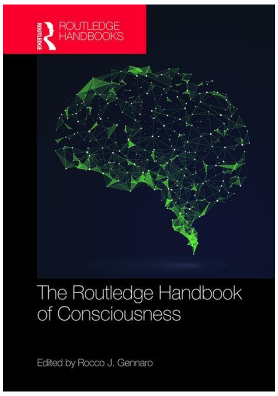

The Routledge Handbook Of Consciousness

Rocco J. Gennaro

The Global Workspace Theory

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9781315676982-10
Bernard J. Baars, Adam Alonzi
Published online on: 26 Mar 2018

How to cite :- Bernard J. Baars, Adam Alonzi. 26 Mar 2018, *The Global Workspace Theory from:* The Routledge Handbook Of Consciousness Routledge
Accessed on: 07 Dec 2023
https://www.routledgehandbooks.com/doi/10.4324/9781315676982-10

9

THE GLOBAL WORKSPACE THEORY

Bernard J. Baars and Adam Alonzi

1 Introduction

A global workspace (GW) is a functional hub of signal integration and propagation in a large population of loosely coupled agents, on the model of "crowd" or "swarm" computing, using a shared "blackboard" for posting, voting on, and sharing hypotheses, so that multiple experts can make up for each others' limitations. Crowd computation has become a major technique for web commerce as well as scientific problem-solving.

In the 1970s Allen Newell's Carnegie-Mellon team showed that a GW architecture was able to solve a difficult practical problem, the task of identifying 1,000 normally spoken words in a normally noisy and distorting acoustical environment, including the many challenges of phonemic and syllabic encoding of slow analogue sound resonances interrupted by fast transients, produced by the inertial movements of many different vocal tracts, each with its distinctive acoustical resonance profile beginning with vocal, soft tissue, nasal, and labiodental turbulence, each with overlapping "coarticulation" of phonemic gestures, with its own idiosyncratic speech styles and dialects, all in an acoustical environment with its own mix of sound-absorbing, masking and echoing surfaces. In real speech this difficult signal identification task is also organized in lexical and morphemic units, with real-world referents, with unpredictable and ambiguous grouping, syntactic, semantic, pragmatic, intonational and emotional organization. Newell's HEARSAY system was able to identify more than 90% of the spontaneous words correctly, even without modern formant tracking, a newer and more effective technique.

HEARSAY was one of the first success stories for the new concept of parallel-distributed architectures, now often called "crowdsourcing" or "swarm computing." The most important point here is the surprising effectiveness of expert crowds using GW-mediated signaling, when none of the individual experts could solve the posted problem by themselves.

One of today's leading speech recognition systems, Apple's SIRI, is still making use of web-based crowdsourcing to identify poorly-defined syllables in numerous languages and dialects, spoken by many different voices in acoustically noisy spaces. SIRI also learns to predict the speaker's vocal tract to improve its detection score. It is still imperfect, but it is commercially viable.

Based on Newell's work, Baars (1988) demonstrated the surprisingly close empirical match between the well-known "central limited capacity" components of the brain associated with consciousness and the notion of a global workspace architecture. The resulting GW theory of conscious cognition has been widely adopted and developed, showing some 15,000 citations since Baars (1988).

A new wave of neuroscience evidence shows that the extended cortex – neo, paleo, and thalamus – can support a dynamic, mobile, context-sensitive and adaptive GW function. Many regions of the cortex support conscious experiences, which can be reported with high accuracy, and which generally compete with each other for limited momentary capacity. However, regions like the cerebellum and the dorsal stream of cortex do not enable conscious contents directly.

Modern computation came of age using very clean electrical signals, digitally distinctive and easy to translate into programming code. The first programs used logical, arithmetic and other symbolic operations that came from centuries of mathematics. "Shannon information" was well-defined and relatively easy to implement in practice, and the mathematical Turing Machine supported formal proofs that almost any determinate function could be implemented by TMs.

The challenge for HEARSAY was quite different from the standard problems of classical computation, and much more biological in spirit, because the real sensory world is not digital but analogue, with poorly-defined stimuli, actions and salience boundaries, many of which must be defined "top-down" based on prior knowledge.

The natural world is not engineered to avoid catastrophic events like head injuries and microparasites; modern humans live in highly protected environments, with none of the pitfalls and dangers we encounter when running over an unimproved natural landscape with poor visual conditions. In contrast, ancient cemeteries often show very high rates of broken human bones and other physical damage, often inflicted by other humans.

Modern buildings use parallel and orthogonal ceilings, floors and walls, making perceptual size estimation and action prediction much easier. Their acoustical properties are typically clean and predictable. Conscious distractions are radically reduced. In the last century the spread of sanitary engineering alone has reduced infectious diseases and doubled the human lifespan.

The world in which our ancestors evolved was fundamentally different. Computational architectures built to deal with unpredictable, high risk events are therefore more biologically realistic. Humans may be among the most adaptable species in the animal kingdom, as shown by the fact that Homo Sapiens has colonized an immense diversity of econiches around the globe in the last 30–40,000 years, beginning with a genetically small and homogeneous "founder population" in north-east Africa some 46,000 years ago.

As they spread out of Africa, humans occupied many hostile environments, using a toolkit that included flint cutting tools, hand axes, hunting bows and flint-tipped arrows, projectile weapons, cooperative hunting and fishing, woven and animal skin clothing sown with bone needles, woven reed matting, and effective social organization. Because the descendants of African founder population were able to rapidly colonize the coastal regions of the Old World, including Australia and New Zealand, it is believed that humans understood practical water travel, using reed bundles and rafts, wood, animal bladders, paddled canoes and sailboats that are still in widespread use today.

In a broad sense, all human biocultural adaptation involves cortex, and novel problems require *conscious* cortical regions and networks, like the ones you are using in this moment. The conscious regions of the cortico-thalamic (CT) system give us the gateway for learning and problem-solving. The proposed reason for the efficiency of conscious cortex in the face of novel challenges is its ability to recruit entirely new coalitions of expert sources to "concentrate" on a single unpredictable question.

The mammalian neocortex is roughly 200 million years old. At a basic level, the cortex is a highly flexible swarm computation architecture, although its frontal half also supports executive functions.

The prefrontal cortex (PFC) interacts with the entire sensory and motor brain, with biocultural motivation and emotions, and appetitive drivers ranging from nutrition to reproductive pheromones.

Emotion theorists have pointed out that "emotions" are dramatic fragments that use a kind of narrative case grammar. We don't just feel "anger," but we experience "anger" toward some perceived violator of the perceiver's social boundaries, such as the murder of a socially protected child. To set the balance right again, the emotional actor often engages in compensatory actions, from an act of protection or revenge, to a negotiated compensation for the loss and humiliation. Thus, emotional acts can easily be strung into entire interpersonal narratives of the kind we have in dreams: A norm-violating provocation followed by just retribution is one very common example of a narrative theme, often seen in ancient epics.

Cooperation and planning are important skills largely made possible by the prefrontal cortex. Experiential hippocampal memory (called "episodic") may record every conscious event. Biological examples of swarm computation are extremely common. Eusocial animals (like ants, naked mole rats, and termites) and slime mold colonies (like *P. polycephalum*, which can solve the famously difficult Traveling Salesman Problem using locally emergent parallel-interactive processing) are prime examples (Jones and Adamatsky 2014). Varieties of swarm computation, including mixed cases of swarm and executive computation, are therefore very common. With the emergence of language, humans learned how to implement executive computation, as in playing chess and calculating arithmetic; however, such sequential computation may be rather recent (approximately 100,000 years ago).

2 Consciously Mediated Processing in the Cortex

Functional specialization of cortical regions was controversial until Broca's and Wernicke's language areas were discovered in the 1800s. The cortex does both swarm and sequential symbolic computation. Using high spatiotemporal resolution imaging tools, we can see individual neurons performing tasks, sometimes phase-locked to population oscillations. The primary projection areas of the senses and motor systems are functional hierarchies, which signal bidirectionally, not strictly top-down or bottom-up. Sometimes single functional neuronal members of a hierarchy can be mobilized by conscious neurofeedback.

Learning throughout the brain appears to occur by the Hebbian rule: "neurons that fire together, wire together." Learned inhibition may occur the same way, using inhibitory (GABA-ergic) neurons. New functional groups are therefore constantly being created, decomposed and reorganized. Neurofeedback signaling is a powerful and general method to induce neuronal learning, using conscious feedback stimuli (tones, flashing lights, etc.). However, there is no evidence that unconscious neurofeedback leads to novel learning. This suggests that learning is consciously mediated, as shown in the case of associative conditioning. Baars (1988) describes how the GWT hypothesis can show how conscious (global) neurofeedback can recruit local neuronal groups to acquire control over local target activity.

Neurons and neuronal cell assemblies can be defined as "expert agents" when they respond selectively to input or output. Conscious experiences may therefore reflect a GW function in the brain. The cortex has many anatomical hubs, but conscious percepts are unitary and internally consistent at any given moment. This suggests that a brain-based GW capacity cannot be limited to only one anatomical hub. Rather, a consciousness-supporting GW should be sought in a mobile, dynamic and coherent binding capacity – a functional hub – for neural signaling over multiple networks.

Two research groups have found conscious (rather than unconscious) visual processing high in the visual hierarchy, including the inferotemporal cortex (IT), superior temporal sulcus

(STS), medial temporal lobe (MTL), lateral occipital complex (LOC) and the PFC. In hearing, Heschel's gyrus seems to involve a consciousness-supporting neuronal hub, and in interoceptive feelings, like nausea and "gut emotions," the anterior insula seems to be involved. External touch is probably mediated by area S1 (the somatosensory homunculus), and related sensory body maps, and the corresponding motor areas influence voluntary movement in various subtle ways.

The theater metaphor is ancient and is associated with more than one theory of consciousness. In GWT focal consciousness acts as the bright spot on the stage, which is directed by the spotlight of attention. The bright spot is surrounded by a "fringe" of vaguely conscious events (Mangan 1993). The stage corresponds to "working memory," the immediate memory system in which we talk to ourselves, visualize places and people, and make plans. Information from the bright spot is globally distributed to two classes of complex unconscious processors: those in the shadowy audience, who primarily receive information from the bright spot; and unconscious contextual systems that shape events within the bright spot, who act "behind the scenes." One example of such a context is the unconscious philosophical assumptions with which we tend to approach the topic of consciousness.

Cross-model conscious integration is extremely common, and is presumably mediated by parietal regions, but the prefrontal cortex is also a "hub of many sensory hubs," intimately connected with the others, and it is difficult to rule out a PFC function in any conscious or voluntary experience. Conscious feelings of knowing (FOKs) are vividly illustrated by Wilder Penfield's (1975) long series of open-brain surgeries on epileptic patients, which found that both sides of the prefrontal lobe (medial and lateral) are involved in feelings of effort, such as tip-of-the-tongue. Tip-of-the-tongue experiences, and their accompanying FOKs, can be induced by asking for the technical names of familiar facts. The question "What are two names for flying dinosaurs?" may elicit strong FOK. Subjects who cannot recall those names still choose accurately and quickly between "pterodactyl" and "brontosaurus." Semantic knowledge may be fully primed in tip-of-the-tongue (TOT) states, before the lexical form of the missing words can be recalled. Such FOK commonly occur when we have compelling and accurate expectations and intentions. They are not limited to language.

Our general hypothesis is that the cortical connectome (the enormous mass of myelinated long-distance fibers emerging from pyramidal cells in the neocortex, paleocortex, and thalamus) supports GW functions: That is, the ability to integrate multiple incoming signals into coherent spatiotemporal coalitions, and to "broadcast" the output signals to activate and recruit large functional cell assemblies in pursuit of some high-level goal. Recent cortical network maps using Diffuse Tractography Imaging (DTI) show classical features of large-scale networks, including small-world organization, optimal signaling efficiency, and robust functioning in the face of local damage.

In humans and macaques, the CT complex underlies reportable conscious percepts, concepts, FOKs, visual images and executive functions. While subcortical areas are sometimes claimed to specify conscious contents, the human evidence is slight and disputed. However, basal ganglia can feed back to cortex via a posterior thalamic pathway, and the thalamus is obviously involved in all cortical input-output signaling. In the case of corticofugal signals (e.g. vocalization, voluntary eye movements, corticospinal tracts, corticovagal output), conscious signaling comes from muscular output leading to sensory input, as in the famous example of the articulatory-auditory feedback loop.

Because cortex and thalamus are so densely interleaved as to constitute a single functional system, we will refer here to the CT system as a whole. CT pathways permit constant reentrant signaling, so that multiple spatiotopic maps, internal topographical representations, can sustain or inhibit each other. The CT system resembles an enormous metropolitan street plan, in which one can travel from any street corner to any other. Almost all cortico-cortical and cortico-thalamic links are bidirectional, so that the normal signaling mode in the CT system is not unidirectional, but resonant. This basic fact has many implications.

Global workspace theory follows the historic distinction between the "focus" of experience vs. the largely implicit background of experience. Extensive evidence shows that visual and auditory consciousness flows from the respective sensory surfaces to frontoparietal and particularly prefrontal regions. The CT core is a great mosaic of multi-layered two-dimensional neuronal arrays. Each array of cell bodies and neurites projects to others in topographically systematic ways. Since all CT pathways are bidirectional, signaling is "adaptively resonant" (reentrant). In this complex, layered two-dimensional arrays are systematically mirrored between cortex and thalamus, region by region.

The CT nexus appears to be the most parallel-interactive structure in the brain, allowing for efficient signal routing from any neuronal array to any other. This connectivity is different from other structures that do not directly enable conscious contents, like the cerebellum. The cerebellum is organized in modular clusters that can run independently of each other, in true parallel fashion. But in the CT core any layered array of cortical or thalamic tissue can interact with any other, more like the worldwide web than a server farm.

CT pathways run in all canonical directions and follow small-world organization, so that each array is efficiently linked to many others. The entire system acts as an oscillatory medium, with markedly different global regimes in conscious and unconscious states. Global workspace dynamics interprets the traditional distinction between the "object" and "ground" of experiences as a directional flow between the moment-to-moment focus of conscious experience vs. the implicit background and sequelae of focal contents. The proposed directionality of broadcasting suggests a testable distinction with information integration theory and dynamic core theory.

3 Dynamic GW vis-à-vis Other Theoretical Proposals

We can widely divide current theories into philosophical and empirically based ones. Some of the philosophical theories are now generating testable hypotheses. Empirical theories can be divided into "localist" vs. "local-global" types. There are no exclusively global theories, since no one denies the evidence for local and regional specialization in the brain. Philosophical theories typically aim to account for subjective experiences or "qualia," a notoriously difficult question. Recently some philosophical perspectives, like "higher order theory" (HOT), have also generated testable proposals about the involvement of brain regions like the PFC. However, brain imaging experiments (e.g. Dehaene and Naccache 2001) have long implicated the frontoparietal cortex in subjective experience.

It is not clear at this time whether philosophically based theories generate novel, testable predictions. However, efforts are underway to test HOT theories. In general, claims to explain subjective qualia are still debated. Zeki (2001) makes the localist claim that conscious percepts of red objects involve "micro-conscious" activation of cortical color regions (visual areas V3/V4). However, most empirical theories combine local and global activities, as briefly discussed above. It is still possible that momentary events may be localized for 100 milliseconds or less, and that full conscious contents emerge over some hundreds of milliseconds. The Dynamic GW theory is a specific version of the "dynamic core" hypothesis proposed by Edelman and Tononi (2000) and, in somewhat different forms, by Edelman (1989) and others. Dynamic Global Workspace theory implies a directional signal flow from binding to receiving coalitions. For each conscious event there is a dominant source and a set of receivers, where the propagated signal is interpreted, used to update local processes, and refreshed via reentrant signaling to the source (Edelman 1989). Conscious sensations arise in a different center of binding and propagation than "feelings of knowing" like the TOT experience, as demonstrated by brain imaging studies (Maril et al. 2001). Directional broadcasting of bound conscious contents is one testable distinction from other proposals (Edelman et al. 2011). Supportive evidence has been reported by Doesburg et al. (2009) and others. Other theories, like Tononi's mathematical measure of complexity, phi, seem less directional (Edelman and Tononi 2000). Llinas and Pare (1991) have emphasized the integration of specific and nonspecific thalamocortical signaling, and Freeman et al. (2003) have developed a conception of hemisphere-wide signaling and phase changes. Nevertheless, current local-global theories are strikingly similar. Whether major differences will emerge over time is unclear.

4 Dynamic Global Workspace as a Local-Global Theory

In 1988, GW theory suggested that "global broadcasting" might be one property of conscious events. Other proposed properties were:

1 Informativeness, that is, widespread adaptation to the novelty of the reportable signal, leading to general habituation (information reduction) of the news contained in the global broadcast. The evidence now supports widespread neuronal updating to novel input.
2 Internal consistency of conscious contents, because mutually exclusive global broadcasts tend to degrade each other. This is a well-established feature of conscious contents, first observed in the nineteenth century and replicated many thousands of times. Binocular rivalry is one well-known example.
3 Interaction with an implicit self-system. Baars (1988) proposed that the observing self is coextensive with implicit frames that shape the objects of consciousness. One major kind of access that has been discussed since Immanuel Kant is the access of the "observing self" to the contents of consciousness. Lou et al. (2010) have shown that self-related brain regions like the precuneus and midline structures from the PAG to orbitofrontal cortex may be mobilized by conscious sensory contents. Baars (1988) proposed that self-other access is a specific variety of framing (contextualizing), and that it is a necessary condition for conscious contents.
4 One of the driving questions of GW theory is how the limited capacity of momentary conscious contents can be reconciled with the widespread access enabled by conscious contents. Why is the conscious portion of an otherwise massively parallel-distributed system a limited and serial process? Would our ancestors not have benefited from the ability to competently perform several tasks at once? A stream of consciousness integrates disparate sources of information, but it is limited to a "single internally consistent content at any given moment" (Baars 1988). The Oxford English Dictionary dedicates 75,000 words to the various definitions of "set." However, a native speaker will, while reading or listening, know almost immediately in what sense the word is being used. We can rapidly detect errors in phonology, syntax, semantics, and discrepancies between a speaker's stated and true intentions, but are not necessarily conscious of how this is done The workspace makes sense of novel and ambiguous situations by calling upon unconscious "expert" processors (see Figure 9.1).

Because almost all neural links in the CT system are bidirectional, reentrant signaling from receivers to broadcasting sources may quickly establish task-specific signaling pathways, in the same way that a fire department might locate the source of a community-wide alarm, and, then,

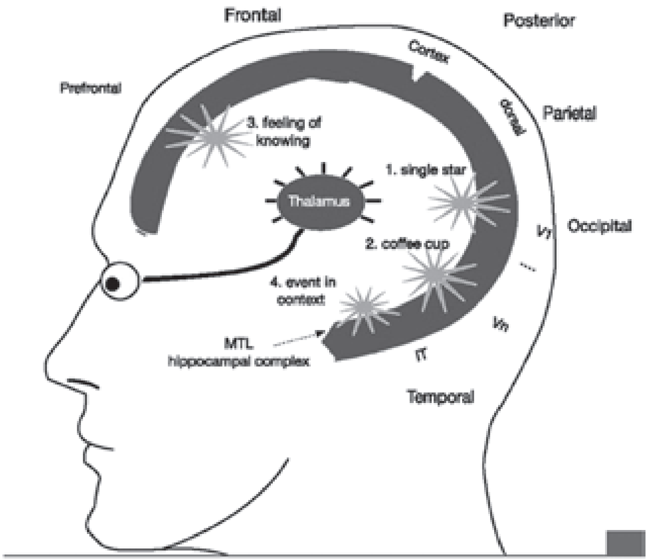

Figure 9.1 Examples of Possible Binding and Broadcasting in the Cortico-Thalamic Core

communicate in a much more task-specific way. Current evidence suggests brief broadcasts, as suggested by the 100 ms conscious integration time of different sensory inputs.

Figure 9.1 shows four examples of possible binding and broadcasting in the CT core (starburst icons). Cortical area V1 and the lateral geniculate nucleus (LGN) – the visual thalamus – can be conceived as two arrays of high-resolution bright and dark pixels, without color. The sight of a single star on a dark night may therefore rely heavily on V1 and its mirror array of neurons in LGN. V1 and LGN interact constantly, with bidirectional signal traffic during waking. The sight of a single star at night reveals some surprising features of conscious vision, including spatial context sensitivity, as in the classical autokinetic effect: single points of light in a dark space begin to wander long subjective distances in the absence of spatial framing cues. The autokinetic effect is not an anomaly, but rather a prototype of decontextualized percepts (Baars 1988). A large literature in perception and language shows scores of similar phenomena, as one can demonstrate by looking at a corner of a rectangular room through a reduction tube that excludes external cues. Any two- or three-way corner in a carpentered space is visually reversible, much like the Necker Cube and the Ames trapezoid. Such local ambiguities exist at every level of language comprehension and production (Baars 1988; Shanahan and Baars 2005).

The dorsal stream of the visual cortex provides egocentric and allocentric "frames" to interpret visual events in nearby space. These parietal frames are not conscious in themselves, but they are required for visual objects to be experienced at all (Goodale and Milner 1992). Injury to the right parietal cortex may cause the left half of visual space to disappear, while contralesional stimulation, like cold water in the left ear, may cause the lost half of the field to reappear. Thus, even a single dot of light in a dark room reveals the contextual properties of conscious perception. Ambiguity and its resolution is a universal need for sensory systems in the natural world, where ambiguity is commonly exacerbated by camouflage, deceptive signaling, distraction, unpredictable movements, ambushes, sudden dangers and opportunities, darkness, fog, light glare, dense obstacles, and constant utilization of cover by predators and prey (Bizley et al. 2012).

Conscious percepts plausibly involve multiple "overlays," like map transparencies, which can be selectively attended. The sight of a coffee cup may involve an object overlaid by color, texture, and reflectance, combining information from LGN, V1, V2, V3/V4, and IT (Crick and Koch 1990). Active cells in those arrays may stream signals across multiple arrays, cooperating, and competing to yield a winner-take-all coalition. Once the winning coalition stabilizes, it may "ignite" a broadcast to other regions.

Conscious vision is strikingly flexible with respect to level of analysis, adapting seamlessly from the sight of a single colored dot to the perception of a dotted (pointillist) painting. An account of conscious vision must therefore explain how a local dot can be perceived in the same visual display as a Georges Seurat painting.

To identify a single star at night, because the highest spatial resolution is attained in the retina, LGN, and V1, the visual cortex must be able to amplify neuronal activity originating in LGN-V1 through attentional modulation. For coffee cups and faces, the relative activity of IT and the fusiform gyrus must be increased. It follows that binding coalitions of visual activity maps can bring out the relative contribution of different feature levels, even for the same physical stimulus (Itti and Koch 2001).

5 Bidirectional Pathways and Adaptive Resonance

Because CT pathways are bidirectional they can support "reentrant signaling" among topographically regular spatial maps. The word "resonance" is often used to describe CT signaling (Wang 2001). It is somewhat more accurate than "oscillation," which applies to true iterative patterns like sine waves. Edelman and coworkers prefer the term "reentry," while others like to use "adaptive resonance." We will use the last term to emphasize its flexible, selective, and adaptive qualities.

Adaptive resonance has many useful properties, as shown in modeling studies like the Darwin autonomous robot series, where it can account for binding among visual feature maps, a basic property of visual perception (Izhikevich and Edelman 2008). Edelman has emphasized that reentry (adaptive resonance) is not feedback, but rather evolves a selectionist trajectory that can search for solutions to biologically plausible problems. Grossberg and others have developed adaptive resonance models for cortical minicolumns and layers.

6 Broadcasting: Any-to-Many Signaling

A few ants can secrete alarm pheromones to alert a whole colony to danger, an example of any-to-many broadcasting among insects. In humans the best-known example is hippocampal-neocortical memory storage of memory traces in the neocortex by way of the hippocampal complex (Nadel et al. 2000; Ryan et al. 2001). Memories of conscious episodes are stored in millions of synaptic alterations in the neocortex (Moscovitch et al. 2005). Computer users are familiar with global memory searches, which are used when specific searches fail. The CT system may enable brain-based global memory searches. "Any-to-many" coding and retrieval can be used to store and access existing information (Nadel et al. 2000; Ryan et al. 2010). It is also useful for mobilizing existing automatisms to deal with novel problems.

Notice that "any-to-many" signaling does not apply to the cerebellum, which lacks parallel-interactive connectivity, or to the basal ganglia, spinal cord, or peripheral ganglia. Crick and Koch have suggested that the claustrum may function as a GW underlying consciousness (Crick and Koch 2005). However, the claustrum, amygdala, and other highly connected anatomical hubs seem to lack the high spatiotopic bandwidth of the major sensory and motor interfaces, as shown by the very high-resolution of minimal conscious stimuli in the major modalities. On the motor side, there is extensive evidence for trainable voluntary control over single motor units and more recently, for voluntary control of single cortical neurons (Cerf et al. 2010). The massive anatomy and physiology of cortex can presumably support this kind of parallel-interactive bandwidth. Whether structures like the claustrum have that kind of bandwidth is doubtful.

We do not know the full set of signaling mechanisms in the brain, and any current model must be considered provisional. Neural computations can be remarkably flexible, and are, to some degree, independent of specific cells and populations. John et al. (2001) has argued that active neuronal populations must have dynamic turnover to perform any single brain function, like active muscle cells. Edelman and Tononi (2000) and others have made the same point with the concept of a dynamic core. GW capacity as defined here is not dependent upon the mere existence of anatomical hubs, which are extremely common. Rather, it depends upon a dynamical capacity, which operates flexibly over the CT anatomy, a "functional hub," so that activated arrays make up coherent "coalitions."

The global neuronal workspace has been used to model a number of experimental phenomena. In a recent model, sensory stimuli mobilize excitatory neurons with long-range cortico-cortical axons, leading to the genesis of a global activity pattern among workspace neurons. This class of models is empirically linked to phenomena like visual backward masking and in attentional blindness (Dehaene and Changeux 2005).

Franklin et al. (2012) have combined several types of computational methods using a quasi-neuronal activation-passing design. High-level conceptual models such as LIDA (Snaider, McCall, and Franklin 2011) can provide insights into the processes implemented by the neural mechanisms underlying consciousness, without necessarily specifying the mechanisms themselves. Although it is difficult to derive experimentally testable predictions from large-scale architectures, this hybrid architecture approach is broadly consistent with the major empirical features discussed in this article. It predicts, for example, that consciousness may play a central role in the classic notion of cognitive working memory, selective attention, learning, and retrieval.

7 Global Chatting, Chanting, and Cheering

Spontaneous conscious mentation occurs throughout the waking state, reflecting repetitive themes described as "current concerns." Conscious mentation is also reported when subjects are awoken from Rapid Eye Movement (REM) dreams and even from slow-wave sleep. The last may reflect waking-like moments during the peaks of the delta wave (Valderrama et al. 2012).

Global brain states can be compared to a football crowd with three states: "chatting," "chanting," and "cheering." Chatting describes the CT activity of waking and REM dreams. It involves point-to-point conversations among spatial arrays in the CT system, which can have very high S/N ratios, though they appear to be random when many of them take place at the same time. Like a football stadium with thousands of coordinated local conversations that are not coordinated globally, the average global activity is a low-level crowd roar, seemingly random, which appears to be fast and low in amplitude.

Nevertheless, as we will see, direct cortical recordings show phase-coupled chatting in the CT core appears to underlie specific cognitive tasks. Thus, chatting activity gives the misleading appearance of randomness *en masse,* but it is in fact highly organized in a task-driven fashion. Because sports arenas show the same properties, the arena metaphor provides us with a useful reminder.

Chanting shows coordinated start-stop crowd activity, about once a second over a prolonged period of time, like the "buzz-pause" rhythm of billions of neurons in the CT core, which results in global delta waves. Chanting sounds like chatting at the peak of the delta wave, followed by simultaneous pausing, which interrupts all conversations at the same time (Massimini et al. 2005).

Finally, a stadium crowd may cheer when a team scores a goal or makes an error. This corresponds to an "event-related" peak of activity. In the brain, the event-related potential (ERP) occurs when a significant or intense stimulus is processed, causing a stereotypical wave pattern to sweep through the brain.

8 Feature and Frame Binding

In GWT frames (previously called contexts) can be thought of as groups of specialists dedicated to processing input in particular ways. As we have seen, there are frames for perception and imagery (where they help shape qualitative experiences), as well as in conceptual thought, goal directed activities and the like (where they serve to access conscious experiences). One of the primary functions of consciousness is to evoke contexts that shape experiences. Some challenges to a dominant frame are more noticeable than others. Consider the following from Eriksen and Mattson (1981):

1 How many animals of each kind did Moses bring on the Ark?
2 In the Biblical story, what was Joshua swallowed by?
3 What is the nationality of Thomas Edison, inventor of the telephone?

While some subjects noticed errors with one or all of these statements, most did not. When asked directly, subjects showed they knew the answers, but it took more severe violations (e.g "How many animals of each kind did Nixon bring on the Ark?") for the majority to see any issues.

Visual features are stimulus properties that we can point to and name, like "red," "bounded," "coffee cup," "shiny," etc. Feature binding is a well-established property of sensory perception. There is much less discussion about what we will call "frame-binding," which is equally necessary, where "frames" are defined as visual arrays that do not give rise to conscious experiences, but which are needed to specify spatial knowledge within which visual objects and events become conscious. Powerful illusions like the Necker Cube, the Ames trapezoidal room, the railroad (Ponzo) illusion are shaped by unconscious Euclidean assumptions about the layout of rooms, boxes, houses, and roads.

The best-known brain examples are the egocentric (coordinate system is centered around the navigator) and allocentric (oriented on something other than the navigator) visuotopic arrays of the parietal cortex. When damaged on the right side, these unconscious visuotopic fields cause the left half of objects and scenes to disappear, a condition called hemi-neglect. Goodale and Milner have shown that even normal visuomotor guidance in near-body space may be unconscious. In vision the dorsal "framing" stream and "feature-based" ventral stream may combine in the medial temporal cortex (MTL) (Shimamura 2010). Baars (1988) reviewed extensive evidence showing that unconscious framing is needed for normal perception, language comprehension and action planning. In sum, normal conscious experiences need both traditional feature binding and frame-binding (Shanahan and Baars 2005).

9 Perceptual Experiences vs. Feelings of Knowing (FOKs)

This Dynamic GW theory figure shows an occipital broadcast (which must mobilize parietal egocentric and allocentric maps as well) evoking spatiotopic activity in the prefrontal cortex, which is known to initiate prefrontal activation across multiple tasks demanding mental effort (Duncan and Owen 2000), and suggests that sensory conscious experiences are bound and broadcast from the classical sensory regions in the posterior cortex, while voluntary effort, reportable intentions, feelings of effort, and the like, have a prefrontal origin, consistent with brain imaging findings.

These findings suggest a hypothesis about sensory consciousness compared to "fringe" FOK, feelings of effort, and reportable voluntary decisions. These reportable but "vague" events have been discussed since William James (1890) who gave them equal importance to perceptual consciousness. Functional magnetic resonance imaging (fMRI) studies show that they predominantly involve prefrontal regions, even across tasks that seem very different.

Because of the small-world connectivity of white matter tracts, different integration and distribution hubs may generate different global wave fronts. The sight of a coffee cup may involve an infero-temporal hub signaling to other regions, while the perception of music may emerge from Heschel's gyrus and related regions. Reportable experiences of cognitive effort might spread outward from a combined dorsolateral prefrontal cortex (dlPFC)/anterior cingulate cortex (ACC) hub.

10 Conscious Events Evoke Widespread Adaptation or Updating

What is the use of binding and broadcasting in the CT system? One function is to update numerous brain systems to keep up with the fleeting present. GW theory suggested that consciousness is required for non-trivial learning (i.e., learning that involves novelty or significance) (Baars 1988). While there are constant efforts to demonstrate robust unconscious learning, after six decades of subliminal vision research there is still little convincing evidence. Subliminal perception may work with known chunks, like facial expressions, but while single-word subliminal priming appears to work, Baars (1988) questioned whether novel two-word primes would work subliminally. The subliminal word pair "big house" might prime the word "tall," while "big baby" might not, because it takes conscious thought to imagine a baby big enough to be called tall. In general, the more novelty is presented, the more conscious exposure is required.

It follows that the Dynamic GW theory should predict widespread adaptive changes after conscious exposure to an event. That is indeed the consensus for hippocampal-neocortical memory coding (Nadel et al. 2012). However, the hippocampal complex is not currently believed to enable conscious experiences. Nevertheless, episodic memory is by definition "memory for conscious events." Conscious events trigger wide adaptation throughout the CT system, and in subcortical regions that are influenced by the CT system. Episodic, semantic, and skill (procedural) processing all follow the same curve of high-metabolic processing to novel, conscious learning followed by a drastic drop in conscious access and metabolic BOLD (blood-oxygen-level dependent) activity after learning.

11 Voluntary Reports of Conscious Events

Conscious contents are routinely assessed by voluntary report, as we know from 200 years of scientific psychophysics. Yet the reason for that fact is far from obvious. Any theory of consciousness must ultimately explain the basic fact that we can voluntarily report an endless range of conscious contents, using an endless range of voluntary actions. Voluntary control is one kind of consciously mediated process. As we learn to ride a bicycle for the first time, each movement seems to come to consciousness. After learning, conscious access drops even as BOLD activity in the CT core declines. We postulate that conscious involvement is necessary for non-trivial acquisition of knowledge and skills, and that the period of conscious access enables permanent memory traces to be established.

While "verbal report" is the traditional phrase, reports do not have to be verbal – any voluntary response will work. Broca's aphasics who cannot speak can point to objects instead. Locked-in (paralyzed) patients, who seem to be comatose, can learn to communicate by voluntary eye movements. Thus "verbal report" should be called "accurate, voluntary report," using any controllable response. Voluntary actions can point to objects and events. A "match to sample" task is commonly used to indicate the similarity of two conscious events, and to specify just noticeable differences. Pointing occurs naturally when mammals orient to a novel or significant stimulus. Children develop pointing abilities using "shared attention" in early childhood.

For simplicity's sake let's assume conscious contents emerge in posterior cortex and voluntary actions emerge in frontal and parietal cortex. We can ask the question in The Dynamic GW theory terms: how is a posterior "binding and broadcasting" event transformed into a frontally controlled action? These facts raise the question of how accurate signal transmission occurs between sensory arrays and frontal executive control. In the case of pointing to a single star on a dark night, the physical minimum of light quanta in the retina can be amplified and transmitted to prefrontal cortex, which can control the movement of a single finger to point to the star.

Even more remarkably, single neurons in the temporal cortex have been shown to be fired at will in surgical patients using intracranial electrodes, providing only that conscious sensory feedback is given during training (Cerf et al. 2010). Thus, the physical minimum to the eye can accurately translate into "any" voluntarily controlled single cell, used as a sensory pointer. Given a million foveal cells for input, and perhaps billions of cortical cells for output, "any-to-any" mapping in the brain can involve remarkably large numbers. With accurate psychophysical performance in both tasks, the signal-to-noise ratio from receptor to effector cell can approach the physical limit. This precision needs explanation in terms of conscious input and voluntary control.

This also suggests an explanation for the standard index of voluntary report. When we report a star on a dark night, posterior broadcasting may lead to frontal binding and ultimately a frontal broadcast. Frontoparietal regents are driven by posterior sensory projections when they become conscious. Because of the striking similarities of spatiotopic coding in frontal and posterior cortices, we can imagine that sensory consciousness can also trigger a new binding, and broadcast an event in the frontal cortex. Voluntary action is therefore an extension of GW dynamics.

Conscious contents enable access to cognitive functions, including sense modalities, working memory, long-term memories, executive decisions and action control. Executive regions of the frontoparietal cortex gain control over distributed unconscious functions. Animals live in a world of unknowns, surrounded by dangers and opportunities that may be fleeting, hidden, camouflaged, surprising, deceptive, and ambiguous. Conscious brains may have evolved to cope with such unknowns (Baars 1988, 2002). Newell and colleagues built the first GW architecture to perform acoustical word recognition, at a time when that task was largely underdetermined (Newell 1990). Their solution was to build a computational architecture, a blackboard model, which would allow many incomplete sources to compete and cooperate to resolve some focal ambiguity. The result was remarkably successful for its time in recognizing nearly 1,000 ordinary words spoken in normal acoustical spaces, complete with hard echoing surfaces, mumbling speakers, and soft, absorbent surfaces, background noises, and the like. Speech recognition is now handled with improved formant tracking, but even today, if semantic unknowns arise in a spoken word stream, a GW architecture may be useful to find the answer. We have no semantic algorithms that interpret word ambiguities across many domains, the way humans routinely do. Baars and Franklin (2003) used GW theory to propose that consciousness enables access between otherwise separate knowledge sources.

GW architectures can also "call" fixed automatisms. For example, in speech recognition word ambiguity may be resolved by a known syntactic rule. A global broadcast of the ambiguous word may recruit routines whose relevance cannot be known ahead of time. We have referred to this as contextualization or frame binding (Baars 1988; Shanahan and Baars 2005). The "frame problem" is a recognized challenge in artificial intelligence and robotics, but it applies equally to living brains. Briefly stated, it is an effort to explain how a "cognitive creature with many beliefs about the world" can regularly update them while remaining "roughly faithful to the world" (Dennett 1978). In GWT this conundrum is solved through the invocation of unconscious context-sensitive and context-shaping processors.

12 Concluding Remarks

The main use of a GW system is to solve problems which any single "expert" knowledge source cannot solve by itself – problems whose solutions are underdetermined. Human beings encounter such problems in any domain that is novel, degraded, or ambiguous. This is obvious for novelty: if we are just learning to ride a bicycle, or to understand a new language, we have inadequate information by definition. Further, if the information we normally use to solve a known problem becomes degraded, determinate solutions again become indeterminate.

What may not be so obvious is that there are problems that are inherently ambiguous, in which all the local pieces of information can be interpreted in more than one way, so that we need to unify different interpretations to arrive at a single, coherent understanding of the information. But there are numerous biological examples of densely vegetated fields and forests that harbor so many hiding places for animals and birds that there is in principle no way to make the visual scene predictable. Many wet jungle regions also have very loud ambient sounds produced by insects, frogs and birds, so that the noise level exceeds the signal emanating from any single individual animal. This situation also applies to the famous human cocktail party effect, where we can understand conversations despite a negative signal-to-noise ratio. Clearly biological sensory systems can thrive in such noisy environments, perhaps using top-down predictions and multimodal signal correlations. Standard sensory studies in humans and animals have generally neglected this ecologically realistic scenario.

Conscious learning is often involved in decomposing such complex signal environments, as in the case of human music conductors, for example, who can rapidly pinpoint wrong notes. In these cases, top-down learning of musical patterns and entire large-ensemble scores is involved, but talented experts spend a lot of conscious time on the learning process, and their spectacular performances do not contradict our observations about the many functions of conscious thought.

References

Baars, B. J. (1988) *A Cognitive Theory of Consciousness*, New York: Cambridge University Press.

Baars, B. J. (2002) "The conscious access hypothesis: origins and recent evidence," *Trends in Cognitive Science* 6: 47–52.

Baars, B. J., and Franklin, S. (2003) "How conscious experience and working memory interact," *Trends in Cognitive Science* 7: 166–172.

Bizley, J. K., Shinn-Cunningham, B. G., and Lee, A. K. (2012) "Nothing is irrelevant in a noisy world: sensory illusions reveal obligatory within and across modality integration," *Journal of Neuroscience* 32: 13402–13410.

Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R. Q., Koch, C., Fried, I. (2010) "On-line, voluntary control of human temporal lobe neurons," *Nature* 467: 1104–1108.

Crick, F., and Koch, C. (1990) "Towards a neurobiological theory of consciousness," *Seminars in Neuroscience* 2: 263–275.

Crick, F., and Koch, C. (2005) "What is the function of the claustrum?" *Philosophical Transactions of the Royal Society B.* 360: 1271–1279.

Dehaene, S., and Changeux, J. P. (2005) "Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness," *PLoS Biology* 3: 0920–0923.

Dehaene, S., and Naccache, L. (2001) "Towards a cognitive permanence of consciousness: basic evidence and a workspace framework," *Cognition.* 79: 1–37.

Dennett, D. (1978) *Brainstorms*, Cambridge, MA: MIT Press.

Doesburg, S. M., Green J. J., McDonald J. J., and Ward L. M. (2009) "Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception," *PLoS ONE* 4: 1–14.

Duncan, J., and Owen, A. M. (2000) "Common regions of the human frontal lobe recruited by diverse cognitive demands," *Trends in Neuroscience* 23: 475–483.

Edelman, G. M. (1989) *The Remembered Present: A Biological Theory of Consciousness*, New York: Basic Books Inc.

Edelman, G. M., Gally, J. A., and Baars, B. J. (2011) "Biology of consciousness," *Frontiers in Psychology* 2: 4.

Edelman, G. M., and Tononi, G. (2000) *A Universe of Consciousness: How Matter Becomes Imagination*, New York: Basic Books Inc.

Erickson, T. D., and Mattson, M. E. (1981) "From words to meaning: A semantic illusion," *Journal of Verbal Learning and Verbal Behavior* 20: 540–551.

Franklin, S., Strain, S., Snaider, J., McCall, R., and Faghihi, U. (2012) "Global workspace theory, its LIDA model and the underlying neuroscience," *Biologically Inspired Cognitive Architectures* 1: 32–43.

Freeman, W. J., Burke, B. C., and Holmes, M. D. (2003) "A periodic phase resetting in scalp EEG of beta-gamma oscillations by state transitions at alpha-theta rates," *Human Brain Mapping* 19: 248–272.

Goodale, M. A., and Milner, A. D. (1992) "Separate visual pathways for perception and action," *Trends in Neuroscience* 15: 20–25.

Itti, L., and Koch, C. (2001) "Computational modelling of visual attention," *Nature Reviews Neuroscience* 2: 194–203.

Izhikevich, E. M., and Edelman, G. M. (2008) "Large-scale model of mammalian thalamocortical systems," *Proceedings of the United States National Academy of Science* 105: 3593–3598.

James, W. (1890) *The Principles of Psychology*, New York: Holt.

John, E. R., Prichep, L. S., Kox, W., Valdés-Sosa, P., Bosch-Bayard, J., Aubert, E., Tom., M. di Michele, F., and Gugino, L.D. (2001) "Invariant reversible QEEG effects of anesthetics," *Conscious Cognition* 10: 165–183.

Jones, J., and Adamatzky, A. (2014) "Computation of the travelling salesman problem by a shrinking blob," *Natural Computing* 13: 1–16.

Llinas, R. R., and Pare, D. (1991) "Of dreaming and wakefulness," *Neuroscience* 44: 521–535.

Lou, H. C., Luber, B., Stanford, A., and Lisanby, S. H. (2010) "Self-specific processing in the default network: a single-pulse TMS study," *Experimental Brain Research* 207: 27–38.

Mangan, B. (1993) "Taking phenomenology seriously: the fringe and its implications for cognitive research," *Consciousness and Cognition* 2: 89–108.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005) "Breakdown of cortical effective connectivity during sleep," *Science* 309: 2228–2232.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., McAndrews, M.P., Levine, B., Black, S., Winocur, G., and Nadel, L. (2005) "Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory," *Journal of Anatomy* 207: 35–66.

Nadel, L., Samsonovich, A., Ryan, L., and Moscovitch, M. (2000) "Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results," *Hippocampus* 10: 352–368.

Newell, A. (1990) *Unified Theories of Cognition*, Cambridge, MA: Harvard University Press.

Penfield, W. (1975) *The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain*, Princeton: Princeton University Press.

Ryan, L., Lin, C. Y., Ketcham, K., and Nadel, L. (2010) "The role of medial temporal lobe in retrieving spatial and nonspatial relations from episodic and semantic memory," *Hippocampus* 20: 11–18.

Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., and Moscovitch, M. (2001) Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11: 707–714.

Shanahan, M., and Baars, B. (2005) "Applying global workspace theory to the frame problem," *Cognition* 98: 157–176.

Shimamura, A. P. (2010) “Hierarchical relational binding in the medial temporal lobe: the strong get stronger,” *Hippocampus* 20: 1206–1216.

Snaider, J., McCall, R., and Franklin, S. (2011) “The LIDA Framework as a General Tool for AGI,” *Fourth Conference on Artificial General Intelligence.* Mountain View, CA: Springer Lecture Notes in Artificial Intelligence.

Valderrama, M., Crépon, B., Botella-Soler, V., Martinerie, J., Hasboun, D., Alvarado-Rojas, C., Baulac, M., Adam, C., Navarro, V., and Le Van Quyen, M. (2012). “Human gamma oscillations during slow wave sleep,” *PLoS ONE* 7: 1–14.

Wang, X. J. (2001) “Synaptic reverberation underlying mnemonic persistent activity,” *Trends in Neuroscience* 24: 455–463.

Zeki, S. (2001) “Localization and globalization in conscious vision,” *Annual Review of Neuroscience* 24: 57–86.

Related Topics

Materialism
Representational Theories of Consciousness
The Multiple Drafts Model
The Information Integration Theory
The Intermediate Level Theory of Consciousness
The Attention Schema Theory of Consciousness
The Neural Correlates of Consciousness
The Biological Evolution of Consciousness