Version 5

A modular network treatment of Baars' Global Workspace consciousness model

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

September 6, 2005

Abstract

Adapting techniques from random and semirandom network theory, this work provides an alternative to the renormalization and phase transition methods used in Wallace's (2005a) treatment of Baars' Global Workspace model. The new formalism predicts dynamics that should be empirically distinguishable from those suggested by the earlier analysis. Nevertheless, like the earlier work, it produces the workspace itself, the tunable threshold of consciousness, and the essential role for embedding contexts, in an explicitly analytic 'necessary conditions' manner which suffers neither the mereological fallacy inherent to brain-only theories nor the sufficiency indeterminacy of neural network or agent-based simulations. This suggests that the new approach, and the earlier, represent different analytically solvable limits in a complicated continuum of possible models. Thus the development significantly extends the theoretical foundations for an empirical general cognitive model (GCM) based on the Shannon-McMillan Theorem. Patterned after the general linear model based on the Central

*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.

Limit Theorem, the proposed technique could be particularly useful in the reduction of experimental data on consciousness.

Key words consciousness, general cognitive model, global workspace, information theory, phase transition, random network.

Introduction

Bernard Baars' Global Workspace (Baars, 1988) has emerged as the first among equals in the Darwinian scientific conflict between models of consciousness in humans (e.g. Dehaene and Naccache, 2001). The central ideas are as follows (Baars and Franklin, 2003):

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

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

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

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

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

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

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

Although this basic canonical structure has been systematically elaborated upon for nearly twenty years by a number of quite eminent researchers, consciousness studies has only recently, in the context of a deluge of data from brain imaging experiments, come to the point of actually digesting the perspective and moving on.

One of the indices of a maturing scientific field is the incorporation of analytic mathematical models into the dialog between theory and experiment. Unfortunately, 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 in his application of information theory to mental function (Dretske, 1981, 1988, 1992, 1993, 1994), give considerably more insight.

Recently Wallace (2005a, b) has attempted a variation on Baars' theme from Dretske's perspective, addressing the necessary conditions which the asymptotic limit theorems of information theory impose on the Global Workspace. Perhaps the central outcome of this work has been 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 becomes almost trivial 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 neatly evades the mereological fallacy which fatally bedevils brain-only theories of human consciousness (Bennett and Hacker, 2003).

Transfer of renormalization 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' (sensu 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.

Here we will explore a different and somewhat simpler analysis, using classic results from random and semirandom network theory which circumvent the complicated renormalization calculations (Erdos and Renyi, 1960; Albert and Barabasi, 2002; Newman, 2003). The unconscious modular 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", we will outline an approach which can serve as the foundation of a different, but roughly parallel, treatment of the Global Workspace. The topological tuning of the threshold for consciousness implied by this new theory should be experimentally distinguishable from the universality class tuning proposed in Wallace (2005a).

The first step is to argue, as before, 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.

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 will be the candidate for the shifting Global Workspace.

While Wallace (2005a) examines the effect of changing the average strength of nondisjunctive weak ties acting across linked unconscious modules, this paper focuses on changing the average *number* of such ties having a fixed strength, a quite complementary perspective.

The third step will be to tune the threshold at which the giant component comes into being, and to tune vigilance, the threshold for accession to consciousness.

Finally, we will again recover the influence of Baars' contexts in a natural manner.

Cognition as language

Cognition is not consciousness. Indeed, most mental, and many physiological, functions, while cognitive in a particular sense, hardly ever become entrained into the Global Workspace of consciousness. For example, 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' in a particular formal sense. The construction is surprisingly straightforward (Wallace, 2000, 2005).

Atlan and Cohen (1998) and Cohen (2000) argue, in the particular 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, from this view, proceeds by functionally combining 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 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 algorithmic composition of internal and external signals.

This path is fed into a highly nonlinear, but otherwise similarly unspecified, nonlinear decision oscillator 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 interest are paths x which trigger pattern recognition-and-response exactly once. That is, given a fixed initial state a_0, such that $h(a_0) \in B_0$, we examine all possible subsequent paths x beginning with a_0 and leading exactly once to the event $h(x) \in B_1$. Thus $h(a_0, ..., a_j) \in B_0$ for all $j < m$, but $h(a_0, ..., a_m) \in B_1$. Wallace (2005a) examines more complicated schemes as well.

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 having $h(a_0) \in B_0$ 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 is that the finite limit

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

(1)

both exists and is independent of the path x.

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 particular sense) is 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).

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 Khinchine (1957), Ash (1990), or Cover and Thomas (1991) for the standard details.

The giant component

The next trick is even more straightforward: a formal equivalence class algebra (and hence a groupoid, (Weinstein, 1996)) 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 defines the vertices of the proposed network of cognitive dual languages. Each vertex then represents a different information source dual to a cognitive process.

We suppose that linkages can fleetingly occur between the ordinarily disjoint cognitive modules defined by this algebra. In the spirit of Wallace (2005a), this is represented by establishment of a non-zero mutual information measure between them: cross-talk.

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 here, initially, 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 – 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. 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 $(\mu - \ln(1+\mu))^{-1} + \mathcal{O}(\log\log n)$. If $\mu = 0$, however, the largest component has size of order $M^{2/3}$. And 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 $\mu = -\alpha^{-1}\ln(1-\alpha) - 1$. Thus, for example, a random graph with approximately $M\ln(2)$ edges will have a giant component containing $\approx M/2$ vertices.

More explicitly, as Corless et al. (1996) discuss, when a graph with M vertices has $m = (1/2)aM$ edges chosen at random, for $a > 1$ it almost surely has a giant connected component having approximately gM vertices, with

$$g(a) = 1 + W(-a\exp(-a))/a, \tag{2}$$

where W is the Lambert-W function defined implicitly by the relation

$$W(x)\exp(W(x)) = x. \tag{3}$$

Figure 1 shows $g(a)$, displaying what is clearly a sharp phase transition at $a = 1$.

Such a phase transition initiates a new, collective, cognitive phenomenon: the Global Workspace 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 defines the richness of the cognitive language of the workspace, will grow as some function of g, as more and more unconscious processes are incorporated into it. Wallace (2005a) examines what, in effect, are the functional forms $H \propto \exp(\alpha g), \alpha\ln[1/(1-g)]$, and $(1/(1-g))^{\delta}$, letting $R = 1/1-g$ define a 'characteristic length' in the renormalization scheme. While these all have explicit solutions for the renormalization calculation (mostly in terms of the Lambert-W function), other, less tractable, expressions are certainly plausible, for example $H \propto g^{\gamma}, \gamma > 0$, γ real.

Given a particular $H(g)$, the universality class tuning of Wallace (2005a) involves adjusting universality class parameters of the phase transition.

Tuning the giant component

In this new class of models, the degree of clustering of the graph of cognitive modules might, itself, be tunable, producing a shifting threshold for consciousness: a topological tuning, which should be observable from brain-imaging studies. Second order iteration might lead to an analog of the hierarchical cognitive model of Wallace (2005a), which is based on universality class tuning.

Albert and Barabasi (2002) describe the intimate relation between network geometry and phase transition behavior as follows:

> "The inseparability of the topology and dynamics of evolving networks is shown by the fact that [the exponents defining universality class in a network] are related by [a] scaling equation... underlying the fact that a network's assembly uniquely determines its topology. However, in no case are these exponents unique. They can be tuned continuously..."

Albert and Barabasi (2002) also discuss how to measure clustering and show the equivalence of the random graph problem to percolation theory, although they fail to exploit the explicit Lambert-W function solution.

By contrast, the development of Wallace (2005a), in focusing on changing the average strength of weak ties between unconscious submodules rather than the average number of fixed-strength weak ties as is done here, tunes the universality class exponents, which may imply subtle shifts in underlying topology.

Following Albert and Barabasi (2002, Section V), we note that real networks differ from random graphs in that their degree distribution, the probability of k linkages between vertices, often follows a power law $P(k) \approx k^{-\gamma}$ rather than the Poisson distribution of random networks,

$P(k) = a^k \exp(-a)/k!, k \geq 0$. Since power laws do not have any characteristic scale, these are consequently termed scale-free networks.

It is possible to extend the Erdos/Renyi threshold results to such 'semi-random' graphs. For example, Luczak (1992) has shown that almost all random graphs with a fixed degree smaller than 2 have a unique giant cluster. Molloy and Reed (1995, 1998) have proved that, for a random graph with degree distribution $P(k)$, an infinite cluster emerges almost surely when

$$Q \equiv \sum_{k \geq 1} k(k-2)P(k) > 0.$$

(4)

Following Volz, (2004), cluster tuning of random networks leads to a counterintuitive result. Define the clustering coefficient C as the proportion of triads in a network out of the total number of potential triads, i.e.

$$C = \frac{3N_\Delta}{N_3},$$

(5)

where N_Δ is the number of triads in the network and N_3 is the number of connected triples of nodes, noting that in every triad there are three connected nodes. Taking the approach of Molloy and Reed (1995), Volz shows quite directly that, for a random network with parameter a, at cluster value C, there is a critical value given by

$$a_C = \frac{1}{1 - C - C^2}.$$

(6)

If $C = 0$ then the giant component forms when $a = 1$ in the absence of clustering. Increasing C *raises* the average number of edges which must be present for a giant component to form. For $C \geq \sqrt{5}/2 - 1/2$, which is precisely the Golden Section, where the numerator in this expression vanishes, no giant component can form, regardless of a. Not all network topologies, then, can actually support a giant component, and hence, in this model, consciousness. This is a matter of some importance with obvious implications ranging from the evolutionary history of consciousness to the nature of sleep.

A fuller exploration of the giant component can be found, e.g. in Newman et al. (2001), especially the discussion leading to their figure 4. In general, 'tuning' of the GC will generate a family of curves similar to figure 1, but with those having threshold to the right of that in the plot 'topping out' at limits progressively less than 1: higher thresholds seem usually imply smaller giant components.

Again, these different exact solutions should be distinguishable through brain imaging or other empirical observations, necessary to restrict the exponentially growing range of theoretical possibilities.

Contexts and the mereological fallacy

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 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. These are, then, to be viewed as among Baars' contexts. 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. 387). Given three interacting information sources, Y_1, Y_2, Z, the splitting criterion, taking Z as the 'external context', is given by

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

(7)

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

This generalizes to

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

(8)

If we assume the brain's Global Workspace/GC to involve a very rapidly shifting 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 and cultural processes will be characterized by even more slowly-acting sources $Z_k, k = 1..n$. Mathematical induction on equation (8) gives a complicated expression for a mutual information splitting criterion which we write as

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

(9)

This encompasses a full biopsychosociocultural structure for individual human consciousness, one in which contexts act as important boundary conditions.

This result does not commit the mereological fallacy which Bennett and Hacker (2003) impute to excessively neurocentric perspectives on consciousness. See Wallace (2005a) for more discussion.

Generalized Onsager relations

Following Wallace (2005a, b), understanding the time dynamics of these systems requires a phenomenology similar to the Onsager relations of non-equilibrium 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

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

(10)

become of central interest.

Equation (10) 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 by Wallace (2005a).

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

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

(11)

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, as Wallace (2005b) argues, may be subject to override by external physiological driving mechanisms.

Imposing a metric for different cognitive dual languages parametized by **K** leads quickly into the rich structures of Riemannian, or even Finsler, geometries (Wallace, 2005b).

Discussion and conclusions

One should be truly reluctant to reify mathematical models of complicated biological, social, and ecological phenomena. As the mathematical ecologist E.C. Pielou (1977, p. 106) states,

> "...[Mathematical] models are easy to devise; even though the assumptions of which they are constructed may be hard to justify, the magic phrase 'let us assume that...' overrides objections temporarily. One is then confronted with a much harder task: How is such a model to be tested? The correspondence between a model's predictions and observed events is sometimes gratifyingly close but this cannot be taken to imply the model's simplifying assumptions are reasonable in the sense that neglected complications are indeed negligible in their effects...
>
> In my opinion the usefulness of models is great... [however] it consists *not in answering questions but in raising them.* Models can be used to inspire new field investigations and these are the only source of new knowledge as opposed to new speculation."

In sum, it is quite unlikely there will ever be an empirically valid mathematical theory of consciousness having the elegant cachet of General Relativity.

Nevertheless, here we have developed an alternative to the renormalization phase transition methodology used in Wallace (2005a) by adapting a perspective based on random or semirandom network theory. While the earlier model focused on changing the average strength of nondisjunctive linkages between elements of a fixed topology of unconscious cognitive processes, the analysis here involves changing the average number of fixed-strength linkages in a possibly shifting topology, or changing the distribution and clustering of linkages. This approach nonetheless reproduces the workspace itself as the giant component, generates the tunable threshold, in which alteration of topology is analogous to universality class tuning, and allows embedding contexts to act as essential boundary conditions to individual conscious experience. It does this in a highly natural, explicitly analytic manner which is not burdened by the mereological fallacy inherent to brain-only theories. In addition, since the model relies strongly on necessary conditions defined by the asymptotic limit theorems of information theory, it does not fall victim to the debilitating sufficiency indeterminacy which afflicts neural network and agent-based computer simulations of higher order cognitive functions (Krebs, 2005).

Ultimately, since the renormalization and network methods give similar results, it seems probable that it will be necessary to combine the techniques, allowing tuning of the strength, number, and distribution of nondisjunctive cross-talk linkages between unconscious cognitive submodules, as well, perhaps, as more subtle indices of the underlying topology of those linkages. Thus Wallace (2005a) and this paper explore different analytically tractable limits of a more complicated regime whose full exploration will require significant time and resources, and will best be done in concert with empirical study to prune down the many possible alternative structures.

This is not an unusual circumstance. Perhaps the best known example is from analytical mechanics where the two-body and many-body problems are both exactly solvable using fairly elementary methods, respectively Newtonian and statistical mechanics. Intermediate cases require a more sophisticated approach, often based on perturbation expansion about an exactly solvable case (e.g. Arnold, 1989).

In any event, the work here further strengthens the possibility of developing an empirical general cognitive model for consciousness and related phenomena which is based on the Shannon-McMillan Theorem, and would be an exact analog to the general linear model that relies on the Central Limit Theorem. A successful GCM, like the GLM, would likely be of considerable utility for the analysis of experimental data, perhaps, given Pielou's remarks, the best possible use for this body of theory.

References

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

Arnold V., 1989, *Mathematical Methods of Classical Mechanics*, second edition, Springer-Verlag, New York.

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.

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.

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

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

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

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

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

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.

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.

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

Hobson J., 2001, *The Dream Machine: Chemically Altered States of Consciousness*, MIT Press, Cambridge, MA.

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

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/.

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

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

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

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.

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

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

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

Volz, E., 2004, Random networks with tunable degree distribution and clustering, *Physical Review E*, 70:056115.

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

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

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

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

Figure Caption

Figure 1. Relative size of the largest connected component of a random graph, as a function of the average number of fixed-strength connections between vertices. W is the Lambert-W function, or the ProductLog in Mathematica, which solves the relation $W(x)\exp[W(x)] = x$. Note the sharp threshold at $a = 1$, and the subsequent topping-out.'Tuning' the giant

component by changing topology generally leads to a family of similar curves, those having progressively higher threshold with correspondingly lower asymptotic limits (e.g. Newman et al., 2001, fig. 4).

RELATIVE SIZE OF LARGEST CONNECTED COMPONENT