

Consciousness: the last 50 years (and the next)

Anil K. Seth

Sackler Centre for Consciousness Science

School of Engineering and Informatics

University of Sussex

Brighton BN1 9QJ

United Kingdom

a.k.seth@sussex.ac.uk

www.anilseth.com

@anilkseth

Abstract. The mind and brain sciences began with consciousness as a central concern. But for much of the twentieth century, ideological and methodological concerns relegated its empirical study to the margins. Since the 1990s, studying consciousness has regained a legitimacy and momentum befitting its status as the primary feature of our mental lives. Nowadays, consciousness science encompasses a rich interdisciplinary mixture drawing together philosophical, theoretical, computational, experimental, and clinical perspectives, with neuroscience its central discipline. Researchers have learned a great deal about the neural mechanisms underlying global states of consciousness, distinctions between conscious and unconscious perception, and self-consciousness. Further progress will depend on specifying closer explanatory mappings between (first-person subjective) phenomenological descriptions and (third-person objective) descriptions of (embodied and embedded) neuronal mechanisms. Such progress will help reframe our understanding of our place in nature, and accelerate clinical approaches to a wide range of psychiatric and neurological disorders.

Keywords: consciousness, neural correlates of consciousness, selfhood, volition, predictive coding.

Introduction

The relationship between subjective conscious experience and its biophysical basis has always been a defining question for the mind and brain sciences. But, at various times since the beginnings of neuroscience as a discipline, the explicit study of consciousness has been either treated as fringe or excluded altogether. Looking back over the past fifty years, these extremes of attitude are well represented. Roger Sperry, pioneer of split-brain operations and of what can now be called ‘consciousness science’ lamented in 1969 that “[m]ost behavioral scientists today, and brain researchers in particular, have little use for consciousness” (Sperry, 1969) (p.532). Presciently, in the same article he highlighted the need for new technologies able to record the “pattern dynamics of brain activity” in elucidating the neural basis of consciousness. Indeed, modern neuroimaging methods have had a transformative impact on consciousness science, as they have on cognitive neuroscience generally.

Informally, consciousness science over the last fifty years can be divided into two epochs. From the mid 1960s until around 1990 the fringe view held sway, though with several notable exceptions. Then, from the late 1980s and early 1990s, first a trickle and more recently a deluge of research into the brain basis of consciousness, a transition catalysed by – among other things – the activities of certain high-profile scientists (for example the Nobel laureates Francis Crick and Gerald Edelman) and by the maturation of modern neuroimaging methods, as anticipated by Sperry.

Today, students of neuroscience – for the most part – feel able to declare (or deny) a primary interest in studying consciousness. There are academic societies and conferences going back more than 20 years and scholarly journals dedicated to the topic. Above all, there is growing a body of empirical and theoretical work drawing ever closer connections between the properties of subjective experience and the operations of the densely complex neural circuits, embodied in bodies embedded in environments, that together give rise to the apparent miracle of consciousness. We can’t yet

know whether today's students will find the solution to the 'problem of consciousness', or whether the problem as currently set forth is simply misconceived. Either way, there is much more to be discovered about the relations between the brain and consciousness, and with these discoveries will come new clinical approaches in neurology and psychiatry, as well as a new appreciation of our place as part of, and not apart from, the rest of nature.

Consciousness: 1960s until 1990

By the mid 1960s, the behaviourism that had dominated 20th Century psychology (especially in America) was in retreat. A new cognitive science was emerging that recognised the existence and importance of inner mental states in mediating between stimulus and response. But consciousness as an explanatory target was still largely off-limits: "We should ban the word 'consciousness' for a decade or two" as cognitive scientist George Miller put it in 1962. As late as 1989, Stuart Sutherland wrote in the *International Dictionary of Psychology*: "Consciousness is a fascinating but elusive phenomenon. It is impossible to specify what it is, what it does, or why it evolved. Nothing worth reading has been written on it." (Sutherland, 1989) (p.95).

Whatever its merits at the time, Sutherland's judgement seems harsh in retrospect. Although lacking a coherent organisation or much coordination, the neuroscience of consciousness in this period saw several substantial advances. These advances, admittedly, were mainly about establishing correspondences between brain regions or activity patterns and properties of conscious experience, rather than addressing fundamental questions such as why consciousness is part of the universe in the first place. Good examples come from neurology and neuropsychology, where surgical interventions, and the study of neurological illnesses, injuries and lesions were revealing the essential dependency of particular aspects of consciousness on specific brain properties. Some of these early studies remain among the most thought provoking even today.

The split-brain (more precisely: callosotomy) studies of Sperry and Michael Gazzaniga are a case in point. Their first experiments, on the World War II veteran W.J., revealed that each cerebral hemisphere could perceive a visual stimulus independently, with only the left hemisphere (in W.J.) being able to provide a verbal report (Gazzaniga, Bogen, & Sperry, 1962). Follow-up studies found that the somatosensory system, the motor system, and many other perceptual and cognitive systems could be similarly ‘split’, while other systems – for example emotion – remained intact (Gazzaniga, 2014). The extent to which a single cranium can house independent conscious subjects is still hotly debated (Gazzaniga, 2014; Pinto et al., 2017; Sasai, Boly, Mensen, & Tononi, 2016), perhaps because the very idea challenges one of our most deeply held assumptions: that consciousness is necessarily unified.

The unity of consciousness was also challenged, more subtly, by psychosurgical lesions of the medial temporal lobe, which were carried out to alleviate intractable epilepsy. The case of Henry Moliason – patient H.M. – is well known to many. Following bilateral removal of the medial temporal lobe in 1953, including both hippocampi, H.M. was cured of his epilepsy but left with a profound anterograde (and substantial retrograde) amnesia (Scoville & Milner, 1957). In a series of studies, neuropsychologists Suzanne Corkin and Brenda Milner found that despite living in what Corkin called the “permanent present tense,” H.M. could learn new motor skills, had intact working and semantic memory, and was generally able to acquire a range of implicit (non-conscious) memories (Corkin, 2013). Only his ability to acquire new explicit, conscious memories was affected. These findings not only charted a new topography of conscious and unconscious kinds of memory, they showed how our apparently unified sense of conscious selfhood – in which episodic memories play a key role – can fragment so that some aspects persist while others are lost.

Another central feature of conscious selfhood is the experience of ‘free will’, or more precisely, experiences of volition (intentions to do this-or-that) and agency (being the cause of events). Here,

the experiments of Benjamin Libet in the 1980s continue to inspire new research and to stoke controversy in equal measure. His studies, designed to measure the timing of conscious decisions to make voluntary movements, were based on a remarkably simple paradigm [(Libet, 1982) see Figure 1A]. Participants press a button at a time of their own choosing, and then report the time they felt the ‘urge’ to move – their *conscious intention* – by noting the position of a dot on an oscilloscope screen.

---- FIGURE 1 ABOUT HERE ----

Libet first observed a previously described build-up of neural electrical activity prior to voluntary movement – the so-called ‘readiness potential’ (Kornhuber & Deecke, 1965). His key innovation was to show that this build-up started several hundred milliseconds before the participant was aware of their intention to move, challenging the assumption that the conscious ‘urge’ was the cause of the voluntary movement. This interpretation has been debated ever since, sparking many fascinating experiments (Haggard, Clark, & Kalogeras, 2002; Schurger, Sitt, & Dehaene, 2012). Libet himself was uncomfortable with the idea that conscious intentions were epiphenomenal, suggesting instead that the time between the conscious ‘urge’, and the actual movement was sufficient to allow a conscious ‘veto’ to take effect. Any conscious ‘veto’, however, is also likely to have identifiable neural precursors – so this in itself doesn’t resolve Libet’s metaphysical quandary (Brass & Haggard, 2007). Perhaps the most convincing interpretation of these puzzling phenomena is that experiences of intention and agency label particular actions – and their consequences – as being self-generated rather than externally imposed, allowing the organism to learn and perhaps make better (voluntary) decisions in the future (Haggard, 2008).

Consciousness: from the 1990s until the present

A convenient date to mark the rehabilitation of consciousness within neuroscience is with Francis Crick and Christof Koch’s landmark 1990 paper, ‘Towards a neurobiological theory of

consciousness’, which opens with the line: “It is remarkable that most of the work in both cognitive science and the neurosciences makes no reference to consciousness (or ‘awareness’), and which goes on to propose a specific theory of visual consciousness based on gamma-band (~40 Hz) oscillations (Crick & Koch, 1990).

Although this specific idea has now fallen from favour, a new industry rapidly developed aimed at uncovering the so-called ‘neural correlates of consciousness’ (NCCs): “the minimal neuronal mechanisms jointly sufficient for any one conscious percept” (Crick & Koch, 1990). The search for NCCs, boosted by the arrival of the now ubiquitous MRI scanner (as well as good-old-fashioned EEG, and invasive neurophysiology in non-human primate studies), gave consciousness research a pragmatic spin. Instead of worrying at the so-called ‘hard problem’ of how conscious experiences could ever arise from ‘mere’ matter (Chalmers, 1996), neuroscientists could get on with looking for brain regions or processes that reliably correlated with particular conscious experiences, or with being conscious at all.

---- FIGURE 2 ABOUT HERE ----

Over the past quarter-century there has been considerable progress in identifying candidate NCCs (Koch, Massimini, Boly, & Tononi, 2016; Metzinger, 2000; Odegaard, Knight, & Lau, 2017), at least in specific contexts like visual or auditory awareness. A classic method has been to compare brain activity for ‘conscious’ and ‘unconscious’ conditions, while keeping sensory stimulation (and as far as possible, everything else) constant. For example, in *binocular rivalry*, conscious perception alternates even though the sensory inputs (different images to each eye) remain the same. Early studies of the brain basis of binocular rivalry used implanted electrodes in monkeys trained to report which of two visual percepts was dominant. These studies found that neuronal responses in early visual areas – in particular V1 – tracked the physical stimulus rather than the percept, while

neuronal responses in ‘higher’ areas – like inferotemporal cortex (IT) – tracked the percept rather than the physical stimulus (Leopold & Logothetis, 1996; Logothetis & Schall, 1989). Subsequent human neuroimaging studies, however, found that neuronal activity in primary visual cortex did correlate with perceptual dominance (Polonsky, Blake, Braun, & Heeger, 2000), and the debate continues as to whether the neuronal mechanisms underlying perceptual transitions lie early in the visual stream, or in higher-order regions such as the parietal or frontal cortices (Blake, Brascamp, & Heeger, 2014). As well as rivalry experiments, so-called ‘masking’ paradigms have also been widely used in consciousness science. These paradigms enable comparison of supraliminal versus subliminal stimulus presentations in a variety of perceptual modalities, with many studies implicating activation of the fronto-parietal network in reportable conscious perception (Dehaene & Changeux, 2011) (Figure 2). These two examples stand for many others; see (Boly et al., 2017; Odegaard et al., 2017) for recent, and conflicting, reviews.

At the same time, a separate strand of research has focussed on transitions in conscious states, both reversible (e.g., sleep and anaesthesia, (Massimini et al., 2005)) and following brain injury (e.g., coma and the vegetative state, (Owen, Schiff, & Laureys, 2009)). Here the challenge is to identify the neural mechanisms that support being conscious at all, rather than those associated with being conscious of this-or-that. One difficulty here is that global transitions of this kind affect the brain and body very generally, so that it is challenging to isolate the neural mechanisms underlying consciousness *per se*. There are additional difficulties in distinguishing so-called ‘enabling’ conditions from those neural mechanisms that actually support conscious states. For example, certain brainstem lesions can abolish consciousness forever, but many believe that the brainstem merely enables conscious states, while the actual ‘generators’ of consciousness may lie elsewhere (Dehaene & Changeux, 2011) (though see (Merker, 2007)).

New theories have accompanied these empirical developments. One of the most influential is Bernard Baars' 'global workspace' theory, which proposes that modular and specialized processors compete for access to a 'global workspace' (Baars, 1988). Mental states become conscious when they are 'broadcast' within the workspace, so that they can influence other processes, including verbal report and motor action. More recent 'neuronal' versions of this theory associate the global workspace with highly interconnected fronto-parietal networks, linking conscious perception to a nonlinear 'ignition' of activity within these networks, a position in line with many neuroimaging studies (Dehaene & Changeux, 2011; Dehaene, Sergent, & Changeux, 2003) (Figure 2).

Workspace theories tend to interpret conscious perception in terms of 'access,' in the sense that a percept is defined as conscious only if it is available for verbal (or other behavioural) report, as well as to other cognitive processes (memory, attention, and so on). An advantage of this view is that conscious status is readily assessable in experiments, since conscious content is by definition reportable. However, another common intuition is that perceptual or 'phenomenal' consciousness is 'richer' than we can report at any time, since reportability is limited by (among other things) the constraints of memory. The distinction between phenomenal consciousness and access consciousness (Block, 2005) remains a productive source of new experiments as well as controversy (Tsuchiya, Wilke, Frassle, & Lamme, 2015).

The road ahead

These are exciting times in consciousness science, and there is only space here to gesture at some promising research directions.

In terms of conscious *level*, new theories and measures have emerged based on 'neuronal complexity' and 'integrated information' (Seth, Barrett, & Barnett, 2011; Tononi, Boly, Massimini, & Koch, 2016). The basic idea is that conscious scenes are both highly *integrated* (each is

experienced as a unified whole) and highly *informative* (each conscious scene is one among a vast repertoire of alternative possibilities), motivating the development of mathematical metrics that combine the same properties. Excitingly, some practical approximations to these measures are showing promise in quantifying ‘residual’ awareness after brain injury, without relying on overt behaviour (Casali et al., 2013).

Research on conscious *content* has continued to focus on brain regions or processes that distinguish conscious from unconscious perceptions. Addressing a recurring concern about the NCC approach, new experimental paradigms are refining our understanding of interactions between the neuronal mechanisms underlying conscious perception, with those underlying behavioural report. So-called “no report” paradigms, which infer perceptual transitions indirectly from (for example) automatic eye movements, are challenging the idea that frontal brain regions are constitutively involved in conscious perception [(Frassle, Sommer, Jansen, Naber, & Einhauser, 2014), though see (van Vugt et al., 2018) for evidence to the contrary]. At the same time, advances in analysis methods like signal detection theory (Green & Swets, 1966) are enabling researchers to draw more rigorous distinctions between objective and subjective aspects of perception, as well as to quantify individual differences in metacognition (cognition about cognition) that may be relevant to consciousness (Barrett, Dienes, & Seth, 2013; Fleming & Dolan, 2012).

Theoretically, ‘predictive coding’ or ‘Bayesian brain’ approaches stand to advance our understanding of the neural basis of conscious perception. These approaches model perception as a process of (possibly Bayesian) inference on the hidden causes of the ambiguous and noisy signals that impinge on our sensory surfaces (Friston, 2009). Inverting some classical views on perception, top-down signals are proposed to convey perceptual predictions, while bottom-up signals convey only or primarily ‘prediction errors’: the discrepancy between what the brain expects and what it gets at each level of processing. This framework provides a powerful interpretation of some

relatively old findings associating conscious perception with the integrity of top-down signalling (Pascual-Leone & Walsh, 2001) and is motivating new studies which explicitly manipulate perceptual expectations, examining how this alters conscious perception (de Lange, Heilbron, & Kok, 2018). Excitingly, these ideas may provide a mechanistic understanding of abnormal perception in some clinical contexts, a good example being the positive symptoms (e.g., hallucinations) of psychosis (Fletcher & Frith, 2009; Powers, Mathys, & Corlett, 2017; Teufel et al., 2015).

There is also increasing focus in consciousness research on experiences of *selfhood*, which encompass basic experiences of embodiment and body ownership (Blanke, Slater, & Serino, 2015), experiences of volition and agency (Haggard, 2008), as well as ‘higher’ aspects of selfhood such as episodic memory and social perception. Here, new developments in virtual and augmented reality (Lenggenhager, Tadi, Metzinger, & Blanke, 2007; Seth, 2013), as well as in characterising interoception (the sense of the body ‘from within’ (Critchley, Wiens, Rotshtein, Ohman, & Dolan, 2004)) are heralding new insights into how our apparently unified experience of being a ‘self’ is constructed, on the fly, from many potentially distinguishable sub-processes – and how breakdowns in this constructive process may underlie a variety of psychiatric conditions.

From this vantage point in time, 50 years after the birth of the British Neuroscience Association, it is fair to say that the scientific study of consciousness has regained its rightful place as a central theme in the mind and brain sciences. A great deal is now known about how embodied and embedded brains shape and give rise to various aspects of conscious level, content, and self. Of course, much more remains to be discovered. Exciting new combinations of theory, experiment, and modelling are helping transform mere correlations into explanations that map from neural mechanism to phenomenology. Accompanying these developments are important clinical applications in neurology and psychiatry, as well as the deep challenges of investigating

consciousness in infancy, in other non-human animals, and perhaps even in future machines.

Whether or not the 'hard' problem of consciousness will yield to these and other developments, the history of the next 50 years will make for fascinating reading.

Declaration of competing interests. The author declares there is no conflict of interest.

Funding. I am grateful to the Dr. Mortimer and Theresa Sackler Foundation which supports the Sackler Centre for Consciousness Science. I also acknowledge support from the Wellcome Trust (Engagement Fellowship) and the Canadian Institute for Advanced Research, Azrieli Programme on Brain, Mind, and Consciousness.

References

- Baars, B. J. (1988). *A cognitive theory of consciousness*. New York, NY: Cambridge University Press.
- Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of Metacognition on Signal-Detection Theoretic Models. *Psychological Methods*, *18*(4), 535-552. doi:10.1037/A0033268
- Blake, R., Brascamp, J., & Heeger, D. J. (2014). Can binocular rivalry reveal neural correlates of consciousness? *Philos Trans R Soc Lond B Biol Sci*, *369*(1641), 20130211. doi:10.1098/rstb.2013.0211
- Blanke, O., Slater, M., & Serino, A. (2015). Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. *Neuron*, *88*(1), 145-166. doi:10.1016/j.neuron.2015.09.029
- Block, N. (2005). Two neural correlates of consciousness. *Trends Cogn Sci*, *9*(2), 46-52.
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *J Neurosci*, *37*(40), 9603-9613. doi:10.1523/JNEUROSCI.3218-16.2017
- Brass, M., & Haggard, P. (2007). To do or not to do: the neural signature of self-control. *J Neurosci*, *27*(34), 9141-9145. doi:10.1523/JNEUROSCI.0924-07.2007
- Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., . . . Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med*, *5*(198), 198ra105. doi:10.1126/scitranslmed.3006294
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. New York, NY: Oxford University Press.
- Corkin, S. (2013). *Permanent present tense*: Allen Lane.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263-275.
- Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nat Neurosci*, *7*(2), 189-195.
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends Cogn Sci*. doi:<https://doi.org/10.1016/j.tics.2018.06.002>
- Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200-227. doi:S0896-6273(11)00258-3 [pii] 10.1016/j.neuron.2011.03.018
- Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc Natl Acad Sci U S A*, *100*(14), 8520-8525.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philos Trans R Soc Lond B Biol Sci*, *367*(1594), 1338-1349. doi:10.1098/rstb.2011.0417
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*, *10*(1), 48-58. doi:nrn2536 [pii] 10.1038/nrn2536
- Frassle, S., Sommer, J., Jansen, A., Naber, M., & Einhauser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J Neurosci*, *34*(5), 1738-1747. doi:10.1523/JNEUROSCI.4403-13.2014
- Friston, K. J. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn Sci*, *13*(7), 293-301.
- Gazzaniga, M. S. (2014). The split-brain: rooting consciousness in biology. *Proc Natl Acad Sci U S A*, *111*(51), 18093-18094. doi:10.1073/pnas.1417892111

- Gazzaniga, M. S., Bogen, J. E., & Sperry, R. W. (1962). Some functional effects of sectioning the cerebral commissures in man. *Proc Natl Acad Sci U S A*, *48*, 1765-1769.
- Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory*. New York, NY: Wiley.
- Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat Rev Neurosci*, *9*(12), 934-946.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci*, *5*(4), 382-385. doi:10.1038/nn827
nn827 [pii]
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci*, *17*(5), 307-321. doi:10.1038/nrn.2016.22
- Kornhuber, H. H., & Deecke, L. (1965). [Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials.]. *Pflugers Arch Gesamte Physiol Menschen Tiere*, *284*, 1-17.
- Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science*, *317*(5841), 1096-1099.
- Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549-553.
- Libet, B. (1982). Brain stimulation in the study of neuronal functions for conscious sensory experiences. *Hum Neurobiol*, *1*(4), 235-242.
- Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, *245*(4919), 761-763.
- Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science*, *309*(5744), 2228-2232.
- Merker, B. (2007). Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav Brain Sci*, *30*(1), 63-81; discussion 81-134. doi:10.1017/S0140525X07000891
- Metzinger, T. (Ed.) (2000). *Neural correlates of consciousness: Empirical and conceptual questions*. Cambridge, MA: MIT Press.
- Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? *J Neurosci*, *37*(40), 9593-9602. doi:10.1523/JNEUROSCI.3217-16.2017
- Owen, A. M., Schiff, N. D., & Laureys, S. (2009). A new era of coma and consciousness science. *Prog Brain Res*, *177*, 399-411. doi:10.1016/S0079-6123(09)17728-2
- Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, *292*(5516), 510-512. doi:10.1126/science.1057099
- Pinto, Y., Neville, D. A., Otten, M., Corballis, P. M., Lamme, V. A., de Haan, E. H., . . . Fabri, M. (2017). Split brain: divided perception but undivided consciousness. *Brain*. doi:10.1093/brain/aww358
- Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat Neurosci*, *3*(11), 1153-1159. doi:10.1038/80676
- Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science*, *357*(6351), 596-600. doi:10.1126/science.aan3458
- Sasai, S., Boly, M., Mensen, A., & Tononi, G. (2016). Functional split brain in a driving/listening paradigm. *Proc Natl Acad Sci U S A*, *113*(50), 14444-14449. doi:10.1073/pnas.1613200113
- Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of*

- Sciences of the United States of America*, 109(42), E2904-2913.
doi:10.1073/pnas.1210467109
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry*, 20(1), 11-21.
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci*, 17(11), 565-573. doi:10.1016/j.tics.2013.09.007
- Seth, A. K., Barrett, A. B., & Barnett, L. (2011). Causal density and integrated information as measures of conscious level. *Philos Trans A Math Phys Eng Sci*, 369(1952), 3748-3767. doi:10.1098/rsta.2011.0079
- Sperry, R. W. (1969). A modified concept of consciousness. *Psychol Rev*, 76(6), 532-536.
- Sutherland, S. (1989). *International Dictionary of Psychology*: Crossroad Classic.
- Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P. R., . . . Fletcher, P. C. (2015). Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proc Natl Acad Sci U S A*, 112(43), 13401-13406. doi:10.1073/pnas.1503916112
- Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci*, 17(7), 450-461. doi:10.1038/nrn.2016.44
- Tsuchiya, N., Wilke, M., Frassle, S., & Lamme, V. A. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends Cogn Sci*. doi:10.1016/j.tics.2015.10.002
- van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, 360(6388), 537-542. doi:10.1126/science.aar7186

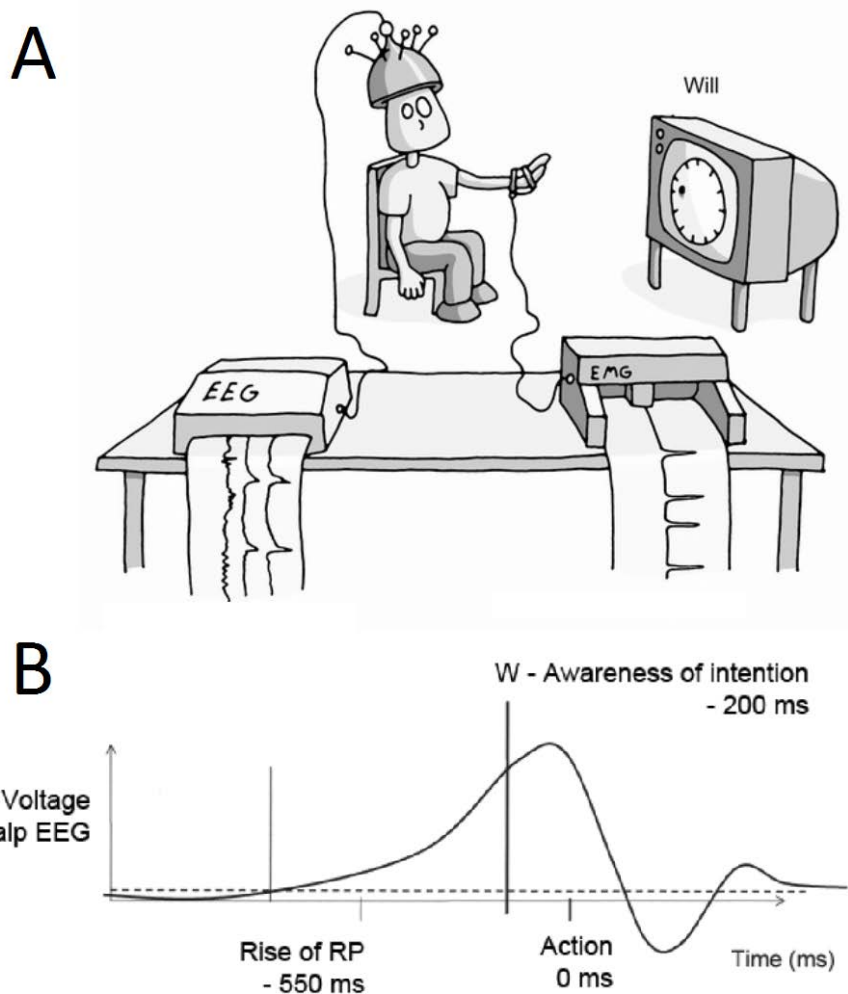


Figure 1. A. The Libet Paradigm. The participant makes a voluntary action and reports the time they felt the ‘conscious urge’ to move, by noting the position of the dot on the screen. Brain signals are measured using EEG and the timing of the actual movement through electromyography (EMG) attached to the wrist. Courtesy of Jolyon Troscianko (<http://www.jolyon.co.uk>) B. Schematized readiness potentials in the scalp EEG. Critically, the readiness potential begins to rise before the participant is aware of their decision to move.

Source: Jolyon Troscianko (<http://www.jolyon.co.uk/illustrations/consciousness-a-very-short-introduction-2/>) [free for academic non-profit use]

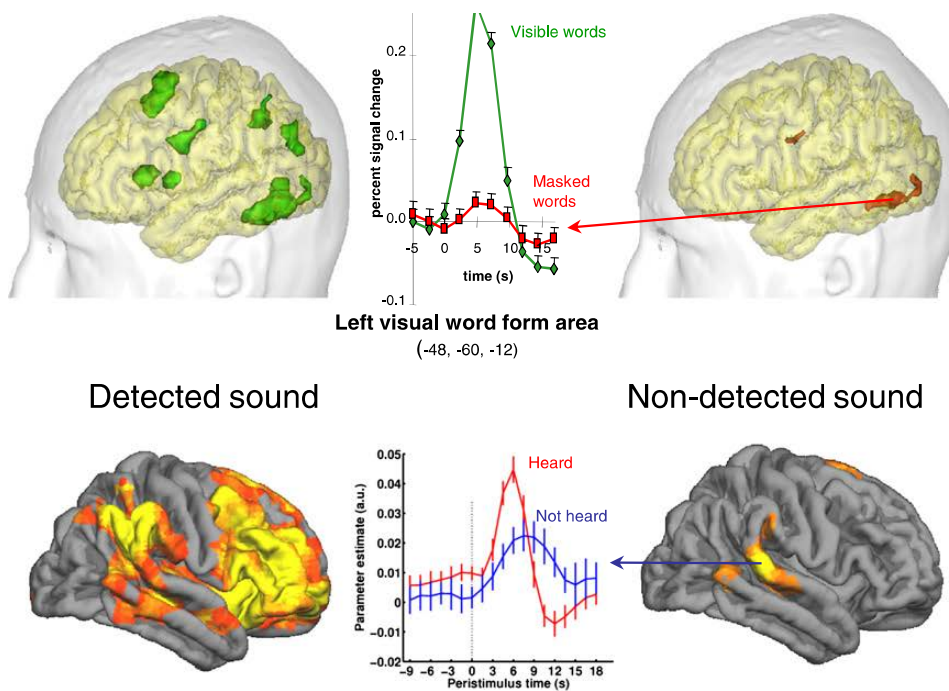


Figure 2. Conscious perception of (A) words or (B) sounds are often associated with widespread activation of the brain, whereas unconscious perception is associated with local activation in specialist processing areas. The data show functional MRI responses time-locked to stimulus presentation.

Source: Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227. [permission needed]