

Current concepts of consciousness with some implications for anesthesia

Bernard J. Baars PhD

THERE is now extensive evidence that consciousness is associated with brainwide distribution and integration of information. Psychological evidence for the hypothesis was summarized by Baars (1988, 1997, 2002) in a framework called Global Workspace (GW) theory, derived from the cognitive architecture tradition. The global hypothesis has implications for perception, emotion, motivation, learning, working memory, voluntary control, and self systems in the brain. Recent brain imaging evidence appears to be broadly supportive, and the hypothesis is now favoured by philosophers like Daniel C. Dennett and Ned Block. The GW hypothesis is directly testable by brain scanning experiments that treat consciousness as a variable. Implications for unconscious states, including general anesthesia, are discussed.

The rediscovery of consciousness

Shortly after 1,900 behaviourists attempted to purge science of mentalistic concepts like consciousness, attention, memory, imagery and voluntary control. "Consciousness," wrote John B. Watson, "is nothing but the soul of theology." But as the facts accumulated over the 20th century, all the traditional ideas were found to be necessary. They were reintroduced with more testable definitions. Memory came back in the 60's; imagery in the 70's; selective attention in the last half century; and consciousness in the last decade or so.

It is roughly true that what we are conscious of is what we can report with accuracy. Conscious brain events are therefore assessed mainly by way of reportability. We now know of numerous brain events that are reportable, and comparable ones that are not. This fact invites experimental testing: why are we conscious of these words at this moment, while a few seconds later they have faded, but can still be called to mind? Why is activity in visual temporal lobe neurons

reportable, while apparently similar activity in parietal regions is not? Why does the thalamocortical system support conscious experiences, while the comparably large cerebellum and basal ganglia do not? These are all testable questions.

The empirical key is to treat consciousness as a controlled variable. We now know of cortical neurons that respond to conscious pictures presented to one eye, while competing unconscious pictures presented to the other eye evoke firing in closely related cells. Goodale and Milner have shown that visual control of manual reaching and grasping appears to be unconscious. Rees and colleagues have demonstrated that of two streams of visual information at the very centre of gaze, one can be conscious and the other not; the conscious one activates distinct brain areas, as assessed by fMRI. The competing unconscious visual stream, which surely reaches cortex, could not be detected. A rich literature has arisen comparing conscious and unconscious brain events in sleep and waking, anesthesia, epileptic states of absence, very specific damage to visual cortex, spared implicit function after brain damage, coma, attentional control, visual imagery, inner speech, memory recall, and more.

Relevance to general anesthesia

The same general pattern of evidence holds true for unconscious states, compared to waking consciousness. In a multicentre study involving 176 surgical anesthesia cases using six different anesthetic agents, John and coworkers showed marked quantitative electroencephalogram changes between conscious, anesthetic, and postanesthetic (conscious) states (John *et al.* 2000). At loss of consciousness, gamma power decreased while lower frequency bands increased in power, especially in frontal leads. Loss of consciousness was accompanied by a very significant drop in coherence between homologous areas of the two hemispheres, and between poste-

From The Neurosciences Institute, San Diego, California, USA.

Address correspondence to: Dr. Bernard J. Baars, The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, California 92121, USA. Phone: 858-626-2000; Fax: 858-626-2099; E-mail: baars@nsi.edu
Canadian Anesthesiologists' Society, 2003. Copyright, B.J. Baars, 2003.

rior and anterior regions of each hemisphere. However, there was hypersynchronous activity within anterior regions, including, the authors suggest, “the orbital and rectus gyri of the prefrontal cortex, the superior frontal gyrus, pre- and paracentral cortical regions, the anterior cingulate gyrus, parahippocampal gyrus and the amygdala, basal ganglia and thalamus” (p. 180).^A These regions have been implicated in previous positron emission tomography studies of anesthesia as well. The same basic changes occurred across all six anesthetics, and reversed when patients regained consciousness.

From the viewpoint of globalist theories, the most readily interpretable finding is the coherence drop in gamma bands after anesthetic loss of consciousness. It suggests a loss of coordination between frontal and posterior cortex, and between homologous regions of the two hemispheres. The authors also suggest that the anteriorization of low frequencies “must exert a profound inhibitory influence on cooperative processes within (frontal) neuronal populations. This functional system then becomes dedifferentiated and disorganized” (p. 180).^A Finally, the decoupling of the posterior, mostly sensory cortex with anterior regions suggests “a blockade of perception” (p. 180).^A

These phenomena appear to be consistent with the GW hypotheses.

The global access hypothesis

The idea that consciousness has an integrative function has a long history. The GW theory suggests a *fleeting memory capacity that enables access between brain functions that are otherwise separate*. This makes sense in a brain that is viewed as a massive parallel distributed system of highly specialized processors. In such a system coordination and control may take place by way of a central information exchange, allowing some specialized processors – such as sensory systems in the brain – to distribute information to the system as a whole. This solution works in large-scale computer architectures, which show typical “limited capacity” behaviour when information flows by way of a GW. A sizable body of evidence suggests that consciousness is the primary agent of such a global access function in humans and other mammals (Baars, 1983, 1988, 1997, 1998).

^A Global workspace theory emerged from the cognitive architecture tradition pioneered by Alan Newell and Herbert A. Simon. Newell and his coworkers were the first to show the utility of a global workspace capacity in a complex system of specialized knowledge sources, which could cooperatively solve problems no single constituent could solve alone. The empirical connection with consciousness was not made in this tradition, however.

The “conscious access hypothesis” therefore implies that consciousness provides a gateway to many capacities of the brain. Table I shows the basic claims made by the GW theory.

A number of scientists and philosophers now support some version of global access. Table II shows a dozen of such comments from different sources, suggesting that the same basic idea emerged independently from different research traditions. The convergence of ideas is striking.

While the GW theory is not the only one to postulate widespread access associated with consciousness, it develops the idea in great detail, exploring its implications for learning, working memory, voluntary control, and self.

A theatre metaphor and brain hypotheses

Like other cognitive architectures, the GW theory may be thought of as a theatre of mental functioning. The theatre metaphor is too simple, but it offers a useful first approximation. Consciousness in the metaphor resembles a bright spot on the stage of immediate memory, directed there by a spotlight of attention, under executive guidance. The rest of the theatre is dark and unconscious.

This approach leads to specific neural hypotheses. For sensory consciousness the bright spot on stage is likely to require the corresponding sensory projection areas of the cortex. Sensory consciousness in different modalities may be mutually inhibitory, within approximately 100-msec time cycles. Sensory cortex can be activated internally as well as externally, resulting in the “internal senses” of conscious inner speech and imagery. Once a conscious sensory content is established, it is distributed widely to a decentralized “audience” of expert networks sitting in the darkened theatre, presumably using corticocortical and corticothalamic fibres. This is the primary functional role of

TABLE I Theoretical claims: brain capacities enabled by conscious events

| | |
|---|--|
| 1 | Conscious perception enables access to widespread brain sources; unconscious sensory processing is much more limited. |
| 2 | Conscious perception, inner speech, and visual imagery enable working memory functions; there is no evidence for unconscious access to working memory. |
| 3 | Conscious events enable <i>almost</i> all kinds of learning: episodic and explicit learning, but also implicit and skill learning. |
| 4 | Conscious perceptual feedback enables voluntary control over motor functions, and perhaps over any neuronal population and even single neurons. |
| 5 | Conscious contents can evoke selective attention. |
| 6 | Consciousness enables access to “self” – executive interpreters, located in part in the frontal cortex. |

TABLE II Some “global workspace” themes since 1983

| |
|--|
| Baars, 1983. “Conscious contents provide the nervous system with <i>coherent, global information</i> .” |
| Edelman, 1989. Global mapping in a reentrant selectionist model of consciousness in the brain (Edelman, 1989). |
| Dennett, 2001. “Theorists are converging from quite different quarters on a version of the global neuronal workspace model of consciousness ... on the eve of the decade of the brain, Baars (1988) had already described a ‘gathering consensus’ in much the same terms: consciousness, he said, is accomplished by a “distributed society of specialists that is equipped with a working memory, called a <i>global workspace</i> , whose contents can be broadcast to the system as a whole” (p. 42). |
| Kanwisher, 2001. “... in agreement with Baars (1988), it seems reasonable to hypothesize that awareness of a <i>particular element of perceptual information must entail not just a strong enough neural representation of information, but also access to that information by most of the rest of the mind/brain</i> .” |
| DeHaene & Naccache, 2001. “We propose a theoretical framework... the hypothesis of a global neuronal workspace... we postulate that this global availability of information through the workspace is what we subjectively experience as the conscious state.” |
| Edelman & Tononi, 2000: 148–9: “When we become aware of something... it is as if, suddenly, many different parts of our brain were privy to information that was previously confined to some specialized subsystem... the wide distribution of information is guaranteed mechanistically by <i>thalamocortical and corticocortical reentry</i> , which facilitates the interactions among distant regions of the brain.” |
| Llinas <i>et al.</i> , 1998. “... the thalamus represents a hub from which any site in the cortex can communicate with any other such site or sites... temporal coincidence of specific and non-specific thalamic activity generates the functional states that characterize human cognition. |
| John, 2000. “Evidence has been steadily accumulating that information about a stimulus complex is distributed to many neuronal populations dispersed throughout the brain.” |
| Varela <i>et al.</i> , 2001. “... the brain... transiently settling into a globally consistent state... (is) the basis for the unity of mind familiar from everyday experience.” |
| Damasio, 1989. “Meaning is reached by time-locked multiregional retroactivation of widespread fragment records. Only the latter records can become contents of consciousness.” |

consciousness: to allow a theatre architecture to operate in the brain, in order to integrate, provide access, and coordinate the functioning of very large numbers of specialized networks that otherwise operate autonomously (Mountcastle, 1978). All the elements of

the GW theory have reasonable brain interpretations, allowing us to generate a set of specific, testable brain hypotheses about consciousness and its many roles in the brain. Some of these ideas have now received considerable empirical support (Baars, 2002).

Inner speech, imagery, and working memory

Both auditory and visual consciousness can be activated endogenously. Inner speech is a particularly important source of conscious auditory-phonemic events, and visual imagery is especially useful for spatial orientation and problem solving. The areas of the left hemisphere involved in outer speech are now known to be involved in inner speech as well (Paulesu, Frith & Frackowiak, 1993). Likewise, mental imagery is known to involve visual cortex.

Likewise, internally generated somatosensory imagery may reflect emotional and motivational processes, including internally generated feelings of pain, pleasure, hope, fear, sadness, etc. Such internal sensations may communicate to other parts of the brain via global broadcasting. Prefrontal executive systems may not have direct access to action control. Rather, they may work by evoking motivational imagery, broadcast from the visual cortex, to control relevant parts of motor cortex, thereby generating appropriate actions. Parts of the brain that play a role in emotion may also be triggered by global broadcasting of conscious contents from sensory cortices and insular cortex. For example, it is established that the midbrain amygdala is needed to recognize facial expressions of fear and anger from the visual system (Le Doux, 1996). Ultimately such areas, working together, shape actions controlled by frontal cortex and subcortical automatic systems like the basal ganglia.

Thus many cortical areas work together to transform goals and emotions into actions (Baars, 1988). Since there are many spatial maps throughout the brain, the trade language of the brain may consist of activated maps, paced by temporal oscillations in the alpha to gamma spectrum. Such oscillations may coordinate the activity of multiple sensory, body space, and external spatial maps.

The attentional spotlight

The sensory “bright spot” of consciousness involves a selective attention system (the theatre spotlight), under dual control of frontal executive cortex and automatic interrupt control from areas such as the brain stem, pain systems, and emotional centres like the amygdala. It is these attentional interrupt systems that allow significant stimuli to “break through” into consciousness in a selective listening task, when the name is spoken in the unconscious channel.

Context vs content of sensory experience

A conscious sensory “bright spot” requires the interaction of sensory analyzers and contextual systems. In vision sensory content seems to be produced by the ventral visual pathway, while contextual systems in the dorsal pathway define a spatial domain within which the sensory event is defined. As pointed out above, parietal cortex is known to include allocentric and egocentric spatial maps, which are not themselves objects of consciousness, but which are required to shape every conscious visual event.

There is a major difference between the symptoms of dysfunctional *content* systems such as the visual ventral stream, compared to disordered context systems. In the case of lesioned content, the subject can generally notice a missing part of normal experience; but for damage to context, the system of experiential expectations is itself damaged, so that one no longer knows what to expect, and hence what is missing. This may be why parietal neglect is so often accompanied by anosognosia, a massive but very specific loss of knowledge about one’s body space. Patients suffering from parietal anosognosia may reject their own body limbs, or see themselves with three hands. Such specific loss of contextual body information is not accompanied by a loss of general intelligence or knowledge about the world. Damage to such contextual knowledge systems may be quite specific.

Self-systems

Cortical activation by a seen object may not be enough to generate subjective consciousness. The activated visual information may need to be sent to self-systems, which serve to maintain constancy of an inner framework across different perceptual situations. When we walk from room to room in a building, we must maintain a complex and multi-leveled organization that can be viewed in the GW theory as a higher-level context. Major goals, for example, are not changed when we walk from room to room, but conscious perceptual experiences do change. Gazzaniga (1996) has found a number of conditions under which split-brain patients encounter conflict between right and left hemisphere executive and perceptual functions. He has proposed the existence of a “narrative self” in the left frontal cortex, based on split-brain patients who are clearly using speech in the left hemisphere to talk to themselves, sometimes trying to force the right hemisphere to obey its commands. When that proves impossible, the left hemisphere will often rationalize or reinterpret the sequence of events so as to repair its understanding of the interhemispheric conflict.

Analogous repairs of reality are observed in other forms of brain damage, such as parietal neglect. They

commonly occur whenever humans are confronted with major, unexpected life changes. The left-hemisphere narrative interpreter may be considered as a higher-level context system that maintains expectations and intentions across many specific situations. While the narrative stream itself is conscious, it is shaped by unconscious contextual executive influences.

If we consider Gazzaniga’s narrative interpreter of the left hemisphere to be one kind of self-system in the brain, it must receive its own flow of sensory input. Visual input from one half of the field may be integrated in one visual hemisphere, as described above, under retinotopic control from area V1. But once it comes together in visual cortex, it needs to be conveyed to frontal areas on the left side of the brain, in order to inform the narrative interpreter of the current state of perceptual affairs. The left prefrontal self system then applies a host of criteria to the input, such as “did I intend this result? Is it consistent with my current and long-term goals? If not, can I reinterpret the input to make sense in my running account of reality?”

It is possible that the right hemisphere has a parallel system that does not speak, internally or externally, but that may be more able to deal with anomalies via irony, jokes, and other emotionally sophisticated strategies. The evidence appears to be good that the right prefrontal cortex can understand such figurative uses of language, while the left does not.

Full consciousness may not exist without the participation of such prefrontal self systems.

New brain evidence for global distribution of conscious contents

On the basis of psychological evidence, the GW theory predicted that conscious contents are widely distributed in the brain (Baars, 1988, 1997). Today that case is supported by a sizable body of brain evidence, at least for sensory consciousness (Baars, 2002). For example, Dehaene and colleagues showed that backward-masked visual words evoked brain activity confined to the well-known visual word recognition areas of cortex (Dehaene *et al.*, 2001). Identical conscious words triggered higher levels of activity in these areas, but more importantly, they evoked far more widely distributed activity in parietal and prefrontal cortex. That result has now been replicated a dozen times, using different brain imaging techniques, different experimental comparisons between conscious and unconscious input (e.g., binocular rivalry, inattention blindness, neglect, extinction), and different modalities (audition, pain, vision, and sensorimotor tasks). In all cases conscious sensory input evoked far wider and more intense brain activity than identical unconscious input.

These findings support the general claim that conscious stimuli mobilize large areas of cortex, presumably to distribute information about the stimuli. This idea is essential to the following discussion. If consciousness serves to mobilize many unconscious specialized networks, the active elements of mental life that always need to be conscious – input, recall, rehearsal, inner speech, visual imagery and report – may be widely distributed in order to recruit specific unconscious functions needed to carry out those tasks.

Evolutionary considerations

A number of neurobiologists suggest that sensory consciousness emerged with early mammals or transitional reptiles (therapsids), who evolved a large thalamocortical complex (e.g., Edelman, Tononi, 2000). Working memory – what may be called “extended consciousness” beyond the immediate contents – involves two endogenous kinds of sensory consciousness, the phonological loop (for inner speech), and the visual sketchpad (for voluntary visual imagery). The phonological loop seems to use the classic speech regions of the left hemisphere, Broca’s and Wernicke’s areas. Wernicke’s area is located near the auditory cortex, and in the case of inner speech it seems to support internalized auditory speech perception. Likewise, visual imagery appears to involve visual cortex, especially when it is vividly conscious.

Working memory involves the purposeful use of such endogenous sensory consciousness. It is not clear when the purposeful use of visual imagery and inner speech evolved, but it plausibly did so *after* the major growth of prefrontal cortex and speech areas associated with hominid evolution – that is, in the last few million years. Prefrontal cortex is key to voluntary goals in humans, and speech is of course the basis of the phonological loop involved in mental rehearsal.

In contrast, spontaneous imagery may have appeared much earlier – such as the visual image of a lion evoked by the sound of a lion’s roar. Such spontaneous imagery may be common among mammalian prey animals, while predators may have spontaneous visual images of prey that can be smelled but not seen. Certainly the dream state, characterized by rich conscious visual imagery, evolved with early mammals.

Thus both external and endogenous sensory capacities may have evolved millions of years *before* the hominid capacities needed for human mental life. These considerations suggest that endogenous sensory consciousness long precedes specifically human mental capacities, and that the distinctive hominid development of human thought depends on voluntary control of these pre-existing functions. This is espe-

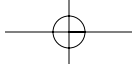
cially true of language and its role in voluntary control of inner speech and purposeful visual imagery.

Summary and conclusions

The GW theory suggests that consciousness enables multiple networks to cooperate and compete in solving problems, such as retrieval of specific items from immediate memory. The overall function of consciousness is to provide widespread access, which in turn may serve functions of coordination and control. Consciousness is the gateway to the brain, enabling control even of single neurons and whole neuronal populations (Baars, 1988). None of these control functions become directly conscious, of course, but conscious feedback seems required to recruit control by prefrontal networks. In the metaphor of the theatre, it is as if each specialized audience member can decide locally whether or not to be driven by input from the bright spot on stage. Executive functions – the director behind the scenes – is also largely unconscious, often using the actor in the spotlight on the stage of working memory capacity to recruit and trigger specific functions.

References (presented in alphabetical order)

- Baars BJ.* Conscious contents provide the nervous system with coherent, global information. *In:* Davidson RJ, Schwartz GE, Shapiro D (Eds.). *Consciousness and Self-Regulation*. NY: Plenum Press; 1983: 41.
- Baars B.* A cognitive theory of consciousness. New York: Cambridge University Press; 1988.
- Baars B.* In the theater of consciousness: The workspace of the mind. New York: Oxford University Press; 1997.
- Baars BJ.* The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Science* 2002; 6: 47–52.
- Damasio AR.* Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition* 1989; 33: 25–62.
- DeHaene S, Naccache L.* Toward a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 2001; 79: 1–37.
- Dennett DC.* Are we explaining consciousness yet? *Cognition* 2001; 79: 221–37.
- Edelman GM.* Neural Darwinism. New York: Basic Books; 1989.
- Edelman GM, Tononi G.* A universe of consciousness. How matter become imagination. New York: Basic Books; 2000.
- John ER.* *Consciousness & Cognition*. 2000; 10: 2.
- Kanwisher N.* Neural events and perceptual awareness, *Cognition* 2001; 79: 89–113.
- Llinas R, Ribary U, Contreras D, Pedroarena C.* The neu-



R6

CANADIAN JOURNAL OF ANESTHESIA

ronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci* 1998; 353: 1841-9.

Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. *Nature Reviews - Neuroscience* 2001; 2: 237.

