

DISCUSSION PAPER

For the ADVANCED SEMINAR in
Consciousness: The WebCourse 2008.

October 4, 2008

*Please do not cite or quote, except in the context of the
Advanced Seminar.*

Bernard J. Baars. Copyright, 2008.

An architecture for the conscious brain:

An Episodic Global Workspace hypothesis.

(Note; Familiarity & novelty, match-mismatch with NC;
two aspects of the same system)

Baars, Ramsøy, Franklin, Cariani. (tentative)

“... the function of the hippocampus (and possibly that
of related ... structures) is to help encode, retain, and

retrieve *experiences* ... the evidence favours a *multiple trace theory* (MTT) of memory...”

---- (Moscovitch et al, 2005). (Italics added)

The medial temporal lobe serves to encode conscious experiences into unconscious memory traces, which are widely distributed throughout the neocortex (Nadel & Moscovitch, ; Eichenbaum; Squire). But the MTL is also a 'convergence zone' for sensory streams, notably the dorsal and ventral streams of the visual cortex. REF We propose that MTL regions may support the emergence of specifically *conscious* percepts based on such convergence. Once these sensory streams are combined in a fast-rising coalition of spatiotopic neurons, they are widely “broadcast” to unconscious neuronal assemblies in both cortical and subcortical organs. The direction of broadcasting may be both frontoparietal, as suggested by evidence found by Dehaene and others, but also posterior, via cortico-cortical and cortico-thalamo-cortical connections. (Dehaene, 2001; Edelman & Tononi et al). One can imagine this event as a standing wave in a liquid breaking down into a fast-diffusing packet of distributed impulses. (Freeman).

In this view, the MTL-neocortical system transforms conscious sensory events into widespread adaptive processes throughout the brain. Those widespread adaptive changes may be called “declarative memories,” but they

could equally well be called “adaptive *updates* on attended events.” (Squire, recent) Longterm memory traces that can be recalled on cue constitute one kind of adaptive update in the brain, but not the only one. Spatial orientation and navigation may be seen as another kind of adaptive updating; this is of course a well-known aspect of MTL functioning. But motivational “updating” may well be yet another aspect of the same system, using conscious sensory input to influence the decision-making roles of the amygdala, limbic system, hypothalamus and frontal lobes. The MTL-NC system is also likely to interact seamlessly with sensorimotor control, endogenous mentation --- such as visual imagery and inner speech --- attentional control, alertness levels and working memory. Even the cerebellum is known to interact with the MTL (Gluck etc 2001)

One may therefore view the MTL conventionally as a declarative memory system, *or* from the perspective of sensory consciousness, triggering massive adaptive activity throughout the brain.

1.xxx The Episodic Global Workspace hypothesis.

(Introduce GWT.

THREE-WAY FIGURE)

Figure 1.0a A three-way relationship between (a) the

episodic or “experiential” memory system of the brain (Moscovitch et al, 2005), (b) the major features of conscious events when compared to similar unconscious ones, and (c) Global Workspace Theory (Baars, 1983, 1988, etc.).

These are also functional features of consciousness as proposed by Global Workspace Theory (GWT) (Baars, 1983, 1988, 2003; Dehaene, 2001; etc; Cariani,). GWT is compatible with other proposals for the brain basis of conscious experience, with the added advantages of simplicity, compatibility with a large body of behavioral and brain evidence, and, as pointed out, intriguing similarities to the known features of the episodic (“experiential”) memory system of the medial temporal lobe (Baars, 2003).

GW architectures have well-established functional properties that resemble the functions attributed to consciousness. (Baars, 1988; 1997, 2003; Baars & Franklin, xxxx) GW systems have been modeled by way of neural networks and artificial intelligence applications (Newman et al; Shanahan & Baars, 200x). They are also compatible with dynamical models of the physiology of the cortex. (Freeman).

This paper will focus on perceptual consciousness in relation to the traditional episodic memory system (Squire recent). The most relevant evidence in this respect is the close experimental comparison between matched conscious

and unconscious streams of information. These are commonly studied today using methods like binocular rivalry, perceptual ambiguities, and visual backward masking.

Figure 1 shows a useful way of thinking about Global Workspace Theory in terms of a *theater metaphor*. Even though GWT was derived from explicit artificial intelligence models developed by Alan Newell and colleagues, and has a long track record in AI applications, robotics, and neural net models, the theater metaphor remains useful today as a way to summarize and visualize its basic properties.

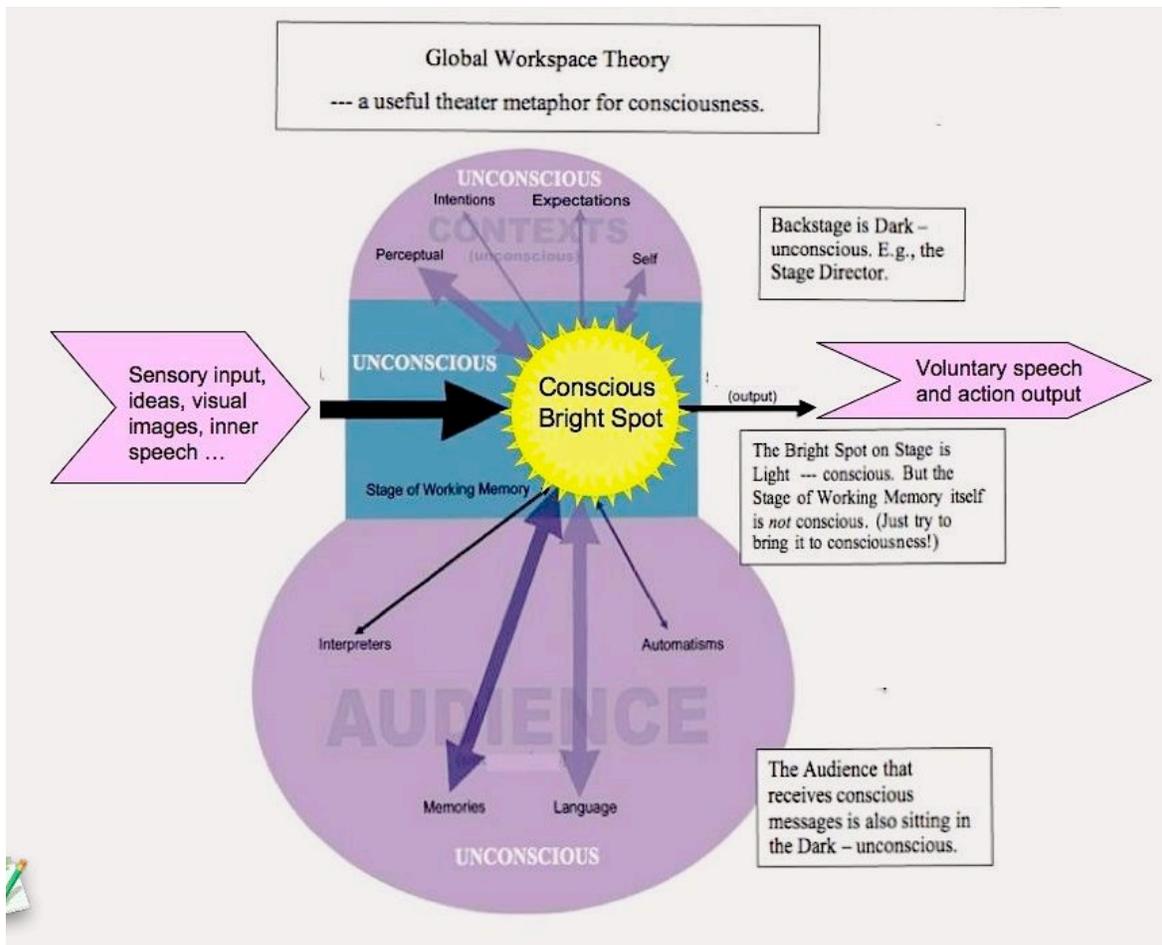


Figure 1. Notice that the conscious “bright spot on stage” is limited to only one content at a time, while many unconscious threads are running, both “backstage” as unconscious contexts and as receiving processors in the unconscious “audience.” Selective attention (not shown) may be analogized to a controllable spotlight.

In the simplest way, a GW architecture can be thought of as a “traffic hub” for information flow in the brain. On the input side of the hub, multiple resonant assemblies of

neurons compete or cooperate with each other for access. Once a the winning coalition of contents gain access they are widely distributed to receiving populations of resonant assemblies, the functional networks of the cortex and its satellites.

Figure 2 (below) shows one way in which this theoretical metaphor might be realized in the brain. In this case the sensory areas are proposed to be the source of global broadcasting. Alternatively, following Dehaene and colleagues, frontal regions might be such a global source. A multi-hub interpretation of global workspace theory suggests that these are not incompatible. Different cortical sources might “broadcast” their contents at different moments in time, running near 10 Hz. (Baars, 1988) For example, an absorbing visual movie, like a car race, might involve broadcasting from visual cortex to frontoparietal and subcortical regions. An intense “tip-of-the-tongue” word search might begin prefrontally, and then spread posteriorly to recruit sensorimotor regions for auditory word form, visual planning, or perceptual search.

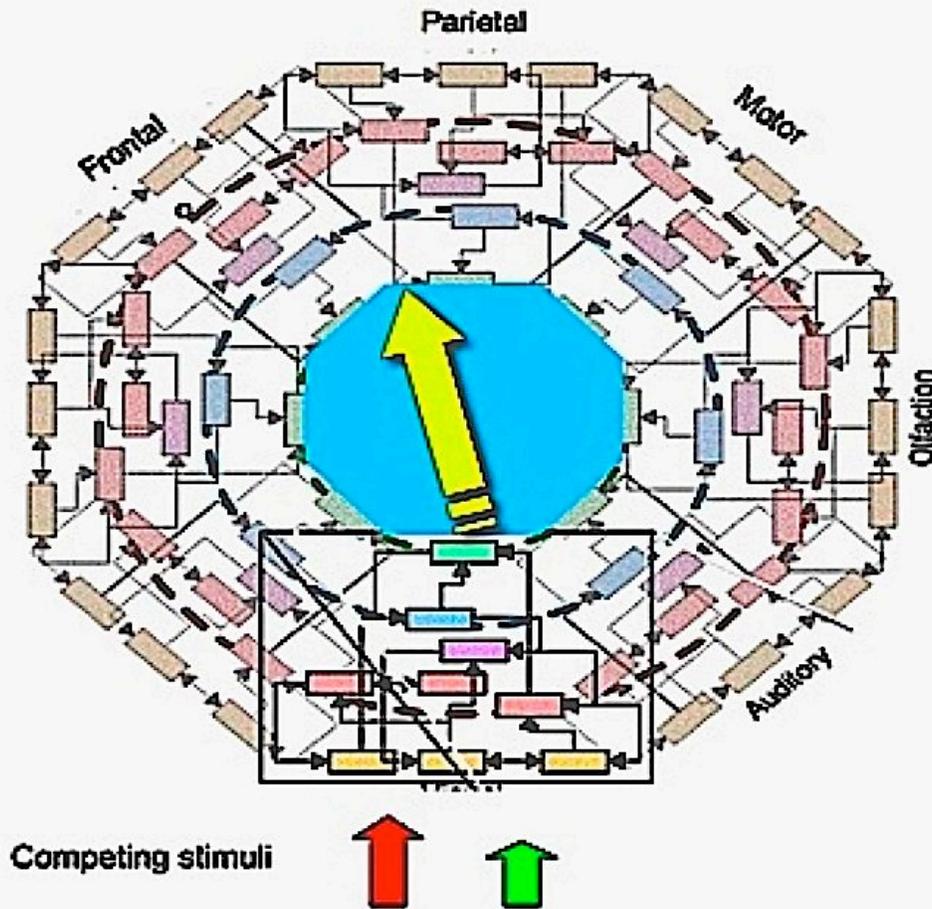


Figure 2. Caption: Notice competing stimuli. Contrastive cases like binocular rivalry, one of our foremost sources of insight into conscious vision (compared to matched unconscious input streams). (Based on Friston, ref, with permission).

Note that there are many hub-like “convergence zones” in the brain, and not all are associated with consciousness; for example... etc. The claim in Global Workspace Theory is that there is likely to be some brain basis for a global workspace cognitive architecture associated with

consciousness (Baars, 1983, 1988, 2003, etc). It is the specific instance(s) involving consciousness that we are seeking to understand. Reportability; contrastive.

1. xxx [Context vs. contents of conscious events.](#)

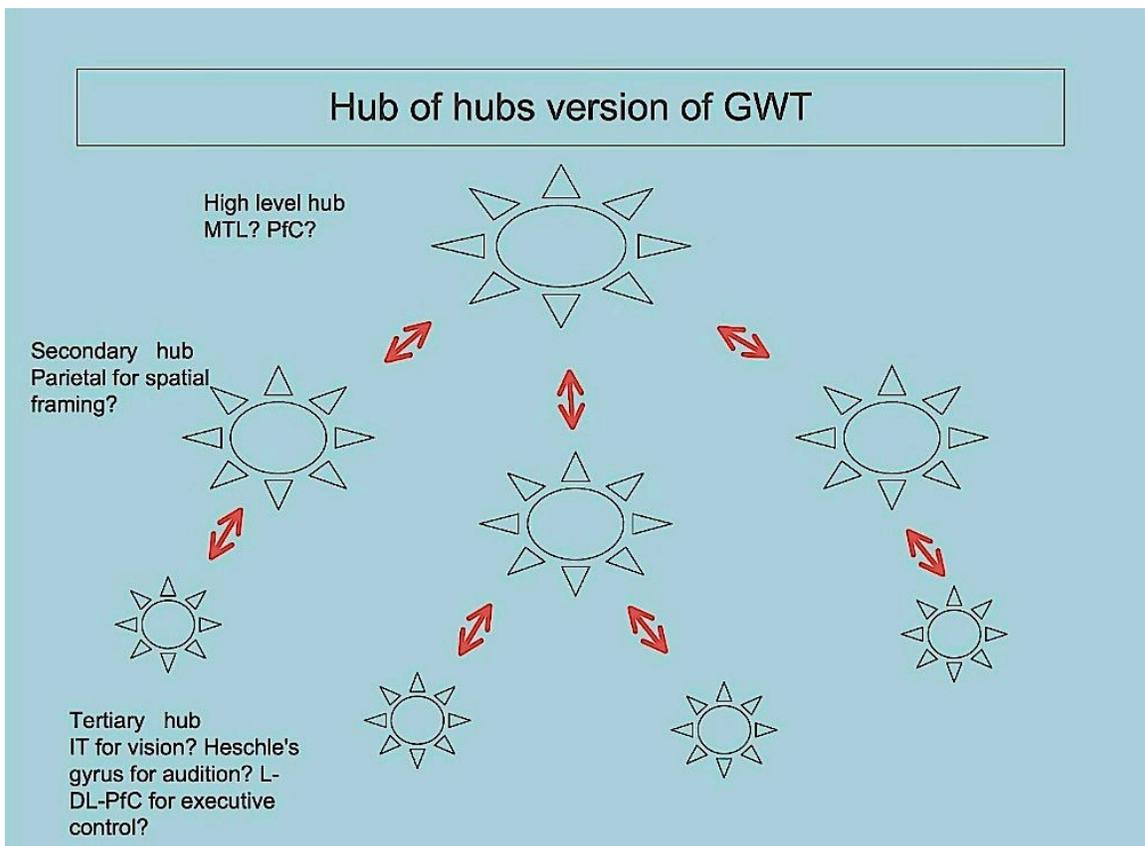
[Reportability. 23,000 articles.](#)

GWT makes a major distinction between conscious (reportable) *contents*, and a large set of unconscious *contexts*, brain events that shape and evoke conscious events without becoming conscious. A well-known example is the dorsal visual stream, which provides spatial frames for visual objects, but which is not believed to give rise to conscious experiences by itself (Goodale & Milner, xxx). Recent research suggests that the dorsal (contextual) visual stream converges with the ventral visual content stream in the MTL, as we might expect from the current hypothesis. (See below)

As a rough approximation THEATER or MOVIE. Hub proposals --- GW most studied functionally. FIGURE FROM Shanahan.

1.xx Small worlds: The brain has multiple connectivity hubs.

Figure 3 shows one conception (etc). Metaphorically, passing around a megaphone in a group of children. Different conscious contents might involve ‘broadcasting’ from different regions of cortex, or even subcortical areas like the dorsomedial thalamus. FOR EXAMPLE... Edelman and Gally () have pointed out the universality of functional redundancy (degeneracy) in biological systems. One of the attractive features of this general approach is that it supports great flexibility in accomodating different neuronal traffic flow in the core brain (Edelman & Tononi, 200x; Friston & Price; Tononi, xxx).



The brain has multiple connectivity hubs, each enabling convergent-to-divergent flow of information (Damasio; Llinas; Edelman, 1989; more recent findings). Some hubs appear to be associated with conscious events, as we will suggest. Others, however, may not; the mere existence of hub-like wiring does not *necessitate* a role in consciousness. FOR EXAMPLE... Dorsal stream. Even though it's cortex. What needs to be demonstrated, therefore, is both an empirical association with consciousness (as assessed by reportable brain events) and also a "convergent integration and divergent distribution" of the conscious events that are associated with the brain structure in question.

These considerations makes this a highly constrained empirical question. (Baars, 1988; 2003).

LIST THE COMING ARGUMENTS.

In addition: Contextual vs. content.

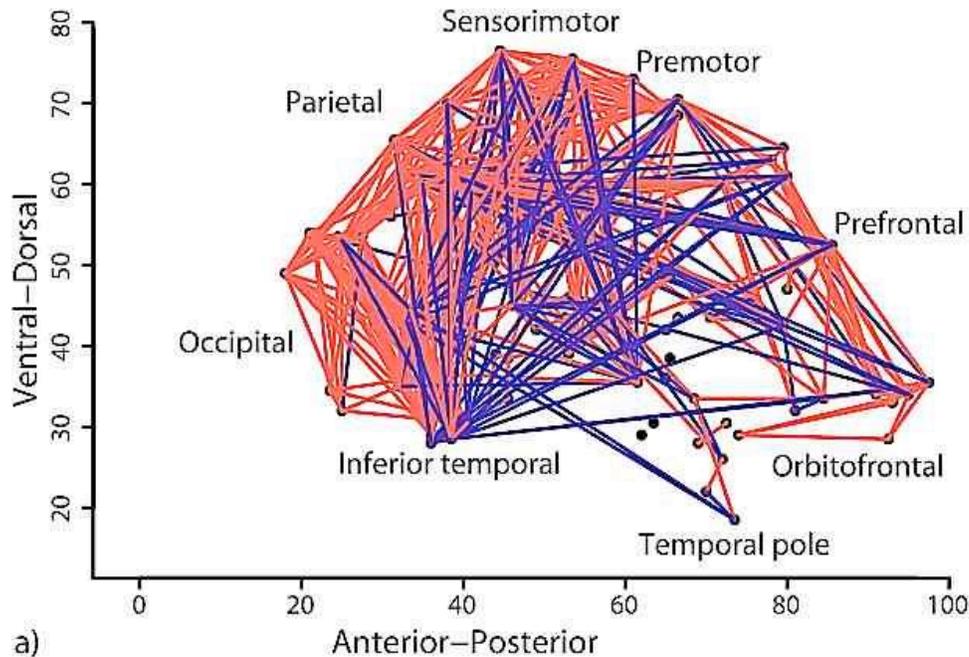


Figure 5. An anatomically more plausible “hub of hub” connectivity diagram. (Achard et al, 2006). Keep in mind that the mere presence of hub connectivity does not by itself show a role in the brain basis of consciousness.

Mention that Singer article that synchrony could also support a hub-like flow of information. It seems likely that both anatomical connectivity, modulated by synaptic chemistry, and controlled moment-to-moment by oscillatory activity may underlie the “hub-like” information flow of conscious contents.

1.xxx A hub of hubs is also a hub.

Such a multi-hub arrangement may operate as a distributed global workspace (GW) architecture. (Baars, Franklin, Shanahan, Cariani, Dehaene). Several brain regions have been proposed for conscious functions from this viewpoint, including thalamic and frontoparietal regions (e.g., Baars, 1988; Newman et al, 1997; Dehaene et al, 2001). The medial temporal lobe (MTL) may be another plausible substrate for conscious sensory experiences, in combination with sensory neocortex. The well-known large-scale projections from MTL to neocortex may be viewed as the "broadcasting" output for this very large system.

Recent studies of the MTL may be moving in this direction. For example, an intracranial recording study of single hippocampal neurons in human patients suggests that hippocampal neurons that were found to fire for both encoding and episodic recall constitute "an internally generated neuronal correlate of the subjective experience of spontaneous emergence of human recollection." (Hagar Gelbard-Sagiv, Malach, Fried, SCIENCE 2008). BETTER QUOTE? INPUT as well as RECALL.

1. xxx A focus on conscious perception.

We focus here on the sensory conscious input for the episodic memory system. However, a similar case can be made for other types of conscious contents, like reportable images, inner speech, concepts, and even vaguely conscious feelings of familiarity and knowing.

MTL is believed to encode conscious episodes into widely distributed unconscious neocortical (NC) traces, a view that resembles the 'global broadcasting' hypothesis of global workspace theory.

Recent articles suggests that MTL may have a number of additional functions, including perceptual ones (REFS). In addition, MTL is well known to encode an animal's spatial location and navigation, and to reflect the distinction between novel vs. familiar aspects of one's surroundings. Peripersonal space, navigation, and novelty are also closely related to immediate conscious experiences, certainly in humans, and arguably in other mammals as well. (REFS) Thus a number of known and plausible functions of MTL relate to immediate conscious experiences.

Historically, scientists have doubted whether MTL enables conscious perception because the most famous human lesion cases appear to have unimpaired conscious experiences. HM, etc. also Clive Wearing playing piano, conducting chorus, etc. However, recent findings have cast

doubt on the standard interpretation of these important case histories. (***) FOR READERS NOT FAMILIAR WITH HM LIT, CLARIFY THIS) GIVE QUOTES FROM:

GIVE QUOTES

Global workspace theory and the conscious brain.

GWT suggests that conscious input is widely distributed via MTL-neocortical connections, and that this 'broadcast' evokes many adaptive brain events beyond memory storage alone. For convenience we label this the "Epi-GW" or *Episodic Global Workspace* hypothesis. In global workspace theory the contents of consciousness result from convergent inputs, which are unified by suppressing incompatible inputs in any 100 ms duty cycle, or by integrating competing inputs into a coherent whole if they can be reconciled in content. Thus a GW architecture can be naturally divided into input processes *prior to* convergent equilibrium, and divergent output or 'broadcasting' after that moment. Recent evidence suggests a "moment of ignition" once a conscious input is specified, followed by accelerated reentrant broadcasting (Del Cul et al, 2007; Revonsuo et al, 200x).

GW input always involves multiple *contextual* systems, which provide the perceptual and semantic "frameworks" within which conscious *contents* are defined. For example, egocentric and allocentric contexts are required to specify

any conscious object in peripersonal space. These contextual systems appear to be enabled by the dorsal stream of the cortex, and according to recent research, by a corresponding dorsal division of the MTL. Once these input streams reach equilibrium they specify a conscious content, which is widely distributed (divergently) to a collection of unconscious specialized networks, such as the unconscious functional networks of the cortex.

The current hypothesis suggests that initial input to Epi-GW may come from the inferior temporal (IT) cortex, which is believed to support sensory feature integration of conscious objects. Competition for single conscious gestalts may be resolved in this region. The entorhinal region of MTL may combine simple perceptual gestalts into more complex scenes, which are mapped into closely related hippocampal activities. Once these input streams come to equilibrium in any single processing cycle of about 100-300 ms, they provide a coherent input to the hippocampus, which may then transmit multiplexed time, space, and sensory codes for additional unconscious processing in the neocortex. Thus IT-Ento may support the input phase of a GW brain function, while Hippo-NC may enable its widely distributed output.

BUT look at Grunwald et al article.

Theta rhythms are believed to coordinate both encoding

and recall of conscious memories via the IT-MTL-NC system. Theta is observed during waking and REM dreaming, especially in tasks requiring high alertness. This suggests that IT-MTL-NC may underlie perceptual conscious experiences and their widespread sequelae. While this may not be the sole neural hub for conscious events, Epi-GW may have been one of the earliest to emerge in vertebrate evolution.

leave semantic gw until later? it doesn't necessarily run through the mtl.

occipitotemporal hub. Perceiving, encoding and retrieving conscious episodes.

SEND TO C&C? Medical Hypotheses? *** Cognition.

Later papers in the series:

II. The parietal hub. Contextual integration in peripersonal space. (I Kahn reports that the dorsal-ventral division is reflected in MTL).

III. The frontal hubs. Executive planning and control, goal competition and Working Memory.

IV. The midline hubs. Executive input to the observing self.

V. A flexible flow of control between brain hubs.

Here is a rough summary of our emerging "MTL-neocortex/global workspace" paper, which seemed to fit your thinking so well in our conversations.

The first thing that captured me was Morris Moscovitch's & Lynn Nadel's Multiple Trace Theory of MTL-NCortical coding, which appears to fit hand-in-glove with GWT. Morris M authored the memory chapter for the cognitive neuroscience text that Nicole Gage and I edited. I had to clean it up and rewrite it pretty deeply, which had the positive result of forcing me to think about the memory literature.

Then Logothetis' work on IT and visual feature integration in the monkey seems to fit so well for the INPUT side of a GWT account of visual consciousness. But IT-MTL is also a multimodal integration area. Then I emailed Howard Eichenbaum, and he clarified the IT-entorhinal relationship, which is roughly "simple, known conscious percepts" vs. "novel and/or associated conscious percepts." Semantic memory comes in as an emergent from episodic memory, according to Nadel-Moscovitch, and semantics equals concepts. A perceptual symbol system approach to semantics also fits this view very nicely.

The great difficulty is HM and other hippocampal lesion patients, who are still clearly conscious. However, it may turn out that there are difficulties with their ability to perform novel integration of conscious percepts, which have perhaps not been diagnosed accurately for HM. If you move the input component of the GWT to the inferotemporal-entorhinal regions, and the episodic encoding to other MTL regions, and the output of a GW architecture to MTL-neocortical multiple-trace coding, essentially you've mapped GWT onto what is called the episodic memory system.

(--- Which is of course a backwards way of talking about conscious experiences.)

What I don't have at this point is something Peter Cariani seem to have a good grasp of: How to look at the hippocampus in terms of timing nets, which have widely distributed (neocortical?) interactions, including memory trace-formation, novelty detection (by mismatch with known information), and memory updating. Those are all functions of conscious experiences, to my way of thinking. So all that makes a lot of sense.

Other notes:

(Both "epi" and "declarative" is used for conscious).

Start with megaphone analogy. Mention small world

networks, which have hubs & highways. Interaction of self hubs with object hubs (see CTC).

--- Could also do interaction with executive interpreter. Two papers make sense. Focus on IT-MT-NC as one.

IN 1988, suggested this as one possibility; quote; current brain evidence seems to support it. SHOW FIGURE. sensory cortex; DL-PfC; precuneus; oscillatory version; dynamical models. Here we focus on one possible hub that is known to be involved in conscious functions, namely the episodic memory system of the brain.

IT-entorhinal viewed as a global workspace in the brain.

Episodic memory encodes the flow of conscious events for longterm storage and retrieval. Both episodic encoding and recall therefore involve conscious experiences, as subjects routinely tell us. The inferior temporal (IT) and medial temporal (MT) regions are involved in perceptual ambiguity resolution, feature integration and memory encoding of conscious events, while longterm episodic and semantic storage itself appears to involve neocortex (NC). These features fit theoretical predictions made by Baars and others regarding a brain-based "global workspace architecture" that supports conscious experience (1983, 1988, 2004). In this paper we explore these parallels. We predict novel features of the IT-MT-NC system, which may be one substrate of conscious experience in the brain.

5 MAJOR CLAIMS AT END OF CTC:

Tie in with recent paper on precuneus?

In 19xx the Logothetis lab discovered that binocular rivalry in the macaque, using multiple unit recordings.

Look at Ency summary of episodic memory.

deal with the HM objection --- not just MTL. Logothetis.

deal with the unconscious processing objection