirect. At a minimum, detailed descriptions of conscious patients after extensive bilateral frontal resections demonstrate that an entirely intact prefrontal cortex is not necessary for consciousness. It remains possible that some prefrontal regions may contribute specific conscious contents, such as feelings of reflection, valuation, and affect (Koch et al., 2016a). In the future, within-state paradigms applied to sleep dreaming (Siclari et al., 2017) and daydreaming (Perogamvros et al., 2017) will be critical to dissociate content-specific NCC from task-related cognitive processes (Koch et al., 2016b).

Measuring brain complexity to detect (and promote) the recovery of consciousness.

Detecting the presence of consciousness in clinical practice relies ultimately on a behavioral input–output paradigm: a neurologist typically probes the patient with sensory stimuli of different modalities or verbal commands and observes his motor responses (Giacino et al., 2004). If the patient reliably produces outputs that are specific for the different inputs, he is considered conscious. The earlier sections of this review already pointed out that this approach can be problematic: unresponsiveness can be paired with consciousness. However, thanks to recent conceptual and technical advances, a similar input–output paradigm can now be applied even to patients who are fully paralyzed and cannot engage in motor behavior (Owen et al., 2006; Cruse et al., 2011; Naci et al., 2014). In this case, subjects are still presented with sensory stimuli or verbal instructions (e.g., “imagine playing tennis”), but the blockage of motor output is bypassed by directly reading-out neuronal responses with fMRI or EEG. These neuroimaging paradigms can be easily interpreted in case of positive results: unresponsive subjects who willfully produce neuronal activations that are consistent and specific for the given input are considered conscious. However, many brain-injured patients may recover consciousness, yet fail to produce the appropriate neuronal responses (Monti et al., 2010; Bardin et al., 2011).

One potential reason for this false-negative result is that consciousness may be present not only independently of executive/motor functions but also in the absence of sensory processing. For example, during dreaming, complex, temporally unfolding episodes can be as intense and vivid as waking experiences, yet consciousness may completely disconnect from the external environment in which consciousness was present, including awake controls and subjects who were disconnected and unresponsive during rapid eye movement (REM) sleep and ketamine anesthesia but retrospectively reported having had vivid conscious experiences upon awakening. Once calibrated on the gold standard of subjective reports, PCI measurements performed at the bedside of brain-injured noncommunicating subjects offered unprecedented sensitivity (94%) in detecting minimal consciousness in patients and allowed identifying a significant percentage (~20%) of vegetative state/unresponsive wakefulness syndrome cases with high brain complexity. In these behaviorally unresponsive patients, PCI values overlapped with the distribution of the benchmark conscious condition, suggesting the presence of disconnected consciousness, which eventually resulted in a higher rate of recovery (Casarotto et al., 2016).

Measuring complexity and studying its mechanisms in the brain may also offer valuable insights on the neuronal bases of loss and recovery of consciousness. TMS is a strong, coarse-
grained perturbation that activates many diverging cortical and corticostriatal connections; therefore, PCI is a global measure that cannot resolve whether there are preferential hubs of complexity within the brain (e.g., in the parietal or in the frontal lobe). Ongoing studies using localized, intracortical single-pulse electrical stimuli and mesoscale (local field potential) recordings may provide deeper insight. For example, a recent intracranial human study (Pigorini et al., 2015) identified a simple mechanism by which the emergence of sustained interaction among cortical areas may be impaired during dreamless sleep: upon receiving an input, cortical neurons tend to respond briefly, then hush and “forget.” This behavior, a potential killer for brain complexity, is due to a basic dynamic, also known as cortical bistability, by which individual neurons tend to fall into a silent hyperpolarized state (down state) after an initial activation (up state) (Compte et al., 2003; Sanchez-Vives et al., 2017). Down states, which are mainly engendered by activity-dependent potassium (K⁺) currents during physiological NREM sleep (Timofeev et al., 2001), may be generally relevant for consciousness and its disorders for at least three reasons. First, they can also occur in pathological conditions due to alterations of the inhibition/excitation balance (Murase et al., 2004) or as a consequence of white matter injury (Timofeev et al., 2000). Second, they can be investigated from the intracellular level in animal models (Steriade et al., 1993) to the whole network level in humans (Cash et al., 2009). Third, they can impair network interactions globally (Lewis et al., 2012; Pigorini et al., 2015; Olcese et al., 2016), but they are in principle reversible.

Notably, a recent microscale study using electrical stimulation and recordings in isolated cortical slices (an extreme case of cortical island) showed that complex causal interactions, as assessed by an adapted version of PCI, can be restored by pharmacological manipulations that reduce neuronal bistability (D’Andola et al., 2017). Can some of these concepts be translated to the bedside of comatose patients? Some patients with vegetative state/unresponsive wakefulness syndrome spontaneously flip from a low-complexity condition to a conscious high-complexity state for no apparent reason (Rosanova et al., 2012). Will we be able to understand and promote this process? In the years to come, it will be crucial to further elucidate the relationships between single-neuron dynamics, overall network complexity, and consciousness through extensive experiments across scales, species, and models.

The conscious rodent brain: ensemble activity and long-range correlation patterns

Animal models provide a unique window to probe the neuron-level mechanisms of many cognitive processes, including consciousness. Rodents, in particular, offer an unprecedented range of tools to access all circuitual elements of the brain, and share many features with the human brain. However, when addressing the question of consciousness in rodents or other animals, it is crucial to ask whether these animals can be conscious at all. It is a valid question regardless of species, applying to nonhuman primates as well as rodents. Although one may adopt an agnostic stance on the topic and just call for further research (Gutfreund, 2017), there is an increasing need to come up with a neuroscientific account of animal consciousness, not necessarily providing definitive answers but rather a “best estimate” facilitating rational decisions on issues of great societal relevance, such as ritual slaughter and pain mitigation in animal health care.

The question of rodent consciousness can be split into two components: (1) do rodent brains have conscious states as opposed to unconscious states (e.g., during dreamless sleep or anesthesia)? (2) What are the contents of consciousness as experienced by rodents? The first question is less difficult to address than the second, as is also the case for nonhuman primates and humans: conscious contents refer to subjective and qualitative aspects of our experiences, which are notoriously hard to get at due to their essentially private nature. In contrast, there are several arguments in favor of rodents having at least a basic capacity to sustain conscious states: (1) the fact that electrophysiological markers indicative of conscious states in humans and primates are also found in rodent brains (e.g., markers for wakefulness and REM sleep, such as desynchronized EEG, as opposed to slow waves in NREM sleep); (2) the strong anatomical homologies between the systems in rodent and human brains deemed important for conscious states (mainly, thalamocortical systems); and (3) the presence of behavioral indicators of conscious states, overlapping in use and applicability between humans and rodents. Such indicators do not only include sleep–wake cycles and markers of wakefulness, such as behavioral reactivity to sensory stimuli and orienting responses, but also the ability to generate nonhabitual (i.e., goal-directed) behaviors (Weiskrantz, 1995; Pennartz, 2015) and distinctive behavioral responses when stimuli are perceived or not (Carandini and Churchland, 2013). A fourth argument is based on recently developed measures of consciousness (Massimini et al., 2005; Ferrarelli et al., 2010; Casali et al., 2013). As in humans and primates, wakeful states in rodents are marked by a higher degree of complexity (Hudetz et al., 2015, 2016) and long-range functional connectivity than during deep, NREM sleep or anesthesia (Olcese et al., 2016). This fourth argument goes beyond “physiological hallmarks” because it is based on inferred functional properties of conscious systems, such as the ability to integrate complex information (Tononi and Edelman, 1998), to generate interactions and joint coding between sensory modalities and memory (Pennartz, 2009), and to represent a near-instantaneous “world model” efficiently and usefully (Pennartz, 2015). While many questions also regarding this argument remain open in rodents, our best, evidence-based guess is that rodent brains are indeed capable of sustaining conscious states, regardless of the precise complexity and richness of represented contents compared with primates. Here we will review recent findings related to this fourth argument, and discuss how they expand our understanding of NCCs in the rodent brain.

The ability to integrate complex information and generate interactions sustaining world representations eventually resides in the brain’s information-coding elements: neurons. Because most previous comparisons between conscious, sleep, and anesthetized states in humans have used recording techniques with macroscopic and mesoscopic resolution (Massimini et al., 2005; Ferrarelli et al., 2010; Park and Friston, 2013; but see Lewis et al., 2012) (EEG, local field potential, or fMRI), it has been difficult to determine how the loss of global functional connectivity (e.g., during transitions from wakefulness to NREM sleep) is associated with cell-level changes in correlation patterns. The few studies on the topic are somewhat inconsistent and report either no change in functional connectivity between conscious and unconscious states (Zhang et al., 2014) or a dramatic loss of communication (Makino and Komiya, 2015; Ishizawa et al., 2016). We investigated this using multiarea tetrode recordings in sensory neocortex and hippocampus of rats undergoing transitions from wakefulness to NREM sleep (for methods, see Bos et al., 2017). We computed conditional mutual information (cMI) to quantify functional connectivity between spike train patterns across all cell pairs, as this measure includes nonlinear correlations that would
escape detection by linear methods (Steuer et al., 2002; MacKay, 2003). Although the results showed area-specific changes in correlation patterns, they confirmed that global functional connectivity drops from wakefulness to NREM sleep (Olcese et al., 2016). This effect was due to a selective loss of interareal interactions between excitatory neurons but not interneurons, whereas intra-areal connectivity was largely preserved for both neuronal types (Fig. 4A). Thus, the drop in global connectivity during NREM sleep is not paralleled by a similar drop in intra-areal connectivity, allowing local computations in relative “islands of activity” to continue. Because cMI does not indicate the direction of information flow between neurons, we also computed the transfer entropy between spiking neurons, and found that this measure allows to distinguish several regimes for directed communication in the brain, depending on the time scale of neural interaction, the anatomical distance between neurons, and, intriguingly, the presence of a neural correlate to the behavioral task performed by the animal before NREM sleep (U.O. et al., unpublished observations).

In addition to spike-based markers distinguishing different brain states, we set out to study differences in ensemble behavior within the awake, conscious state, when animals performed a visual detection task. While correlates of conscious perception have been recently described at the level of membrane potential dynamics and spiking activity (Pinto et al., 2013; Sachidhanantham et al., 2013; Manita et al., 2015; Takahashi et al., 2016), especially involving recurrent activity between cortical areas, it is still poorly understood how perception is reflected in population dynamics. Using a range of visual contrasts, mice were trained to respond to presentations of moving gratings by performing licking actions (however, note that this go/no go paradigm does not exhaustively test for consciousness (Koch, 2004; Montijn et al., 2015; Pennartz, 2015). Ensemble activity in layer 2–3 of mouse V1 was recorded by 2-photon Ca$^{2+}$ imaging, and we asked which of two measures would more accurately correspond to hit versus miss trials: the mean somatic Ca$^{2+}$ response or the population heterogeneity (for explanation, see Fig. 4B,C). We found that heterogeneity allowed a significantly better separation between hit and miss trials than the mean response (Montijn et al., 2015) (Figure 4D). Moreover, applying heterogeneity (but not mean dF/Fo) to the baseline (prestimulus) period, it proved feasible to predict the type of detection response (fast hit response vs miss or slow response). Thus, heterogeneity of single-cell responses within larger ensembles may indicate a neural correlate of detection and/or perception along with arousal and attention (Pinto et al., 2013). We suggest that a state of high heterogeneity arises in association with EEG desynchronization, and is conducive to having a high degree of differentiated, stimulus-specific coding in...
a neural network (Pennartz, 2009; Goltstein et al., 2015; Tononi et al., 2016; compare Renart et al., 2010). In line with this, the presence and orientation of the stimulus were better decodable from the population when the animal made a correct detection response and heterogeneity was high. Strong heterogeneity in sensory cortical cell populations may thus correlate with better representational performance. However, this position does not exhaust the repertoire of potential neural correlates: when analyzing assembly activity as manifested by multineuron activity sequences, we found that visual detection (hit trials) was also paired with a higher consistency in the timing of single-cell activity within the sequence than nondetection (miss trials) (Montijn et al., 2015).

Although rodent behavioral paradigms need to be developed further to yield more specific markers of consciousness, we may conclude that ensemble studies are beginning to reveal population and network phenomena relevant for understanding not only what conscious states consist of (i.e., the cell-level nature of NCCs), but also why such states are conducive for optimal coding of the stimuli and complex scenes of which we can become conscious.

Conclusions and perspectives

In conclusion, consciousness, once regarded as a purely philosophical topic, the “mind-body problem”, has become an active research field in neuroscience. Neuroscientific consciousness research, boosted by novel methods and theoretical advances, has become increasingly accepted and mature, and is now beginning to yield remarkable scientific and clinical progress.

As is often the case within a developing branch of science, concepts and mechanisms that were previously often blended and confused are now being teased apart and more clearly defined. “Content-specific” and “full” NCCs are becoming increasingly well defined. Improved strategies allow conscious experience to be distinguished from related enabling and executive functions, for example, through employment of refined behavioral/no-report paradigms, local inactivation and lesions, and noninvasive brain stimulation. A closely related, critical reevaluation of the prefrontal cortex’ role in consciousness suggests that, although important for behavioral reporting and other executive functions, it is not crucial for conscious experiences per se. Instead, converging evidence points to a posterior “hot zone” for consciousness including parietal, temporal, and occipital areas.

Methods for objective assessment of consciousness are crucial, both experimentally and clinically. The PCI is a highly promising metric for detecting consciousness, inspired by the Integrated information theory. Without requiring any sensory input or behavioral report (i.e., executive functions), PCI seems to be a reliable indicator of the human brain’s capacity for consciousness, being related to the amount of information that can be generated through causal brain interactions. PCI has recently proved highly accurate in a large benchmark population of conscious or unconscious subjects.

This progress also offers insights into neuronal bases of loss and recovery of consciousness, and how these may be manipulated. Thus, converging evidence from animals and humans suggest that neuronal “bistability,” a dynamic property of cortical circuits, can cause unconsciousness by impairing brain complexity. Recently, pharmacological suppression of bistability was shown to restore the complex causal interactions required for consciousness. Rodent models are being developed to allow dissection of molecular, cellular, and circuit mechanisms of consciousness. Ensemble recordings begin to reveal how, in addition to complexity measures, heterogeneity and spike-based measures of functional connectivity can be used to differentiate conscious and unconscious states, and why conscious states promote optimal coding. Population coding studies promise to enable researchers to study consciousness in terms of brain-generated representations, thus supplementing global assessments of integrated information.

Rather than trying to solve all major theoretical problems first, we advocate an interleaved practical and theoretical approach. Experiments across scales, species, and models, spanning from ion channels, neurons, and microcircuits to whole-brain simulations, from unit and ensemble recordings in animals to the patient’s bedside, will be crucial for bridging the gaps between single-neuron dynamics, overall network complexity, and conscious experience. This is a daunting task and the stakes are high, but the recent progress is promising.

References


based index of consciousness independent of sensory processing and behavior. Sci Transl Med 5:198ra105. CrossRef Medline