

Brain, conscious experience and the observing self

Bernard J. Baars¹, Thomas Z. Ramsøy², Steven Laureys³

¹The Neurosciences Institute, San Diego, CA 92121, USA

²Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital, DK-2650 Hvidovre, Denmark

³Cyclotron Research Centre, University of Liege, 4000 Liege, Belgium

ABSTRACT

Conscious perception, like the sight of a coffee cup, seems to involve the brain identifying a stimulus. But conscious input activates more brain regions than are needed to identify coffee cups and faces. It spreads beyond sensory cortex to frontoparietal association areas, which do not serve stimulus identification as such. What is the role of those regions? Parietal cortex support the 'first person perspective' on the visual world, unconsciously framing the visual object stream. Some prefrontal areas select and interpret conscious events for executive control. Such functions can be viewed as properties of the subject, rather than the object, of experience - the 'observing self' that appears to be needed to maintain the conscious state.

Humans seem to have a common intuition of an observing self that has access to conscious sensations, inner speech, images and thoughts. Philosophers such as Gilbert Ryle [1] denounced this idea as fallacious, but current evidence seems broadly supportive. This issue has become more pressing in the past decade as scientists have begun to revisit the basic topic of conscious experience. In brief, one can ask, does normal conscious experience involve an observing self?

Visual consciousness as a test case

Visual consciousness has been studied in depth and it is well established that visual features are identified in the ventral stream of posterior cortex. There, feature-sensitive cells support visual experiences of light, color, contrast, motion, retinal size, location and object identity; small lesions can selectively abolish those conscious properties [2]. However, to recall the experience of a human face, we need the hippocampal system¹. To respond to it emotionally, neurons in amygdala can be activated. But hippocampus and amygdala do not seem to support conscious contents directly. Thus, the ventral visual stream, which is needed for specific conscious contents, seems to influence regions that are not (Box 1).

The direction of influence also goes the other way. When we step from a tossing sailboat onto solid ground, the horizon can be seen to wobble. On an airplane flight at night, passengers can see the cabin tilting on approach to landing, although they are receiving no optical cues about the direction of the plane. In those cases, unconscious vestibular signals (originating in

¹ The hippocampal system evidently reflects conscious visual events with considerable fidelity. In a classic study of episodic memory (i.e. memory for conscious events), subjects were asked simply to pay attention to 10 000 distinct pictures over several days, with only five seconds of exposure per picture. On the following day they showed 96% recognition accuracy. No such memory feats are known for subliminal learning, suggesting that visual stimuli must be conscious in order to evoke spontaneous, highly efficient, episodic learning mediated by the hippocampal system [2,21]

the inner ear) shape conscious vision. In sum, conscious visual brain activities can influence unconscious ones, and vice versa.

Studying consciousness 'as such'

How do we know that conscious activity 'as such' evokes widespread regional interactions? After all, similar unconscious processes might do the same. Fortunately, a growing literature now compares the brain effects of conscious and unconscious stimulation. Precise experimental comparisons allow us to ask what conscious access does *per se*.

Many techniques permit comparisons between conscious and unconscious stimulation. In visual backward masking, a target picture is immediately followed by a scrambled image that does not block the optical input physically, but renders it unconscious. Binocular rivalry has been used for the same reason: it shows that when two competing optical streams enter the two eyes, only one consistent stream can be consciously perceived at any given moment. Most recently, several studies have demonstrated inattention blindness, in which paying attention to one visual flow (e.g. a bouncing basketball) blocks conscious access to another activity at the very center of visual gaze (e.g. a man walking by in a gorilla suit). These studies generally show that unconscious stimuli still evoke local feature activity in sensory cortex² [3].

But what is the use of making something conscious if even unconscious stimuli are identified by the brain? Dehaene and colleagues have shown that although unconscious visual words activate known word-processing regions of visual cortex, the same stimuli, when conscious, trigger widespread additional activity in frontoparietal regions [4,5]. This general result has now been replicated many times, using vision, touch, pain perception, and conscious versus automatic skills [3]. Together, these findings suggest that conscious access to a stimulus involves frontward spread of activation beyond the sensory regions of the posterior cerebrum.

Box 1. Theoretical definitions

Conscious events can be defined in practice as those brain activities that subjects can report with high accuracy under optimal conditions, including minimal distraction and time delay. Unconscious events are those that are known to exist without the ability report them accurately, such as subliminal activation of cortical color cells. Indeed, the word 'accurately reportable' could be used instead of "conscious". However, that would miss out something essential - namely, the fact that accurate reports are about experiences that all intact humans claim to have.

Global workspace theory is a cognitive architecture with an explicit role for consciousness [3,8,9]. Global workspace architectures have been studied in cognitive science, and have practical applications in organizing large, parallel collections of specialized processors, broadly comparable to the brain. In recent years, global workspace theory has been found increasingly useful by neuroscientists [3,4].

It makes minimal assumptions:

² *It is inherently difficult to prove the complete absence of consciousness in state studies. Sleep can vary in arousability from moment to moment, much like vegetative states and even general anesthesia. Some mentation is reported even from slow-wave sleep, and some waking-like functions can be preserved in rare brain damage patients who seem behaviorally unconscious. For most purposes, however, an absolute, stable zero point of consciousness is not needed. There is no question that deep sleep is much less conscious than full, responsive waking.*

- That the brain can be viewed as a collection of distributed specialized networks, most of which do not directly support conscious experiences.
- That potentially conscious brain activities can compete for access to a neuronal global workspace capacity- a fleeting memory whose focal contents integrate multiple sources into a single, coherent brain representation, which is then widely distributed to many unconscious specialized networks. The transfer of information from conscious visual episodes to the (unconscious) hippocampal system is a clear example of such distribution of conscious information in the brain.
- Prefrontal cortex has been suggested as one possible neuronal global workspace [4] but posterior sensory cortices also appears to integrate, briefly retain, and distribute coherent neuronal events that are reportable as conscious [9].
- That some unconscious networks, called contexts, are needed to shape conscious contents. Thus, contextual parietal maps of the visual field, which do not support conscious features, modulate visual feature cells that directly contribute to conscious aspects of seen objects.
- That such contexts can jointly constrain conscious events.
- That intentions and emotions can be viewed as 'goal contexts', shaping consciously reportable actions without themselves becoming conscious at the time.
- That hierarchies of goal contexts can act as executive networks to interpret and act upon conscious events without entering consciousness directly. Areas of prefrontal cortex appear to support such functions.

Franklin and colleagues have implemented global workspace theory in large-scale computer models, to make explicit predictions and test functionality in practical tasks (<http://csrcg.cs.memphic.edu/>). Dehaene and colleagues have recently published a neural net model of dorsolateral prefrontal cortex in these terms [4].

Complementary findings come from studies of unconscious states. In deep sleep, auditory stimulation activates only primary auditory cortex [6]. In vegetative states following brain injury, stimuli that are ordinarily loud or painful activate only the primary sensory cortices [7,8]. Waking consciousness is apparently needed for forward spread of sensory activation to occur.

Global workspace theory and the brain

Global workspace theory emphasizes a two-way flow between conscious and unconscious brain activities [3,9-11]. The theory has been implemented in large-scale computational and neural net models [9-11] and bears a close resemblance to Neural Darwinist models [12]. However, it is helpful to think metaphorically of a theater of mind. In the conscious spotlight on stage - the global workspace - an actor speaks, and his words and gestures are distributed to many unconscious audience members, sitting in the darkened hall. Different listeners understand the performance in different ways. But as the audience claps or boos in response, the actor can change his words, or walk off to yield to the next performer. Finally, behind the scenes, an invisible (unconscious) director and playwright try to exercise executive control over the actor and the spotlight.

Less metaphorically, information appears to flow into a neuronal global workspace to be widely distributed. Such a structure must combine converging inputs - the actors competing for access

to the spotlight - followed by momentary dominance of one coherent input, and then wide distribution of output, in a wave of activity sent to other regions. In the brain, sensory projection areas could function much like a global workspace [3,9,10]. Some prefrontal regions play a role in selecting what enters consciousness (selective attention) and interpreting it to control voluntary action. This simple viewpoint helps to organize the evidence and generates testable hypotheses.

Context and the first-person perspective

Global workspace theory calls unconscious influences that shape conscious experiences 'contexts'. Parietal cortex does not recognize coffee cups; it has no feature cells for visual objects. But it does have unconscious egocentric (body-centered) and allocentric (object-centered) cellular maps, which shape our experience of coffee cups, paintings and our own bodies. Damage to right parietal cortex can cause contralateral neglect, a condition in which the left half of the visual field disappears from consciousness. Neglect patients cannot see the left side of a building they are looking at, and will only eat from the right half of a plate in front of their eyes. Thus the parietal region, which supports no reportable conscious activity in itself, can still profoundly shape conscious vision [13].

Neglect patients can also have disturbing alien experiences of their own bodies, especially of the left arm and leg. Such patients sometimes believe that their left leg belongs to someone else, often a relative, and can desperately try to throw it out of bed. Thus, parietal regions seem to shape contextually both the experience of the visual world and of one's own body. Notice that neglect patients still experience their alien limbs as conscious visual objects (a ventral stream function); they are just alien to oneself.

Vogeley and Fink [14] suggest that parietal cortex is involved in the first-person perspective, the viewpoint of the observing self. When subjects are asked to adopt the visual perspective of another person, functional magnetic-resonance imaging (fMRI) activity peaks in medial parietal, inferior lateral parietal and prefrontal cortex.

Prefrontal self systems

Parts of prefrontal cortex are believed to support other self functions. Damage there can change lifelong personality traits, such as the ability to inhibit antisocial impulses. The case of Phineas Gage is classical, and similar neurological patients are not uncommon [15]. Such personality functions seem contextual in that they rarely become conscious, and then only in passing. Yet they underlie the selection and interpretation of conscious thoughts, speech, emotion and social perception. They could constitute the point of view from which 'we' experience the world.

Table 1. Major properties of four types of unconscious state compared with conscious rest

State	Conscious resting state [19,20]	Deep sleep [22]	General anesthesia [24]	Vegetative state or coma [25]	Epileptic loss of consciousness [23]
Cause	Neuromodulation of the cortex by the brainstem, instructions to avoid deliberate tasks [26]	Physiological: neuromodulation of the forebrain by the brainstem	Pharmacological: a variety of chemical agents	Pathological: trauma, intoxication, anoxia, hypoglycemia	Pathological: slow, synchronized neuronal firing driven by brain

					foci [23]
Behavioral signs	Accurate reportability of attended stimuli; orientation to space, time, and self; visual images, inner speech, abstract thoughts; control of voluntary muscles	No reportability	No reportability	No reportability. Lower brainstem reflexes retained intact [25]	No reportability
Regional metabolism	High in frontoparietal cortex	Low in frontoparietal cortex [27]	Low in frontoparietal cortex [28]	Low in frontoparietal cortex	Low in frontoparietal cortex
EEG voltages	Low-amplitude, irregular high-frequency waves (8-100 Hz), or low-amplitude, regular alpha waves (8-12 Hz)	High-amplitude, regular, low-frequency waves (< 4 Hz)	High-amplitude, regular, low-frequency waves	High-amplitude, regular, low-frequency waves	High-amplitude, spike-wave form, regular, low-frequency waves
Underlying neuronal mechanism (in cortical and thalamic neurons)	Firing irregularly at an average base rate of ~ 10 Hz	Slow, synchronized pausing of base-rate firing [22]	Slow, synchronous pausing of base-rate firing?	Slow, synchronous pausing of base-rate firing?	Slow, synchronous pausing of base-rate firing?
Functional connectivity	High and variable	Low between cortical regions, and between thalamus and cortex [22]	Low between cortical regions, and between thalamus and cortex	Low or absent between cortical regions, and between thalamus and cortex [29,30]	Low between cortical regions, and between thalamus and cortex [23]

Abbreviations: EEG, electroencephalogram.

It is telling that patients with identity disorders such as fugue (a rapid change in personal identity lasting weeks or months) often show amnesia for the eclipsed self. When the patient returns to normal, he or she might report time loss - a period of weeks from which no conscious experiences can be recalled [16]. It is as if each personality serves to organize and interpret conscious events during its time of dominance. While parietal cortex appears to put visual scenes into context, prefrontal regions appear to do the same for more abstract aspects of experience, such as social, emotional and self evaluation.

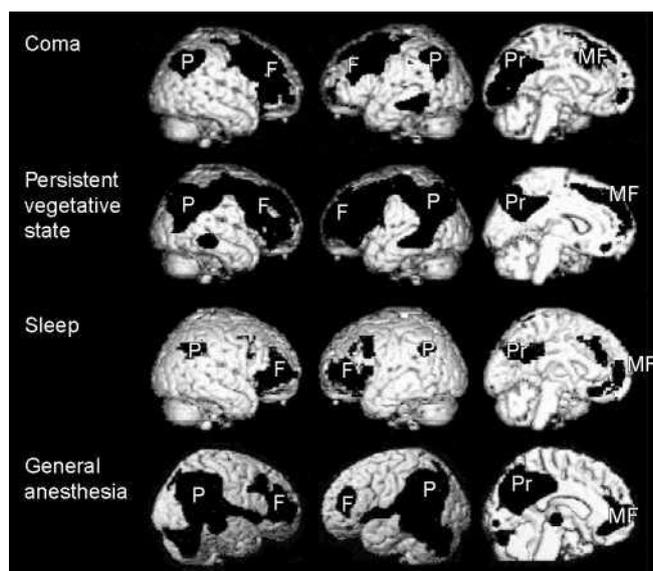
Indeed, conscious experience in general can be viewed as information presented to prefrontal executive regions for interpretation, decision-making and voluntary control. Crick and Koch note that 'it is useful to think of the front or higher/executive part of the cortex as looking at and interacting with the back, or sensory part.' [17] In split-brain patients, a great deal of executive control seems to require the 'narrative interpreter' of the speaking hemisphere. One reason is the pervasive role of inner and outer speech in daily life, to remind ourselves of things to do, to focus on current concerns, and to maintain verbally supported plans and beliefs.

These points suggest a new way to understand conscious and unconscious states.

Is the 'observing self' needed to sustain the conscious waking state?

Conscious waking shows fast oscillatory activity throughout the thalamocortical core, and functional connectivity that can change rapidly as a function of task, content and context [18] (Table 1). Despite this great neuronal variability, recent evidence suggests the existence of a conscious 'baseline state', a resting wakeful state in which no external tasks are required [19,20]. Surprisingly, when subjects are asked simply to rest, their frontoparietal metabolism is consistently higher than in standard cognitive tasks. This might be due to the flow of thoughts that subjects report in the resting state [20]. Spontaneous thoughts seem more self-relevant than standard cognitive tasks, which typically compel subjects to direct their attention away from their personal concerns.

Figure 1. Metabolic activity in four types of unconsciousness, subtracted from conscious controls. Positron emission tomography (PET) scans showing regional decreases in metabolism or blood flow when unconscious states are compared with resting consciousness. Coma, persistent vegetative state, sleep and general anesthesia all show regional decreases in frontoparietal association cortices. Column 1 shows the right lateral aspect of the brain, column 2 the left lateral aspect, and column 3 a medial view of the left hemisphere.



Abbreviations: F, prefrontal; MF, mesiofrontal; P, posterior parietal cortex; Pr, posterior cingulate/precuneus.

Metabolic activity in the conscious resting state is not uniformly distributed. Raichle *et al.* write that medial parietal regions, including 'posterior cingulate cortex and adjacent precuneus can be posited as a tonically active region of the brain that may continuously gather information about the world around, and possibly within, us. It would appear to be a default activity of the brain...' [19]. Mazoyer *et al.* also found high prefrontal metabolism during rest [20]. Notice that these are the same general areas that show additional activity when conscious sensory stimulation is compared with matched unconscious input. We will see that these regions also show markedly lower metabolism in unconscious states.

Four unconscious states: eclipsing the self?²

Table 1 shows the conscious resting state compared with four unconscious states that are

causally very different from each other: deep sleep³, coma/vegetative states, epileptic loss of consciousness⁴, and general anesthesia under six different anesthetic agents⁵.

Surprisingly, despite their very different mechanisms the four states share major common features. These include: (i) widely synchronized slow waveforms that take the place of the fast and flexible interactions needed for conscious functions⁵; (ii) frontoparietal regions becoming hypometabolic; (iii) widely blocked functional connectivity, both corticocortical and thalamocortical; and (iv) behavioral unconsciousness, including unresponsiveness to normally conscious stimuli. The first three of these features lower the probability of waking-type interactions among brain regions.

Figure 1 shows marked hypometabolism in the four unconscious states compared with conscious controls, precisely where we might expect: in frontoparietal regions. Could it be that brain regions that underlie the 'observing self' are thereby disabled?

Summary and future directions

Frontoparietal association areas have many functions beyond those touched on here. However, several lines of evidence suggest that they could have a special relationship with consciousness, even though they do not support the contents of sensory experience. (i) Conscious stimulation in the waking state leads to frontoparietal activation, but unconscious input does not; (ii) in unconscious states, sensory stimulation activates only sensory cortex, but not frontoparietal regions; (iii) the conscious resting state shows high frontoparietal metabolism compared with outward-directed cognitive tasks; and (iv) four causally very different unconscious states show marked metabolic decrements in the same areas.

Although alternative hypotheses must be considered, it seems reasonable to suggest that 'self' systems supported by these regions could be disabled in unconscious states. From the viewpoint of the observing self, this would be experienced as subjective loss of access to the conscious world. Unconscious states might not necessarily block the objects of consciousness; rather, the observing subject might not be at home.

Acknowledgements

B.J.B gratefully acknowledges support from the Neurosciences Institute and the Neurosciences Research Foundation (10640 John Jay Hopkins Drive, San Diego, CA 94549, USA; www.nsi.edu). S.L. is a research associate supported by the Belgian National Fund for Scientific Research (FNRS). We thank Bjorn Merker, Anil Seth, Douglas Nitz and E. Roy John for helpful discussions.

³ At the level of cortical neurons, bursting rates do not change in deep sleep. Rather, neurons pause together at < 4 Hz between bursts [22]. Synchronous pausing could disrupt the cumulative high-frequency interactions needed for waking functions such as perceptual continuity, immediate memory, sentence planning, motor control and self-monitoring. It is conceivable that other unconscious states display similar neuronal mechanisms.

⁴ Although the spike-wave electroencephalogram (EEG) of epileptic seizures looks different from the delta waves of deep sleep and general anesthesia, it is also slow, synchronized and high in amplitude. The source and distribution of spike-wave activity varies in different seizure types. However, the more widespread the spikewave pattern, the more consciousness is likely to be impaired [23]. This is again marked in frontoparietal regions.

⁵ Controversy exists over fast EEG oscillations under the influence of ketamine, and whether its effects should be viewed as primarily dissociative or anesthetic. Ketamine might be an exception to the slow-wave EEG found with other anesthetics.

References

- 1 Ryle, G. (1949) *The Concept of Mind*, Hutchinson, London
- 2 Zeki, S. (2001) Localization and globalization in conscious vision. *Annu. Rev. Neurosci.* 24, 57-86
- 3 Baars, B.J. (2002) *The conscious access hypothesis: origins and recent evidence*. *Trends Cogn. Sci.* 6, 47-52
- 4 Dehaene, S. et al. (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8520-8525
- 5 Dehaene, S. et al. (2001) Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* 4, 752 – 758
- 6 Portas, C.M. et al. (2000) Auditory processing across the sleep-wake cycle: simultaneous EEG and fMRI monitoring in humans. *Neuron* 28, 991-999
- 7 Laureys, S. et al. (2000) Auditory processing in the vegetative state. *Brain* 123, 1589-1601
- 8 Laureys, S. et al. (2002) Cortical processing of noxious somato-sensory stimuli in the persistent vegetative state. *Neuroimage* 17, 732-741
- 9 Baars, B.J. (1988) *A Cognitive Theory of Consciousness*, Cambridge University Press
- 10 Franklin, S. (2000) Deliberation and voluntary action in 'conscious' software agents. *Neural Netw. World* 10, 505-521
- 11 Baars, B.J. and Franklin, S. (2003) How conscious experience and working memory interact. *Trends Cogn. Sci.* 7, 166-172
- 12 Edelman, G.M. (2003) Naturalizing consciousness: a theoretical framework. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5520 – 5534
- 13 Bisiach, E. and Geminiani, G. (1991) Anosognosia related to hemiplegia and hemianopia. In *Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues* (Prigatano, G.P. and Schacter, D.L., eds), Oxford University Press
- 14 Vogeley, K. and Fink, G.R. (2003) Neural correlates of the first-person perspective. *Trends Cogn. Sci.* 7, 38-42
- 15 Anderson, S.W. et al. (1999) Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat. Neurosci.* 2, 1032-1037
- 16 American Psychiatric Association, (2000) *Diagnostic and Statistical Manual of Mental Disorders IV*, American Psychiatric Association
- 17 Crick, F. and Koch, C. (2003) A framework for consciousness. *Nat. Neurosci.* 6, 119-126
- 18 Buechel, C. and Friston, K. (2000) Assessing interactions among neuronal systems using functional neuroimaging. *Neural Netw.* 13, 871-882
- 19 Raichle, M.E. et al. (2001) A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676-682
- 20 Mazoyer, B. et al. (2001) Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* 54, 287 – 298
- 21 Standing, L. (1973) Learning 10,000 pictures. *Q. J. Exp. Psychol.* 25, 207-222
- 22 Steriade, M. (2001) Active neocortical processes during quiescent sleep. *Arch. Ital. Biol.* 139, 37-51

- 23 Blumenfeldt, H. and Taylor, J. (2003) *Why do seizures cause loss of consciousness?* *Neuroscientist* 9, 1-10
- 24 John, E.R. et al. (2001) *Invariant reversible QEEG effects of anesthetics.* *Conscious. Cogn.* 10, 165-183
- 25 Laureys, S. et al. (1999) *Cerebral metabolism during vegetative state and after recovery to consciousness.* *J. Neurol. Neurosurg. Psychiatry* 67, 121
- 26 Hobson, J.A. and Pace-Schott, E.F. (2002) *The cognitive neuroscience of sleep: neuronal systems, consciousness and learning.* *Nat. Rev. Neurosci.* 3, 679-693
- 27 Maquet, P. et al. (2000) *Functional neuroimaging of normal human sleep by positron emission tomography.* *J. Sleep Res.* 9, 207-231
- 28 Kaisti, K.K. et al. (2002) *Effects of surgical levels of propofol and sevoflurane anesthesia on cerebral blood flow in healthy subjects studied with positron emission tomography.* *Anesthesiology* 96, 1358-1370
- 29 Laureys et al. (2000) *Restoration of thalamocortical connectivity after recovery from persistent vegetative state.* *Lancet* 355, 1916
- 30 Laureys, S. et al. (1999) *Impaired functional connectivity in vegetative state: preliminary investigation using PET.* *Neuroimage* 9, 377 - 382