
The Ongoing Search for the Neuronal Correlate of Consciousness

Wolf Singer

A few decades ago the search for the neuronal correlates of consciousness was considered both technically intractable and philosophically questionable. Searching for a material substrate of phenomena accessible only from the first-person perspective appeared to be epistemically problematic. But the development of non-invasive imaging technologies and the availability of intracranial recordings from patients alleviated the imminent technical problems. Progress in the analysis of the connectome of the brain, and the introduction of multisite recordings from the cerebral cortex of animals led to a revision of concepts in the field of cognitive neuroscience, emphasizing principles of distributed processing in recurrent networks with non-linear dynamics, self-organization, and coding in high-dimensional-state space. These advances, together with the growing evidence for epigenetic shaping of brain functions by socio-cultural influences, pave the way for novel theories that attempt to bridge the gap between neuronal processes and subjective states.

Keywords

Binding problem | Cultural evolution | Distributed processing | Epigenetic shaping | Long-range synchronisation of oscillation | Meta-representation | Naturalistic epistemology | Neural correlate of consciousness (NCC) | Oscillations | Perceptual constancy | Small-world architecture | Subconscious processing | Synchrony | Unity of consciousness | Workspace of consciousness

1 Introduction

Progress in brain research, especially in the domain of cognitive neuroscience, renders phenomena that have traditionally been subjects of the humanities amenable to scientific investigation. It has now become possible to investigate the neuronal underpinnings of mental phenomena such as perception, decision making, control of attention, language perception and production, action planning, storage and recall of memories, emotions and moods, desires and aversions, and last but not least consciousness. This research agenda is confronted with a number of fascinating challenges. One is the immense complexity

of the brain processes that underlie these highly-differentiated cognitive functions. Another results from the fact that many of the phenomena whose neuronal correlates are to be investigated are subjective phenomena, accessible and describable only from a first-person perspective. Hence there is an epistemically problematic gap between what is observable from the third-person perspective of scientific inquiry and the explananda that need to be defined in terms of first-person experience. Yet another challenge is that a relatively young scientific discipline is set to enter territories that

Author

Wolf Singer

w.singer@brain.mpg.de

Max Planck Institute for Brain Research (MPI)
Frankfurt a. M., Germany

Commentator

Valdas Noreika

valdas.noreika@mrc-cbu.cam.ac.uk
Medical Research Council
Cambridge, United Kingdom

Editors

Thomas Metzinger

metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt

jennifer.windt@monash.edu

Monash University
Melbourne, Australia

for millennia have been ploughed by great minds who have coined terms, formulated concepts, and constructed belief systems based on evidence extracted from introspection, intuition, and observations that relied exclusively on the natural senses. This raises numerous and on occasions frustrating problems for communication, because bridges have yet to be built between the more recent naturalistic description systems and the highly-differentiated terminology nurtured in the humanities. Some of these problems surface in passionate discussions on the existence of free will, the nature of perception, the constitution of the Self, and intentionality and mental causation—and above all on the question of whether it is even possible to identify neuronal correlates of mental, subjective phenomena.

In this chapter some of these challenges will be discussed from a neurobiological perspective. We shall first review the state of the art in the field of cognitive neuroscience, emphasizing recent changes in our views on the brain. These have been forced upon us by the novel data produced by new and powerful technologies. These insights show the brain to be a highly distributed, self-active system with non-linear dynamics; rather than a hierarchically-organized stimulus-response machine, as has been proposed by behaviourist theories. Subsequently, an excursion will be made into epistemology, to establish the extent to which brain research can contribute to philosophical discussions concerning the nature of perception. The process of perceiving will be interpreted as a constructive act in which sparse sensory signals are matched with a huge amount of stored knowledge; and the various sources of this knowledge will be discussed. This section will set the stage for the following discussion of the putative neuronal correlates of consciousness (NCC), as it will highlight the cognitive constraints and idiosyncrasies of cognitive systems that owe their abilities to evolutionary processes. Before reviewing various theories on the NCC, an attempt will be made to define the explanandum—in full awareness of the futility of this attempt. The review of experimental work in search of the NCC will be followed by a brief

account of our own experimental contributions, and then an attempt will be made to demystify the so-called “hard problem” of consciousness—the problem of finding a naturalistic explanation for the qualia, namely these immaterial connotations of our experiences. My proposal will be that the problem can be alleviated if we consider not only individual brains and biological evolution but also cultural evolution and the social realities that have emerged from socio-cultural interaction between human beings.

As this contribution addresses an interdisciplinary audience I considered it appropriate to not only refer to published work when alluding to experimental findings and concepts but to sometimes provide explicit and detailed background information. To this end I have adapted passages from a few of my own publications, in which I had addressed some of the issues that needed to be recapitulated for the sake of clarity in the present contribution. These passages include descriptions of experimental findings from my lab and descriptions of the state of the art in the neurosciences.

2 The state of the art in cognitive neuroscience: A paradigm shift

2.1 Classical views

As detailed in several previous reviews the neurosciences are about to undergo a paradigm shift towards concepts that consider the brain as a self-organizing complex system with non-linear dynamics that exploits a huge body of stored knowledge to interpret sensory signals, formulate hypotheses and to generate predictive models of the world in order to optimize adapted behavioral responses (Singer 2009, 2013). For many decades, the search for the neuronal underpinnings of cognitive and executive functions has been guided by the behaviourist view that the brain is essentially a highly complex and versatile stimulus-response machine, in which serial processing strategies prevail. This view received further support from early anatomical data that emphasized that feed-forward connections exhibit high topographic precision and possess strong driving synapses, while feedback

connections are diffuse and only modulatory. The rather impressive performance of artificial pattern-recognition systems based on such processing architectures suggested that neuroscientists were on the correct path. Accordingly, they set out to study the responses of neurons to sensory stimulations across the various stages of the processing hierarchy and analyzed activation patterns associated with motor output, hoping that these strategies would eventually lead to reductionist explanations of the neuronal mechanisms that support cognition, memory, decision making, planning, and motor behaviour. The strategy to follow the transformation of activity from the sensory surfaces over the numerous levels of hierarchically-organized processing structures to the respective effector organs proved to be extremely fruitful. Comparison of brains from different species provided compelling evidence that the basic principles according to which neurons function and exchange signals have been preserved throughout evolution with only minor modifications. For the comparatively simple nervous systems of certain invertebrates, this behaviourist approach allowed for near-complete descriptions of the neuronal mechanisms underlying particular behavioural manifestations. This nurtured the expectation that pursuing this research strategy would sooner or later allow us to explain in the same way the more complex behaviour of mammals—and ultimately also the highly-differentiated cognitive functions of primates and human subjects. However, in recent decades the pursuit of this approach has led to an accumulation of evidence that demands a revision of the classical hypothesis, which emphasizes serial feed-forward processing of sensory information within hierarchically-organized architectures.

2.2 Observations forcing an extension of classical views

Advances in the analysis of the cortical connectome, the introduction of multisite recording techniques, and the development of imaging methods assessing whole-brain activity have generated data that necessitate an extension of

classical views, raise novel questions, and likely provide new solutions to old problems.

Anatomical evidence: i) Within processing streams from sensory surfaces to executive organs, feedback projections are in general more numerous than feed-forward projections, emphasizing the importance of top-down control (Felleman & van Essen 1991). ii) Connections linking neurons within distinct cortical areas cross the boundaries between areas (Schwarz & Bolz 1991). Thus, the cerebral cortex appears to be a continuously coupled sheet, the different cortical areas being distinguished mainly by their input and output connections. iii) From primary sensory areas onwards, processing streams diverge into numerous parallel pathways whose nodes are linked by massive reciprocal connections, both within and across modalities (Markov & Kennedy 2013; Markov et al. 2013). iv) The rule that feed-forward connections originate in supra- and feedback projections in infragranular layers does not hold for nearby cortical areas (Markov & Kennedy 2013; Markov et al. 2013). Together with electrophysiological evidence (De Pasquale & Sherman 2011), this threatens the strict distinction between feed-forward driving and feedback modulatory connections. v) Finally, statistical analysis of interareal connectivity suggests an organisation resembling small-world, rich-club networks (see Van den Heuvel & Sporns 2013) that minimize path length between nodes (areas; Van den Heuvel & Sporns 2011; Sporns 2013). However, analysis of projections with cellular resolution suggests as one reason for short path length the surprisingly high degree of connectedness among cortical areas. Statistical analysis suggests that more than 60% of possible links between network nodes are actually realized (Markov & Kennedy 2013).

Functional evidence: i) Even in early sensory areas neurons lose their simple feature-specific responses when challenged with complex stimuli (David et al. 2004; Vinje & Gallant 2000). Moreover, responses are influenced by stimuli in other modalities, by attention, reward expectation, and contents in working memory, thus suggesting contextual modulation not only by intrinsic connections but also by top-down

projections (Engel et al. 2001; Calvert et al. 1997; Iurilli et al. 2012; Muckli & Petro 2013; Stokes et al. 2013). ii) The notion of strictly serial processing from input layer four, via layers three and two, to the output layers five and six of the cerebral cortex needs to be revised in light of evidence that vigorous responses can also be elicited by sensory input when parts of this canonical circuit are disrupted (Constantinople & Bruno 2013). The possibility that supra and infragranular compartments can operate in parallel is further supported by evidence that the two subdivisions engage in oscillatory activity in different frequency bands (gamma in supra- and alpha or beta in infragranular layers; Buffalo et al. 2011; Roopun et al. 2008. iii) Multisite recordings indicate that “spontaneous” fluctuations in the responsiveness of individual neurons are often the reflection of coordinated, highly structured spatio-temporal activity patterns rather than the result of noise (Kenet et al. 2003; Fries et al. 2001a). iv) Widely distributed cortical areas exhibit coherent fluctuations of their spontaneous activity, forming functionally-coupled networks that change in their composition in a state-dependent way (Fox et al. 2005; Hipp et al. 2012; Raichle 2011; Raichle et al. 2001). Thus, the cortex—and in a wider sense the brain—appears to be a highly active, pattern-generating system, rather than just a stimulus-driven device. v) Analysis of whole brain activity with functional magnetic resonance imaging (fMRI) and electroencephalographic (EEG) and magnetoencephalographic (MEG) measurements indicates that virtually all cognitive and executive functions are associated with the activation of networks of often widely-distributed cortical areas (Engen & Singer 2013; Friederici & Gierhan 2013; Hipp et al. 2011; Hodzic et al. 2009; Power & Petersen 2013). This suggests that distributed networks are a substrate of functions rather than individual specialized structures. vi) Finally, analysis of the brain’s dynamic signatures indicates that neuronal populations can engage in oscillatory activity in characteristic frequency bands and synchronize their discharges, such that the respective frequency bands and the composition of coherently active

cell groups depend on central states, attention, cognitive tasks, and goals of action (Buzsáki 2006; Singer 2010).

These novel anatomical and functional data suggest as a prevailing organizational principle distributed processing in densely coupled, recurrent networks with non-linear dynamics, which are capable of supporting high dimensional states. This organization requires a high degree of coordination of distributed processes, suggesting that special mechanisms are implemented to dynamically bind local processes into coherent global states, and to configure functional networks “on the fly” in a context- and goal-dependent way. It has also become clear that the brain is by no means a stimulus-driven system. Rather, it is self-active, permanently generating highly structured, high-dimensional spatio-temporal activity patterns. These patterns are far from being random, and instead seem to reflect the specificities of the functional architecture that is determined by genes, modified by experience throughout post-natal development, and further shaped by learning. These self-generated activity patterns in turn seem to serve as priors with which incoming sensory signals are compared. Perception is now understood as an active, reconstructive process, in which self-generated expectancies are compared with incoming sensory signals. The development of methods that allow simultaneous registration of the activity of large numbers of spatially-distributed neurons revealed a mind-boggling complexity of interaction dynamics—which in turn eludes the capacity of conventional analytical tools and, because of its non-linearity, challenges hypotheses derived from intuition.

In the last decade theoreticians have begun to explore and appreciate the immense computational power of such self-organizing recurrent networks that gave rise to concepts such as “reservoir computing”, “echo-state computing” or “liquid computing” (Buonomano & Maass 2009; Lukoševičius & Jaeger 2009). The evidence that resting-state activity is highly structured, that information is contained in the spatio-temporal relations between the responses of widely distributed neurons, and that stimulus-response functions depend crucially on state

variables generated within the brain are in principle compatible with such advanced concepts of information processing in highly non-linear, high-dimensional dynamic systems; but neurobiological approaches taking such considerations into account are still very rare.

2.3 Persisting explanatory gaps

Our rather detailed knowledge of the response properties of individual neurons in different brain structures, and of the microcircuits that shape these responses, stands in stark contrast to our ignorance of the complex and highly dynamic processes through which the myriads of spatially-distributed neurons interact in order to produce specific behaviours. Evidence from invasive and non-invasive multi-site recordings indicates that most higher brain functions result from the coordinated interaction of large numbers of neurons, which become associated in a context- and goal-dependent way into ad-hoc formed functional networks. These networks are dynamically configured on the backbone of the anatomical connections (for review see [von der Malsburg et al. 2010](#)). Evidence also indicates that these interactions give rise to extremely complex spatio-temporal patterns that are characterized by oscillations in a large number of different frequency bands, which can synchronize, exhibit phase shifts, and even cross frequency coupling ([Uhlhaas et al. 2009](#)). In the light of these novel data, the brain—and in particular the neocortex—appears to be a self-active, self-organizing “complex system” which exhibits non-linear dynamics, is capable of utilizing multiple dimensions for coding (space, amplitude, oscillation frequency, and phase), operates in a tightly-controlled range of self-organized criticality ([Shew et al. 2009](#); edge of chaos), and constantly generates highly-structured, high-dimensional activity patterns that are likely to represent stored information. However, how exactly information is encoded in the trajectories of these high-dimensional and non-stationary time series is largely unknown, and is the subject of increasingly intense research. Moreover, with the exception of a few studies in which selective manipulation of the activity of

defined neuron groups were shown to affect behaviour in a particular way ([Salzman et al. 1992](#); [Houweling & Brecht 2008](#); [Han et al. 2011](#)) most of the available evidence on the relations between neuronal responses and behaviour is still correlative in nature. This makes it difficult to determine whether an observed variable is an epiphenomenon of a hidden underlying process or is causally involved in accomplishing a particular function. Thus, systems neuroscience now faces the tremendous challenge of analyzing the principles of distributed dynamic coding and of obtaining causal evidence for the functional role of specific activation patterns, in order to distinguish between functionally-relevant variables and epiphenomena.

In conclusion, we have to abandon classical notions of the neuronal representation of perceptual objects and, in the same vein, that of motor commands. The consequence is that it becomes once again unclear how the distributed processes that deal with the various properties of a perceptual object—its visual, haptic, acoustic, olfactory and gustatory features—are bound together in order to give rise to a coherent representation or percept. Given this, it may appear more than bold to attempt to identify the neuronal correlates of consciousness—probably the highest and most mysterious of our cognitive functions.

2.4 What neuroscientists believe

Despite the numerous gaps in our understanding of integrated brain functions, neurobiologists agree on a number of general conclusions on the relation between brain processes and behavioural phenomena. The majority of neurobiologists seem to consent that all cognitive and executive functions that we can observe in human beings, including the highest mental activities and consciousness, are the result, not the cause, of neural interactions. Consequently, it is held that mental phenomena follow or emerge from neural interactions and do not precede them. Furthermore, it is assumed that all neural processes obey the known laws of nature. The reason for this is that the behaviour of organisms of low complexity, such as, for example,

molluscs or worms, can be fully explained by registering the activity of their neurons and establishing causal relations between the spatio-temporal patterns of this activity and the respective behaviour. There is, at present, no need to postulate any additional unknown forces, laws, or modes of interaction in order to explain their behaviour. The reason for this is that evolution is a very conservative process. Once an invention has been made that increases fitness it tends to be conserved, unless there is a major change in conditions that makes this invention obsolete or maladapted. Therefore, our nerve cells function in exactly the same way as those of snails. Likewise, the development of structures also follows a very conservative path. Since the first appearance of the cerebral cortex, the six-layered sheet of nerve cells that covers the hemispheres of the brain, no new structures have emerged. There is just more of the same, and this increase in complexity marks the difference between the brain of a human being and that of our nearest neighbours, the great apes. Apparently, this processing substrate and the associated gain of complexity marks the difference between species that failed and those that succeeded in promoting cultural evolution—with all its far reaching consequences. In this context, however, one needs to consider that cultural evolution created a socio-cultural environment of ever-increasing complexity that in turn contributes to the epigenetic shaping of brain architectures. Thus, even if the genetically-determined layout of brain architectures has changed little since the beginning of human civilisation, those features that can be modified by epigenetic shaping are likely to have undergone major modifications. This fact has not always been taken into account in the past; but its implications will be discussed below. But this additional twist concerns the epigenetic modifiability of our brains, and not its basic functional principles.

3 Contributions of neuroscience to philosophy

Once the neurosciences began to investigate the neuronal underpinnings of higher cognitive func-

tions, especially those realized in human brains, an increasing number of questions, traditionally investigated by the humanities, were addressed through empirical studies within the rapidly developing field of cognitive neuroscience. One obvious domain for this investigation was epistemology. Cognitive neuroscience explores from a third-person perspective the mechanisms that mediate our perception and the acquisition of knowledge. Longstanding discussions about the objectivity of cognition, the question of how constructive our perceptual processes really are, and how reliable or idiosyncratic they might be, need to be reconsidered on the basis of neurobiological data. Another question, to which the neurosciences will have to find an answer, is related to the mind-body problem: how can mental phenomena, namely immaterial entities such as the qualia of perception and social realities such as belief and value systems, emerge from the material interactions between nerve cells in human brains? These immaterial phenomena came into this world once the cognitive abilities of *Homo sapiens* initiated the evolution of cultures. They affect our lives as much as the material constraints of the world in which we evolve, but they have a different ontological status to the neuronal processes that brought them into this world. Yet another question that solicits discussion between neuroscientists and philosophers of mind is the nature of consciousness. The question of the constitution of the intentional Self is closely related to this issue, as is the conundrum of the existence of Free Will. If the material processes in individual brains and the social realities resulting from the interactions of humans are the basis and cause of mental phenomena, and if brain processes follow the known laws of nature, then there ought to be unifying description systems that bridge the gap between phenomena assessed from third- and first-person perspectives. If such approaches turn out to be feasible philosophical positions, the postulation of an ontological dualism will have to be modified. This will have far-reaching consequences for our self-understanding and the delineation of the border between “physics” and “metaphysics”.

3.1 An epistemic caveat

It is obvious that our perceptions and imaginations, as well as our ability to reason, are constrained by the cognitive abilities of our brains—and brains, like all other organs, are the product of an evolutionary process. Hence our brains have become adapted to the conditions of the mesoscopic world in which life has evolved. This is the world within the scale of millimeters to meters, it is the world where the laws of classical physics prevail; it is not the world of quantum physics and it is not the world of astrophysics. As a consequence, our cognitive functions have become adjusted to assure survival in this mesoscopic world. Problem-solving in this dangerous and poorly-predictable world requires the application of pragmatic heuristics and hence cognitive abilities that are in all likelihood not optimized to comprehend the essence behind the perceivable phenomena or the “absolute truth” in the Kantian sense. Evolution did not prepare us to directly perceive and understand processes at subatomic or cosmic scales, because they were and are completely irrelevant for our daily struggle for survival. Even more worrying is the possibility that the way in which we reason may also be limited by adaptation to those processes in the narrow range of the world that are relevant for survival and that we can access with our highly selective, specialized senses. In conclusion, it is very likely that our cognition is constrained. And this may apply not only to primary perception, but also to our way of deriving inferences from observables. If this were true it would pose unsurmountable barriers to our attempts to understand, just as it would challenge the consistency of mathematical theories and logical deductions. However, for these very reasons we have no way of knowing whether this is the case.

3.2 The contribution of neuroscience to epistemology

Growing insights into the neuronal mechanisms underlying perception provide compelling support for constructivist positions and emphasize the *epistemic caveats* formulated above. In the

light of neurobiological evidence, perceiving is essentially a constructive process. The sensory categories, for example those according to which we assign qualities to our experiences, are nothing but the idiosyncratic consequence of the layout of our sensory organs. These sample in a highly selective way a narrow range of physico-chemical signals, and this leads to the arbitrary classification of electromagnetic radiation with wavelengths between 400 to 700 nanometers as light, because the photoreceptors in the eye are sensitive to this wavelength range. Radiations with slightly longer wavelengths stimulate our temperature receptors and we categorise the respective sensations as temperature. A similar arbitrariness of category boundaries is observable in other sensory domains. The definition of perceptual objects, for example, is guided by a set of Gestaltrules that our brains apply in order to segment the spatio-temporal continuum of sensory signals into distinct objects—and this holds true for all sensory modalities. Objects are identified as such if they are delineated by spatial or temporal borders and exhibit some intrinsic coherence. This definition is appropriate in the mesoscopic world, but it does not apply to objects at atomic or subatomic scales. If we had no *a priori* definition of the properties of objects, we would not be able to distinguish objects, we would, for example, be unable to extract object-specific features from the continuous two-dimensional brightness distribution that cluttered scenes generate on the retina.

It is now well established by experimental evidence that the sparse sensory signals provided by our highly selective senses are interpreted by the brain on the basis of a vast amount of *a priori* knowledge that is stored in its own functional architecture. Our self-active brains permanently formulate knowledge and context-dependent expectancies, interpret sensory signals as a function of these inferences, and present the result of this constructive process to the workspace of consciousness. Paradoxically, we perceive the world around us as coherent even though our senses extract only a minute fraction of the available signals. Much of what we experience as actually perceived is read out from memory and is the result of reconstruction

and completion. This raises the question of the origins of this knowledge.

3.3 The sources of *a priori* knowledge

It is commonly accepted that all the knowledge a brain can possibly have, and the rules according to which this knowledge is applied for the interpretation of sensory signals and the execution of movements, reside in the functional architecture of the brain. This contradicts the analogy frequently drawn between computers and brains. Computers have processors and separate memories for programmes and for data. In the brain, however, there exist only neurons and connections. Both the stored knowledge and the programs for processing this knowledge reside in the layout of these connections, their polarity—that is, whether they are excitatory or inhibitory—and their graded efficacy. The question of the origin of stored information is thus reduced to the question of which processes determine the functional architecture of the brain.

The most important determinant of the functional architecture of brain—and hence the most important source of knowledge—is, of course, evolution. What makes our brain architectures comparable is evolutionary-acquired information that resides in the genes and determines the layout of the brain's connectome. It is knowledge about the world that is expressed in the functional architecture of brains every time an organism develops. In this sense evolution can be considered a cognitive process. This evolutionary-acquired knowledge pertains essentially to the conditions of the precultural world; and it is implicit—we are not aware of having it because we were not around when it was acquired. Still, we use it to interpret the signals provided by our sense organs and to structure adapted responses.

This inborn knowledge is subsequently complemented by extensive epigenetic shaping of the neuronal architectures, which adapt the developing brain to the actual conditions in which the individual develops. The human brain develops the majority of its connections only after birth, and this process continues approximately until the age of twenty or twenty-five

years. During this developmental period numerous new connections are formed, while many existing connections are removed; and this making and breaking is guided by the neuronal activity itself. Since, after birth, neuronal activity is modulated by interactions with the environment, the development of brain architectures is thus determined by a host of epigenetic factors derived from the natural and social environment. Through this process, the brain acquires knowledge about the specific conditions in which the newborn organism actually evolves, and thereby complements its genetically-inherited knowledge.

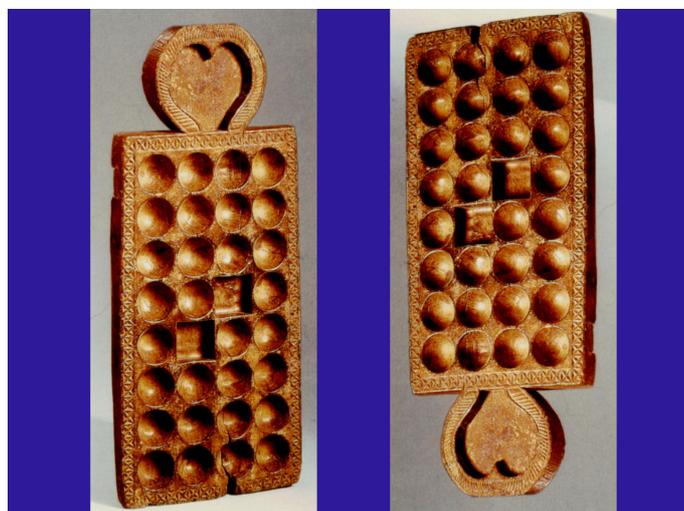


Figure 1: The brain assumes that light comes from above. The circular contours on the left board appear as concavities because the shadows are located at the right upper corner. The right board is actually the same as on the left, just rotated by 180 degrees. Now the shadows are on the lower left border and the contours appear convex.

A considerable part of this developmentally-acquired knowledge also remains implicit because of the phenomenon of childhood amnesia. Children before the age of about four years have only a limited capacity to remember in which context they have experienced and learnt particular contents. The reason for this is that the brain centres required for these storage functions—we call them episodic or biographical or declarative memories—have not yet matured. Thus, while young children learn very efficiently and store contents in a very robust way through structural modifications of their brain architec-

ture, they often have no recollection of the source of this knowledge. Because of this apparent lack of causation, the knowledge acquired in this way is implicit, similarly to evolutionary-acquired knowledge, and therefore often assumes the status of convictions that cannot be questioned.

Like innate knowledge, this acquired knowledge is used to shape cognitive processes and to structure our perceptions. Yet we are not aware that what we perceive is actually the result of such knowledge-based interpretations.

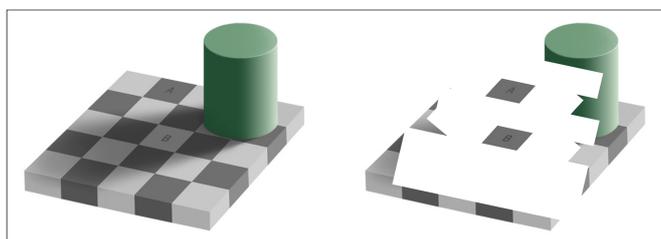


Figure 2: The checker-board illusion by Adelson, illustrating that even brightness perception depends on assumptions derived from context. (For further descriptions see text.)

Finally, there is knowledge acquisition by learning, which accompanies us throughout our lives. This is based on graded changes of the coupling strength of the existing connections between neurons. In the adult brain few new connections are formed and under normal conditions no breaking of connections occurs. The knowledge acquired by these learning processes also biases perception, but this is explicit, and its origins are known. One is usually aware of when and how it has been acquired and can therefore question its validity and by the same token the validity of what is perceived.

3.4 Examples illustrating the influence of priors on perception

The two examples depicted in [figure 1](#) and [figure 2](#) illustrate impressively how *a priori* knowledge structures our primary perceptions. The object in [figure 1](#) is a mould used to produce candies. On the left side one sees the inside of the mould, with its concavities, and on the left we see the rear side with its corresponding con-

vex protrusions. In reality, the pictures are identical, but one is rotated by 180°. The reason for our very different perceptions of the images is that the brain makes the *a priori* assumption that light comes from above—a well adapted assumption in a precultural world with only natural light sources. In this case contours that have the shadow above are interpreted as concave, and those with the shadow below as convex. Thus, an assumption of which we are not aware determines what we perceive. Another striking example is shown in [figure 2](#). It is hard to believe, but surfaces A and B have exactly the same luminance. They appear different because the brain sees the shadow that is caused by the cylinder on the right. Even though the amount of light reflected from surfaces A and B and impinging on the retina is exactly the same, the brain interprets the brightness of the two surfaces as different because it infers the following. Given that there is a shadow, surface B must be brighter than surface A—which has no shadow on it—in order to reflect the same amount of light. Thus, the brain “computes” the inferred brightness of the surfaces, but we are not aware of these computations. We just perceive the result and take it to be real, i.e., we see B as being much brighter than A. These two examples indicate that the brain generates inferences of which we are not aware, that it is permanently reconstructing the world according to *a priori* knowledge, and that we, as perceiving subjects, have to take for granted what the system finally offers us as conscious experience. As expected, this is not only the case with specially designed psycho-physical experiments, but is an essential feature of all our perceptual processes.

The mechanism leading to this “false” perception, to this “illusion”, has of course an important function. Our brain uses this principle to generate perceptual constancy, e.g., to keep colours and contrasts constant despite different illumination conditions. The spectral composition and the intensity of the sunlight change dramatically throughout the day, and therefore the spectral mix of light reflected from a particular, edible berry differs in the morning from that at noon. An animal that relies on colour to

distinguish one edible berry from another, slightly more violet and poisonous berry, cannot rely on an analysis of the “true” or actual spectral composition of reflected light. It first has to assess the spectral composition of the light source—the sunlight—and then must reconstruct the perceived colour. Our brains accomplish this by assessing the actual lighting conditions, by comparing the colours of the sky, of stones, of leaves and barks etc. and then, by using this contextual information, compute the “real” colour of the berries to identify that which is edible. Our brains are capable of assuring colour constancy despite changing illumination conditions, but we are completely unaware of the complexity of the computations assuring constancy and thereby survival in a changing world. In essence, all these operations are based on the evaluation of relations. We rarely perceive absolute values such as those measured by physical devices, be it intensities of stimuli, wave-lengths of sound or light waves, or chemical concentrations. We mostly perceive these variables in relation to others, as differences, increments and contrasts, such that these comparisons are made both across space and time. This is a very economical and efficient strategy because it emphasizes differences, permits coverage of wide ranges of intensities and, as mentioned above, allows for constancy. Given the advantages of these well-adapted heuristics it is at least questionable whether one should confront the resulting perceptions as illusions.

3.5 Conclusion of the excursion into epistemology

Evolution- and experience-dependent development determine and shape the architecture of the brain. Through these processes, knowledge about the world and strategies to use this knowledge for survival and reproduction are implemented in brain architectures. These in turn determine what and how an organism perceives and how it behaves. Because of the selection criteria that guides evolution, the brain adapted to the narrow segment of the world in which life has evolved, and its functions have been optimized to extract and process those signals that

best serve survival and reproduction. Thus the cognitive functions of the brain have probably not been optimized for understanding the deeper structure of the world that assures coherence across scales and cannot be perceived directly. Similarly, the rules according to which we evaluate contingencies and establish associations among events are implemented by specific molecular mechanisms that translate temporal correlations of neuronal activity into lasting changes in the efficacy of neuronal connections. These rules have been preserved virtually unchanged since the evolution of primitive nervous systems, and are at the basis of assignments of causality and the formation of associations. Again, these rules are highly efficient for the generation of models of the mesoscopic world and the formulation of predictions, but they do not apply to processes in the quantum world or to the relativistic dynamics of the universe. Given these specific adaptations of our cognitive functions, one might consider that similar restrictions may also hold for the way we reason. If so, this would present a serious challenge for the generalisation of models and theories based on extrapolation.

4 The contribution of the neurosciences to theories of consciousness

Some of the propositions summarized in the following chapter have been derived from experiments on the neuronal substrate of consciousness that have been described in detail in [Meloni & Singer \(2011\)](#) and [Aru et al. \(2012a, 2012b\)](#). A few decades ago, attempts to identify the neuronal correlates of consciousness (NCC) were considered futile. Because of the rapid development of non-invasive technology for the registration of neuronal activity in the human brain, and because of advances in the analysis of the neuronal underpinnings of higher cognitive functions, the search for NCC has now become a very active field of research in cognitive neuroscience. As expected, this new field is confronted with great challenges that are difficult to overcome. The explanandum is ill-defined; the prerequisites and consequences of conscious processing cannot easily be distinguished by ex-

periment from conscious processing per se; the experience of mental causation and agency is difficult to reconcile with contemporary concepts of self-organization; and finally it is difficult to bridge the epistemic gap between phenomena that are experienced from a first-person perspective and mechanisms described from a third-person perspective. Some of these problems will be addressed in the following paragraphs.

4.1 An attempt to define the explanandum

Most languages have coined a term for consciousness. Thus, it must be a robust phenomenon on which human beings can agree. However, while it is easy to use the term, it is virtually impossible to give a formal definition of what exactly it means. Nevertheless, the implicit understanding of what it is to be conscious seems to be sufficiently clear and widely accepted enough to justify a search for its neuronal correlates and, ultimately, to identify the neuronal mechanisms that enable a subject to be conscious of something. In their seminal paper, [Crick & Koch \(1990\)](#) propose that consciousness is a specific cognitive function and that, as such, it must have neuronal correlates that can be analyzed with the tools of the natural sciences. With the development of non-invasive imaging technologies, the tools became available to actually pursue this project and the search for the neuronal correlates of consciousness (NCC) became a mainstream endeavour.

Before discussing some of the proposed theories for NCC, I shall attempt to give an operational definition of what I mean when referring to awareness and consciousness or, in other terms, what it means to be aware of something or to be conscious. Subjects will be considered aware of something if they are able to report the presence or absence of the content of a cognitive process—irrespective of whether this content is made available by recall from stored memories or drawn from actual sensory experience. Thus, one criterion for awareness is the reportability of the presence of a cognitive content. These reports can in principle consist of any motor response, but to be on the safe side,

it is often requested that the report be verbal. The reason for this is that behavioural responses can be obtained under forced choice conditions that clearly indicate that the brain has processed and recognized the respective sensory material and produced a correct response even though the subject may not have been aware of having perceived the stimulus. There is thus an inherent ambiguity in non-verbal responses. They can but need not necessarily signal awareness, and this constrains research on NCC in animal experiments. Since consciousness is so difficult to define, an attempt will be made to avoid this term. Instead we shall use the adverb “consciously” and the adjective “conscious” in order to further specify particular brain states or aspects of a perceptual process. In addition, research into the hard problem of consciousness ([Chalmers 2000](#)) confronts the problem of explaining the phase transition from neuronal processes to the qualia of subjective experience, but this will be discussed only briefly at the end of this paper—and there in an enlarged context that transcends neurobiological approaches by also taking social interactions into account.

The state of being aware of something has a number of distinct properties that constrain the underlying neuronal mechanisms. One important feature of this state is unity or relatedness: Contents of which one is aware are experienced as simultaneously present and related to each other. Because of the distributed organization of brain processes, mechanisms supporting phenomenal awareness must therefore be able to bind together computational results obtained in multiple specialized and widely distributed processing areas. Another feature of awareness is that the contents that one is aware of change continuously but are bound together in time, appearing as a seamless flow that is coherent in space and time. Finally, subjects are only aware of a small fraction of on-going cognitive operations. Still, even signals of which subjects are not aware are often readily processed and impact behaviour ([Dehaene et al. 1998](#); [van Gaal et al. 2008](#)). Thus there must be gating mechanism that determines which signals are processed consciously, which are processed and control be-

haviour but remain unconscious, and which are excluded from processing altogether. Therefore, the identification of NCC requires a clear delineation between subconscious and conscious processes and an analysis of the mechanisms that gate access to awareness.

4.2 Conscious versus subconscious processing

As mentioned above, an enormous amount of knowledge is stored in the specific architectures of the brain, but we are not aware of most of these “given” heuristics, assumptions, and concepts. These routines determine the outcome of cognitive processes, which often have access to conscious recollection while themselves remaining hidden in the unconscious. We cannot move these implicit hypotheses and rules to the workspace of consciousness by focusing our attention on them, as is possible with most sensory signals and contents stored, for example, in “declarative memory”—the memory in which is inscribed what has been consciously experienced. Excluded from conscious experience are also certain sensory signals—such as those elicited, for example, by pheromones, which are processed by special olfactory subsystems—or the many signals from within the body—such as messages about blood pressures, sugar levels, and so on. It cannot be emphasized enough, however, that signals that are permanently excluded from conscious processing, as well as the facultatively-excluded signals from non-attended sensory stimuli, still have a strong impact on behaviour. Moreover, by influencing attentional mechanisms they can determine which of the stored memories or sensory signals will be transferred to the level of conscious processing. A hungry predator will search for traces of prey rather than mating partners, and so on.

One reason for the gated access of cognitive material to the level of awareness appears to be the limited capacity of the workspace of consciousness. Whether these limitations are due to the inability to attend to large numbers of items simultaneously, or whether they result from the restricted capacity of working memory, or even both, is subject to intense scientific investiga-

tion. The capacity of working memory is limited to about four to seven different items. The phenomenon of “change blindness”, which is the inability to detect local changes in two images presented in quick succession, demonstrates impressively our inability to attend to and consciously process all features of an image simultaneously. Because of these capacity constraints, conscious processing is in essence serial. Items are scrutinized and compared serially and therefore conscious processing is slow. Complex visual scenes are scanned serially and much of what we believe that we perceive simultaneously is actually reconstructed from memory. Which of the many signals finally reach the level of conscious awareness and can then be recalled depends on whether they are attended to, and this in turn is controlled either by external cues, such as the saliency of a stimulus, or by internal motifs, many of which we may actually not be aware of. And then it may occur that even an attentive, conscious search for content stored in declarative memory fails to raise it to the level of awareness. We are all familiar with the temporary inability to remember an episode or a name and have witnessed how a persisting subconscious search process suddenly lifts the content into the workspace of consciousness. It appears that we are not capable of controlling, at all times, which contents enter consciousness.

The differences between conscious and subconscious processes are further emphasized by evidence that the rules governing conscious deliberations and decisions most likely differ from those of subconscious processes. The former are based mainly on rational, logical, and syntactic rules, and the search for solutions is essentially based on serial computations. Arguments and facts are scrutinized one by one, and possible outcomes investigated. This strategy is suitable if variables are well defined, if sufficient time is available, if problems have a structure amenable to analytical treatment, and if precise solutions are required.

Subconscious mechanisms, by contrast, seem to rely more heavily on parallel processing, whereby a large number of variables enter into competition with one another. Then, a “winner takes all” algorithm leads to the sta-

bilization of the activity pattern that is the most likely, given the initial conditions and the heuristics derived either from inborn routines or from past experience. The domains of subconscious processing are situations requiring very fast responses or conditions where large numbers of underdetermined variables have to be considered simultaneously, and weighed against variables that have no or only limited access to conscious processing—such as the wealth of implicit knowledge and heuristics, vague feelings, hidden motives, or drives. The outcome of such subconscious processes manifests itself either in immediate behavioural responses or in what is called “gut feelings”. And it is often not possible to indicate with rational argument why exactly one has responded in such a way and why one feels that something is wrong or right. In experimental settings one can even demonstrate that the rational arguments given for or against a particular response do not correspond to the “real” causes. For the solution of complex problems with numerous entangled variables it often turns out that the subconscious processes lead to better solutions than conscious deliberations—and this is thought to be because of the wealth of heuristics exploitable by subconscious processing. Given the large amount of information and implicit knowledge to which consciousness has no or only sporadic access, and given the crucial importance of subconscious heuristics for decision-making and guidance of behaviour, first identifying the structure of a problem and then deciding whether one should rely on conscious deliberations or listen to the voices of the subconscious appears to be a well adapted strategy.

However, because the two systems operate according to different principles, the solutions to a particular problem may not always agree. Most of the decisions that get us through daily life rely on subconscious processing and follow well-adapted heuristics. If these decision processes do not lead to immediate action, they may still influence subsequent behaviour by manifesting themselves as what we call “gut feelings”. One has no conscious recollection of the reasons that lead to these feelings, but one clearly experiences the reactions of one’s

autonomous nervous system when the results of subconscious processes are in conflict with the outcome of conscious deliberations. In such situations one tends to say: “I decided according to all the rational arguments that I was aware of and took the best decision I could think of, but it somehow feels wrong.” The opposite situation is also possible, “I did what *felt* right to me, but if I think about it, it is absolutely crazy and irrational”. It is only when the two decision systems converge on the same solution one feels good, satisfied, and to some extent “free”.

After this brief excursion into the phenomenology of conscious and subconscious processes, intended to convey some connotations of consciousness, some of the most popular hypotheses about the constitution of consciousness in the brain will be reviewed.

4.3 Some competing hypotheses about the NCC

One class of theories focuses on the philosophical implications of the hard problem of consciousness without attempting to provide detailed descriptions of putative neuronal mechanisms (Searle 1997; Metzinger 2000; Dennett 1992; Chalmers 2000). Solutions to the hard problem have also been sought through transcending current concepts of neuronal processes and incorporating theories borrowed from other scientific disciplines. The most prominent of these approaches assumes that phenomena unravelled by quantum physics also play a role in neuronal processes, and that they might be able to account for the emergence of consciousness from material interactions in the brain (Hameroff 2006; Penrose 1994). None of the predictions of these theories are at present amenable to experimental verification, because there is no evidence that quantum phenomena such as entanglement, superposition and collapse of wave functions, etc., play a role at the macroscopic level of neuronal network functions. Quantum effects do of course exist at the level of molecular and submolecular interactions, but it appears highly unlikely that they are relevant for the macroscopic functions of neurons responsible for information processing. Thus quantum

theories of consciousness attempt to explain one poorly understood phenomenon with still unexplored and unproven mechanisms, and will therefore not be discussed further here.

Another class of theories pursues more modest goals and attempts to examine neuronal mechanisms potentially capable of supporting awareness of cognitive contents. Their aim is to define the neuronal mechanisms supporting the unitary character of awareness, its coherence in space and time, and the control of states that distinguish between conscious and unconscious processing (for review of relevant experimental findings see [Melloni & Singer 2011](#); [Aru et al. 2012a](#)).

The most intuitively plausible solution for the unity of awareness is convergence of the results obtained in distributed processing areas to a singular structure at the top of the processing hierarchy. Theories derived from this intuition predict the activation of specific cortical areas when subjects are aware of stimuli. Consequently, these regions should remain inactive during unconscious processing of the same material. Likewise, lesions of these putative areas should abolish the ability to become aware of perceptual objects. So far, a region with such universal “observer functions” has not been identified, and this option is considered theoretically implausible by some ([Dennett 1992](#)). There is also little—if any—experimental evidence for such a scenario. If brain lesions abolish the functions of sensory areas or regions involved in the recall of memories, patients lose the ability to consciously experience the respective sensory contents or memories, but the ability to process other material consciously remains unaffected. Moreover, behavioural and brain imaging studies have shown that unconscious processing engages very much the same areas as conscious processing, including the frontal and prefrontal cortices ([Lau & Passingham 2007](#); [van Gaal et al. 2008](#)). Thus there is no compelling evidence for specific areas supporting conscious processing. There are prominent examples of the selective elimination from conscious perception of those aspects of the stimulus material that are processed in specific regions without affecting awareness of other con-

tents, such as syndromes of agnosia and blindsight ([Cowey & Stoerig 1991](#)), which result from selective lesions of sensory subsystems.

There are, however, systems and pathways in the brain whose destruction abolishes all conscious experience—but these cannot be considered to be the NCC. Rather, these systems adjust the narrow dynamic range within which the brain has to be kept in order to be operational and to perform the computations that ultimately give rise to awareness. These systems are addressed as modulatory systems; they originate mainly in deep structures of the brain and control global brain states via widely-diverging ascending projections.

Another class of theories favours the notion that the mechanisms supporting awareness of stimulation material are distributed and do not require anatomical convergence ([Rodriguez et al. 1999](#); [Metzinger 2000](#); [Varela et al. 2001](#)). [Baars \(1997\)](#) and [Dehaene et al. \(2006\)](#) propose that there is a workspace of consciousness whose neuronal correlate is a widely distributed network of neurons located in the superficial layers of the cortical mantle. As mentioned above, these neurons are reciprocally coupled through a dense network of cortico-cortical connections that have features of small-world networks. The proposal is that subjects become aware of signals if these are sufficiently salient to ignite coordinated activity within this workspace of consciousness. This is assumed to be the case for signals that either have high saliency because of the high physical energy of the stimuli or those that are made salient due to attentional selection.

Yet another, related proposal is that subjects become aware of contents, irrespective of whether they are triggered by sensory events or recalled by imagery from stored memories, if the distributed neurons coding these contents are organized into assemblies characterized by coherent, temporally-structured activity patterns. In this case, the critical state variable distinguishing conscious from non-conscious processing would be the spatial extent and the precision of coherence of temporally-structured neuronal responses ([Rodriguez et al. 1999](#); [Metzinger 2000](#); [Varela et al. 2001](#)).

In what follows, evidence will be reviewed in support of the latter hypothesis. However, before discussing this evidence we should briefly recall the reasons why temporal coherence should matter in neuronal processing.

4.4 The formation of functional networks by temporal coordination

Because of the small-world architecture of the cortical connectome, any neuron can communicate with any other neuron either directly or via just a few interposed nodes. Thus, efficient and highly flexible mechanisms are required, which permit selective routing of signals and assure that only the neurons that need to interact in order to accomplish a particular task effectively communicate with one another. Evidence from multisite invasive recordings and from non-invasive registration of global activity patterns with magneto-encephalography or functional magnetic resonance imaging indeed indicates that functional sub-networks are configured “on the fly” on the backbone of fixed anatomical connections in a task- and goal-dependent way. One mechanism that can accomplish such fast and selective association of neurons and gate neuronal interaction is the temporal coordination of oscillatory activity (Gray et al. 1989; Fries 2005). Since the discovery (Gray & Singer 1989) that spatially-distributed neurons in the primary visual cortex tend to engage in oscillatory responses in the beta and gamma frequency band when activated by appropriately configured contours, and that these oscillatory responses can synchronize over large distances within and across cortical areas and even hemispheres, numerous studies have confirmed that oscillations and their synchronization in different frequency bands are an ubiquitous phenomenon in the mammalian brain. The pacemakers of these oscillations are reciprocal interactions in local networks of inhibitory and excitatory neurons. The long-distance synchronization of this oscillatory activity appears to be achieved by several mechanisms operating in parallel: Long-range excitatory cortico-cortical connections, long-range inhibitory projections, and pathways ascending from nuclei in the thal-

amus and the basal forebrain (for a review of these see Uhlhaas et al. 2009). When neurons engage in oscillatory activity, they pass through alternating cycles of high and low excitability. At the peak of an oscillation cycle neurons are depolarized, highly susceptible to excitatory input, and capable of emitting action potentials. In the subsequent trough of the cycle, the membrane potential is hyperpolarized and membrane conductance is high because of strong GABAergic inhibition generated by the rhythmically-active inhibitory interneurons. During this phase, neurons are less susceptible to excitatory inputs, because excitatory postsynaptic potentials (EPSPs) are shunted and because the membrane potential is far from the threshold. Hence, neurons are unlikely to respond to presynaptic excitatory drive.

These periodic modulations of excitability can be exploited in order to gate communication among neurons. By adjusting oscillation frequency and phases of coupled neuronal populations, communication among those neurons can either be facilitated or blocked. To form a functional network of distributed neurons it suffices to coordinate their oscillatory activity in such a way that signals emitted by neurons of this network impinge on other members of the network at times when these are highly susceptible to input. One way to achieve this is to entrain the neurons that should be bound into a functional network to engage in oscillations of the same frequency, to synchronize these oscillations, and to adjust the phases such that neurons that ought to be able to communicate can communicate.

Evidence from multi-site recordings indicate that neurons are indeed bound together into sometimes widespread functional networks through synchronization of their oscillatory activity in a task-dependent way (Salazar et al. 2012; Buschman et al. 2012). This supports the hypothesis (Gray et al. 1989; Singer 1999; Fries 2005) that synchronization of oscillatory neuronal activity is a versatile mechanism for the temporary association of distributed neurons and the binding of their responses into functionally coherent assemblies—which as a whole represent a particular cognitive content. Such a dy-

dynamic binding mechanism appears to be an economical and highly flexible strategy for coping with the representations of a virtually unlimited variety of feature constellations characterizing perceptual objects. Taking the unified nature of conscious experience and the virtually infinite diversity of possible contents that can be represented, the formation of distributed representations, through response synchronization, offers itself as a mechanism that allows for the encoding of ever-changing constellations of contents in a unifying format.

Synchronization is also ideally suited to contribute to the selection of contents for access to consciousness. It enhances the saliency of signals by concentrating spike discharges into a narrow temporal window. This increases the coincidence of excitatory postsynaptic potentials (EPSPs) in target cells that receive input from synchronized cell groups. Because coincident EPSPs summate much more effectively than temporally dispersed EPSPs, synchronized inputs are particularly effective in driving postsynaptic target cells. It is thus not unexpected that entrainment of neuronal populations in synchronized gamma oscillation is used for attention-dependent selection of input configurations (Fries et al. 2001b; Fries 2009).

4.5 A prediction relating long-range synchronisation to consciousness

If activation patterns that subjects can become aware of are indeed characterized by globally coherent states of those cortical regions that process the contents actually appearing as unified, we expect these states of awareness to be associated with large-scale synchronization of neuronal activity. Candidate frequency bands are gamma and beta oscillations, as these have been shown to serve the temporal coordination of cortical networks. By contrast, if subjects are not aware of the presented stimulus material, processing should remain confined to smaller sub-networks, which operate in relative isolation and are not integrated into globally coherent states. In this case one should observe only local synchronization of more circumscribed neuronal populations (see Varela et al. 2001).

Finally, adjustments of oscillation frequency and phases fulfil the requirement that assemblies representing consciously processed contents need to be reconfigured at an extremely fast rate. The contents of which subjects are aware can apparently change at a rapid pace, at least four times a second, if one considers that this is the frequency with which the direction of gaze changes during the scanning of natural scenes. Thus, assemblies representing contents that are consciously perceived must be reconfigurable at similarly fast timescales. Evidence suggests that cortical networks operate in a regime of self-organized criticality close to the edge of chaos (Shew et al. 2009). Dynamical systems operating in this range can undergo very rapid state changes, which are characterized by shifts in oscillation frequencies, synchronization, and phase.

4.6 Methodological caveats in the search for the NCC

As discussed previously (Aru et al. 2012a) experiments designed to identify the neuronal correlates of consciousness are often fraught with ambiguities. The most frequently used strategy for the identification of NCC is contrastive analysis. One creates perceptual conditions in which targets are consciously perceived in only a subset of trials, while making sure that physical conditions are kept as constant as possible. This strategy implies that detection tasks have been designed, which operate close to the perceptual threshold. This can be achieved by reducing the physical energy of the stimuli or by masking them. While subjects are engaged in such detection tasks, neuronal responses are measured and then trials are sorted depending on whether the subjects did or did not perceive the stimulus. By subtracting the average responses obtained in the two conditions from one another, those neuronal responses that occur only in the condition of successful detection can be isolated, and these are then commonly interpreted as the neuronal correlate of conscious perception. This seemingly simple approach is not without ambiguity. Thus, noise fluctuations in afferent pathways are likely to lead to signi-

ficant differences in the available sensory evidence, especially because experiments are performed at the perceptual threshold. Therefore, those aspects of neuronal responses that truly reflect NCC may be contaminated by signals resulting from noise fluctuations at processing stages preceding those actually mediating awareness. In addition, once subjects have become aware of stimuli, there are a number of subsequent processing steps that need not necessarily be linked to NCC. These comprise the covert verbalization of stimulus material, the engagement of working memory, the transfer of information into declarative memory, and perhaps also the preparation of covert motor responses. The distinction between these various confounding factors is difficult because all the processes are intimately related to each other. A detailed discussion of this problem is given in [Aru et al. \(2012a\)](#). One distinguishing feature could be the latency of the electrographic signatures of these various processing steps. Noise-dependent fluctuations in sensory evidence should be manifested early on; responses related to NCC proper should have some intermediate latency; and the consequences of having become aware of a stimulus should have the longest latencies. In order to use these latencies as distinguishing criteria, it is of course required that we estimate the precise latency at which the mechanisms leading to conscious perception are likely to be engaged. Assuming that the time required to prepare and execute simple motor responses is generally constant, the interval of interest can be constrained and has been proposed as somewhere between 180 and a few hundred milliseconds, depending on the sensory modality and the difficulty of the detection task. Attempts to use latency criteria for the elimination of confounds is of course restricted to electroencephalographic and magneto-encephalographic data, and cannot be applied to results obtained with functional magnetic resonance imaging because of the limited temporal resolution of this technique.

Another option for the reduction of confounds is to combine manipulations that influence the conscious perception of a stimulus through different mechanisms and to compare

the electrographic responses between conditions ([Aru et al. 2012b](#)). We applied this strategy in investigations of patients with subdurally implanted recording electrodes located over the visual cortex. In one set of trials the visibility of stimuli, in this case faces, was manipulated by changing the sensory evidence of the stimulus material. In another set of trials visibility of the same stimuli was influenced by allowing the subjects to familiarize themselves with some of the stimuli. This also facilitated detectability, but now because of an expectancy-driven top-down process. The reasoning behind this was that neuronal responses reflecting NCC proper should be the same irrespective of whether stimuli are consciously perceived because of enhanced sensory evidence or because of top-down facilitation. As electrographic signature of interest we analyzed the neuronal activity in the gamma band. In a previous study ([Fisch et al. 2009](#)) had shown that category-specific gamma band responses in the visual cortex correlate with conscious perception. Conscious recognition leads to a phasic enhancement of the gamma-band response, thus supporting the notion that conscious perception arises locally within sensory cortices—which is in line with previous conclusions ([Zeki 2001](#); [Malach 2007](#)). In our study we found that the performance and the reports of the subject were clearly modulated both by changing sensory evidence *and* by prior knowledge of the stimuli, as expected; but the gamma-band responses solely reflected the sensory evidence. This suggests that the differential activation of specific areas of the visual cortex, in our case mainly the fusiform face area, reflect processes that prepare access to conscious perception but are not its substrate proper.

Another frequently used paradigm in the search for NCC is interocular rivalry. If the two eyes are presented with stimuli that cannot be fused into one coherent percept, subjects perceive only one of the two stimuli at a time, and these percepts alternate. There are various ways to label the stimuli presented to the two eyes, to trace the responses related to their processing in the brain, and then to see which brain structures have to get involved in order to

support conscious perception. Again, these studies have led to inconclusive results. Some claim that suppression of signals corresponding to the non-perceived stimulus occurs only at very high levels of visual processing, such as, for example, the temporal cortex, which is the highest stage of the ventral processing stream. The conclusion of these studies is that activation of this particular cortical network is a necessary prerequisite for conscious processing (Logothetis et al. 1996; Silver & Logothetis 2004). Others, by contrast, found diverging activity patterns already existing at the level of the thalamus and the primary visual cortex (Haynes et al. 2005; Fries et al. 1997). Recent correlations between the dynamics characterizing binocular rivalry and anatomical features of the primary visual cortex and the commissures linking the primary visual cortices of the two hemispheres provide compelling evidence that the rivalry phenomenon is based on processes occurring within V1 (Genc et al. 2014). However, none of these studies allows us to unambiguously locate the processes that lead to conscious perception. They only contribute to the identification of the earliest levels of processing, in which changes are detectable that correlate with conscious perception.

Interesting and of potential relevance for interpretations given in the next subsection is the observation that the access of sensory signals to conscious processing does not seem to be gated by modulation of the neurons' discharge rate, but rather by changes of the synchronization of their activity—at least in the early stages of processing. What matters is the degree of synchronicity of oscillatory activity in the gamma frequency range. Signals conveyed by well-synchronized neuronal assemblies have access to conscious processing, while signals conveyed by similarly active but purely synchronized neurons fail to do so (Fries et al. 2001c). Of interest in this context is the observation that stimuli access conscious perception more easily if they are attended to, and that attention enhances synchronization of neuronal responses in the gamma frequency band in early visual areas (Fries et al. 2001b). Again, however, this local increase in synchrony is likely to simply enhance

the saliency of the neuronal responses, facilitating their propagation across the cortical networks, and cannot per se be considered a neuronal correlate of consciousness.

4.7 Evidence relating long-range synchronization and consciousness

The results described in what follows were obtained in a study where we presented words that could be perceived in some trials and not in others (by adjusting the luminance of masking stimuli), and simultaneously performed electroencephalographic (EEG) recordings (Melloni et al. 2007). Several measures were analyzed: Time-resolved power changes of local signals; the precision of phase synchronization across recording sites, and over a wide frequency range; and event-related potentials (ERPs). A brief burst of long-distance synchronization in the gamma-frequency range between occipital, parietal, and frontal sensors was the first event that distinguished seen from unseen words at about 180ms poststimulus. In contrast, local synchronization was similar between conditions. Interestingly, after this transient period of synchronization, several other measures differed between seen and unseen words: We observed an increase in amplitude of the P300 ERP for visible words, which most likely corresponds to the transfer of information to working memory. In addition, during the interval period—in which visible words had to be maintained in memory—we observed increases in frontal theta oscillations. Theta oscillations have been related to maintenance of items in short-term memory (Jensen & Tesche 2002; Schack et al. 2005).

To test whether the increase in long-distance synchronization relates to awareness or depth of processing, we further manipulated the depth of processing of invisible words. It has previously been shown that invisible words can be processed up to the semantic and motor level (Dehaene et al. 1998). In a subliminal semantic priming experiment we briefly presented words (which were thus invisible) that were either semantically related or unrelated, alongside a second, visible word on which subjects had to carry out a semantic classification task. Invis-

ible words were processed up to semantic levels, as revealed by modulation of the reaction times, depending on the congruency between invisible and visible words: congruent pairs exhibited shorter reaction times than incongruent ones. We observed increases in power in the gamma frequency range for unseen but processed words. For visible words we additionally observed increases in long-distance synchronization in the gamma-frequency range (Melloni & Rodriguez 2007). Thus, local processing of stimuli is reflected in increases in gamma power, whereas long-distance synchronization seems to be related to awareness of stimuli. This suggests that conscious processing requires a particular dynamical state of the cortical network. The large-scale synchronization that we observed in our study could reflect the transfer of contents into awareness and/or their maintenance. We favour the first possibility, given the transient nature of the effect, and argue that the subsequent theta oscillations might support maintenance. It is conceivable that short periods of long-distance synchronization in the gamma band reflect the update of new contents, while the slower pace of theta oscillations might relate to sustained integration and maintenance of local results in the workspace of consciousness. The interplay between these two frequency bands might underlie the phenomenon of continuous but ever-changing conscious experience (see below).

More recently, Gaillard et al. (2009) have revisited the processing of visible and invisible words. In intracranial recordings in epileptic patients, they observed that invisible words elicited activity in multiple cortical areas, which quickly vanished after 300 ms. In contrast, visible words elicited sustained voltage changes, increases in power in the gamma band, as well as long-distance synchronization in the beta band and long-range Granger causality. In contrast to our study, Gaillard et al. observed a rather late (300–500 ms) rise of long-distance synchronization. However, it is important to note that in the study undertaken by Gaillard et al., phase-synchrony was analyzed for the most part over electrodes within a given cortical area, or at most between hemispheres. It is thus conceiv-

able that earlier synchronization events passed undetected because of incomplete electrode coverage. Despite these restrictions, this study provides one of the most compelling pieces of evidence for a relation between long-distance synchronization and consciousness.

Some results of the experiments on binocular rivalry point in the same direction. Several studies have shown increased synchronization and phase locking of oscillatory responses to the stimulus that was consciously perceived, and controlled behaviour (Cosmelli et al. 2004; Doesburg et al. 2005; Fries et al. 1997; Srinivasan et al. 1999). Cosmelli et al. (2004) extend the findings obtained in human subjects by performing source-reconstruction and analyzing phase-synchrony in source space. These authors observed that perceptual dominance was accompanied by co-activation of occipital and frontal regions, including the anterior cingulate and medial frontal areas. Recently, Doesburg et al. (2009) have provided evidence for a relation between perceptual switches in binocular rivalry and theta- and gamma-band synchronization. Perceptual switches were related to increments in long-distance synchronization in the gamma band between several cortical areas (frontal and parietal) that repeated at the rate of theta oscillations. The authors suggest that transient gamma-band synchronization supports discrete moments of perceptual experience, while theta oscillations structure their succession in time, pacing the formation and dissolution of distributed neuronal assemblies. Thus, long-range gamma synchronization locked to on-going theta oscillations could serve to structure the flow of conscious experience, allowing for changes in content every few hundred milliseconds. Further research is required to clarify the exact relation between the two frequency bands, their respective role in the generation of percepts, and the pacing of changes in perception.

Another paradigm in consciousness research exploits the attentional blink phenomenon. When two stimuli are presented at short intervals among a set of distractors, subjects usually detect the first (S1) but miss the second (S2) when it is presented 200–500 ms after S1. Increases in long-range neuronal synchrony in the beta and gamma frequency ranges have been observed when the S2

is successfully detected (Gross et al. 2004; Nakatani et al. 2005). Furthermore, Gross et al. (2004) observed that successful detection of both S1 and S2 was related to increased long-distance synchronization in the beta range to both stimuli, and this enhanced synchrony was accompanied by higher de-synchronization in the inter-stimulus interval. Thus, de-synchronization might have facilitated the segregation of the two targets, allowing for identification of the second stimulus (also see Rodriguez et al. 1999). Source analysis revealed, as in the case of binocular rivalry, dynamical coordination between frontal, parietal, and temporal regions for detected targets (Gross et al. 2004).

In summary, studies of masking, binocular rivalry, and the attentional blink support the involvement of long-range synchronization in conscious perception. Recent investigations have suggested further that a nesting of different frequencies, in particular of theta and gamma oscillations, could play a role in pacing the flow of consciousness. Furthermore, the study of Gross et al. (2004) suggests that de-synchronization could serve to segregate representations when stimuli follow at short intervals. These results are encouraging and should motivate further search for relations between oscillatory activity in different frequency bands and consciousness, whereby attention should be focused not only on the formation of dynamically-configured networks but also on their dissolution.

4.8 Conclusions on putative mechanisms supporting consciousness

Of the numerous proposals on NCC, those favouring temporal coherence as the mechanism that integrates widely distributed processes appear to be the least controversial. They account best for the apparent discrepancy between the unity of conscious experience and the distributed organization of the brain, because they allow for the dynamic integration of information generated in parallel by spatially segregated processing modules. Large-scale synchronization of oscillatory activity has been identified as a candidate mechanism for the flexible coupling of widely-distributed neuron populations, and hence as a likely NCC. This variable has the advantage that it can be meas-

ured relatively directly in humans who are able to give detailed descriptions about their conscious experience. However, oscillations and synchrony seem to be mechanisms that are as intimately and inseparably related to neuronal processing in general as the modulation of neuronal discharge rates. Thus, without further specification these phenomena cannot stand up as NCC—at least when we disregard the triviality that consciousness does not exist without them. We propose that the spatial scale and the precision and stability of neuronal synchrony might be taken as more specific indicators of whether the communication of information in the brain is accompanied by conscious experience or not. In this framework, conscious experience arises only if information that is widely distributed within or across subsystems is not only processed and passed on to executive structures but also bound together into a coherent, all-encompassing, non-local but distributed meta-representation. This interpretation is compatible with views that take consciousness to be the result of the dynamic interplay of brain subsystems; one that allows for a rapid and highly-flexible integration of information provided by the numerous distributed subsystems that operate in parallel. This view resembles a proposal from Sherrington, formulated in his book *The Integrative Action of the Nervous System* (Sherrington 1906): “[p]ure conjunction in time without necessarily cerebral conjunction in space lies at the root of the solution of the problem of the unity of mind.” The additive value of conscious processing would then be the possibility of establishing in a unified data format ever-changing relations between cognitive contents—irrespective of whether they are read out from memory or induced by sensory signals, and irrespective of the sensory modality providing the signals. By virtue of this dynamic definition of novel relations, non-local meta-representations of specific constellations could be established that have the status of cognitive objects. Just as with any other distributed representation of contents, these could then be stored as distributed engrams by use-dependant modification of synaptic connections, and thus influence future behaviour. Thus, conscious processing would differ from non-conscious processing because it allows for the versatile binding

of the previously unbound into higher-order representations. And if so, “conscious” processing would be functionally relevant and not merely an epiphenomenon. Further arguments supporting the functional role of conscious processing are presented in the following section.

5 Arguments supporting an adaptive value of conscious processing

5.1 Evolutionary considerations

Given the continuity of evolution and the gradual increase in complexity of brains that reached a (perhaps preliminary) maximum in human beings, it appears likely that the ability to be aware of cognitive contents, of one’s own cognitive operations, and finally of oneself as an intentional agent, is not an all-or-none phenomenon but an ability that gradually emerged as we evolved. Paleoanthropological evidence supports this hypothesis by documenting correlations between increasing brain volume and increasingly refined artefacts that reflect gradual increases in cognitive abilities. Not much direct data are available on the evolution of the human brain, because our immediate ancestors are all extinct. Thus, we have to rely on evidence of comparative studies with brains of other species. In less evolved brains the paths from sensory to executive areas of the cerebral cortex are short. These relatively short sensory-motor loops are of course much more elaborate than simple reflex loops, because signals are processed extensively and transmission is made conditional on input from other systems, on past experience, and on context. As evolution proceeds and brains become more and more complex, one observes the addition of new cortical areas. A distinctive feature of this evolutionary process is the way in which these new areas are embedded in already-existing networks. These newly-added areas no longer communicate directly with the periphery, neither on the executive nor on the receptive side (see [figure 3](#)). Instead they receive their input exclusively from the phylogenetically older areas, and also distribute their computational results solely to other cortical areas and not to effector systems. This process is iterative, with more and more areas constantly

added that communicate only with other higher areas. A neuron located in these more recent areas is connected exclusively with other partners in the cerebral cortex. Evidence indicates that all cortical areas, including older and more recent ones, operate according to the same basic principles, because they share the same intrinsic organization. These purely anatomical considerations suggest that the evolutionary-recent areas process the results of the older areas similarly to how they process signals from the outer world. This iteration of “cognitive” processes across several hierarchical levels could thus generate representations of representations, i.e., meta-representations. Information that has already been processed by the already existing areas becomes the object of yet another cortical computation, i.e., of a secondary cognitive operation. These iterative operations can even be circular, because all these areas are interconnected reciprocally. Thus, higher-order areas feed back to lower, order areas and can have their results reprocessed. In principle this recursive process should permit the generation of meta-representations of increasingly higher order. In other words, highly evolved brains can apply their cognitive functions not only to the outer world but also to processes that occur within the brain. Brain processes can therefore become the object of the brain’s own cognitive operations. This could be the basis of phenomenal awareness, the awareness of perceiving, to create a protocol for what one perceives, and in the case of human beings to create symbols for the perceived and for internal states and to communicate them to others. Animals probably share some of these abilities, because their brains are organized in a very similar way.

Curiously, the ability to be aware of the results of cognitive operations provides no clue concerning the computational processes underlying these cognitive functions. We have no insight into the neuronal processes that bring about cognition. We are aware only of the results—just as we are aware of an action without being able to tell which neuronal processes in the motor centers of our brains caused this action. This fact is at the origin of most discrepancies between our intuitions and neurobiological evidence concerning the nature of agency, the experience of Free Will, and

the ontological status of qualia and consciousness (see [final](#) paragraph).

These evolutionary considerations may provide some plausible explanation of the emergence of higher cognitive functions—including consciousness—but they do not suffice to counter the argument that the emergence of consciousness is an epiphenomenon without adaptive value. To address this problem, one must identify functions that can only be realized by conscious processing.

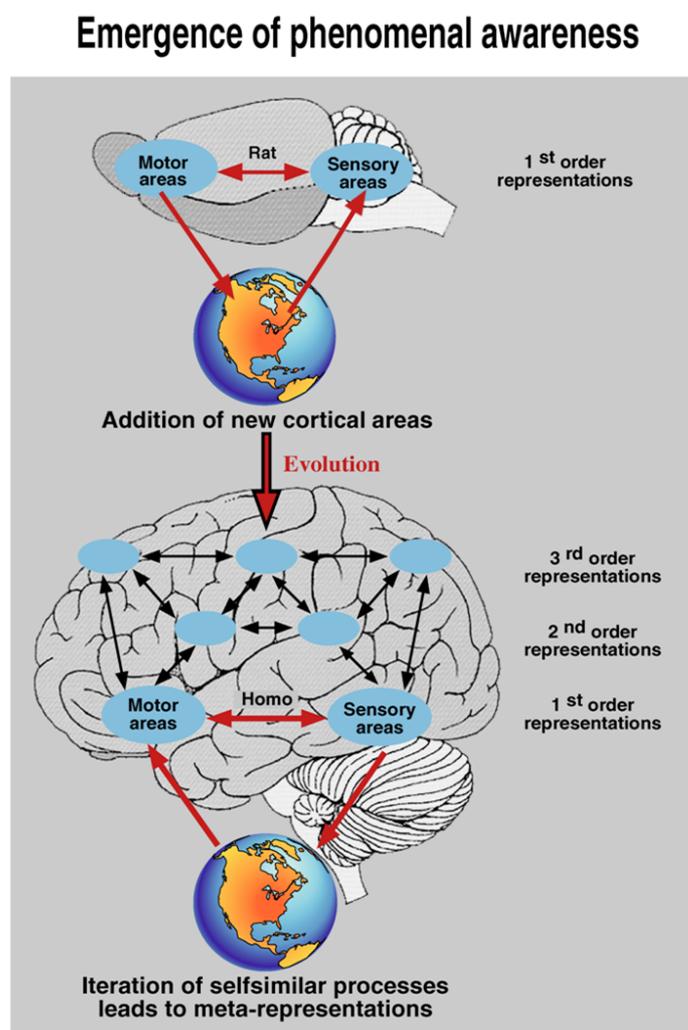


Figure 3: The evolution of complex brains is characterized by a massive increase of cortical areas. This renders responses to stimuli increasingly dependent on intracerebral processes and permits generation of meta-representations.

5.2 Functional considerations

Is the ability to be conscious of one’s own cognitive operations a fitness factor? And would it

make any difference if brains lacked this ability? The philosopher of mind, David Chalmers, once stated: “[n]o, that wouldn’t make a difference. It’s just an epiphenomenon and if we wouldn’t have it we would do as well because the underlying brain processes would be the same and get us through life without us having to be aware of them.” I tend to disagree with this view for a number of reasons. The common denominator, however, is that there is something very special about the nature of conscious processes, and that this uniqueness does indeed constitute a fitness factor—in particular with respect to the ability to develop symbolic communication systems, a theory of mind, differentiated social systems, and, ultimately, culture.

As has been argued above, conscious processing allows for abstraction and symbolic representations due to its versatile binding of virtually all results of lower order cognitive operations in a unified representational space, capitalizing on the lingua franca (the homogeneous data format) of communication among cortical areas. This permits implementation of very effective—we call them rational—strategies for deliberation and decision-making that differ from and complement those of subconscious processes. The question then remains whether the results of these special processes affect brain functions and thereby affect behaviour.

It is difficult to see how the outcome of a conscious deliberation, that is, of an argument-based decision, could not affect future behaviour. A conscious decision leaves traces in declarative memory and so must the consequences of this decision. These traces are inscribed as modifications of the functional architecture of the respective networks. If a decision has averse or beneficial consequences, the experience of these consequences will also alter network properties, and the novel activation patterns generated by the modified networks will enter as a novel argument, or as a change of goal in future deliberations. Eventually, the newly-set goal, which initially had the status of a conscious rational argument, may change its status and become a habit that henceforth influences behaviour without having to appear as an explicit argument in consciousness. It can become one of

the variables that act at a subconscious level. One then refuses another glass of wine not because one recapitulates all the rational arguments against alcohol consumption but simply because it does not feel right to drink too much.

Another, and probably the most important fitness factor of conscious processing is its capacity to support complex societies in cultural evolution. This suggests that there may have been a co-evolution of mechanisms supporting the emergence of conscious behaviour on the one hand and the formation of societies on the other, with the two developments mutually supporting each other.

Cognitive objects represented in consciousness are always bound together into a coherent experience. Whether this is also the case for subconsciously-processed contents is obviously difficult to ascertain, but the following argument suggests that subconscious processing may be less integrated, more modular, and confined to subsystems. We subconsciously orient towards salient stimuli irrespective of their modality, visual, auditory, or tactile qualities, and these stimuli may be analyzed in the subconscious with respect to their behavioural relevance and thus give rise to action. However, if several stimuli compete for processing, usually the most salient will win. Compared to orienting responses, which are guided by conscious processing, there seems to be little evaluation of the embedding context here. Stimuli are processed more independently than they would have been had they entered consciousness, been bound together, and formed a unified coherent percept.

In order to achieve this integration and to relate the signals from various sensory systems to one another at the semantic level, the signals have to be encoded in a sufficiently abstract and homogenous format. As mentioned above, when describing brain evolution, the substrate for this integration of signals preprocessed by segregated sensory systems may be the evolutionary-recent cortical areas. By virtue of integrating and comparing signals from different modalities, it becomes possible to detect the similar in the seemingly different, and hence to extract invariant properties and to arrive at ab-

stract descriptions. Thus the addition of the novel, so called “association areas” of the neocortex prepared the ground not only for a more unified, polymodal representation of cognitive contents but also for symbolic coding—which in turn is a prerequisite for the development of a symbolic language system and abstract reasoning. Thus one might consider consciousness, or the state of conscious processing, as a state where distributed computational results can be bound together into a coherent whole, establishing multiple, simultaneous relations between the various distributed items. This obviously allows for a more abstract, more symbolic, and more comprehensive description of conditions. By itself this is an advanced processing strategy, whose adaptive or fitness value is obvious. However, if this unified, condensed, and abstract information can be routed to a versatile communication system, as seems to be the case with contents that are processed consciously, the evolution of cooperating societies will be greatly facilitated. Not only will it be easier to communicate what one has perceived if the numerous signals from different sensory modalities have already been bound together into coherent wholes but, because of the reflexive nature of awareness, it will also be easier to convey information about one’s internal state. This, in turn, is an important prerequisite for society-building, because it nurtures trust in and predictability of the respective other. A condensed report of the actual contents of one’s conscious state and its storage in the episodic memory of the listener is an effective and parsimonious way to couple brains with one another, to share experiences, and to foster cooperation. This interaction will modify the brains of the communicating partners and thereby act on their behaviour. In a sense this is an example of mental or top-down causation. The results of an information-processing strategy that can only be realized in the workspace of consciousness are stored in declarative memory—in this case not only in a single brain but in those of communicating partners—and henceforth influence future cognitive processes and behaviour. These considerations suggest that it might be useful and perhaps even necessary in consciousness re-

search to consider the phenomena ascribed to consciousness not solely in the context of cognitive functions of individual brains but in the larger context of social interactions. In the following section this strategy is applied in an attempt to approach the “hard problem of consciousness”.

6 Consciousness as a social phenomenon

6.1 Is the hard problem resolvable by considering cultural rather than only biological evolution?

The hard problem in consciousness research, the epistemic difficulty of devising bridging theories between subjective phenomena available only from a first-person perspective and the underlying neuronal mechanisms analyzable only from the third-person perspective, may not be resolvable by considering only the cognitive abilities of isolated brains. In addition, more recent concepts of embodiment that consider the embedding of the nervous system in a body endowed with receptor and effector organs may not suffice. Rather, it may be necessary to reconsider the problem in the context of social phenomena or social realities that emerged during cultural evolution. Through social interactions, realities have been created that can readily be experienced as such but that transcend the reality that existed before humans added cultural to biological evolution. These new realities have the quality of relations. They are immaterial, not tangible, not visible, not directly accessible to our senses—and yet they are perceived as real, as mental objects that humans can agree upon, that can become the object of shared attention and influence behaviour—just like the equally immaterial contents of belief systems.

What seems to pose the epistemic—the hard—problem in philosophy of mind is not so much that we perceive and have feelings and emotions, since we readily grant such abilities to animals and seem to be quite successful in identifying the neuronal substrate of these functions. The real problem appears to be our meta-awareness of having these abilities. It is

this meta-awareness that has mental connotations that appear to be so difficult to relate to neuronal processes. However, this meta-awareness—which is so intimately related with what we address when we talk about consciousness—in all likelihood does not emerge naturally from the functions of an individual brain. Rather, it appears to be the consequence of experiences resulting from social interactions, just as is the case for the experience of individuality, agency, and intentionality. These are attributions that have their roots in interpersonal interactions and are probably appropriated by individuals while they are developing their self-image. Without being embedded in a differentiated socio-cultural environment, without the option of mirroring oneself in the perception of others, without reflexive interactions between persons endowed with a theory of mind, and without an exchange of reports about inner states, formulated in a symbolic language-system, we would probably not be aware of our being conscious.

Thus the phenomenon that we call consciousness has the ontological status of a social reality. It is a construct, just like all the other social realities that our cultures have brought into this world. However, this construct differs in an important respect from other social realities, such as norms, beliefs, and values, in that it is an attribution that we ascribe to ourselves. We learn from social interactions that we are endowed with the ability to be aware of being aware. After conceptualization of this novel experience it becomes an integral part of our self-image—we exchange reports on this shared experience and coin words for it.

Regarded in this way the “hard problem” may not be as hard as it seems. Analyzing individual brains will not unravel the “correlates” of the many semantic connotations of what we term consciousness, because they emerged from a reflexive interaction *between individuals*. But we should eventually be able to identify the neuronal mechanisms underlying the cognitive abilities that allowed human beings to create the socio-cultural environment that itself allowed for the experience of being conscious.

7 Culture-specific, epigenetic shaping of brains and the concept of tolerance

As suggested in the preceding section, inclusion of the social dimension may be indispensable for defining the status of higher cognitive functions such as consciousness. But it may also help us to bridge gaps between naturalistic approaches and first-person phenomenology. Thus, surprisingly, the joint consideration of neurobiological evidence on the dependence of perception of priors and of the epigenetic modification of priors by socio-cultural factors has normative consequences for concepts of tolerance.

Human beings have a similar genetic outfit, and therefore usually agree on what they perceive—in particular when priors suffice for the interpretation of the perceived, which had been acquired during evolution in a pre-cultural world. However, this may not apply to the perception of social realities. Humans raised in different cultures may perceive social realities quite differently. Cultures may differ radically with respect to social conditions, concepts of fairness, justice, and aesthetics, as well as moral criteria, and so on. Thus, epigenetically-installed priors are likely to exhibit considerable culture-specific differences. As a consequence, the ways in which the world is perceived will also differ. Another culture specific variable likely to influence perception is the attachment to caretakers. The nature of early bonding experiences determines whether the world appears as hostile or as peaceful and secure, and whether others can be approached with confidence or scepticism. The ways in which these early attachments to other members are secured differ dramatically among different cultures and, as such, so will the perception of signals that inspire confidence or aggression. Since these early-acquired priors are implicit, and since one is unaware that perceptions are influenced by these early imprinting processes, human beings take what they perceive as “real” and see no reason to question its validity. Thus, two persons raised differently, while observing the same social situation may perceive it in completely different ways. They may come to grossly diverging ethical or moral judgments,

unable to convince the other through argument that he or she is wrong, because both experience what they experience as reality—just as one experiences optical illusions as real. The problem is that, in the case of the perception of social realities, there are no “objective” measuring devices. There are different perceptions, and there is no right or wrong. This has far-reaching consequences for our concepts of tolerance. Solving such problems with majority votes is clearly not a fair solution. Assuming that one’s own position is correct and granting others the right to retain their “wrong” perceptions—so long as they do not disturb the peace—is humiliating and denies them respect. Still, this is often considered to be tolerant behaviour. What should be done, instead, is to grant everybody that her or his perceptions are correct and to postulate that this attitude be reciprocated. Only if this agreement on reciprocity is violated have the dissenting parties the right to exert sanctions.

8 Concluding remarks

In this paper only a selection of concepts currently under discussion in the field of cognitive neuroscience could be reviewed in any detail. The selection criterion was their relatedness to consciousness studies and to some extent their relevance for epistemic issues. This led to the neglect of important domains, some of which have equally relevant bearing on philosophical questions and societal issues. Thus, no consideration was given to developmental aspects that illustrate the close correlation between the maturation of particular brain structures and the emergence of specific cognitive and social functions. Equally neglected were research on executive functions and associated concepts on intentionality, agency, mental causation, and free will. No attempts were made to explore the vast field devoted to the analysis of genuinely “psychic” phenomena, such as memories, emotions, motivations, drives, and aesthetic judgement. Finally, a much more elaborate review would have been required of research that analyzes brain functions that can only be assessed in conditions where brains interact with one an-

other in a social context. This is the domain of the relatively recent field of social neuroscience, which studies phenomena such as confidence, greed, generosity, fairness, parasitism, altruism, compassion, and collective beliefs. This would have been particularly important in view of the attempt made at the end of this contribution to demystify some of the immaterial connotations of consciousness, in particular by redefining some aspects of consciousness as belonging to the domain of social realities, i.e., as a relational construct emerging from reciprocal interactions among brains endowed with the cognitive abilities of humans. The research agendas of these “neglected” fields are the same as those of the domains reviewed above. They all attempt to identify the neuronal mechanisms that are responsible for a particular cognitive or executive function, such that it becomes less and less important whether a cognitive phenomenon is accessible from the third-person perspective—such as the orienting behaviour of an animal—or whether it is accessible only from the first-person perspective—such as in the famous example of the hue of a rose or an emotion. As long as these qualia can be operationalized and rated on some subjective scale, they are amenable to neuroscientific inquiry. In case of the social neurosciences, the agenda is somewhat enlarged, since the objects of studies are phenomena that emerge from social interactions and exist as relational constructs only in interpersonal space. Here the objects of study are the mechanisms underlying the cognitive functions enabling the respective social interactions such as, for example, the ability to have a theory of mind and the mechanisms that permit individual brains to represent social realities. This agenda of the neurosciences may appear bold and, as the reader will have noticed, while we already know much about the component functions in our brains, we are still very far from understanding the distributive processes underlying higher cognitive and executive functions. In fact, the more data that sophisticated analytical tools allow us to accumulate, the more we are humbled by the mind-boggling and no longer intuitively graspable complexity of the brain’s dynamics. However, at least the author and probably the ma-

majority of his colleagues believe that there should be no principle epistemological barriers to the pursuit of this research agenda. The greatest problem in the near future will be that the description of the dynamics of neuronal processes underlying higher functions will take the form of abstract mathematical formulations that lack any resemblance to the experienced or observed result of these functions.

Nevertheless, the naturalistic stance taken by the neurosciences has already in these early days provided some insights, whose relevance goes beyond the research agenda of the neurosciences proper. The data on mechanisms mediating perception discussed at the beginning of this chapter clearly support constructivism and thereby provide arguments for or against particular philosophical positions. Likewise, these epistemic considerations, to the author’s surprise, led to normative consequences in the context of notions of tolerance. Similar normative consequences arise from data on mechanisms responsible for decision-making, motivation, response suppression, conscious versus subconscious processing, personality traits, and so on, as these insights are all consequential for the definition of behavioural norms and the distinction between normal and pathological behaviour. It is foreseeable, therefore, that the neurosciences will become more involved in philosophical, normative, ethical, and societal issues. This should be beneficial for all parties involved, since the communication process is likely to lead to bridging theories, new terms for mutual understanding, and amendments to discipline-specific idiosyncrasies.

References

- Aru, J., Bachmann, T., Singer, W. & Melloni, L. (2012a). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, *36* (2), 737-746. [10.1016/j.neubiorev.2011.12.003](https://doi.org/10.1016/j.neubiorev.2011.12.003)
- Aru, J., Axmacher, N., Do Lam, A. T., Fell, J., Elger, C. E., Singer, W. & Melloni, L. (2012b). Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. *The Journal of Neuroscience*, *32* (43), 14909-14914. [10.1523/JNEUROSCI.2051-12.2012](https://doi.org/10.1523/JNEUROSCI.2051-12.2012)
- Baars, B. J. (1997). In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4* (4), 292-309.
- Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J. & Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the National Academy of Sciences of the USA*, *108* (27), 11262-11267. [10.1073/pnas.1011284108](https://doi.org/10.1073/pnas.1011284108)
- Buonomano, D. V. & Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, *10*, 113-125. [10.1038/nrn2558](https://doi.org/10.1038/nrn2558)
- Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D. & Miller, E. K. (2012). Synchronous oscillatory neuronal ensembles for rules in the prefrontal cortex. *Neuron*, *76* (4), 838-846. [10.1016/j.neuron.2012.09.029](https://doi.org/10.1016/j.neuron.2012.09.029)
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford, UK: Oxford University Press.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D. & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276* (5312), 593-596. [10.1126/science.276.5312.593](https://doi.org/10.1126/science.276.5312.593)
- Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.) *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 17-40). Cambridge, MA: MIT Press.
- Constantinople, C. M. & Bruno, R. M. (2013). Deep cortical layers are activated directly by thalamus. *Science*, *340* (6140), 1591-1594. [10.1126/science.1236425](https://doi.org/10.1126/science.1236425)
- Cosmelli, D., David, O., Lachaux, J.-P., Martinerie, J., Garnero, L., Renault, B. & Varela, F. (2004). Waves of consciousness: ongoing cortical patterns during binocular rivalry. *NeuroImage*, *23* (1), 128-140. [10.1016/j.neuroimage.2004.05.008](https://doi.org/10.1016/j.neuroimage.2004.05.008)
- Cowey, A. & Stoerig, P. (1991). The neurobiology of blindsight. *Trends in Neurosciences*, *14* (4), 140-145. [10.1016/0166-2236\(91\)90085-9](https://doi.org/10.1016/0166-2236(91)90085-9)
- Crick, F. & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, *2*, 263-275.
- David, S. V., Vinje, W. E. & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons. *The Journal of Neuroscience*, *24* (31), 6991-7006. [10.1523/JNEUROSCI.1422-04.2004](https://doi.org/10.1523/JNEUROSCI.1422-04.2004)
- Dehaene, S., Naccache, L., Le Clec', H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.-F. & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395* (6702), 597-600. [10.1038/26967](https://doi.org/10.1038/26967)
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10* (5), 204-211. [10.1016/j.tics.2006.03.007](https://doi.org/10.1016/j.tics.2006.03.007)
- Dennett, D. C. (1992). *Consciousness explained*. London, UK: Penguin.
- De Pasquale, R. & Sherman, S. M. (2011). Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. *The Journal of Neuroscience*, *31* (46), 16494-16506. [10.1523/JNEUROSCI.3664-11.2011](https://doi.org/10.1523/JNEUROSCI.3664-11.2011)
- Doesburg, S. M., Kitajo, K. & Ward, L. M. (2005). Increased gamma-band synchrony precedes switching of conscious perceptual objects in binocular rivalry. *Neuroreport*, *16* (11), 1139-1142. [10.1097/00001756-200508010-00001](https://doi.org/10.1097/00001756-200508010-00001)
- Doesburg, S. M., Green, J. J., McDonald, J. J. & Ward, L. M. (2009). Rhythms of consciousness: Binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS One*, *4* (7), e6142. [10.1371/journal.pone.0006142](https://doi.org/10.1371/journal.pone.0006142)
- Engel, A. K., Fries, P. & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2* (10), 704-716. [10.1038/35094565](https://doi.org/10.1038/35094565)
- Engen, H. G. & Singer, T. (2013). Empathy circuits. *Current Opinion in Neurobiology*, *23* (3), 275-282. [10.1016/j.conb.2012.11.003](https://doi.org/10.1016/j.conb.2012.11.003)
- Felleman, D. J. & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1* (1), 1-47. [10.1093/cercor/1.1.1](https://doi.org/10.1093/cercor/1.1.1)
- Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., Andelman, F., Neufeld, M.Y., Kramer,

- U., Fried, I. & Malach, R. (2009). Neural „ignition“: Enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron*, *64* (4), 562-574. [10.1016/j.neuron.2009.11.001](https://doi.org/10.1016/j.neuron.2009.11.001)
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C. & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the USA*, *102* (27), 9673-9678. [10.1073/pnas.0504136102](https://doi.org/10.1073/pnas.0504136102)
- Friederici, A. D. & Gierhan, S. M. E. (2013). The language network. *Current Opinion in Neurobiology*, *23* (2), 250-254. [10.1016/j.conb.2012.10.002](https://doi.org/10.1016/j.conb.2012.10.002)
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9* (10), 474-480. [10.1016/j.tics.2005.08.011](https://doi.org/10.1016/j.tics.2005.08.011)
- (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience*, *32*, 209-224. [10.1146/annurev.neuro.051508.135603](https://doi.org/10.1146/annurev.neuro.051508.135603)
- Fries, P., Roelfsema, P. R., Engel, A. K., König, P. & Singer, W. (1997). Synchronisation of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences of the USA*, *101* (35), 13050-13055.
- Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R. & Singer, W. (2001a). Rapid feature selective neuronal synchronization through correlated latency shifting. *Nature Neuroscience*, *4* (2), 194-200. [10.1038/84032](https://doi.org/10.1038/84032)
- Fries, P., Reynolds, J. H., Rorie, A. E. & Desimone, R. (2001b). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, *291* (5508), 1560-1563. [10.1126/science.1055465](https://doi.org/10.1126/science.1055465)
- Fries, P., Schröder, J. H., Singer, W. & Engel, A. K. (2001c). Conditions of perceptual selection and suppression during interocular rivalry in strabismic and normal cats. *Vision Research*, *41* (6), 771-783. [10.1016/S0042-6989\(00\)00299-6](https://doi.org/10.1016/S0042-6989(00)00299-6)
- Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L. & Naccache, L. (2009). Converging intracranial markers of conscious access. *PLOS Biology*, *7* (3), e1000061. [10.1371/journal.pbio.1000061](https://doi.org/10.1371/journal.pbio.1000061)
- Genc, E., Bergmann, J., Singer, W. & Koler, A. (2014). Surface area of early visual cortex predicts individual speed of travelling waves during binocular rivalry (forthcoming). *Cerebral Cortex*. [10.1093/cercor/bht342](https://doi.org/10.1093/cercor/bht342)
- Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*, 334-337. [10.1038/338334a0](https://doi.org/10.1038/338334a0)
- Gray, C. M. & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences of the USA*, *86* (5), 1698-1702. [10.1073/pnas.86.5.1698](https://doi.org/10.1073/pnas.86.5.1698)
- Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shaprio, K., Hommel, B. & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the USA*, *101* (35), 13050-13055. [10.1073/pnas.0404944101](https://doi.org/10.1073/pnas.0404944101)
- Hameroff, S. (2006). Consciousness, neurobiology and quantum mechanics. In J. Tuszynski (Ed.) *The Emerging Physics of Consciousness* (pp. 193-253). Berlin, GER: SpringerVerlag.
- Han, X., Chow, B. Y., Zhou, H., Klapoetke, N. C., Chuong, A. S., Rajimehr, R., Yang, A., Baratta, M. V., Winkle, J., Desimone, R. & Boyden, E. S. (2011). A high-light sensitivity optical neural silencer: Development and application to optogenetic control of non-human primate cortex. *Frontiers in Systems Neuroscience*, *5* (18), 1-8. [10.3389/fnsys.2011.00018](https://doi.org/10.3389/fnsys.2011.00018)
- Haynes, J.-D., Deichmann, R. & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, *438*, 496-499. [10.1038/nm1537](https://doi.org/10.1038/nm1537)
- Hipp, J. F., Engel, A. K. & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69* (2), 387-396. [10.1016/j.neuron.2010.12.027](https://doi.org/10.1016/j.neuron.2010.12.027)
- Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M. & Engel, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, *15* (6), 884-890. [10.1038/nn.3101](https://doi.org/10.1038/nn.3101)
- Hodzic, A., Kaas, A., Muckli, L., Stirn, A. & Singer, W. (2009). Cortical responses to invisible objects in the human dorsal and ventral pathways. *NeuroImage*, *45* (4), 1264-1271. [10.1016/j.neuroimage.2009.01.027](https://doi.org/10.1016/j.neuroimage.2009.01.027)
- Houweling, A. R. & Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. *Nature*, *451*, 65-68. [10.1038/nature06447](https://doi.org/10.1038/nature06447)
- Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F. & Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron*, *73* (3), 814-828. [10.1016/j.neuron.2011.12.026](https://doi.org/10.1016/j.neuron.2011.12.026)

- Jensen, O. & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15* (8), 1395-1399. [10.1046/j.1460-9568.2002.01975.x](https://doi.org/10.1046/j.1460-9568.2002.01975.x)
- Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A. & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature*, *425*, 954-956. [10.1038/nature02078](https://doi.org/10.1038/nature02078)
- Lau, H. C. & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *The Journal of Neuroscience*, *27* (21), 5805-5811. [10.1523/JNEUROSCI.4335-06.2007](https://doi.org/10.1523/JNEUROSCI.4335-06.2007)
- Logothetis, N. K., Leopold, D. A. & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, *380*, 621-624. [10.1038/380621a0](https://doi.org/10.1038/380621a0)
- Lukoševičius, M. & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, *3* (3), 127-149. [10.1016/j.cosrev.2009.03.005](https://doi.org/10.1016/j.cosrev.2009.03.005)
- Malach, R. (2007). The measurement problem in human consciousness research. *Behavioral and Brain Sciences*, *30* (5-6), 516-517. [10.1017/S0140525X0700297X](https://doi.org/10.1017/S0140525X0700297X)
- Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Ribeiro Gomes, A. R., Magrou, L., Misery, P., Giroud, P., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., van Essen, D. C. & Kennedy, H. (2013). The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences of the USA*, *110* (13), 5187-5192. [10.1073/pnas.1218972110](https://doi.org/10.1073/pnas.1218972110)
- Markov, N. T. & Kennedy, H. (2013). The importance of being hierarchical. *Current Opinion in Neurobiology*, *23* (2), 187-194. [10.1016/j.conb.2012.12.008](https://doi.org/10.1016/j.conb.2012.12.008)
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W. & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of Neuroscience*, *27* (11), 2858-2865. [10.1523/JNEUROSCI.4623-06.2007](https://doi.org/10.1523/JNEUROSCI.4623-06.2007)
- Melloni, L. & Rodriguez, E. (2007). Non-perceived stimuli elicit global but not large-scale neural synchrony. *Perception*, *36*
- Melloni, L. & Singer, W. (2011). The explanatory gap in neuroscience. *Pontificiae Academiae Scientiarum Acta*, *21*, 61-73.
- Metzinger, T. (Ed.) (2000). *Neural correlates of consciousness: Empirical and conceptual questions*. Cambridge, MA: MIT Press.
- Muckli, L. & Petro, L. S. (2013). Network interactions: non-geniculate input to V1. *Current Opinion in Neurobiology*, *23* (2), 195-201. [10.1016/j.conb.2013.01.0](https://doi.org/10.1016/j.conb.2013.01.0)
- Nakatani, C., Ito, J., Nikolaev, A. R., Gong, P. & van Leeuwen, C. (2005). Phase synchronization analysis of EEG during attentional blink. *Journal of Cognitive Neuroscience*, *17* (12), 1969-1979. [10.1162/089892905775008706](https://doi.org/10.1162/089892905775008706)
- Penrose, R. (1994). *Shadows of the mind: A search for the missing science of consciousness*. New York, NY: Oxford University Press.
- Power, J. D. & Petersen, S. E. (2013). Control-related systems in the human brain. *Current Opinion in Neurobiology*, *23* (2), 223-228. [10.1016/j.conb.2012.12.009](https://doi.org/10.1016/j.conb.2012.12.009)
- Raichle, M. E. (2011). The restless brain. *Brain Connectivity*, *1* (1), 3-12. [10.1089/brain.2011.0019](https://doi.org/10.1089/brain.2011.0019)
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the USA*, *98* (2), 676-682. [10.1073/pnas.98.2.676](https://doi.org/10.1073/pnas.98.2.676)
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B. & Varela, F. J. (1999). Perception's shadow: Long distance synchronization of human brain activity. *Nature*, *397* (6718), 430-433. [10.1073/pnas.98.2.676](https://doi.org/10.1073/pnas.98.2.676)
- Roopun, A. K., Kramer, M. A., Carracedo, L. M., Kaiser, M., Davies, C. H., Traub, R. D., Kopell, N. J. & Whittington, M. A. (2008). Temporal interactions between cortical rhythms. *Frontiers in Neuroscience*, *2* (2), 145-154. [10.1126/science.1099745](https://doi.org/10.1126/science.1099745)
- Salazar, R. F., Dotson, N. M., Bressler, S. L. & Gray, C. M. (2012). Content specific fronto-parietal synchronization during visual working memory. *Science*, *338* (6110), 1097-1100. [10.1126/science.1224000](https://doi.org/10.1126/science.1224000)
- Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *The Journal of Neuroscience*, *12* (6), 2331-2355.
- Schack, B., Klimesch, W. & Sauseng, P. (2005). Phase synchronisation between theta and upper alpha oscillations in a working memory task. *International Journal of Psychophysiology*, *57* (2), 105-114. [10.1016/j.ijpsycho.2005.03.016](https://doi.org/10.1016/j.ijpsycho.2005.03.016)
- Schwarz, C. & Bolz, J. (1991). Functional specificity of the long-range horizontal connections in cat visual cortex: A cross-correlation study. *The Journal of Neuroscience*, *11* (10), 2995-3007.
- Searle, J. R. (1997). *The mystery of consciousness*. London, UK: Granta Books.

- Sherrington, C. S. (1906). *The integrative action of the nervous system*. New York, NY: Charles Scribner's Sons.
- Shew, W. L., Yang, H., Petermann, T., Roy, R. & Plenz, D. (2009). Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *The Journal of Neuroscience*, *29* (49), 15595-15600. [10.1523/JNEUROSCI.3864-09.2009](https://doi.org/10.1523/JNEUROSCI.3864-09.2009)
- Silver, M. A. & Logothetis, N. K. (2004). Grouping and segmentation in binocular rivalry. *Vision Research*, *44* (14), 1675-1692. [10.1016/j.visres.2003.12.008](https://doi.org/10.1016/j.visres.2003.12.008)
- Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, *24* (1), 49-65. [10.1016/S0896-6273\(00\)80821-1](https://doi.org/10.1016/S0896-6273(00)80821-1)
- (2009). Genetic and epigenetic shaping of cognition – prerequisites of cultural evolution. In W. Arber, N. Cabibbo & M. Sánchez Sorondo (Eds.) *The proceedings of the plenary session on scientific insights into the evolution of the universe and of life. 31 October – 4 November 2008, vol Pontificiae Academiae Scientiarum Acta* (pp. 337-347). Vatican City, VA: Pontificia Academia Scientiarum.
- (2010). Neocortical rhythms. An overview. In C. von der Malsburg, W. A. Phillips & W. Singer (Eds.) *Dynamic coordination in the brain. From neurons to mind* (pp. 159-168). Cambridge, MA: MIT Press.
- (2013). Cortical dynamics revisited. *Trends in Cognitive Sciences*, *17* (12), 616-626. [10.1016/j.tics.2013.09.006](https://doi.org/10.1016/j.tics.2013.09.006)
- Sporns, O. (2013). Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, *23* (2), 162-171. [10.1016/j.conb.2012.11.015](https://doi.org/10.1016/j.conb.2012.11.015)
- Srinivasan, R., Russell, D. P., Edelman, G. M. & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *The Journal of Neuroscience*, *19* (13), 5435-5448.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D. & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78* (2), 364-375. [10.1016/j.neuron.2013.01.039](https://doi.org/10.1016/j.neuron.2013.01.039)
- Uhlhaas, P. J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D. & Singer, W. (2009). Neuronal synchrony in cortical networks: history, concept and current status. *Frontiers in Integrative Neuroscience*, *3* (17), 1-19. [10.3389/neuro.07.017.2009](https://doi.org/10.3389/neuro.07.017.2009)
- Van den Heuvel, M. P. & Sporns, O. (2011). Rich club organization of the human connectome. *The Journal of Neuroscience*, *31* (44), 15775-15786. [10.1523/JNEUROSCI.3539-11.2011](https://doi.org/10.1523/JNEUROSCI.3539-11.2011)
- (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, *17* (12), 683-696. [10.1016/j.tics.2013.09.012](https://doi.org/10.1016/j.tics.2013.09.012)
- Van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S. & Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *The Journal of Neuroscience*, *28* (32), 8053-8062. [10.1523/JNEUROSCI.1278-08.2008](https://doi.org/10.1523/JNEUROSCI.1278-08.2008)
- Varela, F., Lachaux, J. P., Rodriguez, E. & Martinerie, J. (2001). The brainweb: Phase synchronisation and large-scale integration. *Nature Reviews Neuroscience*, *2* (4), 229-239. [10.1038/35067550](https://doi.org/10.1038/35067550)
- Vinje, W. E. & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287* (5456), 1273-1276. [10.1126/science.287.5456.1273](https://doi.org/10.1126/science.287.5456.1273)
- von der Malsburg, C., Phillips, W. A. & Singer, W. (2010). *Dynamic coordination in the brain. From neurons to mind*. Cambridge, MA: MIT Press.
- Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience*, *24*, 57-86. [10.1146/annurev.neuro.24.1.57](https://doi.org/10.1146/annurev.neuro.24.1.57)

It's Not Just About the Contents: Searching for a Neural Correlate of a State of Consciousness

A Commentary on Wolf Singer

Valdas Noreika

Global gamma band synchronisation is perhaps the most extensively studied candidate for a Neural Correlate of Consciousness (NCC). Yet despite numerous studies confirming its association with consciousness, it seems to be neither sufficient nor necessary for the presence of all subjective experiences. Analysis of gamma synchronisation studies suggests that it is a correlate of the initial binding of expected, attended, task-dependent contents of consciousness, whereas task-irrelevant contents do not seem to require gamma synchronisation. While discovery of such a content-related NCC is a remarkable achievement for the neurophysiological research of consciousness, it does not fully explain some of the fundamental structural properties of consciousness, namely the temporal and spatial integration of all available experiences into a coherent stream of consciousness. As an alternative, instead of focusing solely on the selected contents of consciousness, the neural mechanisms of the fundamental properties of consciousness could be studied by contrasting states of (un)consciousness. Recent research into the states of consciousness suggests that, for instance, informational complexity is a highly sensitive predictor of the presence of consciousness, possibly reflecting background structural properties of the unity of subjective experiences. As a limiting factor, though, such a state-related NCC does not seem to reflect the phenomenal diversity of the contents of consciousness. Arguably, these limitations could be overcome by devising experimental setups that would simultaneously probe the neural correlates of the contents and the state of consciousness.

Keywords

Contents of consciousness | Gamma band synchronisation | Neural correlate of consciousness (NCC) | Neural correlates of consciousness | Nonconscious states | Spatial binding | State of consciousness | Stream of consciousness | Temporal binding | Unconscious states

1 Introduction

Even though the search for the neural correlates of consciousness is still an unresolved challenge of astonishing complexity (Crick 1994), the continuous efforts to crack the mystery are not expended in vain. Each year brings an increasing number of cognitive neuroscientific studies that reveal yet another piece of the puzzle of the neural basis of subjective experience. However,

it often seems that individual findings are too diverse and sparse to form a coherent picture. In addition to the fundamental problem of the binding of conscious experiences (Singer 2001), we increasingly face the problem of how to bind the findings of consciousness-related studies. The present target paper by Prof. Wolf Singer serves such a discovery-binding function, bring-

Commentator

Valdas Noreika

valdas.noreika@mrc-cbu.cam.ac.uk

Medical Research Council
Cambridge, England

Target Author

Wolf Singer

w.singer@brain.mpg.de

Max Planck Institute for Brain
Research (MPI)
Frankfurt a. M., Germany

Editors

Thomas Metzinger

metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt

jennifer.windt@monash.edu

Monash University
Melbourne, Australia

ing very diverse findings into a unified picture of how the neural correlates with the subjective.

In an impressively erudite manner, [Singer \(this collection\)](#) integrates a very broad range of anatomical and functional findings of the organisational principles of the brain, concluding that the high-level cognitive functions are supported by densely coupled, recurrent neural networks, interacting under the principles of non-linear dynamics. In the proposed framework, perception is treated as an active process, whose self-organisation is initially determined by genes, and later modified by post-natal development, learning, social interactions, and cultural influences. At the neuronal networks level, high-level integration and communication are achieved through synchronisation of oscillations in different electroencephalography (EEG) frequency bands, the most notable of which is gamma band (>30Hz) synchronisation ([Engel et al. 1999](#)). Given that an association between the widespread gamma-band synchronisation and conscious awareness is found in rather different experimental paradigms, such as visual masking ([Melloni et al. 2007](#)), binocular rivalry ([Doeburg et al. 2009](#)), and attentional blink ([Gross et al. 2004](#)), gamma synchronisation is often regarded as the main NCC ([Singer this collection](#)).

Yet the candidature of gamma synchronisation as the correlate of consciousness is challenged by some findings from research into the behavioural states of the brain. If gamma-range activity correlates with consciousness, it should diminish when consciousness ceases. Contrary to this, gamma band activity seems to increase rather than decrease in response to certain general anaesthetics, such as ketamine ([Steriade et al. 1996](#)). Furthermore, gamma synchronisation seems to be absent in some conscious brain states. For instance, it has been reported that large-scale neocortical gamma-band coherence is virtually absent during rapid eye movement (REM) sleep in cats ([Castro et al. 2013](#)), a state typically marked by the most intense dreaming in humans ([Hobson et al. 2000](#)) as well as in felids ([Jouvet 1979](#)).

The target paper briefly mentions neural mechanisms supporting overall brain states, but

dismisses them as modulatory systems that are too general to be considered the NCC ([Singer this collection](#)). In the following, I will argue that the relation between the neural mechanisms of the contents and states of consciousness is not straightforward, and that the puzzle of the neural mechanisms of consciousness cannot be completed without studying the neural mechanisms of conscious states.

2 Contents vs. states of consciousness

An important distinction in consciousness research is that between the contents and a state of consciousness ([Chalmers 2000](#)). The concept of the contents of consciousness refers to individual subjective experiences that occur in phenomenal consciousness, such as reading a word or hearing birdsong, and as such they are sometimes referred to as the phenomenal contents of consciousness ([Revonsuo 2006](#)). Most neural experiments on consciousness, especially in the dominant field of visual awareness studies, are concerned with the neural basis of such specific contents of consciousness, i.e., they select one or two subjective experiences within an overall stream of consciousness. In this type of experiment, participants may be presented with stimuli close to their perceptual threshold ([Del Cul et al. 2009](#)) or they may be instructed to observe ambiguous stimuli that may lead to perception of several alternating contents of consciousness ([Kornmeier & Bach 2012](#)). The brain responses are then contrasted between trials that differ in awareness of these stimuli. Notably, while participants in such experiments report being unaware of some contents of consciousness, they still maintain awareness of other experiences: such as seeing the edges of a computer screen, hearing the background noise of the Magnetic Resonance Imaging (MRI) scanner, or letting their thoughts wander away from the experimental task. Typically, such experiences are ignored as task-irrelevant, and consequently the so-called “unaware” or “unconscious” trials still bear very rich phenomenology.

Contrary to the selective contents of consciousness, the concept of the state of consciousness refers to an overall pattern of subjective

psychological functioning that includes the totality of phenomenal contents of consciousness (Rosenthal 1986; Tart 1972). In addition to the relaxed waking state of consciousness in a healthy volunteer, which could be also regarded as a *baseline state*, altered, unconscious, and non-conscious states can be distinguished. In *altered states of consciousness*, such as dreaming or Lysergic Acid Diethylamide (LSD) psychomodulation, subjective experiences may undergo various perceptual and cognitive alterations, the neural basis of which can be studied by contrasting them with a baseline state of consciousness, e.g., by comparing brain activity before and after hallucinogen intake (Carhart-Harris et al. 2012). Given that there is no widely accepted definition and criterion for an altered state of consciousness (Móro 2010; Revonsuo et al. 2009), a rather common approach is to describe, classify, and study states that are traditionally called *altered*, avoiding a single definition that would grasp the core of all altered states of consciousness (Vaitl et al. 2005).

Contrary to the baseline and altered states of consciousness, unconscious states are deprived of subjective experiences, but they may still maintain the potential to become conscious. For instance, an unconscious state of dreamless sleep may turn into a conscious sleep once a sleeping participant begins to dream (for an alternative interpretation of dreamless sleep, see Thompson this collection). Finally, non-conscious states are those completely deprived of a capacity to support phenomenal consciousness, such as an irreversible coma. In clinical neuroscience, the most extensively studied contrast between a pathological altered state of consciousness and an unconscious or non-conscious state is a comparison between minimally conscious and vegetative state patients (Sitt et al. 2014). When states of (un)consciousness are contrasted, neural representations of specific experiences are typically ignored, making it difficult or even impossible to assess the phenomenal specificity of findings, e.g., if participants were aware of particular external stimuli or what internally generated experiences they had. Nevertheless, research into these states may re-

veal neural patterns that are common to all subjective experiences without individuating them.

It is possible that these two lines of research may eventually reveal rather different, if not independent, NCC systems: a neural correlate of the state of consciousness and a neural correlate of the contents of consciousness (Chalmers 2000). If these exist, any neuroscientific program of consciousness research would be incomplete without searching for a state NCC. Furthermore, even if a separate state NCC did not exist, there is currently no evidence for this, and thus NCC research is incomplete if it does not investigate this possibility. This view is often dismissed on the basis that some of the most plausible candidates for a state NCC, such as the brainstem reticular formation (Merker 2007; Parvizi & Damasio 2001), are rather low-level neural systems, whereas converging evidence shows that consciousness is a cortical process (Singer this collection). Furthermore, it could be argued that a conscious state may be nothing more than the sum of individual experiences, in which case revealing the NCC of specific contents would automatically explain the state NCC. Yet a brief analysis of the fundamental structural properties of consciousness—see the following section—shows that the necessary and sufficient NCC cannot be revealed by an exclusive focus on the contents of consciousness. Notably, the arguments presented in this commentary will be confined to the biological nature of human and animal consciousness, and as such they are not applicable to the problem of machine, extraterrestrial, or silicon-brain consciousness.

3 Unity as the fundamental property of the stream of consciousness

Given that subjective experiences can accompany almost any sensory, cognitive, emotional, and behavioural function of the brain, phenomenal consciousness turns out to be an extremely complex and multi-dimensional process. Nevertheless, introspection shows that despite its qualitative richness, phenomenal consciousness appears to us as a unified and coherent model

of the external and internal environment (Bayne 2010; Revonsuo 2006). The continuity in the diversity of subjective experiences is famously referred to as the stream of consciousness (James 1890). This metaphor points to the unification of experiences occurring at different points in time and space, which is achieved through temporal and spatial binding.

At the cognitive level of description, temporal binding, i.e., integration of subjective experiences over time, is realised through the perception of simultaneity, duration, and successiveness (Kiverstein 2010; Pöppel 1997). Perception of simultaneity may integrate several experiences, e.g., seeing a cat and hearing a birdsong in the park, as occurring at the same time. Perception of duration of selected experiences may extend them in time, e.g., the birdsong might seem to last for a certain period of time. Finally, perception of successiveness may signal the end of one temporally-extended experience, and the beginning of another one, e.g., as the cat reaches the bush and the birdsong ceases, we may notice a cone under the bush. Notably, the change does not typically involve all experiences, and as we are aware of some changing contents, some other experiences continue to endure in time, e.g., we still see the same bush. In addition to the timing-specific functions, temporal binding seems to depend on the iconic memory that contains the just-experienced contents of consciousness, and on the anticipation of subsequent ones, forming the temporally-extended phenomenal experience of now, sometimes referred to as the *specious present* (Dainton 2006; Kelly 1882). Temporal extension of subjective experiences have a simple, yet very important and often overlooked implication for NCC research: if there is a single neural mechanism generating phenomenal consciousness, it should be present as long as we are conscious of at least one single content. Given that our experiences do not cease at a fixed rate, i.e., some are shorter and some longer, and that the change does not happen abruptly for all experiences at once, the NCC should persist for the duration of the stream of consciousness. Thus, a temporary-confined correlate of awareness, such as a negative Event Related Potential (ERP)

waveform briefly peaking at about 200ms from the onset of visual stimuli (Railo et al. 2011), cannot be a sufficient correlate of consciousness, as awareness of visual contents lasts for a much longer period of time.

Spatial binding, i.e., integration of subjective experiences in space, is realised through several complimentary processes, through which each subjective experience occupies a specific location in relation to other experiences, which is sometimes referred to as *location binding* (Treisman 1996). In the baseline state of consciousness, one experience never occurs in isolation from other experiences; and when some experiences cease, we do not experience emptiness, because other experiences fill in their place. Furthermore, individual experiences are spatially integrated not only with respect to each other in phenomenally external space, but also with respect to the common egocentric reference point (Revonsuo 2006). The reference point is typically located in the phenomenal head or chest, and all other experiences are realised in the space as taking a certain distance and angle from this point. While typical phenomenological analysis of 3D space considers visual and auditory experiences, it has recently been shown that emotions and feelings are also experienced as taking certain location with respect to our body parts (Nummenmaa et al. 2014). For instance, anger is overrepresented in hands and arms when compared to sadness. Arguably, even thoughts, which are often regarded as non-spatial entities (Clarke 1995), are usually experienced as occurring within the head rather than somewhere else. Given that the whole phenomenal space is bound together, the NCC should also represent awareness of the whole space rather than, for instance, selected regions on the computer screen. That is, a promising candidate for an NCC should not cease when a specific experience vanishes as long as the spatial and temporal unity of the stream of consciousness is maintained.

So, what type of neural processes should we be looking for when searching for the NCC? If we take the unity of consciousness seriously, we should be looking for a neural process that steadily represents the whole phenomenal space,

and sustains its activity over periods of time longer than the existence of a single experience. Arguably, the neural correlate of unified consciousness cannot be discovered by studying and contrasting only isolated contents of consciousness, as the unity of spatiotemporal interactions simply cannot be derived from solitary experiences. Thus, while continuing to search for the neural mechanisms of the contents of consciousness, the NCC program should be extended by carrying out systematic contrasts between unconscious, baseline, and/or altered states, which would consider the whole stream of consciousness. A possible objection to this proposal is that the unity of consciousness is not fundamental in the strong form of fundamentalism, i.e., some forms of consciousness may still exist despite the possible disintegration of its unity, which seems to happen in states like schizophrenia, sleep onset, or a minimally conscious state. For instance, the stream of consciousness may occasionally undergo a sudden, unpredictable alteration in terms of inner speech and imagery (Noreika et al. 2014). Nevertheless, if one aims to explain the neural mechanisms of normal waking consciousness, the unity thesis, with its NCC-related implications, cannot be ignored. With these considerations in mind, let us examine now the proposal of global gamma synchronisation as the NCC (Singer this collection).

4 Is gamma band synchrony sufficient and necessary for consciousness?

A sufficient and necessary NCC (or perhaps a set of NCCs) should be generic enough to cover all conscious contents and states, and should also be specific enough to cover only conscious contents and states. Notably, gamma synchronisation does not meet the second specificity requirement, as it can be associated with almost any perceptual and cognitive function that depends on the formation of temporary associations of distributed neuronal networks. Among numerous cases, increased gamma-band synchronisation is found to be associated with such tasks as perceptual learning (Gruber et al. 2002), self-paced movement (Pfurtscheller et al.

2003), mental rotation (Bhattacharya et al. 2001), viewing of unpleasant stimuli (Martini et al. 2012), deductive reasoning (Zhang et al. 2014), auditory attention control (Doesburg et al. 2012), face integration (Kottlow et al. 2012), and memory encoding and retrieval (Osipova et al. 2006). Gamma-band synchronisation thus seems to be a generic process that contributes to complex cortical computations involved in most if not all of the higher cognitive functions (Fries 2009).

Singer (this collection) proposes that only global, widespread synchronisation of gamma oscillations is associated with consciousness, whereas local, spatially-restricted synchronisation is not necessarily related to conscious awareness. This might refute studies reporting local synchrony; however, some of the above-mentioned studies found increased global gamma-synchrony when observing unpleasant stimuli (Martini et al. 2012) or carrying out a mental rotation task (Bhattacharya et al. 2001). It could be argued, though, that gamma synchronisation is present in these experiments as a correlate of task-related subjective experiences, such as awareness of memory retrieval. In fact, even though most of these studies did not even mention consciousness or awareness, their participants were not unconscious, and gamma synchronisation could have been associated with the task-dependent subjective experiences. Yet, this line of reasoning is challenged by the simple fact that participants remained conscious in all contrast conditions throughout the experiments. Why would consciousness-related gamma synchronisation increase in some, but not other conditions? This leads us to the question of what exactly gamma synchrony correlates with in studies that specifically manipulate awareness? Let us take a closer look at two key studies, also examined by Singer (this collection).

Melloni et al. (2007) presented pairs of words and asked participants to report on whether both words were the same. Visibility of the first word was manipulated by adjusting the luminance level of the forward and backward masks, which rendered the words visible only in some of the trials. Global gamma-phase synchronisation between the fronto-centro-parietal

electrodes was observed within the 40–182ms time-window after the presentation of the first word only in visible trials, which coincides with the time when conscious perception of the words is expected to emerge. In the latter time-windows, visible words were marked by more localised gamma synchronisation, higher P300 amplitude, and higher amplitude of frontal theta oscillations than invisible words. These findings confirmed that gamma synchronisation is a correlate of visual-semantic awareness, and showed that other electrophysiological processes may also correlate with consciousness.

Doesburg et al. (2009) investigated the role of gamma-phase synchronisation in conscious awareness using a binocular rivalry paradigm, in which a different visual stimulus is presented to each eye. Instead of seeing both stimuli at the same time, people report perceiving only one of the stimuli that continues switching in time. An increase in the gamma-band synchronisation over the fronto-parietal regions was observed in the 600–540ms and 280–220ms time-windows before responses indicating a perceptual switch. Assuming that reaction time was about 250ms, the synchronisation increase coincided with a new percept reaching awareness. Interestingly, gamma synchronisation oscillated at the theta rhythm, suggesting a cross-frequency interaction.

In both of these experiments (Doesburg et al. 2009; Melloni et al. 2007), gamma synchrony peaked around the time when participants began experiencing a new content of consciousness, following which Singer (this collection) draws the well-justified conclusion that gamma synchronisation is associated with a transfer of the new contents into awareness. Given that increased synchronisation may reflect the neural and phenomenal binding required for the fundamental unity of consciousness to emerge, it seems to be an ideal candidate for the NCC. Yet the duration of increased synchronisation is relatively brief and seems to last a much shorter time than the awareness of stimuli. For instance, P300 distinguished visible and invisible words around 300ms post-stimulus, whereas gamma-band synchronisation became local during this time-window (Melloni et al. 2007). Such

brevity of synchronisation suggests that it is involved only in the initial binding of the new contents of consciousness, while a further maintenance of these contents is supported by other neural mechanisms, in particular theta oscillations (Doesburg et al. 2009; Melloni et al. 2007; Singer this collection). Given that the global gamma synchronisation correlates with a spatially- and temporally-local change in the stream of consciousness, its association with an overall unity of consciousness is uncertain and, at least currently, it cannot be accounted as the only or even as the major NCC. If it were such, it would not cease as long as the participant were aware of a particular content of consciousness.

Furthermore, gamma synchrony does not seem to increase in response to each of the new contents of consciousness. In each trial, Melloni et al. (2007) presented a series of stimuli, including a fixation cross, a masking noise, a target word, and a blank screen. Each of these stimuli should have entered consciousness, and even when the target word was unreadable, participants should have perceived something, e.g., an unreadable word, incoherent letters, or a flashing mask. However, the global gamma synchronisation increased only in response to perceived visible words, suggesting that it is a correlate of the initial binding of a selected, expected, attended, coherent, task-relevant content of consciousness. As such, in addition to the lack of specificity, gamma synchrony does not seem to be generic enough to cover all the different contents of consciousness, even within the paradigmatic visual modality. Thus, it seems that the global gamma synchronisation is neither necessary nor sufficient for consciousness to emerge, as subjective experiences may exist without gamma synchronisation, and even when synchronisation is involved in the generation of awareness, other neural processes are needed to maintain its presence.

As discussed in the previous section, the unity of consciousness emerges from the interaction of all experiences available at a time. Gamma synchronisation cannot account for the unity of a state of consciousness, simply because it is involved only in the generation of new task-

dependent contents, and it does not seem to bind these contents within the broader stream of consciousness. Arguably, instead of focusing on selected stimuli, we may be able to detect the neural correlates of the unity of consciousness by contrasting states of consciousness with unconsciousness, since such a contrast would consider the whole stream of phenomenal contents, including their structural unity.

5 Studying the contents and states of consciousness: Let's probe them together!

Perhaps the most powerful contrast conditions for studying a neural correlate of a state of consciousness are comparisons between wakefulness and slow-wave sleep, as well as between wakefulness and general anaesthesia. A notable clinical contrast is a comparison between vegetative state and minimally conscious state patients. Furthermore, new paradigms are available for comparing consciousness with unconsciousness when an overall physiological state of the brain is controlled, such as dreamless vs. dreamful non-rapid eye movement sleep (NREM sleep; Noreika et al. 2009; Siclari et al. 2013) or dreamless vs. dreamful anaesthesia (Noreika et al. 2011). Let us examine several exemplary papers that compare an overall stream of consciousness with its absence.

Sitt et al. (2014) studied auditory-evoked potentials and endogenous fluctuations of EEG signal in 75 vegetative state and 68 minimally-conscious patients. None of the studied evoked potentials (P1, MMN, P3a, P3b, CNV) were able to discriminate patient groups, indicating that task-dependent brain activity does not necessarily distinguish between conscious and unconscious states. Contrary to this, analyses of spontaneous EEG activity showed that unconscious patients had higher power of delta and lower power of theta and alpha oscillations, especially over parietal regions. Furthermore, EEG complexity indices derived from the compressibility of a sequence of data points indicated increased signal complexity over the parietal region in the minimally conscious patients compared to the vegetative state patients. Fi-

nally, electrode connectivity measures derived from information theory showed that vegetative-state patients had lower-weighted symbolic mutual information exchange in the range of theta and alpha oscillations than minimally-conscious patients. Interestingly, none of the EEG connectivity measures in the gamma frequency range, including phase lag index and imaginary coherence, could discriminate patient groups, coinciding with other independent observations that gamma synchrony does not necessarily differentiate conscious and unconscious states of the brain (Castro et al. 2013; Steriade et al. 1996).

The finding that the presence of consciousness is associated with an overall complexity of EEG signal and the magnitude of inter-electrode information exchange (Sitt et al. 2014) seems to support the information integration theory of consciousness (Tononi 2012), which predicts that consciousness depends on information complexity and integration in the system. The information integration theory was recently tested by Casali et al. (2013), who investigated the consciousness-related electrodynamics of the distributed cortical networks in a wide range of states of (un)consciousness, including wakefulness (eyes open, eyes closed), sleep (NREM sleep, REM sleep), anaesthesia (midazolam, xenon, propofol), and consciousness disorders (locked-in syndrome, minimally conscious state, patients who have emerged from a minimally conscious state, vegetative state). In a series of experiments, transcranial magnetic stimulation (TMS) pulses were delivered to different cortical sites, which perturbed spontaneous EEG activity (Massimini et al. 2010). Complexity of such TMS-induced EEG perturbations was then calculated, and its index successfully differentiated the states of consciousness and unconsciousness, even at the individual participant's level (Casali et al. 2013). As predicted, the presence of consciousness was associated with a higher level of information complexity.

In these and similar experiments, the contents of consciousness were not systematically manipulated or controlled for, and conscious participants probably underwent very diverse experiences. Consequently, the reported EEG

complexity as the NCC seems to be independent of particular phenomenal contents, and it may reflect some structural aspects of the whole stream of unified subjective experiences. It seems that phenomenal consciousness emerges in a state of the brain that is capable of generating the required level of information complexity and integration. As requested in the previous sections, such an NCC is stable in time and does not depend on an experience isolated from the rest of phenomenal space. Arguably, this type of study tackles the fundamental unity of consciousness much more directly than typical paradigms for studying the selected contents of consciousness. However, approaching one side of the bridge takes us further away from the other side, and the better characterization we have of the neural mechanisms of the state of consciousness, the less we can say about the neural mechanisms of particular contents of consciousness. For instance, the perturbational complexity index can differentiate conscious and unconscious states, but it is extremely insensitive when it comes to distinguishing between different contents of consciousness. For instance, the values of the complexity index did not systematically differ between the “eyes closed” and “eyes open” conditions in the standard waking state (Casali et al. 2013; Noreika 2014). Arguably, any NCC that cannot distinguish between experiences occurring in the “eyes closed” and “eyes open” conditions cannot be fully satisfactory, as the quality of subjective experiences is the core of the scientific problem of consciousness. Yet even though informational complexity does not reflect qualities of phenomenal contents, it is a promising candidate for an NCC of the background properties of consciousness that enable the emergence of subjective experiences and/or necessitate their structural unity.

We are thus left with studies of the contents NCC, such as focusing on the gamma synchrony, and studies of the state NCC, such as focusing on the information complexity. The first group of studies seems to explain the neural binding of concrete selected contents of consciousness, but it does not have a capacity to address the unity of consciousness. The second group seems to capture neural processes

involved in the whole stream of consciousness, but it ignores differentiation or phenomenal diversity of consciousness. Ideally, research into the NCC would combine both of these complementary approaches. Unfortunately, a systematic combination of the contents- and states-focused paradigms is almost never tested in cognitive neuroscientific studies of consciousness.

The combined *contents-states paradigm* would contrast baseline and altered states, or consciousness and unconsciousness, or the transition between the two, while participants carry out experiments that tackle the neural mechanisms of the contents of consciousness. For instance, one could study binocular rivalry while participants lose consciousness in response to an anaesthetic agent. This could, for instance, provide data to investigate how global gamma synchrony as a correlate of the binding and transfer of new contents to awareness depends on or interacts with a changing level of neuronal information complexity. Another promising avenue is research into awareness-related performance in the transition from wakefulness to sleep (Goupil & Bekinschtein 2011). In a recent attempt, Bareham et al. (2014) demonstrated that healthy individuals show neglect-like loss of awareness of the right side of their space in a drowsy state of consciousness. Thus, spatial awareness and unity seem to depend on the state of alertness, as defined by the relative amplitude of theta and alpha oscillations, which confirms that the contents are not wholly independent of the state. That is, despite the external physical stimuli and environment remaining stable, phenomenal contents may appear, disappear, or reorganise depending on the overall state of consciousness. More such studies are expected to be carried out in future, aiming to integrate the content NCC, the state NCC, and altered states of consciousness research programs under one unified framework of the content-state NCC research.

6 Conclusion

Global gamma-band synchronisation, research into which was largely triggered and continues to be advanced by Prof. Wolf Singer, is one of

the most promising NCCs. Synchronisation seems to increase in most cases when a new, task-dependent content of consciousness is formed. Yet a larger number of complications prevents its acceptance as the main NCC, namely: gamma synchronisation does not persist for as long as the contents of consciousness, some of the contents of consciousness emerge without gamma synchronisation being modulated, and, finally, gamma synchronisation may increase in unconscious or unresponsive states of mind. These complications show that gamma-band synchronisation cannot fully account for the existence of a unified stream of consciousness. Given that consciousness is integrated over cognitive time and space, a sufficient and necessary NCC should persist even when some but not all of the experiences cease to exist in time, or change their location. Nevertheless, even though gamma-band synchronisation seems to be neither necessary nor sufficient for all contents of consciousness to arise, it should be regarded as one of the NCCs specifically involved in the binding of new attended experiences. Future research may also develop more accurate characterization of gamma synchronisation, including its spatial scale, precision, and stability ([Singer this collection](#)), and certain forms of synchrony might be necessarily accompanied by consciousness; yet such evidence is not currently available.

Given that gamma synchronisation cannot be the only NCC, research efforts and resources should be distributed to search for the other NCCs, some of which might be responsible for the maintenance of already-bound single contents, and some of which might contribute to the unity of the whole stream of consciousness. Research paradigms should be developed that allow simultaneous manipulation and testing of both the contents and the states of (un)consciousness. Most likely, none of the discovered NCCs alone will be necessary and sufficient for all forms of subjective experiences to exist. How many of the neural correlates will be sufficient for the stream of consciousness to flow, and whether the sufficient ones will also be necessary, remains to be studied in future. For now, an exciting program of NCC research should

continue searching for the new avenues. Among various proposals, such as a focus on how social interactions and culture modulate neural networks supporting phenomenal contents ([Singer this collection](#)), the present one claims that it's not just about the contents, and that a state of consciousness deserves a treatment of its own.

Acknowledgements

My research is supported by a Wellcome Trust Biomedical Research Fellowship WT093811MA (awarded to Dr. Tristan Bekinschtein).

References

- Bareham, C. A., Manly, T., Pustovaya, O. V., Scott, S. K. & Bekinschtein, T. A. (2014). Losing the left side of the world: Rightward shift in human spatial attention with sleep onset. *Scientific Reports*, *4*, 5092-5092. [10.1038/srep05092](https://doi.org/10.1038/srep05092)
- Bayne, T. (2010). *The Unity of Consciousness*. Oxford, UK: Oxford University Press.
- Bhattacharya, J., Petsche, H., Feldmann, U. & Rescher, B. (2001). EEG gamma-band phase synchronization between posterior and frontal cortex during mental rotation in humans. *Neuroscience Letters*, *311*, 29-32. [10.1016/S0304-3940\(01\)02133-4](https://doi.org/10.1016/S0304-3940(01)02133-4)
- Carhart-Harris, R. L., Erritzoe, D., Williams, T., Stone, J. M., Reed, L. J., Colasanti, A., Tyacke, R. J., Leech, R., Malizia, A. L., Murphy, K., Hobden, P., Evans, J., Feilding, A., Wise, R. G. & Nutt, D.J. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. *Proceedings of the National Academy of Sciences*, *109* (6), 2138-2143. [10.1016/S0304-3940\(01\)02133-4](https://doi.org/10.1016/S0304-3940(01)02133-4)
- Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G. & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, *5* (198), 198ra105-198ra105. [10.1126/scitranslmed.3006294](https://doi.org/10.1126/scitranslmed.3006294)
- Castro, S., Falconi, A., Chase, M. H. & Torterolo, P. (2013). Coherent neocortical 40-Hz oscillations are not present during REM sleep. *European Journal of Neuroscience*, *37* (8), 1330-1339. [10.1111/ejn.12143](https://doi.org/10.1111/ejn.12143)
- Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.) *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (pp. 31-63). Cambridge, MA: MIT Press.
- Clarke, C. J. S. (1995). The Nonlocality of mind. *Journal of Consciousness Studies*, *2* (3), 231-240.
- Crick, F. (1994). *The Astonishing Hypothesis*. New York, NY: Scribner.
- Dainton, B. (2006). *Stream of Consciousness: Unity and Continuity in Conscious Experience (2nd ed.)*. Abingdon, UK: Routledge.
- Del Cul, A., Dehaene, S., Reyes, P., Bravo, E. & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132* (Pt 9), 2531-2540. [10.1093/brain/awp111](https://doi.org/10.1093/brain/awp111)
- Doesburg, S. M., Green, J. J., McDonald, J. J. & Ward, L. M. (2009). Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS One*, *4* (7), e6142-e6142. [10.1371/journal.pone.0006142](https://doi.org/10.1371/journal.pone.0006142)
- (2012). Theta modulation of inter-regional gamma synchronization during auditory attention control. *Brain Research*, *1431*, 77-85. [10.1016/j.brainres.2011.11.005](https://doi.org/10.1016/j.brainres.2011.11.005)
- Engel, A. K., Fries, P., König, P., Brecht, M. & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition*, *8*, 128-151. [10.1006/ccog.1999.0389](https://doi.org/10.1006/ccog.1999.0389)
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience*, *32*, 209-224. [10.1146/annurev.neuro.051508.135603](https://doi.org/10.1146/annurev.neuro.051508.135603)
- Goupil, L. & Bekinschtein, T. A. (2011). Cognitive processing during the transition to sleep. *Archives Italiennes de Biologie*, *150* (2-3), 140-154.
- Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B. & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences*, *101* (35), 13050-13050. [10.1073/pnas.0404944101](https://doi.org/10.1073/pnas.0404944101)
- Gruber, T., Müller, M. M. & Keil, A. (2002). Modulation of induced gamma band responses in a perceptual learning task in the human EEG. *Journal of Cognitive Neuroscience*, *14* (5), 732-744. [10.1162/08989290260138636](https://doi.org/10.1162/08989290260138636)
- Hobson, J. A., Pace-Schott, E. F. & Stickgold, R. (2000). Dreaming and the brain: toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, *23* (6), 904-1121. [10.1017/S0140525X00003976](https://doi.org/10.1017/S0140525X00003976)
- James, W. (1890). *The Principles of Psychology*. New York, NY: Holt.
- Jouvet, M. (1979). What does a cat dream about? *Trends in Neurosciences*, *2*, 280-282. [10.1016/0166-2236\(79\)90110-3](https://doi.org/10.1016/0166-2236(79)90110-3)
- Kelly, E. R. (1882). *The Alternative: A Study in Psychology*. London, UK: Macmillan.
- Kiverstein, J. (2010). Making sense of phenomenal unity: An intentionalist account of temporal experience. *Royal Institute of Philosophy Supplement*, *67*, [10.1017/S1358246110000081](https://doi.org/10.1017/S1358246110000081)
- Kornmeier, J. & Bach, M. (2012). Ambiguous figures—what happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience*, *6*, 51-51. [10.3389/fnhum.2012.00051](https://doi.org/10.3389/fnhum.2012.00051)

- Kottlow, M., Jann, K., Dierks, T. & Koenig, T. (2012). Increased phase synchronization during continuous face integration measured simultaneously with EEG and fMRI. *Clinical Neurophysiology*, *123*, 1536-1548. [10.1016/j.clinph.2011.12.019](https://doi.org/10.1016/j.clinph.2011.12.019)
- Martini, N., Menicucci, D., Sebastiani, L., Bedini, R., Pingitore, A., Vanello, N., Milanese, M., Landini, L. & Gemignani, A. (2012). The dynamics of EEG gamma responses to unpleasant visual stimuli: From local activity to functional connectivity. *NeuroImage*, *60* (2), 922-932. [10.1016/j.neuroimage.2012.01.060](https://doi.org/10.1016/j.neuroimage.2012.01.060)
- Massimini, M., Ferrarelli, F., Murphy, M. J., Huber, R., Riedner, B. A., Casarotto, S. & Tononi, G. (2010). Cortical reactivity and effective connectivity during REM sleep in humans. *Cognitive Neuroscience*, *1*, 176-183. [10.1080/17588921003731578](https://doi.org/10.1080/17588921003731578)
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W. & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *Journal of Neuroscience*, *27*, 2858-2865. [10.1523/JNEUROSCI.4623-06.2007](https://doi.org/10.1523/JNEUROSCI.4623-06.2007)
- Merker, B. (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, *30*, 63-81. [10.1017/S0140525X07000891](https://doi.org/10.1017/S0140525X07000891)
- Móró, L. (2010). Hallucinatory altered states of consciousness. *Phenomenology and the Cognitive Sciences*, *9*, 241-252. [10.1007/s11097-010-9162-2](https://doi.org/10.1007/s11097-010-9162-2)
- Noreika, V. (2014). *Alterations in the States and Contents of Consciousness: Empirical and Theoretical Aspects*. Turku, Finland: University of Turku.
- Noreika, V., Valli, K., Lahtela, H. & Revonsuo, A. (2009). Early-night serial awakenings as a new paradigm for studies on NREM dreaming. *International Journal of Psychophysiology*, *74*, 14-18. [10.1016/j.ijpsycho.2009.06.002](https://doi.org/10.1016/j.ijpsycho.2009.06.002)
- Noreika, V., Jylhäkangas, L., Móró, L., Valli, K., Kaskinoro, K., Aantaa, R., Scheinin, H. & Revonsuo, A. (2011). Consciousness lost and found: Subjective experiences in an unresponsive state. *Brain and Cognition*, *77* (3), 327-334. [10.1016/j.bandc.2011.09.002](https://doi.org/10.1016/j.bandc.2011.09.002)
- Noreika, V., Canales-Johnson, A., Koh, J., Taylor, M., Massey, I. & Bekinschtein, T. A. (Under Review). Intrusions of a drowsy mind: Electroencephalographic correlates of phenomenological unpredictability.
- Nummenmaa, L., Glerean, E., Hari, R. & Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, *111*, 646-651. [10.1073/pnas.1321664111](https://doi.org/10.1073/pnas.1321664111)
- Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E. & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience*, *26* (28), 7523-7531. [10.1523/JNEUROSCI.1948-06.2006](https://doi.org/10.1523/JNEUROSCI.1948-06.2006)
- Parvizi, J. & Damasio, A. (2001). Consciousness and the brainstem. *Cognition*, *79* (1-2), 135-160. [10.1016/S0010-0277\(00\)00127-X](https://doi.org/10.1016/S0010-0277(00)00127-X)
- Pfurtscheller, G., Graimann, B., Huggins, J. E., Levine, S. P. & Schuh, L. A. (2003). Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement. *Clinical Neurophysiology*, *114* (7), 1226-1236. [10.1016/S1388-2457\(03\)00067-1](https://doi.org/10.1016/S1388-2457(03)00067-1)
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, *1* (2), 56-61. [10.1016/S1364-6613\(97\)01008-5](https://doi.org/10.1016/S1364-6613(97)01008-5)
- Railo, H., Koivisto, M. & Revonsuo, A. (2011). Tracking the processes behind conscious perception: A review of event-related potential correlates of visual consciousness. *Consciousness and Cognition*, *20* (3), 972-983. [10.1016/j.concog.2011.03.019](https://doi.org/10.1016/j.concog.2011.03.019)
- Revonsuo, A. (2006). *Inner presence. Consciousness as a Biological Phenomenon*. Cambridge, MA: MIT Press.
- Revonsuo, A., Kallio, S. & Sikka, P. (2009). What is an altered state of consciousness? *Philosophical Psychology*, *22* (2), 187-204. [10.1080/09515080902802850](https://doi.org/10.1080/09515080902802850)
- Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, *49* (3), 329-359. [10.1007/BF00355521](https://doi.org/10.1007/BF00355521)
- Siclari, F., LaRocque, J. J., Postle, B. R. & Tononi, G. (2013). Assessing sleep consciousness within subjects using a serial awakening paradigm. *Frontiers in Psychology*, *4* (542), 542-542. [10.3389/fpsyg.2013.00542](https://doi.org/10.3389/fpsyg.2013.00542)
- Singer, W. (2001). Consciousness and the binding problem. *Annals of the New York Academy of Sciences*, *929*, 123-146. [10.1111/j.1749-6632.2001.tb05712.x](https://doi.org/10.1111/j.1749-6632.2001.tb05712.x)
- (2015). The Ongoing Search for the Neuronal Correlate of Consciousness. In T. Metzinger & J. M. Windt (Eds.) *Open MIND* (pp. 1-30). Frankfurt a. M., GER: MIND Group.
- Sitt, J. D., King, J.-R., Karoui, I. E., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S. & Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain*, *137* (Pt 8), 2258-2270. [10.1093/brain/awu141](https://doi.org/10.1093/brain/awu141)
- Steriade, M., Contreras, D., Amzica, F. & Timofeev, I. (1996). Synchronization of fast (30-40 Hz) spontaneous

- oscillations in intrathalamic and thalamocortical networks. *Journal of Neuroscience*, 16 (8), 2788-2808.
- Tart, C. T. (1972). States of consciousness and state-specific sciences. *Science*, 176, 1203-1210.
[10.1126/science.176.4040.1203](https://doi.org/10.1126/science.176.4040.1203)
- Thompson, E. (2015). Dreamless sleep, the embodied mind, and consciousness: The relevance of a classical Indian debate to cognitive science. In T. Metzinger & J. M. Windt (Eds.) *Open MIND* (pp. 1-20). Frankfurt a. M., GER: MIND Group.
- Tononi, G. (2012). The integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie*, 150 (2-3), 56-90.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6 (2), 171-178.
[10.1016/S0959-4388\(96\)80070-5](https://doi.org/10.1016/S0959-4388(96)80070-5)
- Vaitl, D., Birbaumer, N., Gruzelier, J., Jamieson, G. A., Kotchoubey, B., Kübler, A., Lehmann, D., Miltner, W. H.R., Ott, U., Pütz, P., Sammer, G., Strauch, I., Strehl, U., Wackermann, J. & Weiss, T. (2005). Psychobiology of altered states of consciousness. *Psychological Bulletin*, 131 (1), 98-127.
[10.1037/0033-2909.131.1.98](https://doi.org/10.1037/0033-2909.131.1.98)
- Zhang, L., Gan, J. Q. & Wang, H. (2014). Optimized gamma synchronization enhances functional binding of fronto-parietal cortices in mathematically gifted adolescents during deductive reasoning. *Frontiers in Human Neuroscience*, 8 (430), eCollection 2014-eCollection 2014. [10.3389/fnhum.2014.00430](https://doi.org/10.3389/fnhum.2014.00430)

State or content of consciousness?

A Reply to Valdas Noreika

Wolf Singer

An attempt is made to distinguish between brain states required to support consciousness and the neuronal underpinnings of conscious versus non-conscious processing in an awake, attentive brain, respectively. It is argued that brain states supporting consciousness are characterised by high dimensional dynamics exhibiting a high degree of complexity, implying that conscious states are graded. Different mechanisms determine whether signals are processed at the conscious or sub-conscious level. Thus, there is no unique neuronal correlate of consciousness.

Keywords

Brain dynamics | Complexity | Conscious processing | Conscious state | Content of consciousness | Dimensionality

Author

Wolf Singer

w.singer@brain.mpg.de

Max Planck Institute for Brain Research (MPI)
Frankfurt a. M., Germany

Commentator

Valdas Noreika

valdas.noreika@mrc-cbu.cam.ac.uk

Medical Research Council
Cambridge, United Kingdom

Editors

Thomas Metzinger

metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt

jennifer.windt@monash.edu

Monash University
Melbourne, Australia

1 Introduction

My sincere thanks go to Valdas Noreika for having identified with succinct clarity the weaknesses in our current attempt to identify the neuronal correlates of consciousness (NCC). I would have sincerely appreciated these comments before finalising my manuscript, as they would have forced me to distinguish more clearly between the neuronal underpinnings of the conscious state and the neuronal correlates of conscious versus unconscious processing.

Noreika is absolutely right in pointing out that the search for the mechanisms permitting access to conscious processing falls short of

identifying the NCC proper and, likewise, that the determination of variables required for the maintenance of a conscious state is insufficient if pursued without considering the contents of conscious processing. The mere fact that one can distinguish between the “conscious state” and the conditions required for “conscious processing”, yet also consider both as targets in the search for the NCC, suggests that the explanandum is ill-defined. Presently, both studies devoted to the distinction between conscious and unconscious processing and those investigating the brain states required for conscious pro-

cessing are considered as investigations of the NCC, although they clearly target different neuronal mechanisms. Thus, studies on consciousness are fraught with the problem of a lack of a clear definition of “the” consciousness for which we wish to find a neuronal correlate. Another problem is that we are still far from fully understanding the neuronal mechanisms underlying higher cognitive functions. Behavioural studies suggest, for example, that perception involves probabilistic Bayesian-matching operations in which sensory evidence is compared with stored knowledge about the probability of occurrence and the features of the respective perceptual objects. However, it is entirely unknown where and how the huge amount of priors are stored, how the specific priors can be retrieved on the fly within the few hundreds of milliseconds sufficient for recognition, and how the matching operations are realized in neuronal networks. Thus, at the present stage it is even impossible to precisely define the signatures of neuronal activity that could be considered the result of a perceptual process or as the neuronal representation of a percept.

In the light of these uncertainties, the distinctions between conscious and unconscious processing or between states compatible with conscious and unconscious processing, respectively, appear to be exploited primarily in order to learn more about mechanisms underlying pattern recognition, decision making, and intentionality, rather than serving the search for the neuronal underpinnings of the ill-defined phenomenon that we address as “consciousness”. In contrast to NCC research, these more humble approaches have been quite successful, probably because the explananda are well-defined and can be operationalised.

2 The conscious state

The analysis of the neuronal prerequisites required for the maintenance of consciousness has a long history and has only recently been considered part of consciousness research. The reason for this is that the criteria used for distinctions between conscious and non-conscious states or altered states of consciousness can be

tested in both humans and animals. Examples of these criteria are reactivity to noxious stimuli, the ability to move intentionally, and the ability to accomplish a number of well-defined cognitive tasks, involving attention, short and long term memory, recognition, and decision making. Thus, the plethora of studies performed both on animals and humans on the neuronal mechanisms underlying arousal, attention, wakefulness, sleep, anaesthesia, and coma all contribute to our understanding of the neuronal prerequisites of states permitting conscious processing. Accordingly, it is well-established that brain functions characteristic of the conscious state require that neuronal networks operate in a critical dynamical range. This range is regulated by half a dozen globally-acting modulatory systems that originate in deep and evolutionary ancient brain structures. The adjustable neuronal parameters are essentially the balance between excitatory and inhibitory effects and the time- and length constants of dendritic integration. These adjustments lead to marked modifications of the system’s dynamics. These modulations are reflected by changes in the prevailing frequencies of oscillatory activity, the degree and spatial granularity of synchronisation (also addressed as correlation length), and the propagation of signals across the network.

Classical brain theories have not attributed much attention to the significance of these dynamic variables for processing and assume that loss of consciousness in sleep and anaesthesia is essentially due to reduced excitability and signal transmission. However, in more recent theories, brain dynamics are thought to play a crucial role in information processing. This novel framework provides much more specific explanation of the breakdown of consciousness in sleep, anaesthesia, and coma. These theories posit that oscillations and the concomitant variables, such as synchronisation, phase locking, phase relations, and cross frequency coupling, are relevant for signal selection by attention, binding operations, and the representation of nested semantic relations (for review see [Singer 1999](#); [Buzsáki et al. 2013](#)). In addition, these complex dynamics have been proposed as a substrate for the generation of the high-di-

mensional coding space required for the storage and superposition of priors, the matching of stored information with sensory evidence, and the segregation of patterns for classification (for review see [Singer 2013](#)). The basis of these operations is the transformation of low-dimensional input patterns into high-dimensional dynamic states, in order to perform the necessary computations in this space and to then retransform the results into low-dimensional output signals. The advantages of performing computations in high-dimensional dynamic space are currently explored in the conceptual framework of “reservoir computing” or “liquid state or echo state machines” ([Bertschinger & Natschläger 2004](#); [Buonomano & Maass 2009](#); [Jaeger 2001](#)).

Recent analysis of the properties of recurrent networks, such as those realized in neuronal systems and in particular the cerebral cortex, indicate that such high-dimensional dynamic states can indeed be generated in delay-coupled networks ([Lazar et al. 2009](#); [Buonomano & Maass 2009](#); [Soriano et al. 2013](#); for review see [Singer 2013](#)). In the present context it is important to recall that the dynamics required for such computations can emerge only when the networks are in the appropriate state. The optimal state has been identified as the edge of chaos, slightly below self-organised criticality, the so-called SOC state, because in this state the dimensionality or the complexity of the system are very high. Computationally this range is optimal because it offers a maximum of possible bifurcation points and storage capacity. ([Plenz & Thiagarajan 2007](#)). In this conceptual framework, computational results should consist of substates with reduced dimensionality. Experimental evidence indicates that the high-dimensional resting states are actually reduced by sensory input, imagery, recall of memories, or focused attention. These processes are all associated with enhanced correlation between neuronal responses due to the induction of synchronized high-frequency oscillations—where enhancing correlations reduces dimensionality (for review see [Singer 2013](#)). The notion that SOC states are optimal prerequisites for processing also fits with the robust evidence that states compatible with consciousness are characterized

by “desynchronized” brain activity, i.e., states characterized by uncorrelated activity, such as are typical for wakefulness and arousal. If, and evidence suggests this to be the case (for review see [Singer 1999, 2013](#)), establishment of lower-dimensional synchronous substates, e.g., the formation of transiently-synchronized assemblies of neurons, is an integral part of the computations, then dynamic states characterized by global, large scale synchrony would be inappropriate as background for computations underlying higher cognitive functions.

As outlined in the target paper and above, higher cognitive functions require fine-grained binding operations among semantically-related contents that need to be encoded in ad hoc-formed neuronal assemblies. Such concatenation of multiple assemblies by partial correlations and perhaps also cross-frequency coupling would be impossible in networks that are already highly synchronized to begin with and hence exhibit low complexity and dimensionality. The well-established notion that deep sleep, anaesthesia, and most forms of coma are associated with brain states that exhibit slow oscillations synchronized over considerable distances agrees with this interpretation. In agreement with the prediction that low-dimensional brain states are incompatible with sophisticated processing are also the recent stimulation experiments cited by Noreika. It is to be expected that stimulation of a dynamic system that is in a low-dimensional state and at an overall reduced level of excitability will elicit only a spatially-restricted responses of low complexity—in particular if the stimulus is itself very low-dimensional, as is the case for a TMS pulse.

Considering more recent theories on brain functions, it appears as if the prerequisite or the NCC of a conscious state is a dynamic state that assures a high degree of complexity and high-dimensionality of resting-state dynamics. It is only in this state that the higher cognitive functions can be realized that one expects from a conscious brain.

It should be noted, however, that this operational definition of consciousness makes no inferences about the subjective contents of consciousness or the awareness of particular qualia

of experience. According to this definition, consciousness is simply a brain state that allows animals and humans to accomplish higher cognitive functions that include not only perception but also decision making, planning of actions, generation of procedural and episodic memories, and last but not least intentionality and reasoning. Thus, one would expect consciousness, defined in this way, to be a graded phenomenon. If the state of the brain changes towards reduced complexity and dimensionality, there should be a graded deterioration of functions. Those requiring integration of widely-distributed assemblies should become impeded first, while simple reactions to salient sensory stimuli would persist for much longer. This seems to be in perfect agreement with the gradual deterioration of cognitive functions as the brain state shifts from high levels of alertness to drowsiness and sleep.

3 Conscious versus subconscious processing

As Noreika points out, “consciousness” defined by the status of phenomenal content is something very different from a conscious state, as this connotation of consciousness can only be investigated in human subjects. The reason for this is that the distinguishing criterion is the degree of subjective awareness of a cognitive content, and this variable can only be assessed through verbal report. It is simply not possible to know whether a monkey trained to press a lever to signal that it has recognized a particular pattern has the subjective experience that we equate with conscious perception. The monkey brain has the same mechanisms as humans for the allocation of attention, the selection of objects for perception, and the routing of experiences to the different storage modes (working memory, procedural and episodic memory). Thus it is very likely that monkeys are aware of their perceptions in a similar way to us, and that the distinction between conscious and non-conscious processing holds for them as well—but we have no way of knowing. Conditioned lever presses in response to stimuli do not require conscious perception of the stimuli, just as

stopping at a red light while being engaged in a conversation does not require conscious recollection of having perceived the light. It is for this reason that the criterion for conscious processing is the reportability of the perceived stimulus, and hence this aspect of consciousness can only be studied in humans.

Attempts to identify the differences between the neuronal processes that accompany non-conscious and conscious processing, respectively, are of course interesting in their own right. The expectation is that they will provide answers to the question of why certain processes are reportable and have access to working and episodic memory while others are excluded, or the question of why certain forms of reasoning and decision-making require conscious deliberations while others do not. However, as pointed out so stringently by Noreika, these attempts fall short of identifying the NCC proper, and at best cover some aspects of conscious processing while being fraught with problems. The most difficult problems are related to the distinction between the processes that are essential for subjective awareness and reportability and those that are the consequence of having become aware of something or that simply provide favourable conditions for becoming aware, such as the allocation of attention or the saliency of stimuli. So far the only neuronal signatures distinguishing between reportable and non-reportable processes have been found to be transitory, lasting at most a few hundred milliseconds. Noreika argues rightly that this disqualifies these events as NCCs because the stream of consciousness is continuous and the awareness of contents can persist for quite some time.

4 Conclusion and outlook

We need to be more cautious when using the term NCC and to define precisely, each time we perform a search for underlying neuronal mechanisms, which of the many aspects of “consciousness” we actually intend to investigate. We need to differentiate between processes assuring access to conscious processing, which are expected to be transient, and processes necessary for sustaining the stream of consciousness

that has longer time-constants. And finally, we need to distinguish processes assuring sustained awareness of contents that are most likely related to the transfer of material to short- and long-term memories. If we proceed in this way, subdividing “consciousness” into subfunctions including reportability and defining these as explananda, some of the present problems may dissolve. However, the consequence is that we shall have to give up the search for “the” overarching NCC.

If we pursue this agenda, it is to be expected that correlates will be found for all aspects of consciousness except those associated with the “hard” problem, which appears to be a specific human problem. As I argued in the target paper, searching for the neuronal correlates of qualia in individual brains is unlikely to be successful because the immaterial and therefore somewhat mysterious connotations of qualia are likely to have the status of social realities. What we can achieve, however, is an identification of brain processes that underlie those cognitive functions required for generating social realities. These would be the ability to engage in social interaction, to develop a theory of mind, to find symbolic descriptions of internal states, and to reach consensus on the “reality” of these through communication with others.

To conclude this brief reply to the extremely inspiring commentary on my target paper, I want to express my sincere gratitude to Noreika for having pointed out the critical issues in our research on the NCC. The reply forced me to engage with this research again and helped me substantially in clarifying my own position in the debate.

References

- Bertschinger, N. & Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. *Neural Computation*, 16 (7), 1413-1436.
[10.1162/089976604323057443](https://doi.org/10.1162/089976604323057443)
- Buonomano, D. V. & Maass, W. (2009). State-dependent computations: Spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10, 113-125.
[10.1038/nrn2558](https://doi.org/10.1038/nrn2558)
- Buzsáki, G., Logothetis, N. & Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron*, 80 (3), 751-764.
[10.1016/j.neuron.2013.10.002](https://doi.org/10.1016/j.neuron.2013.10.002)
- Jaeger, H. (2001). The “echo state” approach to analysing and training recurrent neural networks - with an Erratum note. *German National Research Center for Information Technology, GMD Report*, 148
- Lazar, A., Pipa, G. & Triesch, J. (2009). SORN: A self-organizing recurrent neural network. *Frontiers in Computational Neuroscience*, 3 (23), 1-9.
[10.3389/neuro.10.023.2009](https://doi.org/10.3389/neuro.10.023.2009)
- Plenz, D. & Thiagarajan, T. C. (2007). The organizing principles of neuronal avalanches: Cell assemblies in the cortex? *Trends in Neurosciences*, 30 (3), 99-110.
[10.1016/j.tins.2007.01.005](https://doi.org/10.1016/j.tins.2007.01.005)
- Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, 24 (1), 49-65.
[10.1016/S0896-6273\(00\)80821-1](https://doi.org/10.1016/S0896-6273(00)80821-1)
- (2013). Cortical dynamics revisited. *Trends in Cognitive Sciences*, 17 (12), 616-626.
[10.1016/j.tics.2013.09.006](https://doi.org/10.1016/j.tics.2013.09.006)
- Soriano, M. C., Garcia-Ojalvo, J., Mirasso, C. R. & Fischer, I. (2013). Complex photonics: Dynamics and applications of delay-coupled semiconductor lasers. *Review of Modern Physics*, 85 (1), 421-470.
[10.1103/RevModPhys.85.421](https://doi.org/10.1103/RevModPhys.85.421)