Version 9.1

Darwin's rainbow: Evolutionary radiation and the spectrum of consciousness

Rodrick Wallace, Ph.D.
The New York State Psychiatric Institute
Robert G. Wallace, Ph.D.
Department of Ecology and Evolutionary Biology
University of California, Irvine *

September 20, 2006

Abstract

Evolution is littered with paraphyletic convergences: many roads lead to functional Romes. We propose here another example – an equivalence class structure factoring the broad realm of possible mathematical realizations of the Baars Global Workspace consciousness model. The construction suggests many different physiological systems can support rapidly shifting, sometimes highly tunable, temporary assemblages of interacting unconscious cognitive modules. The discovery implies various animal taxa exhibiting behaviors we broadly recognize as conscious are, in fact, simply expressing different forms of the same underlying phenomenon. Mathematically, we find much slower, and even multiple simultaneous, versions of the basic structure can operate over very long timescales, a kind of paraconsciousness often ascribed to group phenomena. The variety of possibilities, a veritable rainbow, suggests minds today may be only a small surviving fraction of ancient evolutionary radiations - bush phylogenies of consciousness and paraconsciousness. Under this scenario, the resulting diversity was subsequently pruned by selection and chance extinction. Though few traces of the radiation may be found in the direct fossil record, exaptations and vestiges are scattered across the living mind. Humans, for instance, display an uncommonly profound synergism between individual consciousness and their embedding cultural heritages, enabling efficient Lamarkian adaptation.

Key Words animal consciousness, cognition, culture, global workspace, homotopy, information theory, paraconsciousness, phase transition, renormalization, topology.

Introduction

High order mental function has long been suspected in animal species other than man. The *American Zoologist* (vol. 40, no. 6, 2000), for example, published the proceedings of a wide-ranging symposium on animal consciousness held by the Society for Integrative and Comparative Biology in January of 1999. From that volume, Cartmill (2000) quotes Darwin (1889),

> "...The difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind. We have seen that the senses and intuitions, the various emotions, attention, curiosity, imitation, reason etc., of which man boasts, may be found in an incipient, or even sometimes in a well-developed condition, in the lower animals..."

In the same volume Griffin (2000) writes:

> "A striking aspect of this symposium is that almost all the participants take it for granted that some animals *are* conscious – not all animals all the time, but at least some animals on some occasions... Instead of the resistance to considering such questions that used to be customary among behavioral scientists, there was stimulating and constructive discussion of the *content* of animal consciousness...
>
> This change in climate of opinion does not make the analysis of animal consciousness easy. All the difficulties and uncertainties which led behaviorists to ban the subject from scientific psychology remain in place. But difficulties are not impossibilities... Insofar as animals are conscious the content of their conscious experiences probably differs in many ways from human consciousness, so that investigating it requires more than merely inquiring whether particular kinds of human awareness occur in other species..."

Indeed, Griffin (1976, 1992) and Griffin and Speck (2004) review a vast range of ethological observations suggestive of consciousness in many animal species.

More recently, animal sentience was the topic of a full issue of *Applied Animal Behavior Science*, (Vol. 10, nos. 1-2, October, 2006), in which Pepperberg writes

---

*Address correspondence to Rodrick Wallace, PISCS Inc., 549 W. 123 St., Suite 16F, New York, NY, 10027, USA. Telephone (212) 865-4766, email rdwall@ix.netcom.com. Affiliations are for identification only.

"...[D]ata... demonstrate that many species of animal posses cognitive capacities that we have until now considered unique to humans and other primates."

Simmonds (2006), in that same volume, states "...[V]arious complex behaviors and social structures ...support the notion that cetaceans should be regarded as intelligent animals."

A parallel perspective has even emerged within the academic consciousness industry, appearing in such journals as *Consciousness and Cognition* and at meetings of the Association for the Scientific Study of Consciousness. For example Seth et al. (2005) explore criteria for consciousness in mammals, and Edelman et al. (2005) discuss possible hallmarks of consciousness in non-mammalian species, particularly birds and cephalopods.

All in all, rapidly-operating consciousness appears to be an ancient, and highly effective, evolutionary adaptation. As Edelman et al. (2005) put the matter

"It is plausible that complex brains capable of rich and flexible behavioral repertoires began evolving in two very different radiations between 530 and 540 million years ago... (Morris, 2000; Knoll and Carroll, 1999)."

Edelman et al. (2005) also point out, however, that mammals, birds, and cephalopods have very different neural anatomies. While an analog to the reentrant thalamocortical structure which seems to underlie human/mammal consciousness has been identified in birds, the cephalopod brain remains an enigma, in spite of striking ethological observations suggestive of consciousness.

Evolutionary psychologists have long speculated on the nature of mind. Heyes (2003), for example, writes

"The task of evolutionary psychology is not to show that natural selection can influence cognitive processes but to establish exactly what kind of effects natural selection, and developmental selection, do and do not tend to have. Just as natural selection tends to be conservative with respect to respiratory pigments (e.g. hemoglobin) and revisionist with respect to respiratory structures (e.g. skin, gills, and lungs), it is likely that some properties of behavior-control systems are more susceptible than others to phylogenetic change and therefore that they show greater variation across species and in the course of development."

Recently Wallace (2005a, b; 2006a) presented several different Dretske-style 'necessary conditions' communication theory formulations of Bernard Baars' Global Workspace model of consciousness (Baars, 1988, 2005), the first among equals in the current Darwinian competition between perspectives on human consciousness (e.g. Dehaene and Naccache, 2001). The limitation to necessary conditions implies that the Wallace approach, which specifically involves analytically tractable asymptotic limits in a much larger continuum of possible treatments, is formally analogous to developing empirical regression models, based, however, on the Rate Distortion and Shannon-McMillan Theorems rather than on the Central Limit Theorem.

The obvious inference is that evolutionary process, unlike an applied mathematician, will, in Hayes sense, over a few tens or hundreds of millions of years, find many explicit solutions to complicated intermediate cases. The question then remains regarding what may be conserved and what may vary.

Since the Wallace argument is based on equivalence classes of mathematical models, Pielou's (1977, p. 106) caution is worth repeating in this context:

"...[Mathematical models] are easy to devise; even though the assumptions of which they are constructed may be hard to justify, the magic phrase 'let us assume that...' overrides objections temporarily. One is then confronted with a much harder task: How is such a model to be tested? The correspondence between a model's predictions and observed events is sometimes gratifyingly close but this cannot be taken to imply the model's simplifying assumptions are reasonable in the sense that neglected complications are indeed negligible in their effects...

In my opinion the usefulness of models is great... [however] it consists *not in answering questions but in raising them.* Models can be used to inspire new field investigations and these are the only real source of new knowledge as opposed to new speculation."

This paper outlines how a broad spectrum of markedly different structures can instantiate Baars' Global Workspace model of consciousness. Of particular interest is the possibility of slowly-acting global workspace assemblies of unconscious cognitive modules, even of simultaneous multitasking, in both conscious and other organisms, and their social assemblies. The question then arises, of course, as to just what evolution has made of this vast realm of possibilities.

Formal Theory

1. The basic Global Workspace model

Bernard Baars' Global Workspace Theory (Baars, 1988, 2005) is rapidly becoming the de facto standard model of human consciousness (e.g. Dehaene and Naccache, 2001; Dehaene and Changeaux, 2005). The central ideas are as follows (Baars and Franklin, 2003):

(1) The brain can be viewed as a collection of distributed specialized networks (processors).

(2) Consciousness is associated with a global workspace in the brain – a fleeting memory capacity whose focal contents are widely distributed (broadcast) to many unconscious specialized networks.

(3) Conversely, a global workspace can also serve to integrate many competing and cooperating input networks.

(4) Some unconscious networks, called contexts, shape conscious contents, for example unconscious parietal maps modulate visual feature cells that underlie the perception of color in the ventral stream.

(5) Such contexts work together jointly to constrain conscious events.

(6) Motives and emotions can be viewed as goal contexts.

(7) Executive functions work as hierarchies of goal contexts.

Although this basic approach has been the focus of many researchers for nearly two decades, academic consciousness studies have only recently, under the relentless pressure of a deluge of empirical results from brain imaging experiments, begun digesting the perspective and preparing to move on.

Currently popular agent-based and artificial neural network (ANN) treatments of cognition, consciousness and other higher order mental functions, taking Krebs' (2005) view, are little more than sufficiency arguments, in the same sense that a Fourier series expansion can be empirically fitted to nearly any function over a fixed interval without providing real understanding of the underlying structure. Necessary conditions, as Dretske argues (Dretske, 1981, 1988, 1993, 1994), give considerably more insight. Perhaps the most cogent example is the difference between the Ptolemaic and Copernican models of the solar system: one need not always expand in epicycles, but can seek the central motion. Dretske's perspective provides such centrality. Keplerian and Newtonian treatments, unfortunately, still lie ahead of us: Atmanspacher (2006) has likened the current state of consciousness theory to that of physics four hundred years ago.

Wallace (2005a) has addressed Baars' theme from Dretske's viewpoint, examining the necessary conditions which the asymptotic limit theorems of information theory impose on the Global Workspace. A central outcome of this work has been the incorporation, in a natural manner, of constraints on individual consciousness, i.e. what Baars calls contexts. Information theory methods, extended by the obvious homology between information source uncertainty and the free energy density of a physical system, make it possible to formally account for the effects on individual consciousness of parallel physiological modules like the immune system, embedding structures like the local social network, and, most importantly, the all-encompassing cultural heritage which so uniquely marks human biology (e.g. Richerson and Boyd, 2004). This embedding evades the mereological fallacy which fatally bedevils brain-only theories of human consciousness (Bennett and Hacker, 2003).

Transfer of phase change approaches from statistical physics to information theory via the same homology generates the punctuated nature of accession to consciousness in a similarly natural manner. The necessary renormalization calculation focuses on a phase transition driven by variation in the average strength of nondisjunctive weak ties (Granovetter, 1973) linking unconscious cognitive submodules. A second-order universality class tuning allows for adaptation of conscious attention via rate distortion manifolds which generalize the idea of a retina. A version of the Baars model (including contexts) emerges as an almost exact parallel to hierarchical

regression, based, however, on the Shannon-McMillan rather than the Central Limit Theorem.

Wallace (2005b; 2006a) recently proposed a somewhat different approach, extended here, which uses classic results from random and semirandom network theory (Erdos and Renyi, 1960; Albert and Barabasi, 2002; Newman, 2003) applied to a modular network of cognitive processors. The unconscious modular network structure of the brain is, of course, not random. However, in the spirit of the wag who said "all mathematical models are wrong, but some are useful", the method serves as the foundation of a different, but roughly parallel, treatment of the Global Workspace to that given in Wallace (2005a), and hence as another basis for a benchmark model against which empirical data can be compared. In fact it will become clear that there can be intermediated classes of models as well as those acting at very slow rates (Wallace, 2006b), suggesting an equivalence class structure for Global Workspace models, as well as a classification by underlying operating rate, matters leading to the central focus of this work.

The first step is to argue for the existence of a network of loosely linked unconscious cognitive modules, and to characterize each of them by the richness of the canonical language – information source – associated with it. This is in some contrast to attempts to explicitly model neural structures themselves using network theory, e.g. the neuropercolation approach of Kozma et al. (2004, 2005), which nonetheless uses many similar mathematical techniques. Here, rather, the central focus is on the necessary conditions imposed by the asymptotic limits of information theory upon any realization of a cognitive process. A certain class of cognitive processes, in this formulation, is to be associated with a canonical dual information source which will be constrained by the Rate Distortion Theorem, or, in the zero-error limit, the Shannon-McMillan Theorem. It is interactions between nodes in this abstractly defined network which will be of interest here, rather than whatever mechanism or biological system, or mixture of them, actually constitute the underlying cognitive modules. Thus various, very different, biological or other structures can, in this formulation, all instantiate functionally similar Global Workspace models.

This proves to be a central point.

The second step, in some contrast to Wallace (2005a), is to examine the conditions under which a giant component (GC) suddenly emerges as a kind of phase transition in a network of such linked cognitive modules, to determine how large that component is, and to define the relation between the size of the component and the richness of the cognitive language associated with it. This level of approximation subsumes both Baars' 'fleeting memory capacity' which acts as an analog to Newell's blackboard computing model, and the specialized modules which have been recruited by broadcast, into a single object, and is one way to produce the large-scale brain connectivity which is the sine qua non of consciousness, in conformance with a large and growing body of brain imaging studies (e.g. Wallace, 2005b; 2006a).

Again, at least logically, many different possible substrates

could conceivably produce similar functional modules.

Implicit, then, is the possibility of a number of different mechanisms which achieve such large-scale structure. Wallace (2005a), for example, explores phase transitions centering around an inverse temperature analog involving the average strength of weak ties between modules. Intermediate models are possible. The giant component approach, however, seems a particularly simple example, which is elaborated here. Empirical comparisons of consciousness between different animal taxa, for example insects, fish, reptiles, birds, and mammals, would likely be particularly illuminating, as different fundamental mechanisms constituting and linking cognitive modules may have evolved in each.

The third step, following Wallace (2005b; 2006a), is to use renormalization parameters to tune the threshold at which the giant component comes into being, along with its topological structure, via an iteration involving a tunable rate distortion manifold retina-analog. This generalizes Newell's blackboard model to give a highly flexible version of Baars' 'fleeting memory capacity' across many possible underlying neural topologies.

Wallace (2005a), by contrast, uses 'universality class tuning' to direct the phase transitions associated with changing the average strength of weak ties between modules.

These are, as said, clearly different analytically tractable asymptotic limits in a much larger domain of possible modeling approaches.

Although both kinds of linkage are sufficient to produce large-scale brain connectivity, a basic kind of consciousness which may be characteristic of many animal taxa, another iteration seems required to produce higher mental function. Some second level models may be more amenable to such iteration than others, again a likely matter of empirical study across animal families.

The information theoretic modular network treatment can be enriched by introducing a groupoid formalism which is roughly similar to recent analyses of linked dynamic networks described by differential equations (e.g. Stewart et al., 2003, Stewart, 2004; Weinstein, 1996; Connes, 1994). Internal and external linkages between information sources break the underlying groupoid symmetry, and introduce more structure, the global workspace and the effect of contexts, respectively. The analysis provides a foundation for further mathematical exploration of linked cognitive processes.

2. Cognition as an information source

Cognition is not consciousness. Most mental, and many physiological, functions, while cognitive in a formal sense, hardly ever become entrained into the Global Workspace of consciousness: one seldom is able to consciously regulate immune function, blood pressure, or the details of binocular tracking and bipedal motion, except to decide 'what shall I look at', 'where shall I walk'. Nonetheless, many cognitive processes, conscious or unconscious, appear intimately related to language, broadly speaking. The construction is fairly straightforward (Wallace, 2000, 2005a, b).

Atlan and Cohen (1998) and Cohen (2000) argue, in the context of immune cognition, that the essence of cognitive function involves comparison of a perceived signal with an internal, learned picture of the world, and then, upon that comparison, choice of one response from a much larger repertoire of possible responses.

Such choice inherently and fundamentally involves information and information transmission since it always generates a reduction in uncertainty (e.g. Ash, 1990, p. 21).

Cognitive pattern recognition-and-response proceeds by an algorithmic combination of an incoming external sensory signal with an internal ongoing activity – incorporating the learned picture of the world – and triggering an appropriate action based on a decision that the pattern of sensory activity requires a response.

More formally, a pattern of sensory input is mixed in an unspecified but systematic algorithmic manner with a pattern of internal ongoing activity to create a path of combined signals $x = (a_0, a_1, ..., a_n, ...)$. Each a_k thus represents some functional composition of internal and external signals. Wallace (2005a) provides two neural network examples.

This path is fed into a highly nonlinear, but otherwise similarly unspecified, decision oscillator, h, which generates an output $h(x)$ that is an element of one of two disjoint sets B_0 and B_1 of possible system responses. Let

$$B_0 \equiv b_0, ..., b_k,$$

$$B_1 \equiv b_{k+1}, ..., b_m.$$

Assume a graded response, supposing that if

$$h(x) \in B_0,$$

the pattern is not recognized, and if

$$h(x) \in B_1,$$

the pattern is recognized, and some action $b_j, k+1 \leq j \leq m$ takes place.

The principal objects of formal interest are paths x which trigger pattern recognition-and-response. That is, given a fixed initial state a_0, examine all possible subsequent paths x beginning with a_0 and leading to the event $h(x) \in B_1$. Thus $h(a_0, ..., a_j) \in B_0$ for all $0 < j < m$, but $h(a_0, ..., a_m) \in B_1$.

For each positive integer n, let $N(n)$ be the number of high probability grammatical and syntactical paths of length n which begin with some particular a_0 and lead to the condition $h(x) \in B_1$. Call such paths 'meaningful', assuming, not unreasonably, that $N(n)$ will be considerably less than the number of all possible paths of length n leading from a_0 to the condition $h(x) \in B_1$.

While the combining algorithm, the form of the nonlinear oscillator, and the details of grammar and syntax, are all unspecified in this model, the critical assumption which permits inference on necessary conditions constrained by the asymptotic limit theorems of information theory is that the finite limit

$$H \equiv \lim_{n \to \infty} \frac{\log[N(n)]}{n}$$

(1)

both exists and is independent of the path x.

Define such a pattern recognition-and-response cognitive process as *ergodic*. Not all cognitive processes are likely to be ergodic, implying that H, if it indeed exists at all, is path dependent, although extension to nearly ergodic processes, in a certain sense, seems possible (Wallace, 2005a).

Invoking the spirit of the Shannon-McMillan Theorem, since choice involves an inherent reduction in uncertainty, it is then possible to define an adiabatically, piecewise stationary, ergodic (APSE) information source $\mathbf{X}$ associated with stochastic variates X_j having joint and conditional probabilities $P(a_0, ..., a_n)$ and $P(a_n|a_0, ..., a_{n-1})$ such that appropriate joint and conditional Shannon uncertainties satisfy the classic relations

$$H[\mathbf{X}] = \lim_{n \to \infty} \frac{\log[N(n)]}{n} =$$

$$\lim_{n \to \infty} H(X_n|X_0, ..., X_{n-1}) =$$

$$\lim_{n \to \infty} \frac{H(X_0, ..., X_n)}{n}.$$

This information source is defined as *dual* to the underlying ergodic cognitive process (Wallace, 2005a).

Recall that the Shannon uncertainties $H(...)$ are cross-sectional law-of-large-numbers sums of the form $-\sum_k P_k \log[P_k]$, where the P_k constitute a probability distribution. See Khinchin (1957), Ash (1990), or Cover and Thomas (1991) for the standard details.

3. The cognitive modular network symmetry groupoid

A formal equivalence class algebra can be constructed by choosing different origin points a_0 and defining equivalence by the existence of a high probability meaningful path connecting two points. Disjoint partition by equivalence class, analogous to orbit equivalence classes for dynamical systems, defines the vertices of the proposed network of cognitive dual languages. Each vertex then represents a different information source dual to a cognitive process. This is not a representation of a neural network as such. It is, rather, an abstract set of 'languages' dual to the cognitive processes instantiated by various forms of biological wetware, mechanical dryware, or their direct or systems-level hybrids. This is a functional classification.

The structure is a groupoid, in the sense of Weinstein (1996). States a_j, a_k in a set A are related by the groupoid morphism if and only if there exists a high probability grammatical path connecting them, and tuning across the various possible ways in which that can happen – the different cognitive languages – parametizes the set of equivalence relations and creates the groupoid. This assertion requires some development.

Note that not all possible pairs of states (a_j, a_k) can be connected by such a morphism, i.e. by a high probability, grammatical and syntactical cognitive path, but those that can define the groupoid element, a morphism $g = (a_j, a_k)$ having the natural inverse $g^{-1} = (a_k, a_j)$. Given such a pairing, connection by a meaningful path, it is possible to define 'natural' end-point maps $\alpha(g) = a_j, \beta(g) = a_k$ from the set of morphisms G into A, and a formally associative product in the groupoid $g_1 g_2$ provided $\alpha(g_1 g_2) = \alpha(g_1), \beta(g_1 g_2) = \beta(g_2)$, and $\beta(g_1) = \alpha(g_2)$. Then the product is defined, and associative, i.e. $(g_1 g_2)g_3 = g_1(g_2 g_3)$.

In addition there are natural left and right identity elements λ_g, ρ_g such that $\lambda_g g = g = g \rho_g$ whose characterization is left as an exercise (Weinstein, 1996).

An orbit of the groupoid G over A is an equivalence class for the relation $a_j \sim Ga_k$ if and only if there is a groupoid element g with $\alpha(g) = a_j$ and $\beta(g) = a_k$.

The isotopy group of $a \in X$ consists of those g in G with $\alpha(g) = a = \beta(g)$.

In essence a groupoid is a category in which all morphisms have an inverse, here defined in terms of connection by a meaningful path of an information source dual to a cognitive process.

If G is any groupoid over A, the map $(\alpha, \beta) : G \to A \times A$ is a morphism from G to the pair groupoid of A. The image of (α, β) is the orbit equivalence relation $\sim G$, and the functional kernel is the union of the isotropy groups. If $f : X \to Y$ is a function, then the kernel of f, $ker(f) = [(x_1, x_2) \in X \times X : f(x_1) = f(x_2)]$ defines an equivalence relation.

As Weinstein (1996) points out, the morphism (α, β) suggests another way of looking at groupoids. A groupoid over A identifies not only which elements of A are equivalent to one another (isomorphic), but *it also parametizes the different ways (isomorphisms) in which two elements can be equivalent*, i.e. all possible information sources dual to some cognitive process. Given the information theoretic characterization of cognition presented above, this produces a full modular cognitive network in a highly natural manner.

To anticipate slightly, for individual consciousness, empirical study implies that only a single, rapidly-operating, giant component seems possible. The groupoid argument, however, allows parametization across the possible ways in which a state may belong to several different information sources dual to different cognitive processes. That is, at least in theory, there can be several, simultaneous, multitasking global workspaces. The simplest way for this to occur is to operate the system at a very slow rate, so that no individual cognitive module is forced to work faster than really can, in terms of Rate Distortion Theorem limitations on information sent between workspaces. This suggests that consciousness itself can be embedded in slower-acting global workspace phenomena, or that, for some organisms, the global workspace itself, or an assemblage of them, may operate very slowly. That is, tem-

poral segregation forms another way of creating equivalence classes of states connected by meaningful paths.

The groupoid approach has become quite popular in the study of networks of coupled dynamical systems which can be defined by differential equation models, e.g. Stewart et al. (2003), Stewart (2004). This work extends the technique to networks of interacting information sources which, in a dual sense, characterize cognitive processes, and cannot at all be described by the usual differential equation models. These latter, it seems, are much the spiritual offspring of 18th Century mechanical clock models. Animal cognitive and conscious processes involve neither computers nor clocks, but remain constrained by the limit theorems of information theory, and these permit scientific inference on necessary conditions, in the sense of Dretske.

Note that the groupoid formalism can also be extended to equivalence classes of the paths generated by individual modules. This will be shown to produce a directed homotopy groupoid classification of paths rather than of states defining topological properties of the underlying manifold. Different animal families and temporal realms can be expected to express different homotopy groupoids.

Wallace (2006b) formally extends the model in such a way that several global workspaces may exist simultaneously, as the individual cognitive modules multitask. This is unlikely to be normal behavior, however, on the few hundred millisecond timescale which characterizes individual animal consciousness.

4. Internal forces breaking the symmetry groupoid

The symmetry groupoid, as constructed for unconscious cognitive submodules in a kind of information space, is parametized across that space by the possible ways in which states a_j, a_k can be equivalent, i.e. connected by a meaningful path of an information source dual to a cognitive process. These are different, and in this approximation, non-interacting unconscious cognitive processes.

But symmetry groupoids, like symmetry groups, are made to be broken: by internal cross-talk akin to spin-orbit interactions within a symmetric atom, and by cross-talk with slower, external, information sources, akin to placing a symmetric atom in a powerful magnetic or electric field.

As to the first process, suppose that linkages can fleetingly occur between the ordinarily disjoint cognitive modules defined by the network groupoid. In the spirit of Wallace (2005a), this is represented by establishment of a non-zero mutual information measure between them: a cross-talk which breaks the strict groupoid symmetry developed above.

Wallace (2005a) describes this structure in terms of fixed magnitude disjunctive strong ties which give the equivalence class partitioning of modules, and nondisjunctive weak ties which link modules across the partition, and parametizes the overall structure by the average strength of the weak ties, to use Granovetter's (1973) term. By contrast the approach of Wallace (2005b), extended here as an example, is to simply look at the average number of fixed-strength nondisjunctive links in a random topology. These are, again, obviously two analytically tractable limits of a much more complicated regime.

Since nothing is known about how the cross-talk connections can occur, at first assume they are random and construct a random graph in the classic Erdos/Renyi manner. Suppose there are M disjoint cognitive modules – M elements of the equivalence class algebra of languages dual to some cognitive process – now taken to be the vertices of a possible graph.

For M very large, following Savante et al. (1993), when edges (defined by establishment of a fixed-strength mutual information measure between the graph vertices) are added at random to M initially disconnected vertices, a remarkable transition occurs when the number of edges becomes approximately $M/2$. Erdos and Renyi (1960) studied random graphs with M vertices and $(M/2)(1 + \mu)$ edges as $M \to \infty$, and discovered that such graphs almost surely have the following properties (Molloy and Reed, 1995, 1998; Grimmett and Stacey, 1998; Luczak, 1990; Aiello et al., 200; Albert and Barabasi, 2002):

[1] If $\mu < 0$, only small trees and unicyclic components are present, where a unicyclic component is a tree with one additional edge; moreover, the size of the largest tree component is $(\mu - \ln(1 + \mu))^{-1} + \mathcal{O}(\log \log n)$.

[2] If $\mu = 0$, however, the largest component has size of order $M^{2/3}$.

[3] If $\mu > 0$, there is a unique giant component (GC) whose size is of order M; in fact, the size of this component is asymptotically αM, where $\mu = -\alpha^{-1}[\ln(1 - \alpha) - 1]$, which has an explicit solution for α in terms of the Lambert W-function. Thus, for example, a random graph with approximately $M \ln(2)$ edges will have a giant component containing $\approx M/2$ vertices.

Such a phase transition initiates a new, collective, cognitive phenomenon, emergently defined by a set of cross-talk mutual information measures between interacting unconscious cognitive submodules. The source uncertainty, H, of the language dual to the collective cognitive process, which characterizes the richness of the cognitive language of the workspace, will grow as some monotonic function of the size of the GC, as more and more unconscious processes are incorporated into it. Wallace (2005b) provides details.

Others have taken similar network phase transition approaches to assemblies of neurons, e.g. neuropercolation (Kozma et al., 2004, 2005), but their work has not focused explicitly on modular networks of cognitive processes, which may or may not be instantiated by neurons or particular neural structures. Restricting analysis to such functional, as opposed to structural, modular networks finesses much of the underlying conceptual difficulty, and permits use of the asymptotic limit theorems of information theory and the import of techniques from statistical physics, a matter we will discuss later.

5. External forces breaking the symmetry groupoid

A higher order information source, associated with the GC of a random or semirandom graph, can be constructed from the interlinking of unconscious cognitive modules by mutual information. Similarly, external information sources, for example in humans the cognitive immune and other physiolog-

ical systems, and embedding sociocultural structures, can be represented as slower-acting information sources whose influence on the GC can be felt in a collective mutual information measure. This constitutes a structured environment, to be viewed as among Baars' contexts (Baars, 1988, 2005; Baars and Franklin, 2003). The collective mutual information measure will, through the Joint Asymptotic Equipartition Theorem which generalizes the Shannon-McMillan Theorem, be the splitting criterion for high and low probability joint paths across the entire system.

The tool for this is network information theory (Cover and Thomas, 1991, p. 388). Given three interacting information sources, Y_1, Y_2, Z, the splitting criterion, taking Z as the 'external context', is given by

$$I(Y_1, Y_2|Z) = H(Z) + H(Y_1|Z) + H(Y_2|Z) - H(Y_1, Y_2, Z),$$

(2)

where $H(..|..)$ and $H(.., ..., ..)$ represent conditional and joint uncertainties (Khinchin, 1957; Ash, 1990; Cover and Thomas, 1991).

This generalizes to

$$I(Y_1, ... Y_n|Z) = H(Z) + \sum_{j=1}^{n} H(Y_j|Z) - H(Y_1, ..., Y_n, Z).$$

(3)

If the Global Workspace/Giant Component involves a very rapidly shifting, and indeed highly tunable, dual information source X, embedding contextual cognitive modules like the immune system will have a set of significantly slower-responding sources $Y_j, j = 1..m$, and external social, cultural and other environmental processes will be characterized by even more slowly-acting sources $Z_k, k = 1..n$. Mathematical induction on equation (3) gives a complicated expression for a mutual information splitting criterion of the general form

$$I(X|Y_1, .., Y_m|Z_1, .., Z_n).$$

(4)

This encompasses a fully interpenetrating biopsychosocio-cultural structure for individual consciousness in humans, one in which Baars' contexts act as important, but slow and relatively flexible, boundary conditions, defining the underlying topology available to the far more rapidly shifting global workspace (Wallace, 2005a, b).

Such a structure does not run afoul of the mereological fallacy which Bennett and Hacker (2003) impute to excessively neurocentric perspectives on consciousness in humans, that is, the mistake of imputing to a part of a system the characteristics which require functional entirety. The underlying concept of this fallacy should extend to any cognitive structures interacting with their environments (Wallace 2006a).

Note that slowly-acting embedding physiological or other global workspaces, in the sense of Wallace (2006b), would constitute contexts in this sense.

6. Punctuation

As a number of researchers have noted – see Wallace, (2005a) for discussion – equation (1),

$$H \equiv \lim_{n \to \infty} \frac{\log[N(n)]}{n},$$

is homologous to the thermodynamic limit in the definition of the free energy density of a physical system. This has the form

$$F(K) = \lim_{V \to \infty} \frac{\log[Z(K)]}{V},$$

(5)

where F is the free energy density, K the inverse temperature, V the system volume, and $Z(K)$ is the partition function defined by the system Hamiltonian.

Wallace (2005a) shows at some length how this homology permits the natural transfer of renormalization methods from statistical mechanics to information theory. In the spirit of the Large Deviations Program of applied probability theory, this produces phase transitions and analogs to evolutionary punctuation in systems characterized by APSE information sources. These biological phase changes appear to be ubiquitous in natural systems and can be expected to dominate constructed systems as well. Wallace (2002) uses these arguments to explore the differences and similarities between evolutionary punctuation in genetic and learning plateaus in neural systems.

7. Topology I

The network development above is predicated on there being a variable average number of fixed-strength linkages between components. Clearly, the mutual information measure of cross-talk is not inherently fixed, but can continuously vary in magnitude. This can be addressed by a parametized renormalization. In essence the modular network structure linked by mutual information interactions has a topology depending on the degree of interaction of interest. Define an interaction parameter ω, a real positive number, and look at geometric

structures defined in terms of linkages which are zero if mutual information is less than, and 'renormalized' to unity if greater than, ω. A given ω will define a regime of giant components of network elements linked by mutual information greater than or equal to it.

The fundamental conceptual trick is to invert the argument: A given topology for the giant component will, in turn, define some critical value, ω_C, so that network elements interacting by mutual information less than that value will be unable to participate, i.e. will be locked out and not be consciously perceived. Thus ω is a tunable, syntactically-dependent, detection limit, and depends critically on the instantaneous topology of the giant component. That topology is, fundamentally, the basic tunable syntactic filter across the underlying modular symmetry groupoid, and variation in ω is only one aspect of a much more general topological shift. More detailed analysis is given below in terms of a topological rate distortion manifold.

Suppose the giant component at some 'time' k is characterized by a set of parameters $\Omega_k \equiv \omega_1^k, ..., \omega_m^k$. Fixed parameter values define a particular giant component having a particular topological structure (Wallace, 2005b). Suppose that, over a sequence of 'times' the giant component can be characterized by a (possibly coarse-grained) path $x_n = \Omega_0, \Omega_1, ..., \Omega_{n-1}$ having significant serial correlations which, in fact, permit definition of an adiabatically, piecewise stationary, ergodic (APSE) information source in the sense of Wallace (2005a). Call that information source $\mathbf{X}$.

Suppose, again in the manner of Wallace (2005a), that a set of (external or else internal, systemic) signals impinging on consciousness, i.e. the giant component, is also highly structured and forms another APSE information source $\mathbf{Y}$ which interacts not only with the system of interest globally, but specifically with the tuning parameters of the giant component characterized by $\mathbf{X}$. $\mathbf{Y}$ is necessarily associated with a set of paths y_n.

Pair the two sets of paths into a joint path $z_n \equiv (x_n, y_n)$, and invoke some inverse coupling parameter, K, between the information sources and their paths. By the arguments of Wallace (2005a) this leads to phase transition punctuation of $I[K]$, the mutual information between $\mathbf{X}$ and $\mathbf{Y}$, under either the Joint Asymptotic Equipartition Theorem, or, given a distortion measure, under the Rate Distortion Theorem.

$I[K]$ is a splitting criterion between high and low probability pairs of paths, and partakes of the homology with free energy density described in Wallace (2005a). Attentional focusing then itself becomes a punctuated event in response to increasing linkage between the organism or device and an external structured signal, or some particular system of internal events. This iterated argument parallels the extension of the General Linear Model into the Hierarchical Linear Model of regression theory.

Call this the Hierarchical Cognitive Model (HCM).

Note that the groupoid treatment above parametizes the ways in which states can be linked by meaningful paths, and, logically, permits a state to be a member of more than one simultaneous global workspace, i.e. multitasking. Slowly rate

systems are more likely to be able to multitask than faster ones – multiple conscious states are likely to be highly pathological. Pure slow global workspaces, that is, single or multitasking paraconsciousness, seems possible in this kind of system, as does the embedding of consciousness itself within slowly-acting paraconscious structures. Thus, as stated, rate-of-operation can become another equivalence class criterion for groupoid construction and splitting.

The HCM version of Baars' global workspace model stands in some contrast to other current work.

Tononi (2004), for example, takes a complexity perspective on human consciousness, in which he averages mutual information across all possible bipartitions of the thalamocortical system, and, essentially, demands an infomax clustering solution. Other clustering statistics, however, may serve as well or better, as in generating phylogenetic trees, and the method does not seem to produce conscious punctuation in any natural manner.

Dehaene and Changeux (2005) take an explicit Baars global workspace perspective on consciousness, but use an elaborate neural network simulation to generate a phenomenon analogous to inattentional blindness. While their model does indeed display the expected punctuated behaviors, as noted above, Krebs (2005) unsparingly labels such constructions with the phrase "neurological possibility does not imply neurological plausibility", suggesting that the method does little more than fit a kind of Fourier series construction to high level mental processes.

The approach here attempts a central motion model of consciousness, focusing on modular networks defined by function rather than by structure.

8. Cognitive quasi-thermodynamics

A fundamental homology between the information source uncertainty dual to a cognitive process and the free energy density of a physical system arises, in part, from the formal similarity between their definitions in the asymptotic limit. Information source uncertainty can be defined as in equation (1). This is quite analogous to the free energy density of a physical system, equation (5).

Feynman (1996) provides a series of physical examples, based on Bennett's work, where this homology is, in fact, an identity, at least for very simple systems. Bennett argues, in terms of irreducibly elementary computing machines, that the information contained in a message can be viewed as the work saved by not needing to recompute what has been transmitted.

Feynman explores in some detail Bennett's microscopic machine designed to extract useful work from a transmitted message. The essential argument is that computing, in any form, takes work, the more complicated a cognitive process, measured by its information source uncertainty, the greater its energy consumption, and the ability to provide energy to the brain is limited. Inattentional blindness, one can argue, emerges as an inevitable thermodynamic limit on processing capacity in a topologically-fixed global workspace, i.e. one which has been strongly configured about a particular task.

Understanding the time dynamics of cognitive systems

away from phase transition critical points requires a phenomenology similar to the Onsager relations of nonequilibrium thermodynamics. If the dual source uncertainty of a cognitive process is parametized by some vector of quantities $\mathbf{K} \equiv (K_1, ..., K_m)$, then, in analogy with nonequilibrium thermodynamics, gradients in the K_j of the *disorder*, defined as

$$S \equiv H(\mathbf{K}) - \sum_{j=1}^{m} K_j \partial H / \partial K_j$$

(6)

become of central interest.

Equation (6) is similar to the definition of entropy in terms of the free energy density of a physical system, as suggested by the homology between free energy density and information source uncertainty described above.

Pursuing the homology further, the generalized Onsager relations defining temporal dynamics become

$$dK_j / dt = \sum_i L_{j,i} \partial S / \partial K_i,$$

(7)

where the $L_{j,i}$ are, in first order, constants reflecting the nature of the underlying cognitive phenomena. The L-matrix is to be viewed empirically, in the same spirit as the slope and intercept of a regression model, and may have structure far different than familiar from more simple chemical or physical processes. The $\partial S / \partial K$ are analogous to thermodynamic forces in a chemical system, and may be subject to override by external physiological driving mechanisms (Wallace, 2005c).

Equations (6) and (7) can be derived in a simple parameter-free covariant manner which relies on the underlying topology of the information source space implicit to the development. Suppose that different physiological cognitive phenomena have, in the sense of Wallace (2000, 2005, Ch. 3), dual information sources. The question of interest involves the local properties of the system near a particular reference state. Impose a topology on the system, so that, near a particular 'language' A, dual to an underlying cognitive process, there is (in some appropriate sense) an open set U of closely similar languages $\hat{A}$, such that $A, \hat{A} \subset U$. It may be necessary to coarse-grain the physiological responses to define these information sources. The problem is to proceed in such a way as to preserve the underlying essential topology, while eliminating 'high frequency noise'. The formal tools for this can be found, e.g., in Chapter 8 of Burago et al. (2001).

Since the information sources dual to the cognitive processes are similar, for all pairs of languages $A, \hat{A}$ in U, it is possible to:

[1] Create an embedding alphabet which includes all symbols allowed to both of them.

[2] Define an information-theoretic distortion measure in that extended, joint alphabet between any high probability (i.e. grammatical and syntactical) paths in A and $\hat{A}$, written as $d(Ax, \hat{A}x)$ (Cover and Thomas, 1991). Note that these languages do not interact, in this approximation.

[3] Define a metric on U, for example,

$$\mathcal{M}(A, \hat{A}) = |\lim \frac{\int_{A,\hat{A}} d(Ax, \hat{A}x)}{\int_{A,A} d(Ax, A\hat{x})} - 1|,$$

(8)

using an appropriate integration limit argument over the high probability paths. Note that the integration in the denominator is over different paths within A itself, while in the numerator it is between different paths in A and $\hat{A}$.

Consideration suggests $\mathcal{M}$ is a formal metric, having $\mathcal{M}(A, B) \geq 0, \mathcal{M}(A, A) = 0, \mathcal{M}(A, B) = \mathcal{M}(B, A), \mathcal{M}(A, C) \leq \mathcal{M}(A, B) + \mathcal{M}(B, C)$.

Other approaches to constructing a metric on U may be possible.

Since H and $\mathcal{M}$ are both scalars, a 'covariant' derivative can be defined directly as

$$dH / d\mathcal{M} = \lim_{\hat{A} \to A} \frac{H(A) - H(\hat{A})}{\mathcal{M}(A, \hat{A})},$$

(9)

where $H(A)$ is the source uncertainty of language A.

Suppose the system to be set in some reference configuration A_0.

To obtain the unperturbed dynamics of that state, impose a Legendre transform using this derivative, defining another scalar

$$S \equiv H - \mathcal{M} dH / d\mathcal{M}.$$

(10)

The simplest possible Onsager relation here – again an empirical equation like a regression model – is just

$$d\mathcal{M}/dt = L\,dS/d\mathcal{M},$$

(11)

where t is the time and $dS/d\mathcal{M}$ represents an analog to the thermodynamic force in a chemical system. This is seen as acting on the reference state A_0. For

$$dS/d\mathcal{M}|_{A_0} = 0,$$

$$d^2S/d\mathcal{M}^2|_{A_0} > 0$$

(12)

the system is quasistable, a Black hole, if you will, and externally imposed physiological forcing mechanisms will be needed to effect a transition to a different state.

Conversely, changing the direction of the second condition, so that

$$dS^2/d\mathcal{M}^2|_{A_0} < 0,$$

leads to a repulsive peak, a White hole, representing a possibly unattainable realm of states.

Explicit parametization of $\mathcal{M}$ introduces standard – and quite considerable – notational complications (e.g. Burago et al., 2001; Auslander, 1967): Imposing a metric for different cognitive dual languages parametized by $\mathbf{K}$ leads to Riemannian, or even Finsler, geometries (Wallace, 2005c), including the usual geodesics.

One can apply this formalism to the example of the giant component, with the information source uncertainty/channel capacity taken as directly proportional to the component's size, which increases monotonically with the average number of (renormalized) linkages, a, after the critical point. $H(a)$ then rises to some asymptotic limit: the homology between information source uncertainty and free energy density suggests that raising the cognitive capacity of the giant component, making it larger, requires energy. Beyond a certain point, the system just runs out of steam. Altering the topology of the network, no longer focusing on a particular demanding task, would allow detection of cross-talk signals from other submodules, as would the intrusion of a signal above the renormalization limit ω.

The manner in which the system runs out of steam involves a maxed-out, fixed topology for the giant component of consciousness. As argued above, the renormalization parameter ω then becomes an information/energy bottleneck. To keep the giant component at optimum function in its particular topology, i.e. focused on a particular task involving a necessary set of interacting cognitive submodules, a relatively high limit must be placed on the magnitude of a mutual information signal which can intrude into consciousness.

Consciousness, according to this model, is tunable, and signals outside the chosen syntactical/grammatical bandpass are often simply not strong enough to be detected, broadly accounting for the phenomena of inattentional blindness.

The basic focus mechanism can be modeled in far more detail.

9. A simple rate distortion manifold

The iteration above – analogous to expanding the General Linear Model to the Hierarchical Linear Model – which involved paths in parameter space, can itself be significantly extended. This produces a generalized tunable retina model which can be interpreted as a 'Rate Distortion manifold', a concept which further opens the way for import of a vast array of tools from geometry and topology (e.g. Glazebrook, 2006).

Suppose, now, that threshold behavior in conscious reaction requires some elaborate system of nonlinear relationships defining a set of renormalization parameters $\Omega_k \equiv \omega_1^k, ..., \omega_m^k$. The critical assumption is that there is a tunable zero order state, and that changes about that state are, in first order, relatively small, although their effects on punctuated process may not be at all small. Thus, given an initial m-dimensional vector Ω_k, the parameter vector at time $k+1$, Ω_{k+1}, can, in first order, be written as

$$\Omega_{k+1} \approx \mathbf{R}_{k+1}\Omega_k,$$

(13)

where $\mathbf{R}_{t+1}$ is an $m \times m$ matrix, having m^2 components.

If the initial parameter vector at time $k = 0$ is Ω_0, then at time k

$$\Omega_k = \mathbf{R}_k \mathbf{R}_{k-1} ... \mathbf{R}_1 \Omega_0.$$

(14)

The interesting correlates of consciousness are, in this development, *now represented by an information-theoretic path defined by the sequence of operators* $\mathbf{R}_k$, each member having m^2 components. The grammar and syntax of the path defined by these operators is associated with a dual information source, in the usual manner.

The effect of an information source of external signals, $\mathbf{Y}$, is now seen in terms of more complex joint paths in Y and

R-space whose behavior is, again, governed by a mutual information splitting criterion according to the JAEPT.

The complex sequence in m^2-dimensional *R*-space has, by this construction, been projected down onto a parallel path, the smaller set of m-dimensional ω-parameter vectors $\Omega_0, ..., \Omega_k$.

If the punctuated tuning of consciousness is now characterized by a 'higher' dual information source – an embedding generalized language – so that the paths of the operators $\mathbf{R}_k$ are autocorrelated, then the autocorrelated paths in Ω_k represent output of a parallel information source which is, given Rate Distortion limitations, apparently a grossly simplified, and hence highly distorted, picture of the 'higher' conscious process represented by the *R*-operators, having m as opposed to $m \times m$ components.

High levels of distortion may not necessarily be the case for such a structure, *provided it is properly tuned to the incoming signal.* If it is inappropriately tuned, however, then distortion may be extraordinary.

Consider a single iteration in more detail, assuming now there is a (tunable) zero reference state, $\mathbf{R}_0$, for the sequence of operators $\mathbf{R}_k$, and that

$$\Omega_{k+1} = (\mathbf{R}_0 + \delta\mathbf{R}_{k+1})\Omega_k,$$

(15)

where $\delta\mathbf{R}_k$ is 'small' in some sense compared to $\mathbf{R}_0$.

Note that in this analysis the operators $\mathbf{R}_k$ are, implicitly, determined by linear regression. Thus it is possible to invoke a quasi-diagonalization in terms of $\mathbf{R}_0$. Let $\mathbf{Q}$ be the matrix of eigenvectors which Jordan-block-diagonalizes $\mathbf{R}_0$. Then

$$\mathbf{Q}\Omega_{k+1} = (\mathbf{Q}\mathbf{R}_0\mathbf{Q}^{-1} + \mathbf{Q}\delta\mathbf{R}_{k+1}\mathbf{Q}^{-1})\mathbf{Q}\Omega_k.$$

(16)

If $\mathbf{Q}\Omega_k$ is an eigenvector of $\mathbf{R}_0$, say Y_j with eigenvalue λ_j, it is possible to rewrite this equation as a generalized spectral expansion

$$Y_{k+1} = (\mathbf{J} + \delta\mathbf{J}_{k+1})Y_j \equiv \lambda_j Y_j + \delta Y_{k+1}$$

$$= \lambda_j Y_j + \sum_{i=1}^{n} a_i Y_i.$$

(17)

$\mathbf{J}$ is a block-diagonal matrix, $\delta\mathbf{J}_{k+1} \equiv \mathbf{Q}\mathbf{R}_{k+1}\mathbf{Q}^{-1}$, and δY_{k+1} *has been expanded in terms of a spectrum of the eigenvectors of* $\mathbf{R}_0$, with

$$|a_i| \ll |\lambda_j|, |a_{i+1}| \ll |a_i|.$$

(18)

The point is that, provided $\mathbf{R}_0$ has been tuned so that this condition is true, the first few terms in the spectrum of this iteration of the eigenstate will contain most of the essential information about $\delta\mathbf{R}_{k+1}$. This appears quite similar to the detection of color in the retina, where three overlapping non-orthogonal eigenmodes of response are sufficient to characterize a huge plethora of color sensation. Here, if such a tuned spectral expansion is possible, a very small number of observed eigenmodes would suffice to permit identification of a vast range of changes, so that the rate-distortion constraints become quite modest. That is, there will not be much distortion in the reduction from paths in *R*-space to paths in Ω-space. Inappropriate tuning, however, can produce very marked distortion, even inattentional blindness.

Reflection suggests that, if consciousness indeed has something like a grammatically and syntactically-tunable retina, then appropriately chosen observable correlates of consciousness may, at a particular time and under particular circumstances, actually provide very good local characterization of conscious process. Large-scale global processes are another matter, and inappropriate focus can lead to large errors in this analysis.

Note that Rate Distortion Manifolds can be quite formally described using standard techniques from topological manifold theory (Glazebrook, 2006; Bak et al., 2006). The essential point is that a rate distortion manifold is a topological structure which constrains the 'stream of consciousness' much the way a riverbank constrains the flow of the river it contains. This is a fundamental insight, which can be pursued further.

10. Topology II

The groupoid treatment of modular cognitive networks above defined equivalence classes of *states* according to whether they could be linked by grammatical/syntactical high probability 'meaningful' paths. Next is the precisely complementary question regarding *paths*: For any two particular given states, is there some sense in which it is possible to define equivalence classes across the set of meaningful paths linking them?

This is of particular interest to the iterated hierarchical model which involves tuning the renormalization parameters characterizing the dancing, flowing, punctuated accession to consciousness.

A closely similar question is central to recent algebraic geometry approaches to concurrent, i.e. highly parallel, computing (e.g. Pratt, 1991; Goubault and Raussen, 2002; Goubault, 2003), which we adapt.

For the moment restrict the analysis to a giant component system characterized by two renormalization parameters, say ω_1 and ω_2, and consider the set of meaningful paths connecting two particular points, say a and b, in the two dimensional ω-space plane of figure 1. The generalized quasi-Onsager arguments surrounding equations (6), (7) and (12) suggests that there may be regions of fatal attraction and strong repulsion, Black holes and White holes, which can either trap or deflect the path of consciousness.

Figures 1a and 1b show two possible configurations for a Black and a White hole, diagonal and cross-diagonal. If one requires path monotonicity – always increasing or remaining the same – then, following, e.g. Goubault (2003, figs. 6,7), there are, intuitively, two direct ways, without switchbacks, that one can get from a to b in the diagonal geometry of figure 1a, without crossing a Black or White hole, but there are three in the cross-diagonal structure of figure 1b.

Elements of each 'way' can be transformed into each other by continuous deformation without crossing either the Black or White hole. Figure 1a has two additional possible monotonic ways, involving over/under switchbacks, which are not drawn. Relaxing the monotonicity requirement generates a plethora of other possibilities, e.g. loopings and backwards switchbacks, whose consideration is left as an exercise. It is not clear under what circumstances such complex paths can be meaningful, a matter for further study.

These ways are the equivalence classes defining the topological structure of the two different ω-spaces, analogs to the fundamental homotopy groups in spaces which admit of loops (e.g. Lee, 2000). The closed loops needed for classical homotopy theory are impossible for this kind of system because of the 'flow of time' defining the output of an information source – one goes from a to b, although, for nonmonotonic paths, intermediate looping would seem possible. The theory is thus one of directed homotopy, dihomotopy, and the central question revolves around the continuous deformation of paths in ω-space into one another, without crossing Black or White holes. Goubault and Rausssen (2002) provide another introduction to the formalism.

These ideas can, of course, be applied to lower level cognitive modules as well as to the iterated hierarchical cognitive model where they are, perhaps, of more central interest.

Note that this analysis can be extended without difficulty to the formulation of Wallace (2005a), in terms of universality class tuning.

Discussion

We have shown here consciousness may come in a variety of proverbial colors.

Wallace (2005a) used a 'mean field' approximation for a GW consciousness model which focused on tuning the universality class constants associated with an average strength of weak ties linking unconscious cognitive modules. Different renormalization strategies gave rise to different retina-like rate distortion manifolds of tuned universality classes. The focus here has, by contrast, been on the number of weak ties linking cognitive modules which exceed a topologically-tuned threshold. These two limits each produce very large, and quite different, sets of regression-style statistical models, all of which instantiate the Global Workspace of consciousness. Obviously, intermediate structures are possible, including those which operate at markedly different rates, but these are harder to analyze. The evident conjecture, then, is that each essential variation in structure will have an associated model equivalence class, according to another groupoid likely related to another underlying topological structure (i.e. Topology III). To reiterate, another equivalence class structure would be defined by the rates at which the global workspaces operate, typically very fast vs. very slow, in which case the latter are taken as contexts for the former.

The simplest realm is that of a shifting giant component of strongly crosstalking unconscious modules linked in first order. The next step, according to this development, is the emergence of a tunable Rate Distortion Manifold retina analog directing attention. Average strength of weak ties, average number of weak ties, and intermediate conditions each would seem to have associated classes of possible models: topology upon topology.

If one views the basic phenomenon as arising from a shifting structure of crosstalking unconscious cognitive modules operating in the realm of a few hundred milliseconds, then, according to this development, evolution will have had a vast playing field of possibilities, all producing global workspace behaviors - consciousness. Simple systems might well be found in many insects, higher order iterations across a considerable range of animal species. One clear conjecture is that such iterations will almost always be susceptible to some form of inattentional blindness, as the syntactic/grammatical bandpass of a tuned rate distortion manifold is always limited.

Again, in Pielou's sense (1977, p. 106), the mathematical modeling exercise of this paper suggests that many different physiological structures can produce unconscious cognitive modules which may be linked by a great variety of mechanisms into various forms of shifting Global Workspace architectures producing conscious behaviors on the scale of a few hundred milliseconds, although iterated versions, since they involve focus, should suffer some form of inattentional blindness.

Further, it seems possible that slowly acting global workspace architectures linking unconscious cognitive modules might well be found in many living things, what might well be called paraconsciousness, although consciousness as-we-know-it is clearly limited to realms of a few hundred milliseconds. It may even be possible for conscious animals themselves to entertain, as contexts, other global workspace processes which operate on much slower timescales than consciousness itself. Certainly human institutions, in which we are all embedded, constitute structures which have multiple global workspaces, albeit operating over very long times (Wal-

lace, 2006b). Appropriate organismal colonies may provide some other examples.

A menagerie of possible global workspace models has implications for the mind's origins.

Evolutionary history is littered with eras of innovative flux. Multicellularity began among the Ediacara 650 million years ago and later in the Tommotian phase of the Cambrian Explosion, with a diversity of body plans unrivaled since. The Tetrapoda's late Devonian land invasion included more than one multiple-digit plan. *Ichthyostega, Acanthostega*, and *Tulerpeton* emerged from the muck with six, seven, even eight digits twiddling from each limb. Varieties of great apes far outnumbering present-day remnants brachiated through the Miocene. By industrializing *Gallonanserae* into huge cities of chickens we have today helped bring about a golden age of influenza evolution an exploding phylogeny exploring multiple hilltops across Sewell Wright's adaptive landscape. Only one or two branches may prove the next pandemic.

We hypothesize similar evolutionary radiations of types of mind, wherein bouts of divergent innovations in consciousness and paraconsciousness arose together. One can imagine multiple attempts at building limbic-like systems and, subsequently, cortexes, atop the reptilian brain. There are a variety of ways to interconnect the three. Perhaps there were out-and-out novel structures in addition. In this scenario, as in the above examples, the resulting evolutionary bushes are subsequently pruned by some combination of failure to adapt to changing circumstances and blind chance.

At first glance we are left with little recourse in excavating such mental fossils. The Cretaceous meteorite left no dinosaur minds behind. But brain endocasts may offer clues. We could also surmise that at some point the brain of the occasional hominid or mastodon found in ice might be preserved well enough for dissection or other analysis. Finally, living mammalian and other heads are likely filled with evolutionary artifacts and exaptations jury-rigged together by selection and historical accident (Gould, 1991). The epistemological criteria by which one can identify the mental equivalents of the human appendix or whale hip bone are available for experimental test (Skoyles, 1999).

Human consciousness, unlike that of most animals, is strongly, if not almost uniquely, synergistic with the rich dual heritage system of embedding culture (Richerson and Boyd, 2004). This suggests a culturally-dictated plasticity (e.g. Heine, 2001) which may have enabled us to readily adapt to changing circumstances and the vicissitudes of blind chance (e.g. Wallace and Wallace, 1999). Indeed, exaptation of basic animal consciousness – phenomenal, access, or whatever – to such synergism may well constitute the unique late hominid advantage which has enabled small, highly disciplined bands of us to become the most fearsome predators on the planet (Wallace, 2006b).

The development presented here carries the burden of seeming a 'theory of everything', in that the different equivalence classes of global workspace models appear to cover a vast swath of higher level cognitive phenomena across a considerable range of organisms and their assemblages. From Pielous viewpoint the 'let us assume that...' of equation (1), the requirement that an asymptotically piecewise stationary ergodic information source be 'dual' to a given cognitive process, overrides objections temporarily and allows construction of this grand castle-in-air. Two points:

First, few cognitive phenomena are likely to be ergodic in this sense, and development of a quasi-ergodic version of the theory is a starkly nontrivial mathematical enterprise. Absent the mathematical smoke-and-mirrors, what remains again in Pielou's sense is the set of questions these models raise regarding the ubiquity of conscious and paraconscious phenomena across many organisms and their natural groupings. In this spirit, the apparent ease with which cognitive modules may link together to form shifting, or even tunable, giant components, or engage in information dynamic phase transitions, suggests that individual consciousness in animals, and individual or group paraconsciousness for organisms and their assemblies, may have been quite common over a long period for very fundamental reasons.

Second, the development is most fundamentally a statistical model based on the Rate Distortion Theorem, in much the same sense that the General Linear Model is a statistical model based on the Central Limit Theorem. It is no more a theory of everything than is a fitted regression equation. The primary scientific utility of such things, as always, lies in empirical comparisons between models fitted to different systems, and in the structure of model residuals.

Statistical and mathematical models best serve as subordinate partners in the ongoing scientific conversation among theory, observation, and experiment. This work is no exception.

The discovery of large, markedly different, equivalence classes of mathematical models which all exhibit Global Workspace behaviors suggests it would be fruitful to widen our perspectives on animal consciousness even beyond the broad limits laid down by the symposium held by the Society for Integrative and Comparative Biology in 1999 and explored more recently by *Applied Animal Behavior Science*.

The road to a paleontology of consciousness is wide and open.

References

Aiello W., F. Chung, and L. Lu, 2000, A random graph model for massive graphs, in *Proceedings of the 32nd Annual ACM Symposium on the Theory of Computing*.

Albert R., and A. Barabasi, 2002, Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74:47-97.

Ash R., 1990, *Information Theory*, Dover Publications, New York.

Atlan H., and I. Cohen, 1998, Immune information ,self-organization and meaning, *International Immunology*, 10:711-717.

Atmanspacher H., 2006, Review of *Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model*, http://www.igpp.de/english/tda/pdf/wallace.pdf.

Auslander L., 1967, *Differential Geometry*, Harper and Row, New York.

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

Baars B., and S. Franklin, 2003, How conscious experience and working memory interact, *Trends in Cognitive Science*, doi:10.1016/S1364-6613(03)00056-1.

Baars, B., 2005, Global workspace theory of consciousness: toward a cognitive neuroscience of human experience, *Progress in Brain Research*, 150:45-53.

Bak A., R. Brown, G. Minian, and T. Porter, 2006, Global actions, groupoid atlases, and related topics, *Journal of Homotopy and Related Structures*, 1:1-54. Available from the ArXiv depository.

Bennett M., and P. Hacker, 2003 *Philosophical Foundations of Neuroscience*, Blackwell Publishing, London.

Burago D., Y. Burago, and S. Ivanov, 2001, *A Course in Metric Geometry*, American Mathematical Society, Providence, RI.

Cartmill M., 2000, Animal consciousness: some philosophical, methodological, and evolutionary problems, *American Zoologist*, 40:835-846.

Cohen I., 2000, *Tending Adam's Garden: Evolving the Cognitive Immune Self*, Academic Press, New York.

Connes A., 1994, *Noncommutative Geometry*, Academic Press, San Diego.

Corless R., G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, 1996, On the Lambert W function, *Advances in Computational Mathematics*, 4:329-359.

Cover T., and J. Thomas, 1991, *Elements of Information Theory*, John Wiley and Sons, New York.

Darwin C., 1889, *The Descent of Man, and Selection in Relation to Sex*, 2nd ed., Appleton, New York.

Dehaene S., and L. Naccache, 2001, Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework, *Cognition*, 79:1-37.

Dehaene S., and J. Changeux, 2005, Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness, *PLOS Biology*, 3:e141.

Dretske F., 1981, *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA.

Dretske F., 1988, *Explaining Behavior*, MIT Press, Cambridge, MA.

Dretske, F., 1993, Mental events as structuring causes of behavior, in *Mental Causation* (ed. by A. Mele and J. Heil), pp. 121-136, Oxford University Press.

Dretske F., 1994, The explanatory role of information, *Philosophical Transactions of the Royal Society A*, 349:59-70.

Edelman D., B. Baars, and A. Seth, 2005, Identifying hallmarks of consciousness in non-mammalian species, *Consciousness and Cognition*, 14:169-187.

Erdos P., and A. Renyi, 1960, On the evolution of random graphs, reprinted in *The Art of Counting*, 1973, 574-618 and in *Selected Papers of Alfred Renyi*, 1976, 482-525.

Feynman, R., 1996, *Feynman Lectures on Computation*, Addison-Wesley, Reading, MA.

Freeman, W., 2003, The wave packet: an action potential of the 21st Century, *Journal of Integrative Neurosciences*, 2:3-30.

Glazebrook, J., 2006, Rate distortion manifolds as model spaces for cognitive information. In preparation.

Goubault, E., and M. Raussen, 2002, Dihomotopty as a tool in state space analysis, *Lecture Notes in Computer Science*, Vol. 2286, April, 2002, pp. 16-37, Springer, New York.

Goubault E., 2003, Some geometric perspectives in concurrency theory, *Homology, Homotopy, and Applications*, 5:95-136.

Gould S., 1991, Exaptation: a crucial tool for evolutionary psychology, *Journal of Social Issues*, 47:43-65.

Granovetter M., 1973, The strength of weak ties, *American Journal of Sociology*, 78:1360-1380.

Griffin D., 1976, *The Question of Animal Awareness*, The Rockefeller University Press, New York.

Griffin D., 1992, *Animal Minds*, University of Chicago Press, Chicago.

Griffin D., 2000, Scientific approaches in animal consciousness, *American Zoologist*, 40:889-892.

Griffin D., and G. Speck, 2004, New evidence of animal consciousness, *Animal Cognition*, 7:5-18.

Grimmett G., and A. Stacey, 1998, Critical probabilities for site and bond percolation models, *The Annals of Probability*, 4:1788-1812.

Hayes C., 2003, Four routes of cognitive evolution, *Psychological Review*l, 110:713-727.

Heine S., 2001, Self as a cultural product: an examination of East Asian and North Amereican selves, *Journal of Personality*, 69:881-906.

Khinchin A., 1957, *The Mathematical Foundations of Information Theory*, Dover Publications, New York.

Knoll A, and S. Carroll, 1999, Early animal evolution: emerging views from comparative biology and geology, *Science*, 284:2129-2137.

Kozma R., M. Puljic, P. Balister, B. Bollobas, and W. Freeman, 2004, Neuropercolation: a random cellular automata approach to spatio-temporal neurodynamics, *Lecture Notes in Computer Science*, 3305:435-443.

Kozma R., M. Puljic, P. Balister, and B. Bollobas, 2005, Phase transitions in the neuropercolation model of neural populations with mixed local and non-local interactions, *Biological Cybernetics*, 92:367-379.

Krebs, P., 2005, Models of cognition: neurological possibility does not indicate neurological plausibility, in Bara, B., L. Barsalou, and M. Bucciarelli (eds.), *Proceedings of CogSci 2005*, pp. 1184-1189, Stresa, Italy. Available at http://cogprints.org/4498/.

Lee J., 2000, *Introduction to Topological Manifolds*, Springer, New York.

Luczak T., 1990, *Random Structures and Algorithms*, 1:287.

Molloy M., and B. Reed, 1995, A critical point for random graphs with a given degree sequence, *Random Structures and Algorithms*, 6:161-179.

Molloy M., and B. Reed, 1998, The size of the giant component of a random graph with a given degree sequence, *Combinatorics, Probability, and Computing*, 7:295-305.

Morris S., 2000, Perspectives: the Cambrian 'explosion': slow fuse or megatonnage? *Proceedings of the National*

Academy of Sciences, 97:4426-4429.

Newman M., S. Strogatz, and D. Watts, 2001, Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, 64:026118, 1-17.

Newman M., 2003, Properties of highly clustered networks, arXiv:cond-mat/0303183v1.

Pepperberg I., 2006, Cognitive and communicative abilities of Grey parrots, *Applied Animal Behavior Science*, 100:77-86.

Pielou, E., 1977, *Mathematical Ecology*, John Wiley and Sons, New York.

Pratt V., 1991, Modeling concurrency with geometry, *Proceedings of the 18th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, 311-322.

Richerson P., and R. Boyd, 2004, *Not by Genes Alone: How Culture Transformed Human Evolution*, Chicago University Press.

Savante J., D. Knuth, T. Luczak, and B. Pittel, 1993, The birth of the giant component, arXiv:math.PR/9310236v1.

Seth A., B. Baars, and D. Edelman, 2005, Criteria for consciousness in humans and other mammals, *Consciousness and Cognition*, 14:119-139.

Shannon C., and W. Weaver, 1949, *The Mathematical Theory of Communication*, University of Illinois Press, Chicago, IL.

Simmonds M., 2006, Into the brains of whales, *Applied Animal Behavior Science*, 100:103-116.

Skoyles, J., 1999, Neural plasticity and exaptation, *American Psychologist*, 54:438-439.

Stewart I., M. Golubitsky, and M. Pivato, 2003, Symmetry groupoids and patterns of synchrony in coupled cell networks, *SIAM Journal of Applied Dynamical Systems*, 2:609-646.

Stewart I., 2004, Networking opportunity, *Nature*, 427:601-604.

Tononi G., 2004, An information integration theory of consciousness, *BMC Neuroscience*, 5:42.

Wallace R., and R.G. Wallace, 1999, Organisms, organizations, and interactions: an information theory approach to biocultural evolution, *BioSystems*, 51:101-119.

Wallace R., 2000, Language and coherent neural amplification in hierarchical systems: renormalization and the dual information source of a generalized spatiotemporal stochastic resonance, *International Journal of Bifurcation and Chaos*, 10:493-502.

Wallace R., 2005a, *Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model*, Springer, New York.

Wallace R., 2005b, A global workspace perspective on mental disorders, *Theoretical Biology and Medical Modelling*, 2:49, http://www.tbiomed.com/content/2/1/49.

Wallace R., 2005c, The sleep cycle: a mathematical analysis from a global workspace perspective, http://cogprints.org/4517/

Wallace R., 2006a, Pitfalls in biological computing: canonical and idiosyncratic dysfunction of conscious machines, *Mind and Matter*, 4:91-113.

Wallace R., 2006b, Institutional paraconsciousness and its pathologies, http://cogprints.org/5117/.

Weinstein A., 1996, Groupoids: unifying internal and external symmetry, *Notices of the American Mathematical Association*, 43:744-752.

Figure Captions

Figure 1a. Diagonal Black and White holes in the two dimensional ω-plane. Only two direct paths can link points a and b which are continuously deformable into one another without crossing either hole. There are two additional monotonic switchback paths which are not drawn.

Figure 1b. Cross-diagonal Black and White holes as in 1a. Three direct equivalence classes of continuously deformable paths can link a and b. Thus the two spaces are topologically distinct. Here monotonic switchbacks are not possible, although relaxation of that condition can lead to 'backwards' switchbacks and intermediate loopings.