University of Memphis

University of Memphis Digital Commons

2010

Two varieties of Unconscious Processes

S. Franklin

B. Baars

Follow this and additional works at: https://digitalcommons.memphis.edu/ccrg_papers

Two Varieties of Unconscious Processes

Stan Franklin and Bernard J. Baars

Abstract

Unconscious processes come in two varieties, the preconscious, whose contents may become conscious, and the never-conscious, whose contents may not. In this chapter we make use of Global Workspace Theory and its LIDA model to catalog never-conscious and preconscious processes, and offer an explanation of the functionally of the distinction. The LIDA/GWT model suggests that the functional distinction between never-conscious and preconscious processes derives from one of the major purposes of an agent's consciousness mechanism, which is to select the most salient portion of the agent's current situation to which to attend, that is to broadcast globally, in order to choose the best next action.

Introduction

Cognitive processes[1] come in two varieties, conscious and unconscious, thought to be distinguished by the contents of conscious processes being broadly accessible throughout the nervous system (Baars, 2002). This conscious access hypothesis is consistent with a theory of conscious content being accessed throughout the thalamocortical core (Edelman & Tononi, 2000). A dynamical

---

[1] In this paper we use the terms "process" and "content" in a strictly computational fashion. A process is thought of as an algorithm or a program as distinct from any content that might be produced by it.

systems approach leads to the same hypothesis (Freeman, 2003). Philosophers have also discussed this hypothesis (Block, 2007). Finally, the conscious access hypothesis is a central tenet of Global Workspace Theory (see below, and Baars, 1988).

But unconscious processes also have their varieties. These are fundamental to our understanding of conscious events, because they provide the necessary comparison conditions. Unconscious brain processes often resemble conscious ones very closely, and it is the _differences_ between them that provide clues about the nature of consciousness.

The contents of *preconscious processes* may remain unconscious or may become part of the conscious contents, while we will introduce the term "*Never-conscious*" to refer to those brain events that never come to consciousness --- such as automatisms, neuronal processes in the cerebellum, and long-term procedural memories. Deheane et al (2006) have suggested using *subliminal processing* for this purpose, but that term has been in common use for more than 50 years in the field of subliminal (sub-threshold) sensory perception. We therefore prefer to use "*Never-conscious*," abbreviated as N-conscious.

The LIDA model of Global Workspace Theory (Franklin & Patterson, 2006; Franklin et al, 2005) (see below) enables us to draw the distinction between N-conscious vs. preconscious processes. The sections below will describe the LIDA model, discuss N-conscious and preconscious processes, and offer a case for the functionality of the distinction.

Global Workspace Theory

Global workspace theory (GWT, Baars 1988) was originally conceived as a neuropsychological model of conscious and unconscious processes, but has been broadened into a higher-level theory of human cognitive processing, supported by empirical evidence (Baars 2002), particularly when it is combined with the LIDA computational model of cognition. GWT views the nervous system as a distributed parallel system with many different specialized processes. Some coalitions of these processes enable an agent[2] to make sense of the sensory data coming from the current environmental situation. Other coalitions incorporating the results of the processing of sensory data compete for attention. The winner occupies what Baars calls a global workspace; the settled contents of which are broadcast to all other processes. In combination with other criteria, the contents of the global workspace are proposed to be conscious (see Baars, 1988). This conscious broadcast serves to recruit other, unconscious, processes to be used to select an action with which to deal with the current situation. GWT is therefore a theory of how consciousness functions within cognition. Unconscious contexts influence this competition for consciousness.

GWT is constantly being tested against new bodies of evidence. For example, any adequate theory must account for both conscious and unconscious brain conditions. The unconscious comparison conditions include:

---

[2] This term refers to an autonomous agent in the sense of Franklin & Graesser (1997), which includes humans and most animals, as well as some software agents and mobile robots. Here're we're concerned mostly with human agents.

1. Large brain regions that are constantly active, but never result in direct conscious output, such as the cerebellum and basal ganglia. The cerebellum, for example, may have as many neurons as the cerebral cortex. But cortex has traditionally been regarded as the source of conscious contents, while these other regions are not. Cortex has long been believed to provide the detailed contents of consciousness. A large number of studies show that lesions to visual cortex results in selective types of blindness such as "blindsight," face blindness (prosopagnosia), colour blindness or selective object blindness. Local or even global lesions in the cerebellum have no such effects, though the cerebellum may indirectly affect cortically-supported conscious contents. Since cortex and cerebellum have comparable numbers of neurons, it is clear that the sheer number of neurons involved in brain structures does not necessarily give rise to conscious experiences.

2. Other large regions of the cortex are *not* believed to give rise to conscious contents, notably the "dorsal visual stream," which includes most of the parietal lobes (Goodale & Milner, 2004). It follows that merely being represented in cortex, even in regions sensitive to visual features, does not necessarily give rise to conscious contents. The dorsal stream may be best viewed as a "contextual" information stream that interacts with conscious objects, events, and scenes, all represented directly in the ventral stream. Dorsal visuotopic areas represent peripersonal space, object space, and the hand-eye coordination controlled by sensory and

motor cortex, as in the case of reaching for an object or posting a letter in the slot of a mailbox. When the right parietal cortex is damaged, the left side of conscious visual space disappears from consciousness, though some unconscious processes continue. Such "parietal neglect" can be reversed through the irrigation of the left ear canal with cold water. The entire pattern of results suggests that the dorsal stream does not supply conscious object or event information by itself, but that it may provide a necessary condition for visual features and object represented in the ventral stream, including the MTL, to become conscious. This may sound like a subtle distinction, but it is absolutely indispensible. Basal brain nuclei that project to cortex for neuromodulation, for example, are crucial for the state of consciousness, without which we do not have specific experiences at all. Yet basal neuromodulating nuclei are not believed to directly support specific conscious contents.

3. The entire cortex is relatively less conscious for about six hours of the 24-hour cycle, during slow-wave sleep (SWS). Some researchers maintain that true unconsciousness occurs only during the so-called "DOWN" states of the delta wave during Slow-Wave Sleep. It is certainly well established that some mental contents are often reported when people are awoken from SWS . Destexhe et al (2007) have therefore suggested that even during SWS people may be fleetingly conscious during the UP states or peak of each slow oscillation. That question is under current study. If it is true, then totally unconscious events may only occur intermittently even

during SWS, for a total of perhaps three hours per 24 hour cycle. Other evidence indicates that Slow-Wave Sleep becomes deeper through the night, and that wider brain regions decrease their metabolic activity over the hours of sleep. The greatest degree and duration of genuine unconsciousness may therefore occur in the latter half of the night.

4. It has long been known that novel skills are more available to consciousness than the identical behaviors after practice. Riding a bicycle is a common example. A great deal of evidence indicates that most cortical activity declines rapidly with the loss of conscious access to repeated skills (e.g., Schneider, 2009). Automatic (habitual) processes are believed to be handled by the basal ganglia and cerebellum when they become highly predictable. Cortical synaptic changes are no doubt involved as well. Once long-term synaptic changes have grown and stabilized it is possible that they are also less likely to give rise to conscious contents directly. However, this question is currently being studied by means of direct brain stimulation, and we can expect clearer results in the near future.

5. A number of experimental ways of manipulating conscious perceptual access have been carefully studied over half a century, including subliminal stimulation, visual backward masking, the attentional blink and selective attention. In many cases the unconscious comparison condition has been demonstrated to activate the same brain regions that are involved in the conscious condition, though to a lesser degree (e.g.,

Dehaene, et al, 2006). Conscious contents, when compared to unconscious ones, generally evoke far more forward cortical activity, in frontoparietal regions. Dehaene and colleagues take this to mean that global broadcasting takes place in the conscious case, arguably from the prefrontal regions using long axons. However, it is still possible that the sensory regions themselves may constitute global workspaces, using long fibres to send their information forward to frontoparietal cortex and across the corpus callosum. Baars & Franklin (2003) have suggested that the conscious broadcast (which may take 100 ms or multiples thereof) recruits areas involved in classical working memory, such as the dorso-lateral prefrontal cortex and the language areas of the speaking hemisphere. Classical working memory operates over 10-30 seconds, therefore much longer than the hypothesized global broadcast associated with GWT. Since brain oscillations are widely believed to carry signals among cortical regions, GWT may be associated with alpha or theta rhythms. It is generally thought that faster gamma activity may be multiplexed on such near-10 Hz waveforms. Alpha and theta activity may therefore have a pacing or grouping function to support high-gamma, long-distance resonance among cortical regions that together may be the physiological substrate of conscious moments.

6. A number of disease states lead to characteristic impairments of conscious states and experiences. These include certain epileptic seizures and epileptic "behavioral automatisms," loss of consciousness

due to fainting, sleep disorders like sleepwalking, vegetative and minimally conscious states, analgesia and anesthesia, psychotic hallucinations and delusions, the dementias and so on. An adequate theory of conscious contents should be able to generate testable hypotheses about these conditions.

The LIDA Model and its Architecture

The LIDA model is a comprehensive, conceptual and computational[3] model covering a large portion of human cognition (Franklin & Patterson, 2006; Franklin, et al., 2007). Besides GWT, the model implements and fleshes out a number of psychological and neuropsychological theories including situated cognition (Varela, et al, 1991), perceptual symbol systems (Barsalou, 1999), working memory (Baddeley & Hitch, 1974), memory by affordances[4] (Glenberg, 1997). long-term working memory (Ericsson & Kintsch, 1995), and the H-CogAff architecture (Sloman, 1999). The comprehensive LIDA model includes a broad array of cognitive modules and processes, a database of which, including known possible neural correlates can be found online at <ccrg.cs.memphis.edu/tutorial/correlates.html>.

---

[3] At this writing the LIDA model is only partially implemented. We claim it as a computational model since each of its modules and most of its processes have been designed for implementation.

[4] Gibson (1979) introduced the term *affordance*, which is often interpreted to mean that information about the available uses of an object is evoked by the object itself. We are using it in the sense that the agent can derive such information from the object.

The LIDA Cognitive Cycle

The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. Every "autonomous agent" (Franklin & Graesser, 1997), human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). Sophisticated agents, such as humans, process (make sense of) the input from such sampling in order to facilitate their decision making. Neuroscientists call this three-part process the action-perception cycle (Freeman, 2002). The agent's "waking life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle constitutes a unit of sensing, processing and acting. A cognitive cycle can be thought of as a cognitive "moment." Higher-level cognitive processes are composed of many of these cognitive cycles, so that each one can be considered to be a cognitive "atom."

Just as atoms have inner structure, the LIDA model hypothesizes a rich inner structure for its cognitive cycles (Baars & Franklin, 2003, Franklin, et al, 2005). During each cognitive cycle the LIDA agent first makes sense of (see below) its current situation as best as it can by updating its representation of both external and internal features of its world. By a competitive process to be described below, it then decides what portion of the represented situation is most in need of attention. This portion is broadcast, making it the current content of consciousness, and enabling the agent to choose an appropriate action and execute it.

Figure 1 shows the process in more detail. It starts in the upper left corner and
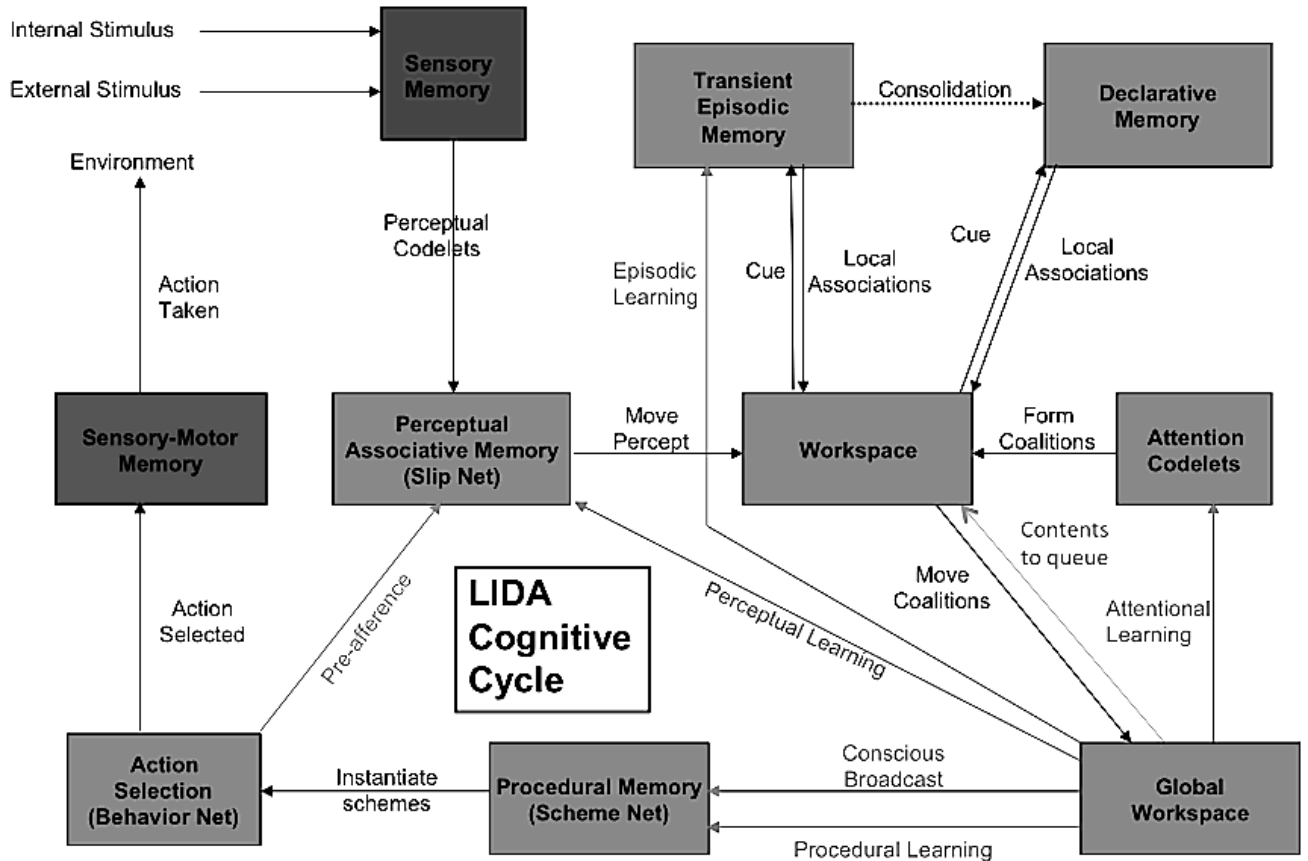
proceeds roughly clockwise.

Figure I: LIDA Cognitive Cycle Diagram. The boxes represent modules, not
processes, and it is the processes that are preconscious or n-conscious. Some
modules involve both kinds of processes.

The cycle begins with sensory stimuli from external and internal sources in the

agent's environment. Low-level feature detectors in sensory memory begin the

process of making sense of the incoming stimuli. These low-level features are

passed on to perceptual associative memory where higher-level features, such

as objects, categories, relations, events, situations, etc. are recognized. These entities, which have been recognized preconsciously, make up the percept that passes to the workspace, where a model of the agent's current situation is assembled. This percept serves as a cue to two forms of episodic memory, transient and declarative. Responses to the cue consist of local associations, that is, remembered events from these two memory systems that were associated with the various elements of the cue. In addition to the current percept, the workspace contains recent percepts and the models assembled from them that haven't yet decayed away.

A new model of the agent's current situation is assembled from the percepts, its associations, and the undecayed elements of the preceding model. This assembling process will typically require structure-building codelets[5]. These are small, special purpose processors, each of which has some particular type of structure it is designed to build. To fulfil their task these codelets may draw upon perceptual and sensory memory, to enable the recognition of relations and situations. The newly assembled model constitutes the agent's understanding of its current situation within its world. It has made sense of the incoming stimuli. All the contents of this workspace, including its current situational model, are preconscious, and may be selected to become conscious as we shall now see.

For an agent operating within a complex, dynamically changing environment, this current model may well be much too much for the agent to consider all at once in

---

[5] The term codelet refers generally to any small, special purpose processor or running piece of computer code.

deciding what to do next. It needs to selectively attend to a portion of the model. Portions of the model therefore compete for attention. They take the form of coalitions of structures obtained from the current situational model. Such coalitions are formed by attention codelets, whose function is to bring certain structures to consciousness. Once one of the coalitions wins the competition, in effect, the agent has decided what to attend to --- i.e., what to make conscious via the global workspace.

The purpose of this processing stage is to help the agent decide what to do next. To this end, a representation of the contents of the winning coalition is broadcast globally from the global workspace (hence the name Global Workspace Theory). Though the contents of this conscious broadcast are available globally, a primary recipient is procedural memory, which stores templates of possible actions including their contexts and possible results. Procedural memory also stores an activation value for each template, which attempts to measure the likelihood of an action taken within its context producing the expected result. Templates whose contexts intersect sufficiently with the contents of the conscious broadcast instantiate copies of themselves with their variables specified to the current situation. Instantiated templates remaining from previous cycles may also continue to be available. These instantiations are passed to the action selection mechanism, which chooses a single action from one of these instantiations. The chosen action then goes to sensory-motor memory, from which it is executed by an appropriate algorithm. The action taken affects the environment, and the cycle is complete.

The LIDA model hypothesizes that all human cognitive processing is via a continuing iteration of such cognitive cycles. The unconscious elements of these cycles are proposed to occur asynchronously, with each cognitive cycle taking roughly 100-200 ms. These cycles cascade, that is, several cycles may have different processes running simultaneously in parallel. This cascading must, however respect the serial nature of conscious processes that are necessary to maintain the stable, coherent image of the world (Franklin, 2005; Merker, 2005). The cascading cycles, which partially overlap, allows a rate of cycling in humans of five to ten cycles per second. A cognitive "moment" is thus quite short! There is considerable evidence from cognitive psychology and neuroscience that is consistent with such cognitive cycling in humans (Massimini, et al., 2005; Sigman & Dehaene, 2006; Uchida, et al, 2006; Willis & Todorov, 2006).

Some of the strongest evidence comes from the study of alpha and theta rhythms, which are involved in visual immediate memory, and in hippocampal-neocortical integration of new episodic and semantic memories. Alpha and theta cycle in the 10-Hz neighbourhood and Steriade has indeed suggested that they are not categorically different. The same set of neurons may generate the slightly different pace of alpha and theta bands.

So far these questions are not yet settled, but the evidence appears consistent with our hypotheses.

Preconscious and N-conscious Processes

Figure 1 illustrates the various cognitive modules and processes as hypothesized by the LIDA model, and the relationships between them. Beginning at the upper left, Sensory Memory consists of the initial representations of sensory input together with the products of the primary level feature detectors (Koch, 2004). Some of the earlier contents of Sensory Memory may be N-conscious, in that they can never become conscious. The contents of Sensory Memory are sent to Sensory-Motor Memory to enable the execution of the action selected during each cognitive cycle. In the visual system, for example, the sensory-motor component of these contents travel along the dorsal stream (Goodale & Milner, 2004)

Several subsequent modules in the figure, Perceptual Associative Memory, the Workspace, and the two episodic memories, are concerned with making sense of the current situation. Any of the contents of their processes may be included in the winning coalition in the global workspace. Thus, all these contents are preconscious by the definitions given above.

The workings of the attention codelets and the global workspace (competition for consciousness and the conscious broadcast) never become conscious and are thus N-conscious. The same is true of the workings and the contents of Procedural Memory, Action Selection and Sensor-Motor Memory. All are N-conscious.

The LIDA/GWT model suggests that the functional distinction between N-conscious and preconscious process derives from one of the major roles of an

agent's consciousness mechanism, to select the most salient portion of the agent's current situation to which to attend during each cognitive cycle. Only those processes whose contents might be part of that salient portion may be preconscious. Other critically important processes during any cognitive cycle serve many quite different functions, while their contents, if any, need never come to consciousness. For example, the action selection process during a single cognitive cycle is consciously mediated in that it requires momentary use of the current conscious contents. However the mechanism by which it selects the next action is entirely N-consciousness.

All higher-level mental activities, being multi-cyclic, must employ N-conscious, preconscious, and conscious processes during each of their cognitive cycles. During volition or conscious decision making, for example, a possible action will "pop into mind," that is, become conscious, perhaps in a fleeting visual image of the intended action outcome. This conscious thought is the result of both N-conscious and preconscious process operating during the cognitive cycle that produced the conscious event (Franklin, 2000).

Whether an unconscious cognitive process is N-conscious or preconscious depends entirely on its function. Does it carry content that might become part of a conscious broadcast? If so, it is preconscious by definition; if not, it is N-conscious.

Summary and Conclusions.

We have briefly sketched the GWT and LIDA models, and how they support information processing in autonomous agents, such as humans and animals. Consciousness is believed to involve momentary broadcasts of contents from one region or process in the brain, to be received by many others. Such broadcast events are believed to occur over approximately 100-200 ms bursts, consistent with our knowledge of alpha and theta activity. Distinctive contents of each burst may be multiplexed or modulated by faster gamma activity, which is often taken to mean 60-120 Hz. But "conscious bursts" are themselves embedded in cognitive cycles, which are unconscious. These cycles may be divided into preconscious elements, which can potentially provide conscious contents, and N-conscious, or "Never-conscious" elements, which are crucial to activating, and directing attention codelets to specific contents, but which never become conscious in themselves. There is a great deal of evidence that the brain supports both kinds of currently unconscious events.

The GWT-LIDA model therefore suggests there are two different kinds of unconscious events that occur quite fast and efficiently in any given cognitive cycles. The unconscious elements of such cycles may overlap in time, but the conscious ones face standard limited-capacity constraints and are therefore forced to occur serially.

References

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

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

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

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47–89). New York: Academic Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

Block, N. (2007). Consciousness, Accessibility and the Mesh between Psychology and Neuroscience. *Behavioral and Brain Sciences, 30*, 481-548.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*, 204–211.

Destexhe, Alain, Stuart W. Hughes, Michelle Rudolph and Vincenzo Crunelli, Are corticothalamic 'up' states fragments of wakefulness? Trends in Neurosciences, 2007 Jul;30(7):334-42.

Edelman, Gerald M., & Tononi, G. (2000). *A Universe of Consciousness*. New York: Basic Books.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Franklin, S. (2000). Deliberation and Voluntary Action in 'Conscious' Software

Agents. *Neural Network World, 10*, 505–521

Franklin, S. (2005). Evolutionary Pressures and a Stable World for Animals and

Robots: A Commentary on Merker. *Consciousness and Cognition, 14*,

115–118.

Franklin, S., Baars, B. J., Ramamurthy, U., & Ventura, M. (2005). The Role of

Consciousness in Memory. *Brains, Minds and Media, 1*, 1–38, pdf.

Franklin, S., & Graesser, A. C. (1997). Is it an Agent, or just a Program?: A

Taxonomy for Autonomous Agents *Intelligent Agents III* (pp. 21–35).

Berlin: Springer Verlag.

Franklin, S., & Patterson, F. G. Jr. (2006). The LIDA Architecture: Adding New

Modes of Learning to an Intelligent, Autonomous, Software Agent *IDPT-*

2006 Proceedings (Integrated Design and Process Technology): Society

for Design and Process Science.

Franklin, S., Ramamurthy, U., D'Mello, S. K., McCauley, L., Negatu, A., L., R.

S., et al. (2007, November 9-11). *LIDA: A Computational model of*

global workspace theory and developmental learning. Paper presented

at the AAAI Fall Symposium on AI and Consciousness: Theoretical

Foundations and Current Approaches, Arlington, VA.

Freeman, W. J. (2002). The limbic action-perception cycle controlling goal-

directed animal behavior. *Neural Networks, 3*, 2249-2254.

Freeman, W. J. (2003). Evidence from human scalp electroencephalograms of

global chaotic itinerancy. *Chaos, 13*, 1067–1077.

Gibson, James J. (1979). *The Ecological Approach to Visual Perception*.

    Mahwah, New Jersey: Lawrence Erlbaum Associates.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*,

    1–19.Goodale, Melvyn A., & Milner, D. (2004). *Sight Unseen*. Oxford:

    Oxford University Press.

Koch, C. (2004). *The Quest for Consciousness: A neurobiological approach*.

    Englewood, Colorado Roberts & Co.

Massimini, M., Ferrarelli, F., Huber, R., Esser, Steve K., Singh, H., & Tononi, G.

    (2005). Breakdown of Cortical Effective Connectivity During Sleep.

    Science, 309, 2228–2232.

Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition

    to consciousness in animal evolution. *Consciousness and Cognition, 14*,

    89–114.

Schneider, W. (2009). Automaticity and Consciousness. In W. Banks (Ed.),

    Encyclopedia of Consciousnes (pp. 83-92). St. Louis: Elsevier

Sigman, M., & Dehaene, S. (2006). Dynamics of the Central Bottleneck: Dual-

    Task and Task Uncertainty. *PLoS Biol., 4*.

Sloman, A. (1999). What Sort of Architecture is Required for a Human-like

    Agent? In M. Wooldridge & A. S. Rao (Eds.), *Foundations of Rational*

    Agency (pp. 35–52). Dordrecht, Netherlands: Kluwer Academic

    Publishers.

Uchida, N., Kepecs, A., & Mainen, Zachary F. (2006). Seeing at a glance,

    smelling in a whiff: rapid forms of perceptual decision making. *Nature*

    Reviews Neuroscience, 7, 485–491.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*.

    Cambridge, MA: MIT Press.

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a

    100-Ms Exposure to a Face. *Psychological Science, 17*, 592–599.