Version 3

The sleep cycle: a mathematical analysis from a global workspace perspective

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

July 29, 2005

Abstract

Dretske's technique of invoking necessary conditions from information theory to describe mental process can be used to derive a version of Hobson's AIM treatment of the sleep/wake cycle from a mathematical formulation of Baars' Global Workspace model of consciousness. One implication of the analysis is that some sleep disorders may be recognizably similar to many other chronic, developmental dysfunctions, including autoimmune and coronary heart disease, obesity, hypertension, and anxiety disorder, in that these afflictions often have roots in utero or adverse early childhood experiences or exposures to systematic patterns of structured stress. Identification and alteration of such factors might have considerable impact on population-level patterns of sleep disorders, suggesting the possibility of a public health approach rather than current exorbitantly expensive case-by-case medical intervention.

Key words AIM, consciousness, global workspace, information theory, phase transition, NREM, REM, sleep.

Introduction.

Sleep and consciousness present complementary enigmas across a broad spectrum of intellectual disciplines. Philosophers, psychologists, psychiatrists, neuroscientists and others have long speculated about their nature, content, purpose, and the relations between them. This has not been a smooth processes, however. For most of the 20th Century, behaviorist ideology in psychology precluded the scientific study of consciousness, and at least one essential characteristic of sleep – the Rapid Eye Movement (REM) stage – was only discovered in the 1950's.

Currently two complementary candidates for 'standard models' have emerged describing the phenomena, Hobson's activation-input-modulation (AIM) treatment of the sleep-wake cycle (Hobson, 2001) and Baars' global workspace (GW) analysis of consciousness (Baars, 1988). The purpose of this paper is to extend a recent formal mathematical treatment of the GW model (Wallace, 2005) toward Hobson's perspective. The basis for this attempt is Dretske's recognition that the asymptotic limit theorems of information theory provide a set of necessary conditions for all mental process, just as they do for any phenomena involving the transmission and transformation of information (Dretske, 1981, 1988, 1993, 1994). The essential trick is to implement an obvious homology between information source uncertainty and the free energy density of a physical system. This permits importation of renormalization methods for information systems which undergo the punctuated 'phase changes' characteristic of accession to consciousness. It also permits development of a kind of analog to the phenomenology of Onsager relations in a physical system, allowing description of behavior away from such critical points, in the context of a set of topological rate distortion manifolds which generalize the idea of an adaptable retina (Wallace, 2005, Ch. 4). Taking the model to second order in tuning the universality class associated with conscious punctuation gives the Baars result. Here we use this formalism to produce, in what we claim is a highly natural manner, something much like Hobson's treatment of the sleep/wake cycle.

The reader should be broadly familiar with the different stages of sleep, their role in fixation of memory (e.g. Rauchs et al., 2005), and Hobson's work (Hobson, 2001), as well as with the general ideas of Baars' global workspace model (Baars, 1988). With regard to the latter, essential features include the shifting serial broadcast/recruitment of information among unconscious modules, and punctuated behavior (e.g. Sergeant and Dehaene, 2004). Note that REM sleep is now viewed as a state of consciousness which, compared to alert waking, is deficient in both neuromuscular function (i.e. sleep paralysis) and analytic ability, while particularly rich in emotional cognition. Finally, both waking up and going to sleep involve complicated physiological processes to effect a transition between states. Since the late 1940's researchers have understood the importance of the reticular activating mechanism in the change from sleep to waking (Evans, 2003; Moruzzi and Magoun, 1949). More recently, an arousal inhibitory mechanism, a thalamo-cortical process, has been recognized which transfers the body from waking to sleep (Evans, 2003). The necessity of mechanisms to effect these transitions will prove to be central to the subsequent argument.

*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. Affiliations are for identification only.

The mathematical development will be limited to first order and to the minimum needed to go from the GW to the AIM model.

Extending the global workspace model.

Figure 1, from Hobson (2001) via Mahowald (2003), is a schematic of the activation-input-modulation model of sleep. It is a three dimensional classification of sleep/wake states based on (1) degree of central nervous system activation, the A axis, (2) classification of input into endogenous (internal) vs. exogenous (external), the I axis, and (3) modulation, M, based on aminergic vs. cholinergic-dominated neurochemistry. The normal trajectory, the dotted line, is the orderly progression believed necessary to efficiently fix memory (Rauchs et al., 2005).

To anticipate the argument slightly, we shall argue that the CNS activation axis has its analog in the information dynamics version of global workspace theory through the probability of interaction between distinct neural modules, which generally process internal or external signals unconsciously, but can be recruited into the global workspace by long-range neural structures. Coma and (deep, slow-wave) NREM sleep are states in which such global recruitment does not, in fact, occur. The information theory analysis differs from Hobson's model in that this recruitment is highly punctuated – in effect a phase transition – once the probability of interaction between distant modules exceeds a threshold. Hence the sudden accession of some external or internal state to consciousness.

Modulation, from the information theory perspective, can likely be related to the 'richness', in a technical sense to be explored more fully below, of conscious cognitive process: REM sleep consciousness, since it is deficient in somatic function and analytic ability when compared to waking consciousness, is consequently far less rich, in this technical sense, than waking consciousness, although dreams have great and convincing, indeed delusional, emotional valence.

The trick will be to extend the information theory treatment of the GW model to include an independent axis analogous to Hobson's endogenous/exogenous Input axis.

In reality, of course, these AIM axes can be rotated, so that the independent parameters of the information theory treatment could then represent appropriate topological transformations or superpositions of Hobson's A, I, and M quantities. Hobson's degree of CNS activation seems the most likely to have an exact parallel in the information theory treatment. In essence, then, the information theory perspective must be extended to a three dimensional form topologically similar, in a general sense, to the AIM model. This requires some development.

Wallace (2000, 2005, Ch. 3) argues that cognitive process can always be associated with a 'dual' information source emitting a stream of internally correlated signals having a grammar and syntax which make it a kind of language subject to the asymptotic limit theorems of information theory. The system may have to be 'coarse grained', i.e. reduced from continuous to discrete, to allow the use of information theory arguments, in a precise analog to symbolic dynamics. Note that cognition does not imply consciousness, which must be addressed through a second order iteration exactly similar to a hierarchical linear regression. That second order model, however, can also undergo phase transitions much like those we describe here.

The dual information source of a cognitive process has a source uncertainty H which is assumed to be parametized by the 'strength of weak ties' between subcomponents of a network structure upon which the dual language is 'spoken'. Again, see Wallace (2005) for a full discussion.

Strong ties are those relations between individual elements of a network of interacting modules which disjointly partition the structure into subcomponents. Weak ties are interactions which do not disjointly partition the network. These are, respectively, analogous to local and mean field couplings in a physical system.

Higher probability of weak ties, i.e. nondisjoint coupling, is associated with higher overall channel capacity, and hence higher information source uncertainty: greater possible language richness. A disjointly partitioned network will not transmit information across its segregated subgroups.

Figure 2 shows H as a function of $X = 1/P$ where P is the probability of weak ties between network/module subcomponents. In typical psychophysics style, $H(X)$ is assumed to take a reverse S-shaped form. The phase transition at X_0 arises from the homology between information source uncertainty and free energy density of a physical system (Wallace, 2005, Ch. 2, Section 7). To the left of X_0 the components can transmit information across the whole structure, to the right, the system has crystallized into its individual submodules, and ceases large-scale broadcast. This is a model of the accession to consciousness, where dispersed functional structures are recruited into a shifting global workspace.

The second stage of the argument invokes a phenomenology similar to the Onsager relations of physical theory, which determine the behavior of physical systems under nonequilibrium thermodynamic conditions. If the source uncertainty H is parametized by a vector of quantities $\mathbf{K} \equiv (K_1, ..., K_m)$, then we are interested in the *disorder*,

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

The traditional phenomenological 'generalized Onsager relations', defining system dynamics, are then

$$dK_j/dt = \sum_i L_{j,i} \partial S / \partial K_i, \quad (2)$$

where the $L_{i,j}$ are constants depending on the nature of the underlying structures. They are to be viewed empirically, in the same way as the slope and intercept parameters of a regression model, and may not at all have the simple canonical forms expected of an ordinary thermodynamic system.

These equations are taken as determining the time-dynamics of the unconstrained system under the 'forces' $\partial S/\partial K_i$. The trick, of course, lies in characterizing the constraints.

For cognitive, as opposed to physical process, Wallace (2005) argues that the general tendency will be for the system to be driven away from peaks in S: cognitive process does not like residing on a cliff, but prefers to be either at the top or bottom of an S-shaped, or reverse S-shaped curve.

In figure 2, then, there should be a strong tendency for a system already at high H to remain there, with S acting as a barrier to transition to disjoint status. Conversely, a system with $X > X_0$, a disjoint structure, would face a barrier to becoming coherent. This implies that active physiological mechanisms should be required to go from consciousness to unconsciousness, or vice versa: in the context of sleep, the arousal inhibitory mechanism and the reticular activating mechanism, respectively.

The most direct way to produce an analog to the AIM model from this structure is to impose a second reverse S-shaped curve on figure 2, on an axis *perpendicular to the* $H - X$ *plane*. This would represent the decline in richness of cognitive function as one goes from waking consciousness to REM consciousness. Figure 3a and 3b show the full model, based on a particular 'double reverse S' form of H in 3a and the resulting S in 3b. H, S, and X are as above, while Y is an appropriate parameter representing a shift in consciousness from waking to REM sleep.

Note the barriers between the states of figure 3 marked W, R, and N, presumably waking consciousness, REM sleep, and dissociated, unconscious NREM sleep. This suggests the necessity of active mechanisms, not only for going from sleep to awake and back, but for going from waking to REM sleep: Hobson's AIM model is particularly rich in its description of intermediate pathological states like trance, waking hallucinations, the effect of psychedelic drugs, and the like, and is probably applicable as well to hypnosis. The analysis suggests searching for some physiological phenomenon similar to, but separate from, the reticular activating and arousal inhibitory mechanisms. This must implement any direct transition from waking consciousness to REM sleep.

Discussion and conclusions.

Clearly equation (2), by itself, is insufficient to describe the normal sleep/wake cycle, which is not at all automatic, and requires distinct physiological mechanisms – reticular activating, arousal inhibiting and possibly others – to effect transition between the quasi-stable states W, N, and R, and to ensure the orderly progression so important to the fixation of memory. Wallace (2004, 2005, Ch. 6) has invoked higher order 'Zero Mode Identification' (ZMI) cognitive modules analogous to an adaptable retina in the operation of a broad spectrum of cognitive physiological and psychological processes ranging from the immune system and Hypothalamic-Pituitary-Adrenal (HPA) axis to emotional and conventional cognitive function. Chronic 'autocognitive' disease is then viewed as the identification of an excited mode, supposed to be only temporarily activated, as the normal, zero-state of the system. Systematic misidentification is supposed to account for autoimmune dysfunction, certain anxiety disorders, and the like (Wallace, 2004, 2005).

It is tempting, then, to postulate a generalized ZMI structure for the sleep/wake cycle, but with *three* rather than one zero-modes. Call it, for want of a better name, the Sleep Mode Identification (SMI) module, a 'retina' with three focal zones, if you will. Normal operation of the SMI is the progression $W \rightarrow N \rightarrow R \rightarrow W$ among quasi-stable states. $W \rightarrow R$ would be characteristic of hypnosis and related conditions, while recurrent 'stalling' at an intermediate – usually unstable – focal point would define chronic sleep disorders, in much the same way that chronic inflammation characterizes coronary heart disease (CHD) and chronic excitation of the blood pressure regulatory mechanism defines hypertension.

It is widely believed that a broad group of chronic 'autocognitive' disorders like hypertension, obesity, CHD, other autoimmune diseases, anxiety disorder, and the like, are developmental in nature, with roots in utero or early childhood. Chronic mind/body dysfunction, from this perspective, instantiates an image of embedding structured psychosocial or other stressors on child development (Wallace, 2004, 2005). These considerations suggest that some sleep disorders are also likely to be developmental in nature, with similar origins, possibly progressive onset, and often constituting internalized images of structured external stressors. As with many conditions, sleep disorders are likely to be comorbid with other mind/body dysfunctions.

Recent studies of insomnia seem consistent with this analysis. Bixler et al. (2002) found, for a large study sample in Central Pennsylvania, that the principal determinants of insomnia in a logistic regression model were depression, female status, and non-Caucasian ethnicity. Indeed, non-Caucasians were found to suffer insomnia at nearly twice the rate of Caucasians, 12.9 vs. 6.6 percent. Female and minority status in the US, of course, subject a person to substantial gender and racial discrimination over the entire life course. Similarly, Ohayon et al. (2002) have found working on a rotating daytime shift to cause significant sleep disturbances, resulting in higher rates of on-the-job sleepiness, work-related accidents, and sick leaves. Assignment to shift work is often a quintessential consequence of employee disempowerment. These examples indeed suggest that external stressors can write an image of themselves on sleep pattern, as well as on other physiological and psychological processes, by a variety of mechanisms.

With regard to comorbidity, Vitiello et al. (2002) found that sleep disorders in older adults are not a consequence of age itself, but rather cosegregate with other medical and psychiatric disorders and health-related burdens.

In sum, it does indeed seem possible to relate Hobson's AIM model to Baars' Global Workspace model in a 'natural' manner, based on Dretske's necessary conditions analysis as extended by the obvious homology between information source uncertainty and the free energy density of a physical system. Independent of the details of the mathematical modeling exercise, this perspective suggests that some chronic sleep disorders may be similar to many other chronic mental and physical disorders, not only by virtue of representing failure of supervisory 'mode identification' cognitive modules, but by being broadly developmental in nature. If true, some sleep disorders might significantly respond, at the population level, to highly efficient and relatively inexpensive public health interventions which do not simply shift disease management burden to those afflicted, but rather prevent suffering altogether.

References

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

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. 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.

Evans B., 2003, Sleep, consciousness and the spontaneous and evoked electrical activity of the brain. Is there a cortical integrating mechanism? *Neurophysiologie Clinique*, 33:1-10.

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

Mahowald M., 2003, Understanding consciousness, *Perspectives in Biology and Medicine*, 42:297-304.

Moruzzi G. and H. Magoun, 1949, Brain stem reticular formation and activation of the EEG, *Electroencephology and Clinical Neurophysiology*, 1:455-473.

Rauchs G., B. Desgranges, J. Foret, and F. Eustache, 2005, The relationships between memory systems and sleep stages, *Journal of Sleep Research*, 14:123-140.

Sergeant C. and S. Dehaene, 2004, Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. In press, *Psych. Sci.*

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., 2004, Comorbidity and anti comorbidity: autocognitive developmental disorders of structured psychosocial stress, *Acta Biotheoretica*, 50:101-116.

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

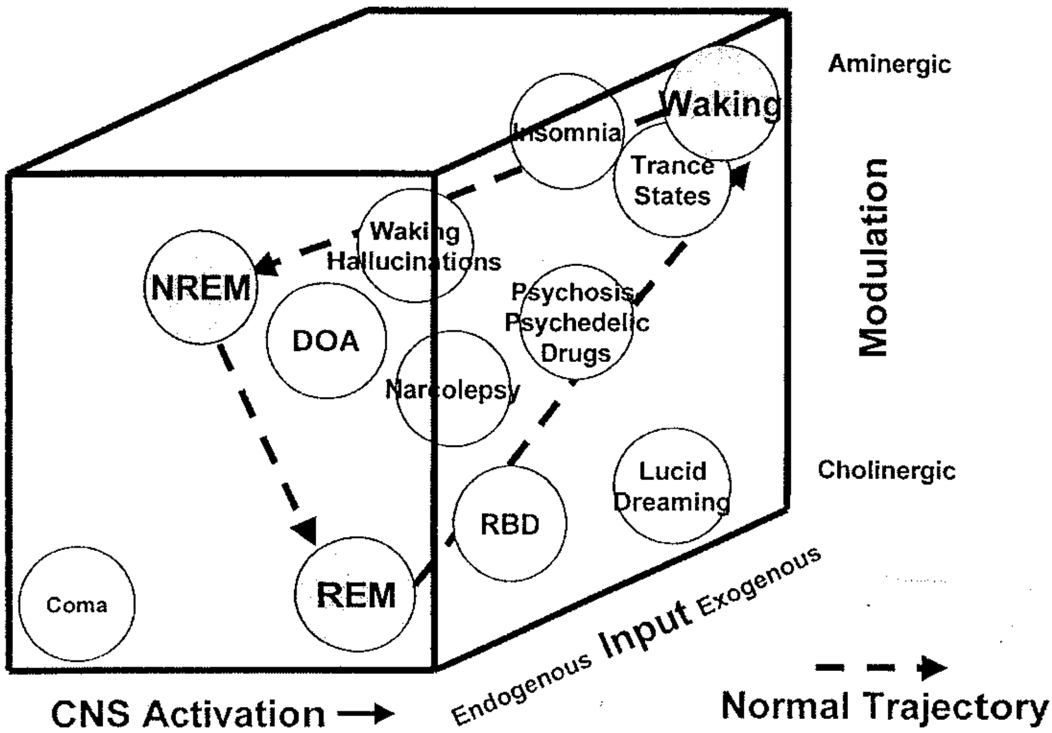
Figure captions.

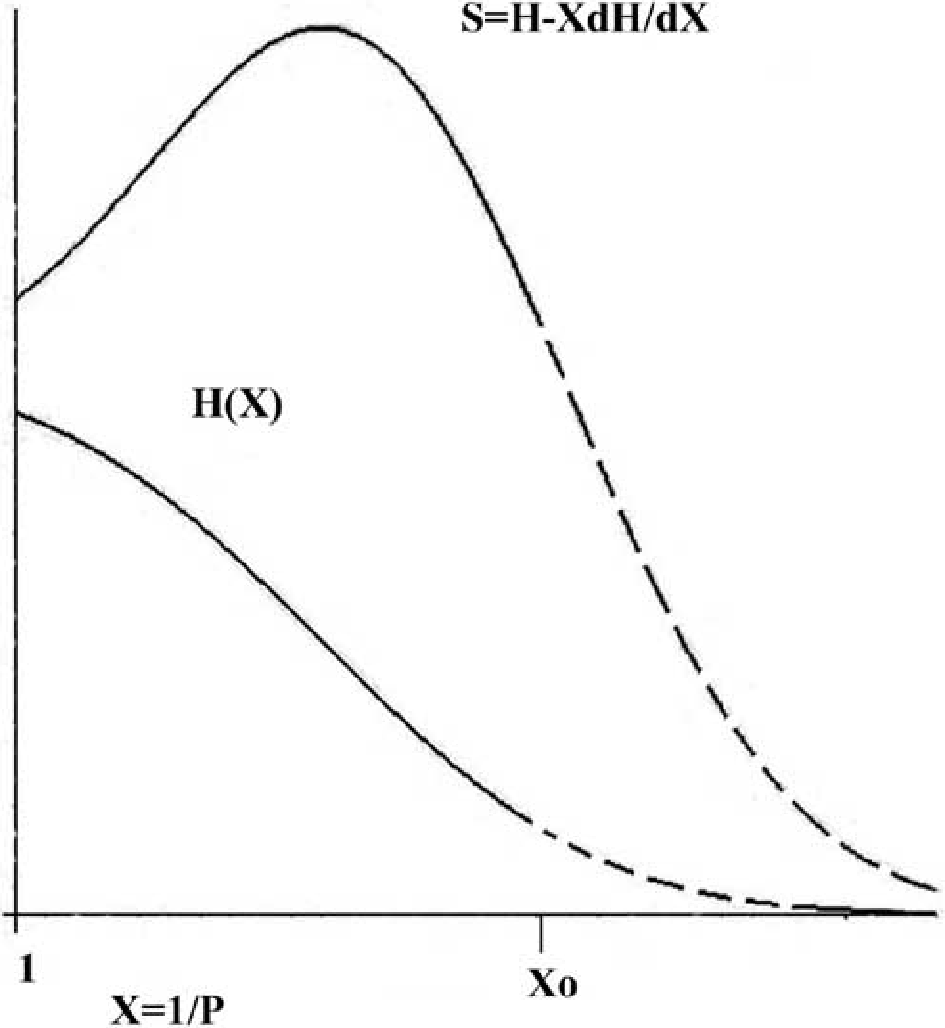
Figure 1. Hobson's AIM model (Hobson, 2001) via Mahowald (2003). Three dimensions characterize the model: (A) Level of CNS activation (low vs. high), (I) Sensory input, (internal vs. external) and (M), neurotransmitter modulation, (aminergic vs. cholinergic). Normally only small, well defined positions of this space are traversed, wakefulness, NREM sleep, and REM sleep. Under pathological circumstances, however, any portion may be occupied, resulting in altered states of consciousness and behavior, for example phases of narcolepsy are explained as admixtures of wakefulness and REM sleep.

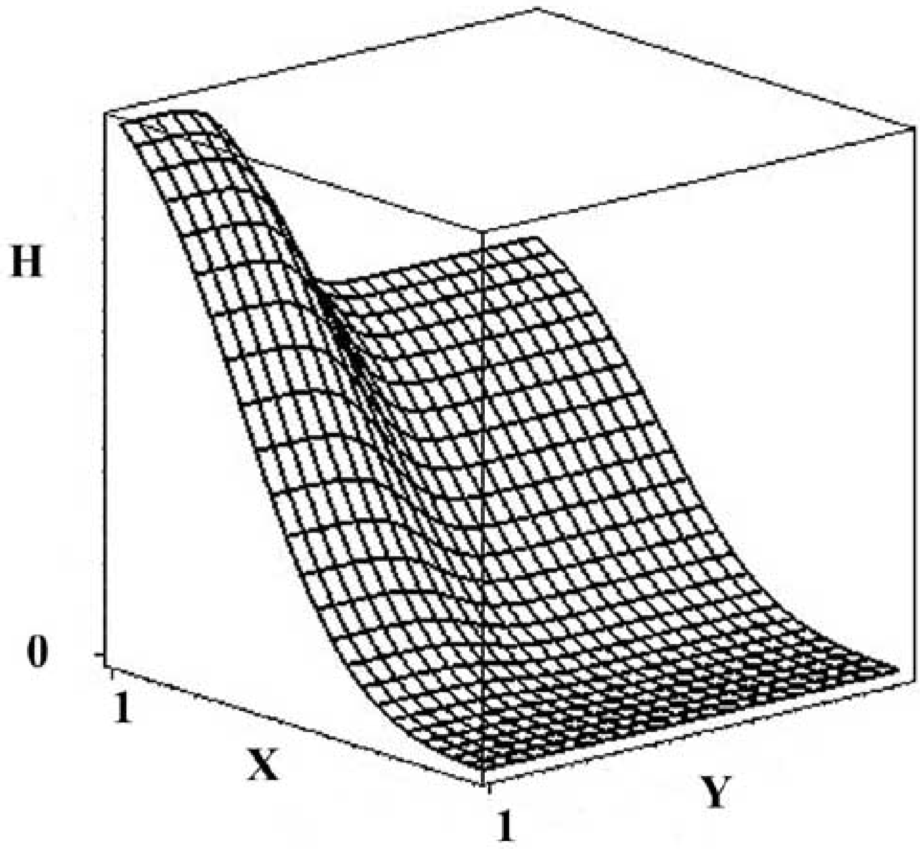
DOA = disorders of arousal (sleepwalking, sleep terrors, confusional arousals). RBD = REM sleep behavior disorder.

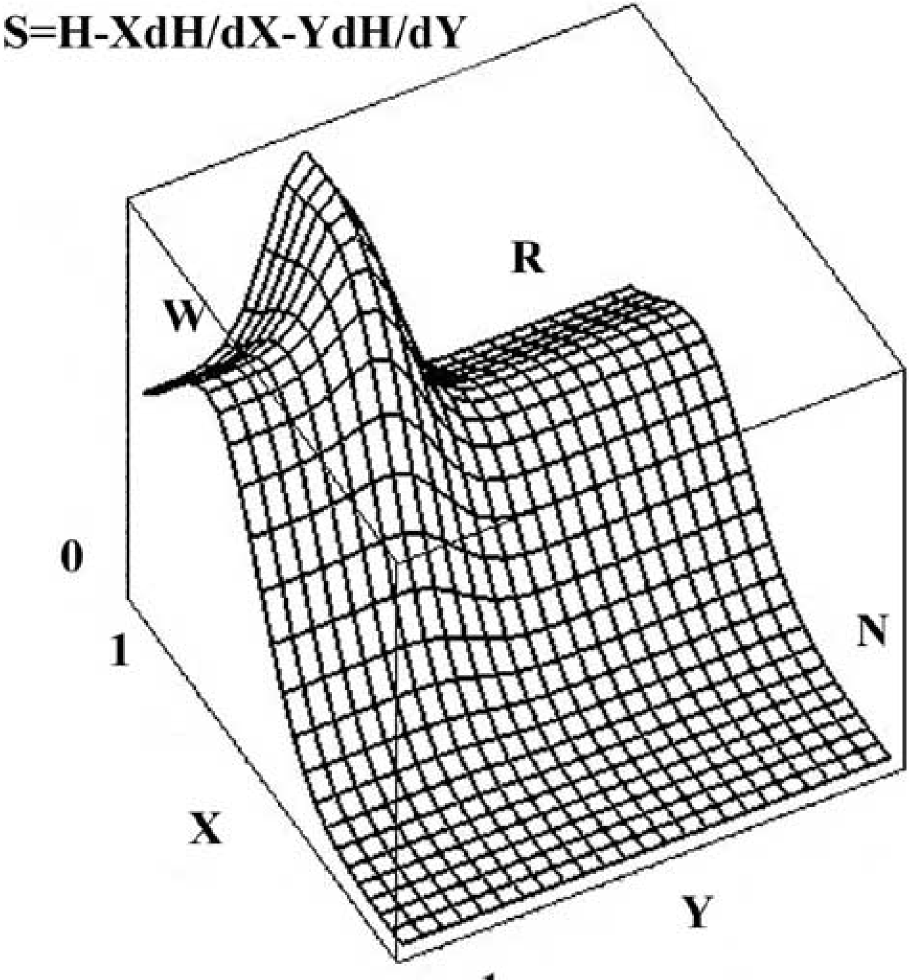
Figure 2. Two dimensional model of consciousness according to Global Workspace theory. $X = 1/P$ is the inverse probability of nondisjunctive 'weak' ties between cooperating, distant, neural modules. $H(X)$ is the source uncertainty dual to the cognitive process associated with consciousness, and $S = H - XdH/dX$ is the disorder construct. Note that X_0 is the critical point at which the workspace falls apart into component parts: consciousness ceases. According to theory, S constitutes a barrier which must be overcome by active physiological mechanisms. H is likely to be analogous to Hobson's M.

Figure 3. (a) Extension of H in figure 2 by addition of a second reverse S-shaped curve perpendicular to the H-X axis. This new curve represents the decline in 'richness' of consciousness between waking and REM sleep. (b) $S = H - XdH/dX - YdH/dY$ from H in (a). Note that three states, marked W, N and R lie in regions of relative stability, requiring, in this formulation, active physiological mechanisms for transitions between them. We infer the existence of a cognitive 'Sleep Mode Identification' module – SMI – analogous to a tunable retina whose 'focus' determines sleep progression. The normal course is $W \rightarrow N \rightarrow R$. Sleep pathologies – SMI 'focus' on intermediate states – represent, in this model, rate-distorted images of embedding structured stressors written on the SMI by developmental process.


Aminergic
Waking
Insomnia
Trance States
Modulation
Waking Hallucinations
NREM
DOA
Psychosis Psychedelic Drugs
Narcolepsy
Cholinergic
Lucid Dreaming
RBD
REM
Coma
Endogenous Input Exogenous
Normal Trajectory
CNS Activation


S=H-XdH/dX
H(X)
1
Xo
X=1/P


H
0
1
X
Y
1


S=H-XdH/dX-YdH/dY
W
R
0
1
X
Y
N
1