Version 9

Institutional paraconsciousness and its pathologies

Rodrick Wallace, Ph.D.
The New York State Psychiatric Institute*

August 21, 2006

Abstract

This analysis extends a recent mathematical treatment of the Baars consciousness model to analogous, but far more complicated, phenomena of institutional cognition. Individual consciousness is limited to a single, tunable, giant component of interacting cognitive modules, instantiating a Global Workspace. Human institutions, by contrast, support several, sometimes many, such giant components simultaneously, although their behavior remains constrained to a topology generated by cultural context and by the path-dependence inherent to organizational history. Such highly parallel multitasking - institutional paraconsciousness – while clearly limiting inattentional blindness and the consequences of failures within individual workspaces, does not eliminate them, and introduces new characteristic dysfunctions involving the distortion of information sent between global workspaces. Consequently, organizations (or machines designed along these principles), while highly efficient at certain kinds of tasks, remain subject to canonical and idiosyncratic failure patterns analogous to, but more complicated than, those afflicting individuals. Remediation is complicated by the manner in which pathogenic externalities can write images of themselves on both institutional function and therapeutic intervention, in the context of relentless market selection pressures. The approach is broadly consonant with recent work on collective efficacy, collective consciousness, and distributed cognition.

Key words bandpass, cognition, community, culture, directed homotopy, global workspace, groupoid, institution, information theory, random network, rate distortion, therapeutic intervention, topology.

INTRODUCTION

Humans, in disciplined groups, are the most fearsome predators on Earth. Humans, in large-scale organization, have recast both the topography and ecological dynamics of the entire planet. Human organizations, at all scales, are cognitive, taking the perspective of Atlan and Cohen (1998), in that they perceive patterns of threat or opportunity, compare those patterns with some internal, learned or inherited, picture of the world, and then choose one or a small number of responses from a vastly larger repertory of what is possible to them.

Both individuals and organizations operate within the constraints and affordances of human culture, which, to take the perspective of the evolutionary anthropologist Robert Boyd, at the individual level, "...is as much a part of human biology as the enamel on our teeth" (e.g. Richerson and Boyd, 2004).

One starting point for understanding the necessity of including culture in the study of cognition or consciousness at any scale lies in the observations of Nisbett et al. (2001), and others, following the tradition of Markus and Kitayama (1991), regarding fundamental differences in perception between test subjects of Southeast Asian and Western cultural heritage across an broad realm of experiments. East Asian perspectives are characterized as holistic and Western as analytic. Nisbett et al. (2001) find:

(1) Social organization directs attention to some aspects of the perceptual field at the expense of others.

(2) What is attended to influences metaphysics.

(3) Metaphysics guides tacit epistemology, that is, beliefs about the nature of the world and causality.

(4) Epistemology dictates the development and application of some cognitive processes at the expense of others.

(5) Social organization can directly affect the plausibility of metaphysical assumptions, such as whether causality should be regarded as residing in the field vs. in the object.

(6) Social organization and social practice can directly influence the development and use of cognitive processes such as dialectical vs. logical ones.

Nisbett et al. (2001) conclude that tools of thought embody a culture's intellectual history, that tools have theories build into them, and that users accept these theories, albeit unknowingly, when they use these tools.

Heine (2001) puts the matter as follows:

> "Cultural psychology does not view culture as a superficial wrapping of the self, or as a framework within which selves interact, but as something that is intrinsic to the self. It assumes that without culture there is no self, only a biological entity deprived of its potential... Cultural psychology maintains that the process of becoming a self is contingent on individuals interacting with and seizing meanings from the cultural environment..."

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

Clearly, culture must have a similarly intimate relation with the cognitive functioning of human organizations.

Recently the scientific study of human consciousness has again become permissible, after nearly a century of silence enforced by ideological diktat, and Baars' Global Workspace Theory (GWT), (Baars, 1988, 2005) is rapidly emerging as the first among equals in the Darwinian competition between theoretical approaches (e.g. Dehaene and Naccache, 2001). Wallace and colleagues (e.g. Wallace, 2005a, b, 2006; Glazebrook, 2006) have developed the first comprehensive mathematical model of GWT, using a Dretske-like information theory formalism (Dretske, 1981, 1988, 1993, 1994), extended by techniques from statistical physics, the Large Deviations Program of applied probability, and the topological theory of highly parallel computation.

Although individual human consciousness is, at present, socially constructed as a great scientific 'mystery', institutional cognition is, in fact, a far more complex and varied phenomenon, significantly less constrained by biological evolution, and far more efficient in many important respects. The ability to engage in culturally-sculpted organizational cognition, in fact, may be as fundamental an evolutionary adaptation for human survival as individual consciousness. The dual heritage mechanisms of genes and culture serve at both individual and collective scales of human endeavor (Richerson and Boyd, 2004).

According to the cultural anthropologists, the structures, functions, and innate character of institutional cognition are greatly variable and highly adaptable across social and physical geography, and across history. Individual human consciousness, by contrast, although profoundly shaped by culture, remains constrained by the primary biological necessity of single-tasking, leading to the striking phenomenon of inattentional blindness (IAB) when the Rate Distortion Manifold of consciousness become necessarily focused on one primary process to the virtual exclusion of others which might be expected to intrude (e.g. Mack, 1998; Dehaene and Changeux, 2005; Matsuda and Nisbett, 2006).

Simons and Chabris (1999) detail a particularly spectacular example of IAB. A videotape was made of a basketball game between teams in white and black jerseys. Experimental subjects who viewed the tape were asked to keep silent mental counts of either the total number of passes made by one or the other of the teams, or separate counts of the number of bounce and areal passes. During the game, a figure in a full gorilla suit appears, faces the camera, beats its breast, and walks off the court. About one half of the experimental subjects completely failed to notice the Gorilla during the experiment. See Simons (2000) for an extended discussion, and Wayand et al. (2005) for more recent experiments.

Other case histories, involving an aircraft crew which became fixated on an unexpectedly flashing control panel light during a landing, or a man walking a railroad track while having a cell phone conversation, are less benign.

Generalizing a second order treatment of Baars' Global Workspace model of individual consciousness to organizational structures will suggest the contrasting possibility of an analogous collective multitasking, effectively an institutional paraconsciousness, although that is far more complex mathematically than the individual case. There will emerge, however, an institutional analog to individual inattentional blindness, and additional failure modes specific to the complication of communication between multiple workspaces, as well as those related to the failure of individual workspaces within the organization. Remediation appears severely limited by the effects on it of the externalities so often responsible for the failures themselves.

Next is a recapitulation of recent work on individual consciousness as a kind of second order iteration of simple cognition, followed by the nontrivial extensions needed to describe institutional multiple workspaces.

FORMAL THEORY

1. The Global Workspace model of individual consciousness

The central ideas of Baars' Global Workspace Theory of individual consciousness 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 central focus of many researchers for two decades, consciousness studies has only recently, in the context of a deluge of empirical results from brain imaging experiments, begun digesting the perspective and preparing to move on.

Theory, however, sadly lags experiment. As Atmanspacher (2006) has put it,

> "To formulate a serious, clear-cut and transparent formal framework for cognitive neuroscience is a challenge comparable to the early stage of physics four centuries ago."

Currently popular agent-based and artificial neural network (ANN) treatments of cognition, consciousness and other higher order mental functions, to take 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.

Wallace (2005a, b; 2006) addresses 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 that work is the incorporation, in a natural manner, of constraints on individual consciousness, i.e. what Baars calls contexts. Using information theory methods, extended by an obvious homology between information source uncertainty and free energy density, it is 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. The Baars model emerges as an almost exact parallel to hierarchical regression, based, however, on the Shannon-McMillan rather than the Central Limit Theorem.

Wallace (2005b) recently proposed a somewhat different approach, using 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.

The first step is to argue for the existence of a network of loosely linked cognitive unconscious 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, we look at the necessary conditions imposed by the asymptotic limits of information theory on any realization of a cognitive process, be it biological wetware, silicon dryware, or some direct or systems-level hybrid. All cognitive processes, in this formulation, are 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, both of which are described further in the Mathematical Appendix. It is interactions between nodes in this abstractly defined network which will be of interest here, rather than whatever mechanisms, social or biological system, or mixture of them, actually constitute the underlying cognitive modules.

The second step 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 is the candidate for Baars' shifting Global Workspace of consciousness.

While Wallace (2005a) explores the effect of changing the average strength of nondisjunctive weak ties acting across linked unconscious modules, Wallace (2005b) focuses on changing the average *number* of such ties having a fixed strength, a complementary perspective whose extension, via a kind of renormalization, leads to a far more general approach.

The third step, following Wallace (2005b), is to tune the threshold at which the giant component comes into being, and to tune vigilance, the threshold for accession to consciousness.

Wallace's (2005b) information theory 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 equation models (e.g. Golubitsky and Stewart, 2006; Stewart et al., 2003, Stewart, 2004; Weinstein, 1996; Connes, 1994; Bak et al., 2006). 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.

The generalization necessary for the study of institutional cognition is to examine the conditions under which cognitive modules may multitask, engaging in more than one giant component at the same time. This is something which the individual human mind does not do well, and that individual consciousness does not do at all. The obvious tradeoff, of course, is the very rapid flow of individual consciousness, a matter of a few hundred milliseconds, as opposed to the much slower, if considerably more comprehensive, operations of institutional cognition.

2. Cognition as language

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 individual 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.

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, we 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 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

$$H \equiv \lim_{n \to \infty} \frac{\log[N(n)]}{n} \quad (1)$$

both exists and is independent of the path x.

We call such a pattern recognition-and-response cognitive process *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, it is possible to define an adiabatically, piecewise stationary, ergodic 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, or of some circuit in silicon. It is, rather, an abstract set of 'languages' dual to the cognitive processes instantiated by either biological wetware, social process, or their hybrids.

This 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 $a_j \sim G a_k$ 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.

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. Golubitsky and Stewart, 2006; Stewart et al. (2003), Stewart (2004)). Here we have outlined how to extend 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. Cognitive and conscious processes in humans involve neither computers nor clocks, but remain constrained by the limit theorems of information theory, and these permit scientific inference on necessary conditions.

4. Internal forces breaking the symmetry groupoid

The symmetry groupoid, as we have constructed it for cognitive modules, 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 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 putting 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 nonzero 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), which we outline here, is to simply look at the average number of fixed-strength nondisjunctive links in a random topology. These are obviously the two analytically tractable limits of a much more complicated regime.

Since we know nothing about how the cross-talk connections can occur, we will – 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 – which we now take 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. At the level of the individual mind, unconscious cognitive modules link up to become the Global Workspace of consciousness, 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. Restricting analysis to such 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

Just as a higher order information source, associated with the GC of a random or semirandom graph, can be constructed out of the interlinking of unconscious cognitive modules by mutual information, so too external information sources, for example in humans the cognitive immune and other physiological 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. For machines or institutions these would be the onion-like '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 we assume the Global Workspace/Giant Component to involve 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 which we write as

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

(4)

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

This result does not commit the mereological fallacy, of which Bennett and Hacker (2003) accuse the many 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 machines or organizations interacting with their environments. See Wallace (2006) for further discussion.

6. Punctuation phenomena

As a number of researchers have noted, in one way or another, – 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 piecewise, adiabatically stationary, ergodic information sources. These biological phase changes appear to be ubiquitous in natural systems and can be expected to dominate machine and organizational behaviors 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. Institutional paraconsciousness

The random 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 we address 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. Suppose we 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. Any 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 at this point 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. We hence are assuming that the ω is a tunable, syntactically-dependent, detection limit, and depends critically on the instantaneous topology of the giant component defining, for the human mind, the global workspace of consciousness. 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.

There is considerable empirical evidence from fMRI brain imaging experiments to show that individual human consciousness involves a single global workspace, a matter leading necessarily to the phenomenon of inattentional blindness. Cognitive submodules within institutions, – individuals, departments, formal and informal workgroups – by contrast, can do more than one thing, and indeed, are usually required to multitask. Clearly this will lessen the probability of inattentional blindness, but, we will find, does not eliminate it, and introduces other failure modes examined in more detail later.

We must postulate a set of crosstalk information measures between cognitive submodules, each associated with its own giant component having its own special topology.

Suppose the set of giant components 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 set having a particular set of topological structures. Suppose that, over a sequence of 'times' the set of giant components 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 internal) signals impinging on the set of giant components, 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 set of giant components 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 by the institution then itself becomes a punctuated event in response to increasing linkage between the organization 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). For individual consciousness, there is only one giant component. For an institution, there will be a larger, and often very large, set of them.

This leads to the possibility of new failure modes related to impaired communication between Giant Components.

That is, a complication specific to high order institutional cognition lies in the necessity of information transfer between giant components. The form and function of such interactions will, of course, be determined by the nature of the particular institution, but, synchronous or asynchronous, contact between giant components is circumscribed by the Rate Distortion Theorem. That theorem, reviewed in the Mathematical Appendix, states that, for a given maximum acceptable critical average distortion, there is a limiting maximum information transmission rate, such that messages sent at less than that limit are guaranteed to have distortion less than the critical maximum. Too rapid transmission between parallel global workspaces – information overload – violates that condition, and guarantees large average distortion. This is a likely failure mode which appears unique to multiple workspace systems which, we will argue, may otherwise have a lessened probability of inattentional blindness.

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 our ability to provide energy to the brain is limited. Inattentional blindness, we will 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. We suppose that different cognitive phenomena have, in the sense of Wallace (2000, 2005, Ch. 3), dual information sources, and are interested in the local properties of the system near a particular reference state. We impose a topology on the system, so that, near a particular 'language' A, dual to an underlying cognitive process, there is (in some sense) an open set U of closely similar languages $\hat{A}$, such that $A, \hat{A} \subset U$. Note that it may be necessary to coarse-grain the system's 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 $\hat{A}$, which we write 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 – again an empirical equation like a regression model – in this case becomes

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

(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 forcing mechanisms will be needed to effect a transition to a different state. We shall explore this circumstance below in terms of the concept of ecosystem resilience.

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.

9. The simplest rate distortion manifold

The second order 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.

Suppose, now, that threshold behavior for institutional 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 institutional attention 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.

It is of interest to examine 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. It is thus 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 institutional inattentional blindness, in spite of multitasking.

Note that higher order Rate Distortion Manifolds are likely to give better approximations than lower ones, in the same sense that second order tangent structures give better, if more complicated, approximations in conventional differentiable manifolds (e.g. Pohl, 1962).

Indeed, Rate Distortion Manifolds can be quite formally described using standard techniques from topological manifold theory (Glazebrook, 2006). The essential point is that a rate distortion manifold is a topological structure which constrains the 'stream of institutional consciousness' as well as the pattern of communication between institutional giant components, much the way a riverbank constrains the flow of the river it contains. This is a fundamental insight, which we pursue further.

10. The topology of cognition

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. One can ask 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 second order hierarchical model which, in effect, describes a universality class tuning of the renormalization parameters characterizing the dancing, flowing, tunably 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 attention 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 institutional cognition.

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 second order hierarchical cognitive model of institutional cognition where they are, perhaps, of more central interest.

Empirical study will likely show how the influence of cultural heritage or developmental history defines quite different dihomotopies of attentional focus in human organizations. That is, the topology of blind spots and their associated patterns of perceptual completion in human organizations will be culturally or developmentally modulated. It is this developmental cultural topology of multitasking organization attention which, acting in concert with the inherent limitations of the rate distortion manifold, generates the pattern of organizational inattentional blindness.

Such considerations, and indeed the Black Hole development of equation (12), suggest that a multitasking organization which becomes trapped in a particular pattern of behavior cannot, in general, expect to emerge from it in the absence of some forcing mechanism.

This form of behavior is central to ecosystem resilience theory.

Ecosystem theorists, in fact, recognize several different kinds of resilience (e.g. Gunderson, 2000). The first, which they call 'engineering resilience', since it is particularly characteristic of machines and man-machine interactions, involves the rate at which a disturbed system returns to a presumed single, stable, equilibrium condition, following perturbation. From that limited perspective, a resilient system is one which quickly returns to its one stable state.

Not many biological or social phenomena seem resilient in this simplistic sense.

Holling's (1973) particular contribution was to recognize that sudden transitions between different, at best quasi-stable, domains of relation among ecosystem variates were possible, i.e. that more than one 'stable' state was possible for real ecosystems. Gunderson (2000) puts the matter as follows:

> "One key distinction between these two types of resilience lies in assumptions regarding the existence of multiple [quasi-]stable states. If it is assumed that only one stable state exists or can be designed to exist, then the only possible definition and measures for resilience are near equilibrium ones – such as characteristic return time... The concept of ecological resilience presumes the existence of multiple stability domains and the tolerance of the system to perturbations that facilitate transitions among stable states. Hence, ecological resilience refers to the width or limit of a stability domain and is defined by the magnitude of disturbance that a system can absorb before it changes stable states... The presence of multiple [quasi-]stable states and transitions among them [has] been [empirically] described in a [large] range of ecological systems..."

The topology of institutional cognition provides a tool for study of resilience in human organizations or social systems. The obvious conjecture is that the set of directed homotopy equivalence classes described above formally classifies quasi-equilibrium states, and thus characterizes the different possible ecosystem resilience modes.

PATHOLOGIES OF INDIVIDUAL CONSCIOUSNESS

A central focus of this work is characterization of failure modes in multiple-workspace institutional cognition. Insight regarding the depth and breadth of the possible complexities can be gained from a recapitulation of failures in a system having but a single global workspace, i.e. the human mind. The result is not reassuring.

Mental disorders in humans are not well understood. Indeed, such classifications as the *Diagnostic and Statistical Manual of Mental Disorders - fourth edition*, (DSM-IV, 1994), the standard descriptive nosology in the US, have been characterized as prescientific by Gilbert (2001) and others. Arguments from genetic determinism fail, in part because of an apparently draconian population bottleneck which, early in our species' history, resulted in an overall genetic diversity less than that observed within and between contemporary chimpanzee subgroups. Arguments from psychosocial stress fare better, but are affected by the apparently complex and contingent developmental paths determining the onset of schizophrenia – one of the most prevalent serious mental disorders – dementias, psychoses, and so forth, some of which may be triggered in utero by exposure to infection, low birthweight, or other stressors.

Gilbert suggests an extended evolutionary perspective, in which evolved mechanisms like the 'flight-or-fight' response are inappropriately excited or suppressed, resulting in such conditions as anxiety or post traumatic stress disorders. Nesse (2000) suggests that depression may represent the dysfunction of an evolutionary adaptation which down-regulates foraging activity in the face of unattainable goals.

Kleinman and Good, however, (1985, p. 492) have outlined some of the cross cultural subtleties affecting the study of depression which seem to argue against any simple evolutionary interpretation:

> "When culture is treated as a constant (as is common when studies are conducted in our own society), it is relatively easy to view depression as a biological disorder, triggered by social stressors in the presence of ineffective support, and reflected in a set of symptoms or complaints that map back onto the biological substrate of the disorder... However, when culture is treated as a significant variable, for example, when the researcher seriously confronts the world of meaning and experience of members of non-Western societies, many of our assumptions about the nature of emotions and illness are cast in sharp relief. Dramatic differences are found across cultures in the social organization, personal experience, and consequences of such emotions as sadness, grief, and anger, of behaviors such as withdrawal or aggression, and of psychological characteristics such as passivity and helplessness or the resort to altered states of consciousness. They are organized differently as psychological realities, communicated in a wide range of idioms, related to quite varied local contexts of power relations, and are interpreted, evaluated, and responded to as fundamentally different meaningful realities... Depressive illness and dysphoria are thus not only interpreted differently in non-Western societies and across cultures; they are *constituted* as fundamentally different forms of social reality."

More generally, Kleinman and Cohen (1997) find that

> "[S]everal myths... have become central to psychiatry... The first is that the forms of mental illness everywhere display similar degrees of prevalence... [Second is] an excessive adherence to a principle known as the pathogenic/pathoplastic dichotomy, which holds that biology is responsible for the underlying structure of a malaise, whereas cultural beliefs shape the specific ways in which a person experiences it. The third myth maintains that various unusual culture-specific disorders whose biological bases are uncertain occur only in exotic places outside the West... In an effort to base psychiatry in 'hard' science and thus raise its status to that of other medical disciplines, psychiatrists have narrowly focused on the biological underpinnings of mental disorders while discounting the importance of such 'soft' variables as culture and socioeconomic status..."

Further, serious mental disorders in humans are often comorbid among themselves – depression and anxiety, compulsive behaviors, psychotic ideation, etc. – and with serious chronic physical conditions such as coronary heart disease, atherosclerosis, diabetes, hypertension, dyslipidemia, and so on. These too are increasingly recognized as developmental in nature (e.g. Wallace, 2005a), and are frequently compounded by behavioral problems like violence or substance use and abuse. Indeed, smoking, alcohol and drug addiction, compulsive eating, and the like, are often done as self-medication for the impacts of psychosocial and other stressors, constituting socially-induced 'risk behaviors' which synergistically accelerate a broad spectrum of mental and physical problems.

Recent research on schizophrenia, dyslexia, and autism, supports a 'brain connectivity' model for these disorders which is of considerable interest from a global workspace perspective, since large-scale brain connectivity is essential for the operation of consciousness, a principal, and very old, evolutionary adaptation in higher animals.

Burns et al. (2003), on the basis of sophisticated diffusion tensor magnetic resonance imaging studies, find that schizophrenia is a disorder of large-scale neurocognitive networks rather than specific regions, and that pathological changes in the disorder should be sought at the supra-regional level. Both structural and functional abnormalities in frontoparietal networks have been described and may constitute a basis for the wide range of cognitive functions impaired in the disorder, such as selective attention, language processing and attribution of agency.

Silani et al. (2005) find that, for dyslexia, altered activation observed within the reading system is associated with altered density of grey and white matter of specific brain regions, such as the left middle and inferior temporal gyri and left arcuate fasciculus. This supports the view that dyslexia is associated with both local grey matter dysfunction and with altered [larger scale] connectivity among phonological/reading areas.

Villalobos et al. (2005) explore the hypothesis that large-scale abnormalities of the dorsal stream and possibly the mirror neuron system, may be responsible for impairments of joint attention, imitation, and secondarily for language delays in autism. Their empirical study showed that those with autism had significantly reduced connectivity with bilateral inferior frontal area 44, which is compatible with the hypothesis of mirror neuron defects in autism. More generally, their results suggest that dorsal stream connectivity in autism may not be fully functional.

Courchesne and Pierce (2005) suggest that, for autism, connectivity within the frontal lobe is excessive, disorganized, and inadequately selective, whereas connectivity between frontal cortex and other systems is poorly synchronized, weakly responsive and information impoverished. Increased local but reduced long-distance cortical-cortical reciprocal activity and coupling would impair the fundamental frontal function of integrating information from widespread and diverse systems and providing complex context-rich feedback, guidance and control to lower-level systems.

Coplan (2005) has observed a striking pattern of excessive frontal lobe self-connectivity in certain cases of anxiety disorder, and Coplan et al. (2005) find that maternal stress can affect long-term hippocampal neurodevelopment in a primate model.

As stated, brain connectivity is the sine qua non of the Global Workspace model of individual human consciousness, and further analysis suggests that these disorders cannot be fully understood in the absence of a functional theory of consciousness, and in particular, of a detailed understanding of the elaborate regulatory mechanisms which must have evolved over the past half billion years to ensure the stability of that most central and most powerful of adaptations.

Distortion of consciousness is not simply an epiphenomenon of the emotional dysregulation which many see as the 'real' cause of mental disorder. Like the pervasive effects of culture, distortion of consciousness lies at the heart of both the individual experience of mental disorder and the effect of it on the embedding of the individual within both social relationships and cultural or environmental milieu. Distortion of consciousness in mental disorders inhibits both routine social interaction and the ability to meet internalized or expected cultural norms, a potentially destabilizing positive feedback. Distortion of consciousness profoundly affects the ability to learn new, or change old, skills in the face of changing patterns of threat or opportunity, perhaps the most critical purpose of the adaptation itself. Distortion of consciousness, particularly any decoupling from social and cultural context, is usually a threat to long-term individual survival, and those with mental disorders significantly affecting consciousness typically face shortened lifespans.

Human communities, as natural as neighborhoods, or as intentional as a multinational corporation or an army, have, according to the perspective of this work, multiple, effectively simultaneous, global workspaces which must not only function individually, but in concert with, while perhaps in competition for resources with, other similar modules.

Granovetter's (1973) Strength of Weak Ties argument is that nondisjunctive relationships, those which do not disjointly partition a community, are essential for that community's efficient functioning. Strong ties are those which do disjointly partition a group, for example religious affiliation, age cohort, ethnicity, national origin, skin color in some cases, language, and so on. For an institutional setting these might be the classification by Division, Department, Work Group, informal office-politics clique, and so on. From the perspective of GWT, Granovetter's weak ties permit both the formation of individual workspaces, and enable those workspaces to communicate effectively.

Understanding failures within and between institutional global workspaces seems predicated on an understanding of weak tie structure and dynamics, which are themselves embedded in larger social and cultural contexts. Clearly, individual workspace failures will often be subject to monitoring and control by parallel workspaces, limiting the damage, as it were, in a manner impossible for human consciousness. Thus institutions may suffer less inattentional blindness and less consequence from individual workspace failure, by virtue of parallel operations, but these problems will not be eliminated. Powerful subgroups not subject to contextual constraint seem a particular problem.

Given that some form of failure is inevitable for such systems, how are they to be corrected, and what are the limits to such correction?

INSTITUTIONAL THERAPEUTICS

Institutions, by virtue of their many highly parallel, multitasking global workspaces, seem inherently less prone to errors from inattentional blindness or from misfunction of individual workspaces. Within organizations, several workspaces typically examine a problem, choose possible modes of action, and, essentially, negotiate and reach consensus. If cross communication between them is not too distorted, a 'good' institutional decision usually emerges. Persistent failure by an individual workspace, or a particular subset of them, can be addressed by external correction, i.e. therapeutic intervention by other entities. Pathological context, which is often responsible for individual workspace failure, can, however, become convoluted with the intervention itself, resulting in therapeutic failure. It is possible to model this process in more detail.

Recall that the essential characteristic of cognition in this formalism involves a mapping, $h(x)$, of a (convolutional) path $x = a_0, a_1, ..., a_n, ...$ onto a member of one of two disjoint sets, B_0 or B_1. Thus respectively, either (1) $h(x) \in B_0$, implying no action taken, or (2), $h(x) \in B_1$, and some particular response is chosen from a large repertoire of possible responses. There is an evident problem in defining these two disjoint sets, suggesting that some higher order, i.e. executive, cognitive module is needed to determine what constitutes B_0, the set of normal actions and procedures, those not constituting explicit intervention. Again, this is because there is no low energy mode for information systems. That is, virtually all states are more or less high energy, high information content or transmission, states. Thus there is no natural way to identify a ground state using the physicist's favorite variational or other minimization arguments.

Suppose that higher order executive cognitive module, which can be described as a kind of Zero Mode Identification, interacts with an embedding, highly structured quasi-language of systemic perturbation - market forces and failures, disasters, structured noise, and the like. Instantiating a Rate Distortion image of that embedding stress, the ZMI begins to include one or more members of the set B_1 into the set B_0, or vice versa, when a circumstance requiring action is ignored. Recurrent hits on that aberrant state would be experienced as episodes of institutional pathology, over or under reaction.

Empirical tests of this hypothesis quickly involve real-world regression models of the interrelations among measurable markers of success and failure, leading to the Rate Distortion Manifold arguments above.

Different eigenmodes Y_k of the RDM regression model characterized by the zero mode matrix $\mathbf{R}_0$ can be taken to represent the shifting-of-gears between different languages defining the sets B_0 and B_1. That is, different eigenmodes of the RDM would correspond to different required (and possibly mixed) characteristic systemic responses.

If there is a state (or set of states) Y_1 such that $\mathbf{R}_0 Y_1 = Y_1$, then the unitary kernel Y_1 corresponds to the condition 'no response required', i.e. the set B_0.

Suppose pathology becomes manifest, i.e.

$$\mathbf{R}_0 \to \mathbf{R}_0 + \delta\mathbf{R} \equiv \hat{\mathbf{R}}_0,$$

so that some chronic 'excited state' becomes the new unitary kernel, and

$$Y_1 \to \hat{Y}_1 \neq Y_1$$

$$\hat{\mathbf{R}}_0 \hat{Y}_1 = \hat{Y}_1.$$

Next, assume other, perhaps embedding, corporate global workspaces induce a sequence of therapeutic counterperturbations – deliberate therapeutic interventions – $\delta\mathbf{T}_k$ according to the pattern

$$[\hat{\mathbf{R}}_0 + \delta\mathbf{T}_1]\hat{Y}_1 = Y^1,$$

$$\hat{\mathbf{R}}_1 \equiv \hat{\mathbf{R}}_0 + \delta\mathbf{T}_1,$$

$$[\hat{\mathbf{R}}_1 + \delta\mathbf{T}_2]Y^1 = Y^2$$

$$...$$

(19)

so that, in some sense,

$$Y^j \to Y_1.$$

(20)

That is, the system, as monitored by the RDM, is driven to its original condition.

It may or may not be possible to have $\hat{\mathbf{R}}_0 \to \mathbf{R}_0$. That is, actual remediation may not be possible, in which case palliation or control is the therapeutic aim.

The essential point is that the pathological state represented by $\hat{\mathbf{R}}_0$ and the sequence of therapeutic interventions $\delta\mathbf{T}_k, k = 1, 2, ...$ are interactive and reflective, depending on the regression of the set of vectors Y^j to the desired state Y_1, much in the same spirit as Jerne's immunological idiotypic hall of mirrors.

The therapeutic problem revolves around minimizing the difference between Y^k and Y_1 over the course of treatment. That difference represents the inextricable convolution of treatment failure with adverse reactions to the course of treatment itself, and failure of compliance, often attributed through social construction by 'provider' to 'patient', i.e. failure of the therapeutic alliance.

It should be obvious that the treatment sequence $\delta\mathbf{T}_k$ itself is a cognitive path of interventions which has, in turn, a dual information source in the sense previously invoked.

Treatment may, then, interact in the usual Rate Distortion manner with the pathogenic patterns of structured perturbation - market pressures, failures, disasters, red noise and the like – which are, themselves, signals from an embedding information source. Thus treatment failure, adverse reactions, and noncompliance will, of necessity, embody a distorted image of embedding structured perturbations which may indeed be responsible for the primary misfunction.

This coupling would most likely occur in a highly punctuated manner, depending in a quantitative way on the degree of interlinking of the three-fold system of affected individual workgroup, therapeutic interaction, and treatment mode, with that perturbation.

Clearly this is only one example of a much larger spectrum of possibilities. Empirical study would seem necessary at this point to prune down the search tree, as it were, making further theoretic analysis practical.

One disturbing implication of this analysis is the apparent difficulty of correcting institutional paraconsciousness once it becomes overtly pathological. The general experience of greatly shortened lifespan for most individuals suffering pathologies of global workspace connectivity - schizophrenia, dementia, autism, and the like - suggests that the relentless impact of market forces, which are effectively evolutionary selection pressures, guarantees rapid extinction or merger as the most likely outcome for any systematic, large-scale, organizational cognitive lapses. Certainly the last century has not been kind to states which possessed, or attempted to gain, colonial empires of various forms, nor to the regions and peoples which suffered occupation under colonial regimes (Memmi, 1969; Fanon, 1962).

DISCUSSION AND CONCLUSIONS

The simple groupoid defined by an institution's inherent cognitive modular structure can be broken by intrusion of (rapid) crosstalk within it, and by the imposition of (slower) crosstalk from without – market forces and the embedding culture. The former initiates a set of topologically-determined giant component global workspaces, in a punctuated manner, while the latter deform the underlying topology of the entire system, the directed homotopy limiting what paths can actually be traversed. Broken symmetry creates richer structure in systems characterized by groupoids, just as it does for those characterized by groups.

Multitasking institutional attention acts through a Rate Distortion manifold, a kind of retina-like filter for grammatical and syntactical meaningful paths. Signals outside the topologically constrained tunable syntax/grammar bandpass of this manifold are subject to lessened probability of punctuated conscious detection: organizational inattentional blindness. Culture and path-dependent developmental history will, according to this model, profoundly affect the phenomenon by imposing additional topological constraints defining the 'surface' along which this second order behavior can (and cannot) glide.

Glazebrook (2006) has suggested that, lurking in the background of this basic construction, is what Bak et al. (2006) call a groupoid atlas, i.e. an extension of topological manifold theory to groupoid mappings. Also lurking is identification and exploration of the natural groupoid convolution algebra which so often marks these structures (e.g. Weinstein, 1996; Connes, 1994).

Consideration suggests, in fact, that a path may be meaningful according to the groupoid parametization of all possible dual information sources, and that tuning is done across that parametization via a rate distortion manifold.

Implicit, however, are the constraints imposed by embedding cultural heritage or institutional history, which may further limit the properties of $\mathbf{R}_0$, i.e. hold it to a developmentally determined topology.

Here we have attempted to reexpress this trade-off in terms of a syntactical/grammatical version of conventional signal theory, i.e. as a tuned meaningful path form of the classic balance between sensitivity and selectivity, as particularly constrained by the directed homotopy imposed by cultural heritage on a basic institutional experience that is itself the outcome of historical process.

Overall, this analysis is analogous to, but more complicated than, Wallace's information dynamics instantiation of Baars' Global Workspace theory (Wallace, 2005a, b; 2006). Intuitively, one suspects that the higher the dimension of the second order attentional Rate Distortion Manifold, that is, the greater the multitasking, the broader the effective bandwidth of attentional focus, and the less likely is inattentional blindness. For a conventional differentiable manifold, a second or higher order tangent space would give a better approximation to the local manifold structure than a simple plane (Pohl, 1962).

In a similar manner, the highly parallel, multitasking set of global workspaces which constitute institutional paraconsciousness should be less prone to the consequences of individual workspace failure, provided cross communication between them is relatively undistorted.

It is not difficult to introduce the evolutionary selection pressures of market forces into this model, using the approach of Wallace (2002).

Nonetheless, inattentional blindness, while constrained by multitasking, is not eliminated by it. This suggests that higher order institutional cognition, the generalization of individual consciousness, is subject to canonical and idiosyncratic patterns of failure analogous to, but perhaps more subtle than, the kind of disorders described in Wallace (2005b, 2006). Indeed, while machines designed along these principles – i.e. rapid, multitasking Global Workspace devices – could be spectacularly efficient at many complex tasks, ensuring their stability might be even more difficult than for intelligent, and hence conscious, machines designed as analogs of the human mind. The relatively slow pace of institutional paraconsciousness may well provide significant opportunities for self or externally-induced correction which would be much more difficult for machines operating in the range of a few hundred milliseconds.

Further, the necessity of interaction – synchronous or asynchronous – between institutional giant components suggests the possibility of failures governed by the Rate Distortion Theorem. Forcing rapid communication between institutional global workspaces ensures high levels of average distortion. Recent, and very elegant, ethnographic work by Cohen et al. (2006) and Laxmisan et al. (2006) regarding systematic medical error in emergency rooms focuses particularly on 'handover' problems at shift change, where incoming medical staff are rapidly briefed by outgoing staff. Systematic information overload in such circumstances seems almost routine, and is widely recognized as a potential error source.

Analogs with mental illness in humans suggest that failures within individual workspaces, while limited by the interactional context of other workspaces, will remain serious sources of overall institutional failure, particularly when involving powerful, higher authority modules or work groups not subject to peer-level contextual constraint. Therapeutic intervention can itself reflect the very external perturbations causing individual or large-scale workspace failures. Institutions suffering significant workspace pathologies, like similarly afflicted individuals, are likely to face truncated lifespans in the face of market selection pressures, other relentless externalities, and the deteriorating internal dynamics of corporate dementia.

The hierarchical cognitive model appropriate to institutional paraconsciousness is considerably more complicated than that for individual human consciousness, which, perhaps in a tradeoff permitting rapid, accurate, response to environmental stimulation, seems biologically limited to a single shifting, tunable giant component structure. Human institutions, by contrast appear able to entertain several, and perhaps many, such global workspaces simultaneously, although these must, for reliability, operate at a much slower rate than is possible for individual consciousness.

Shared culture seems to provide far more than merely a shared language for the establishment of the human organizations which enable our adaptation to, or alteration of, our varied environments. It also may provide the stabilizing mechanisms needed to overcome many of the canonical and idiosyncratic failure modes inherent to such organizations – inattentional blindness and pathologies within and between individual workspaces.

In sum, institutional paraconsciousness, most especially in its emergent second order manifestation through generalized weak ties, is far more complicated, if necessarily much slower, than individual consciousness, although both are confined by the topological contexts of culture and developmental history. The proper understanding of such phenomena, however, requires the use of cutting edge methods.

Some comparison with other perspectives is warranted. One starting point is Robert Sampson's 'collective efficacy' approach to community function (Sampson, Raudenbush and Earls, 1997; Sampson, 2004). A large and growing body of sociological research, beginning with Granovetter (1973), emphasizes the essential role of nondisjunctive 'weak' social ties within a community, that is, ties which operate across such classifications as age cohort, ethnicity, religion, occupation, institutional membership, and so on. Collective efficacy is a recent reworking of the basic idea (Sampson, 2004):

> "...[C]ollective efficacy [means] an emphasis on shared beliefs in a neighborhood's capability for action to achieve an intended effect, coupled with an active sense of engagement on the part of residents. Some density of social networks is essential... [b]ut the key theoretical point is that *networks have to be activated to be ultimately meaningful*. Collective efficacy therefore helps to elevate the 'agentic' aspect of social life over a perspective centered on the accumulation of stocks of [social capital]. This is consistent with a redefinition of social capital in terms of expectations for action within a collectivity... [in sum] *social networks foster the conditions under which collective efficacy may flourish, but they are not sufficient for the exercise of control*."

From the viewpoint of this analysis, what Sampson invokes is a limited version of community cognition, the ability of a neighborhood to perceive patterns of threat or opportunity, to compare those perceived patterns with an internal, shared, picture of the world, and to choose one or a few collective actions from a much larger repertory of those possible, and to carry them out. Disjunctive or 'strong' social ties define some of the underlying cognitive modules – collective and individual – within the neighborhood. Weak ties, then, are those which link such modules – individual or collective – across the community. Individuals, defined subgroups, or formal organizations, may have multiple roles within that community, permitting the formation of multiple global workspaces, if the strength of the various weak ties linking them is sufficient. Institutional cognition, in the sense of this work, emerges as a dynamic, collective phenomenon. Cultural constraints and developmental trajectory serve as 'contexts' to both stabilize and direct the resulting cognitive processes, which may still fail through inattentional blindness, resource limitation, or other pathologies including failure of communication between institutional global workspaces.

This is, however, not Sampson's static, cross-sectional, structure, but, rather, is deeply constrained, not just by shared culture, but by the path dependent historic development of the community itself. Our own work (e.g. Wallace and Wallace, 1997; D. Wallace and R. Wallace, 1998, 1998a; Wallace and Fullilove, 1999; Wallace et al., 1996) and that of others (e.g. Fullilove, 2004) demonstrates that 'planned shrinkage', 'urban renewal', or other disruptions of weak ties akin to ethnic cleansing, can place neighborhoods onto decades-long trajectories of social disintegration which short-circuit effective community cognition. This is, indeed, a fundamental political purpose of such programs.

There are other treatments of collective consciousness. Thomas Burns and his collaborators have, at times, focused particularly on the role of social process in understanding consciousness. Burns and Engdalh (200x) write

> "What is particularly striking about the academic consciousness industry is the absence of sociology...
>
> A collective has the capacity in its collective representations and communications about what it can (and cannot) do, or should do (or should not do). It monitors its activities, its achievements and failures, and... analyzes and discusses itself as a defined and developing collective agent... a collective has potentially a rich basis not only for talking about, discussing, agreeing (or disagreeing) about a variety of objects... but it also has a means to conceptualize and develop alternative types of social relationships, effective forms of leadership, coordination and control, and... new normative orders and institutional arrangements... These potentialities enable systematic, directed problem solving, and the generation of variety and complex strategies. In particular selective environments, these make for major evolutionary advantages...
>
> [However]..[c]ollective representations and reflectivity and directed problem-solving based on them may prevent human groups from experiencing or discovering the un-represented or un-named; unrecognized or poorly defined problems cannot be dealt with... Reflective and problem-solving powers may then be distorted, the generation of alternatives and varieties narrow and largely ineffective, and social innovation and transformation misdirected and possibly self-destructive..."

The criticism of the academic consciousness industry is most apt.

Burns and Engdahl (200x, 1998a, b) appear to describe institutional cognition and something much like inattentional blindness. Our innovation is to propose that the particular evolutionary advantage of such cognition is its potential ability to entertain several global workspaces simultaneously, and thus raise overall action capacity while reducing, but not eliminating, inattentional blindness, although introducing failures involving miscommunication between institutional global workspaces.

From a related viewpoint, that of distributed cognition, Patel (1998) describes a paradigm shift regarding the role of technology in medicine:

> "The notion of distributed cognition suggests that the immediate physical and social resources outside the person participate in cognition, not just as a source of input and a receiver of output, but as a vehicle of thought... the individual and the environment should be viewed as dynamically interacting, resulting in cognitive performance and learning...
>
> Increasingly, researchers have come to characterize cognition as a distributed process... [I]ntelligence can be seen as distributed in designed artifacts such as computer-user interfaces; in representations, such as diagrams; and through communication in social contexts... The idea of intelligence (i.e. knowledge and cognition) being distributed in a group, or in artifacts, customs, and situations, is interesting because it provides a framework for addressing a number of theoretical and empirical questions."

As discussed briefly, Patel and colleagues (Cohen et al., 2006 and references therein; Laxmisian et al., 2006) use a distributed cognition model to analyze emergency room process and function, and found what can be interpreted as 'rate distortion' error effects at shift change handover of patients between clinical teams, something emergency room personnel have characterized as a 'telephone game' (Cohen et al., 2006). This appears to be a prime example of a canonical failure mode involving rate-driven, distorted communication between the institutional global workspaces which generalize individual consciousness.

We end where we began. Humans in disciplined teams are the most fearsome predators on Earth. Partly, this is because the institutional paraconsciousness of such groups, while not operating at the few hundred milliseconds of individual consciousness, is still very fast, while having both improved ability to act and decrease in the likelihood of being blindsided or suffering workspace failure.

Institutional paraconsciousness seems far more complicated than individual consciousness, but prone, as well, to particular collective error modes.

Given the massive overt and structural violence of the last century, this is no small matter, and our propensity to very large-scale institutional paraconsciousness, and the evident difficulty in correcting demented organizations, may, over the long term, ultimately be experienced as a seriously defective evolutionary adaptation. Deeper understanding of these phenomena is clearly a priority.

MATHEMATICAL APPENDIX

The Shannon-McMillan Theorem

According to the structure of the underlying language of which a message is a particular expression, some messages are more 'meaningful' than others, that is, are in accord with the grammar and syntax of the language. The Shannon-McMillan or Asymptotic Equipartition Theorem, describes how messages themselves are to be classified.

Suppose a long sequence of symbols is chosen, using the output of the random variable X above, so that an output sequence of length n, with the form

$$x_n = (\alpha_0, \alpha_1, ..., \alpha_{n-1})$$

has joint and conditional probabilities

$$P(X_0 = \alpha_0, X_1 = \alpha_1, ..., X_{n-1} = \alpha_{n-1})$$

$$P(X_n = \alpha_n | X_0 = \alpha_0, ..., X_{n-1} = \alpha_{n-1}).$$

(21)

Using these probabilities we may calculate the conditional uncertainty

$$H(X_n | X_0, X_1, ..., X_{n-1}).$$

The uncertainty of the *information source*, $H[\mathbf{X}]$, is defined as

$$H[\mathbf{X}] \equiv \lim_{n \to \infty} H(X_n | X_0, X_1, ..., X_{n-1}).$$

(22)

In general

$$H(X_n | X_0, X_1, ..., X_{n-1}) \leq H(X_n).$$

Only if the random variables X_j are all stochastically independent does equality hold. If there is a maximum n such that, for all $m > 0$

$$H(X_{n+m} | X_0, ..., X_{n+m-1}) = H(X_n | X_0, ..., X_{n-1}),$$

then the source is said to be of *order* n. It is easy to show that

$$H[\mathbf{X}] = \lim_{n \to \infty} \frac{H(X_0, ... X_n)}{n+1}.$$

In general the outputs of the $X_j, j = 0, 1, ..., n$ are *dependent*. That is, the output of the communication process at step n depends on previous steps. Such serial correlation, in fact, is the very structure which enables most of what is done in this paper.

Here, however, the processes are all assumed stationary in time, that is, the serial correlations do not change in time, and the system is *stationary*.

A very broad class of such self-correlated, stationary, information sources, the so-called *ergodic* sources for which the long-run relative frequency of a sequence converges stochastically to the probability assigned to it, have a particularly interesting property:

It is possible, in the limit of large n, to divide all sequences of outputs of an ergodic information source into two distinct sets, S_1 and S_2, having, respectively, very high and very low probabilities of occurrence, with the source uncertainty providing the splitting criterion. In particular the Shannon-McMillan Theorem states that, for a (long) sequence having n (serially correlated) elements, the number of 'meaningful' sequences, $N(n)$ – those belonging to set S_1 – will satisfy the relation

$$\frac{\log[N(n)]}{n} \approx H[\mathbf{X}].$$

(23)

More formally,

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

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

$$= \lim_{n\to\infty} \frac{H(X_0, ..., X_n)}{n+1}.$$

(24)

Using the internal structures of the information source permits *limiting attention only to high probability 'meaningful' sequences of symbols*.

The Rate Distortion Theorem

The Shannon-McMillan Theorem can be expressed as the 'zero error limit' of the Rate Distortion Theorem (Dembo and Zeitouni, 1998; Cover and Thomas, 1991), which defines a splitting criterion that identifies high probability pairs of sequences. We follow closely the treatment of Cover and Thomas (1991).

The origin of the problem is the question of representing one information source by a simpler one in such a way that the least information is lost. For example we might have a continuous variate between 0 and 100, and wish to represent it in terms of a small set of integers in a way that minimizes the inevitable distortion that process creates. Typically, for example, an analog audio signal will be replaced by a 'digital' one. The problem is to do this in a way which least distorts the *reconstructed* audio waveform.

Suppose the original stationary, ergodic information source Y with output from a particular alphabet generates sequences of the form

$$y^n = y_1, ..., y_n.$$

These are 'digitized,' in some sense, producing a chain of 'digitized values'

$$b^n = b_1, ..., b_n,$$

where the b-alphabet is much more restricted than the y-alphabet.

b^n is, in turn, *deterministically retranslated* into a reproduction of the original signal y^n. That is, each b^m is mapped on to a unique n-length y-sequence in the alphabet of the information source Y:

$$b^m \to \hat{y}^n = \hat{y}_1, ..., \hat{y}_n.$$

Note, however, that many y^n sequences may be mapped onto the *same* retranslation sequence $\hat{y}^n$, so that information will, in general, be lost.

The central problem is to explicitly minimize that loss.

The retranslation process defines a new stationary, ergodic information source, $\hat{Y}$.

The next step is to define a *distortion measure*, $d(y, \hat{y})$, which compares the original to the retranslated path. For example the *Hamming distortion* is

$$d(y, \hat{y}) = 1, y \neq \hat{y}$$

$$d(y, \hat{y}) = 0, y = \hat{y}.$$

(25)

For continuous variates the *Squared error distortion* is

$$d(y, \hat{y}) = (y - \hat{y})^2.$$

(26)

There are many possibilities.

The distortion between paths y^n and $\hat{y}^n$ is defined as

$$d(y^n, \hat{y}^n) = \frac{1}{n}\sum_{j=1}^{n} d(y_j, \hat{y}_j).$$

(27)

Suppose that with each path y^n and b^n-path retranslation into the y-language and denoted y^n, there are associated individual, joint, and conditional probability distributions

$$p(y^n), p(\hat{y}^n), p(y^n|\hat{y}^n).$$

The *average distortion* is defined as

$$D = \sum_{y^n} p(y^n) d(y^n, \hat{y}^n).$$

(28)

It is possible, using the distributions given above, to define the information transmitted from the incoming Y to the outgoing $\hat{Y}$ process in the usual manner, using the Shannon source uncertainty of the strings:

$$I(Y, \hat{Y}) \equiv H(Y) - H(Y|\hat{Y}) = H(Y) + H(\hat{Y}) - H(Y, \hat{Y}).$$

If there is no uncertainty in Y given the retranslation $\hat{Y}$, then no information is lost.

In general, this will not be true.

The *information rate distortion function* $R(D)$ for a source Y with a distortion measure $d(y, \hat{y})$ is defined as

$$R(D) = \min_{p(y,\hat{y}); \sum_{(y,\hat{y})} p(y)p(y|\hat{y})d(y,\hat{y}) \leq D} I(Y, \hat{Y}).$$

(29)

The minimization is over all conditional distributions $p(y|\hat{y})$ for which the joint distribution $p(y, \hat{y}) = p(y)p(y|\hat{y})$ satisfies the average distortion constraint (i.e. average distortion $\leq D$).

The *Rate Distortion Theorem* states that $R(D)$ *is the maximum achievable rate of information transmission which does not exceed the distortion* D. Cover and Thomas (1991) or Dembo and Zeitouni (1998) provide details.

More to the point, however, is the following: Pairs of sequences $(y^n, \hat{y}^n)$ can be defined as *distortion typical*; that is, for a given average distortion D, defined in terms of a particular measure, pairs of sequences can be divided into two sets, a high probability one containing a relatively small number of (matched) pairs with $d(y^n, \hat{y}^n) \leq D$, and a low probability one containing most pairs. As $n \to \infty$, the smaller set approaches unit probability, and, for those pairs,

$$p(y^n) \geq p(\hat{y}^n|y^n) \exp[-nI(Y, \hat{Y})].$$

(30)

Thus, roughly speaking, $I(Y, \hat{Y})$ embodies the splitting criterion between high and low probability pairs of paths.

For the theory of interacting information sources, then, $I(Y, \hat{Y})$ can play the role of H in the dynamic treatment that follows.

The rate distortion function can actually be calculated in many cases by using a Lagrange multiplier method – see Section 13.7 of Cover and Thomas (1991).

ACKNOWLEDGEMENTS

The author thanks Drs. D. Wallace and V. Patel for useful discussions.

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 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.

Burns J., D. Job, M. Bastin, H. Whalley, T. Macgillivary, E. Johnstone, and S. Lawrie, 2003, Structural disconnectivity in schizophrenia: a diffusion tensor magnetic resonance imaging study, *British Journal of Psychiatry*, 182:439-443.

Burns T., and E. Engdahl, 200x, Toward a social theory of human consciousness, *Karl Jaspers Forum*, http://www.kjf.ca/N-61-BUR.htm.

Burns T., and E. Engdahl, 1998a, Collective consciousness and its socio-cultural foundations, *Journal of Consciousness Studies*, 5:67-85.

Burns T., and E. Engdahl, 1998b, Individual selves, self-awareness and reflectivity, *Journal of Consciousness Studies*, 5:166-184.

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

Cohen T., B. Blatter, C. Almeida, E. Shortliffe, and V. Patel, 2006, A cognitive blueprint of collaboration in context: Distributed cognition in the psychiatric emergency department, *Artificial Intelligence in Medicine*, 37:73-83.

Coplan J., 2005, Personal communication.

Coplan J., M. Altemus, S. Matthew, E. Smith, B. Scharf. P. Coplan, J. Kral, J. Gorman, M. Owens, C. Nemeroff, and L. Rosenblum, 2005, Synchronized maternal-infant elevations of primate CSF CRF concentrations in response to variable foraging demand, *CNS Spectrums*, 10:530-536.

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.

Courchesne E., and K. Pierce, 2005, *Current Opinion in Neurobiology*, 15:225-230.

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

Dembo A., and O. Zeitouni, 1998, *Large Deviations: Techniques and Applications*, Second edition, Springer, New York.

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.

DSMIV, 1994, *Diagnostic and Statistical Manual*, fourth edition, American Psychiatric Association, Washington, DC.

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.

Fullilove, M., 2004, *Root Shock*, Ballantine Books, New York.

Gilbert P., 2001, Evolutionary approaches to psychopathology: the role of natural defenses, *Austrailian and New Zealand Journal of Psychiatry*, 35:17-27.

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

Golubitsky M., and I. Stewart, 2006, Nonlinear dynamics and networks: the groupoid formalism, *Bulletin of the American Mathematical Society*, 43:305-364.

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.

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

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

Gunderson L., 2000, Ecological resilience - in theory and application, *Annual Reviews of Ecological Systematics*, 31:425-439.

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

Holling C., 1973, Resilience and stability of ecological systems, *Annual Reviews of Ecological Systematics*, 4:1-23.

Laxmisian A., F. Hakimzada, O. Sayan, R. Green, J. Zhang, V. Patel, 2006, The multitasking clinician; Decision-making and cognitive demand during and after team handoffs in emergency care. Submitted.

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

Kleinman A. and B. Good, 1985, *Culture and Depression: Studies in the Anthropology of Cross-Cultural Psychiatry of Affect and Depression*, University of California Press, Berkeley, CA.

Kleinman A., and A. Cohen, 1997, Psychiatry's global challenge, *Scientific American*, xx:86-89.

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.

Mack A., 1998, *Inattentional Blindness*, MIT Press, Cambridge, MA.

Markus H., and S. Kitayama, 1991, Culture and the self-implications for cognition, emotion, and motivation, *Psychological Review*, 98:224-253.

Matsuda T., and R. Nisbett, 2006, Culture and change blindness, *Cognitive Science: A Multidisciplinary Journal*, 30:381-399.

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.

Nesse R., 2000, Is depression an adaptation? *Archives of General Psychiatry*, 57:14-20.

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.

Nisbett R., K. Peng, C. Incheol, and A. Norenzayan, 2001, Culture and systems of thought: holistic vs. analytic cognition, *Psychological Review*, 108:291-310.

Patel, V., 1998, Individual to collaborative cognition: a paradigm shift? *Artificial Intelligence in Medicine*, 12:93-96.

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

Pohl W., 1962, Differential geometry of higher order, *Topology* 1:169-211.

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.

Sampson R., S. Raudenbush, and F. Earls, 1997, Neighborhoods and violent crime: A multilevel study of collective efficacy, *Science*, 277:xxx-xxx.

Sampson R., 2004, Networks and neighborhoods, in *Network Logic: Who Governs in an Interconnected World?*, ed. by H. McCarthy, P. Miller, and R. Skidmore, Demos, London.

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

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

Siliani G., U. Frith, J. Demonet, F. Fazio, D. Perani, C. Price, C. Frith, and E. Paulesu, 2005, Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study, *Brain*, 128(Pt.10):2453-2461.

Simons, D., and C. Chabris, 1999, Gorillas in our midst: sustained inattentional blindness for dynamic events, *Perception*, 28:1059-1074.

Simons D., 2000, Attentional capture and inattentional blindness, *Trends in Cognitive Sciences*, 4:147-155.

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.

Villalobos M., A. Mizuno, B. Dahl, N. Kemmotsu and R. Muller, 2005, Reduced functional connectivity between VI and Inferior frontal cortex associated with visumotor performance in autism, *Neruoimage*, 25:916-925.

Wallace D., and R. Wallace, 1998, *A Plague on Your Houses*, Verso, New York.

Wallace D. and R. Wallace, 1998a, Scales of geography, time, and population: the study of violence as a public health problem, *American Journal of Public Health*, 88:1853-1858.

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., 2002, Adaptation, punctuation, and information: a rate distortion approach to non-cognitive 'learning plateaus' in evolutionary process, *Acta Biotheoretica*, 50:101-116.

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

Wallace R., M. Fullilove, and A. Flisher, 1996, AIDS, violence and behavioral coding: information theory, risk behavior, and dynamic process on core-group sociogeographic networks, *Social Science and Medicine*, 47:339-352.

Wallace R., and D. Wallace, 1997, Community marginalization and the diffusion of disease and disorder in the United States, *British Medical Journal*, 314:1341-1345.

Wallace R., and R. Fullilove, 1999, Why simple regression models work so well describing risk behaviors in the USA, *Environment and Planning A*, 31:719-734.

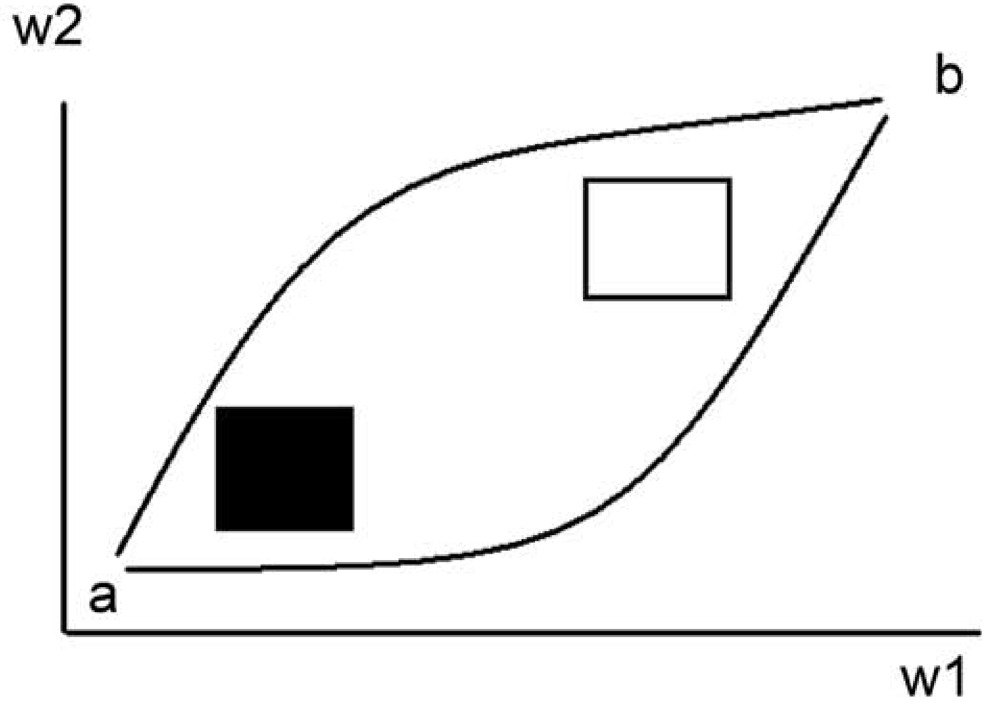
Wayand J., D. Levin and D. Alexander Varakin, 2005, Inattentional blindness for a noxious multimodal stimulus, *American Journal of Psychology*, 118:339-352.

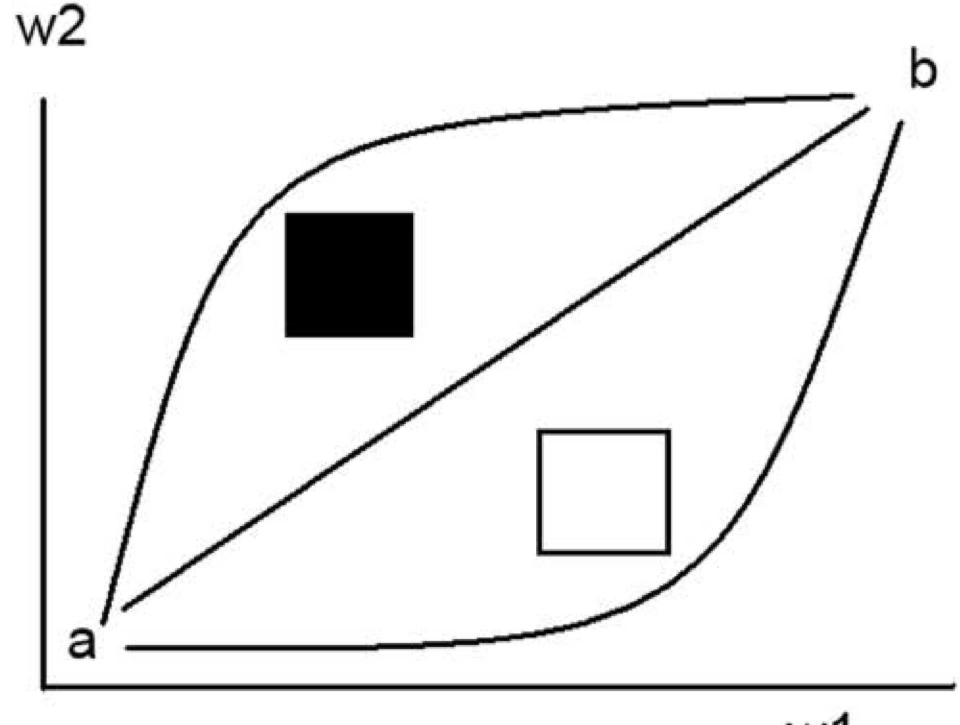
Weinstein A., 1996, Groupoids: unifying internal and external symmetry, *Notices of the American Mathematical Association*, 43:744-752.

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.


w2
b
a
w1


w2
b
a
w1