Version 12

The cultural epigenetics of psychopathology: The missing heritability of complex diseases found?

Rodrick Wallace
Division of Epidemiology
The New York State Psychiatric Institute*

October 14, 2009

Abstract

We extend a cognitive paradigm for gene expression based on the asymptotic limit theorems of information theory to the epigenetic epidemiology of mental disorders. In particular, we recognize the fundamental role culture plays in human biology, another heritage mechanism parallel to, and interacting with, the more familiar genetic and epigenetic systems. We do this via a model through which culture acts as another tunable epigenetic catalyst that both directs developmental trajectories, and becomes convoluted with individual ontology, via a mutually-interacting crosstalk mediated by a social interaction that is itself culturally driven. We call for the incorporation of embedding culture as an essential component of the epigenetic regulation of human mental development and its dysfunctions, bringing what is perhaps the central reality of human biology into the center of biological psychiatry. Current US work on gene-environment interactions in psychiatry must be extended to a model of gene-culture-environment interaction to avoid becoming victim of an extreme American individualism that threatens to create paradigms particular to that culture and that are, indeed, peculiar in the context of the world's cultures. The cultural and epigenetic systems of heritage may well provide the 'missing' heritability of complex diseases now under so much intense discussion.

Key Words: biological psychiatry; cognitive paradigm; gene expression; holonomy; homotopy; information source; mental disorder; spontaneous symmetry breaking; topology

*Address correspondence to: R. Wallace, 549 W. 123 St., 16F, New York, NY, 10027. wallace@pi.cpmc.columbia.edu. Affiliation is for identification only.

1 Introduction

1.1 Mental disorders and culture

Human mental disorders are not well understood. Official classifications as the *Diagnostic and Statistical Manual of Mental Disorders - fourth edition*, (DSM-IV, 1994), the standard descriptive nosology in the US, have even been characterized as 'prescientific' by P. Gilbert (2001) and others. Johnson-Laird et al. (2006) claim that current knowledge about psychological illnesses is comparable to the medical understanding of epidemics in the early 19th century. Physicians realized then that cholera, for example, was a specific disease, which killed about a third of the people whom it infected. What they disagreed about was the cause, the pathology, and the communication of the disease. Similarly, according to Johnson-Laird et al., most medical professionals these days realize that psychological illnesses occur (cf. DSMIV), but they disagree about their cause and pathology. Notwithstanding DSMIV, Johnson-Laird et al. doubt whether any satisfactory a priori definition of psychological illness can exist because it is a matter for theory to elucidate.

Atmanspacher (2006) argues that formal theory of high level cognitive process is itself at a point similar to that of physics 400 years ago, in that the basic entities, and the relations between them, have yet to be delineated.

More generally, simple arguments from genetic determinism regarding mental disorders fail, in part because of a draconian population bottleneck that, early in our species' history, resulted in an overall genetic diversity less than that observed within and between contemporary chimpanzee subgroups. Manolio et al. (2009) describe this conundrum more generally in terms of 'finding the missing heritability of complex diseases'. They observe, for example, that at least 40 loci have been associated with human height, a classic complex trait with an estimated heritability of about 80 %, yet they explain only about 5 % of phenotype variance despite studies of tens of thousands of people. This result, they find, is typical across a broad range of supposedly heritable diseases, and call for extending beyond current genome-wide assoication approaches to illuminate the genetics of complex diseases and enhance its potential to enable effective disease prevention or treatment.

Arguments from psychosocial stress fare better (e.g., Brown et al., 1973; Dohrenwend and Dohrenwend, 1974; Eaton, 1978), particularly for depression (e.g., Risch et al., 2009), but are affected by the apparently complex and contingent developmental paths determining the onset of schizophrenia, dementias, psychoses, and so forth, some of which may be triggered in utero by exposure to infection, low birthweight, or other functional teratogens.

P. 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) outline something of the cross cultural subtleties affecting the study of depression that seem to argue against any simple evolutionary or genetic interpretation. They state that, 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, they continue, 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.

Since publication of that landmark study, a number of comprehensive overviews have been published that support its conclusions, for example Bebbington (1993), Jenkins, Kleniman and Good (1990), *Journal of Clinical Psychiatry* (Supplement 13), and Manson (1995). As Marsella (2003) writes, it is now clear that cultural variations exist in the areas of meaning, perceived causes, onset patterns, epidemiology, symptom expression, course, and outcome, variations having important implications for understanding clinical activities including conceptualization, assessment, and therapy.

Kleinman and Cohen (1997) argue forcefully that several myths have become central to psychiatry. The first is that the forms of mental illness everywhere display similar degrees of prevalence. The 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.

Heine (2001) describes an explicit cultural psychology that views the person as containing a set of biological potentials interacting with particular situational contexts that constrain and afford the expression of various constellations of traits and patterns of behavior. He says that, unlike much of personality psychology, cultural psychology focuses on the constraints and affordances inherent

to the cultural environment that give shape to those biological potentials. Human nature, from this perspective, is seen as emerging from participation in cultural worlds, and of adapting oneself to the imperatives of cultural directives, meaning that our nature is ultimately that of a cultural being.

Heine describes how 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, so that without culture there is no self, only a biological entity deprived of its potential. Individual selves, from Heine's perspective, are inextricably grounded in a configuration of consensual understandings and behavioral customs particular to a given cultural and historical context, so that understanding the self requires an understanding of the culture that sustains it. Heine argues, then, that the process of becoming a self is contingent on individuals interacting with, and seizing meanings from, the cultural environment.

Heine warns that the extreme nature of American individualism means that a psychology based on late 20th century American research not only stands the risk of developing models that are particular to that culture, but of developing an understanding of the self that is peculiar in the context of the world's cultures.

Indeed, as Norenzayan and Heine (2005) point out, for the better part of a hundred years, a considerable controversy has raged within anthropology regarding the degree to which psychological and other human universals do, in fact, actually exist independent of the particularities of culture.

Arnett (2008), in a paper provocatively titled *The Neglected 95 %*, similarly argues that US psychological research focuses too narrowly on Americans, who comprise less than 5 percent of the world's population, and on perhaps another 7 percent in Western countries. He asserts that the majority of the world's population lives in under vastly different conditions, underlying doubts of how representative American psychological research can be, and finds the narrowness of American research to be a consequence of a focus on a philosophy of science that emphasizes fundamental processes and ignores or strips away cultural context.

Henrich et al. (2009), in a wide-ranging review, extend the considerations of Norenzayan and Heine, finding that Western, educated, industrialized and democratic (WEIRD) subjects, across domains of visual perception, fairness, categorization, spatial cognition, memory, moral reasoning, and self-concepts, constitute frequent outliers compared with the rest of the species. They conclude that addressing questions of *human* psychology will require tapping broader subject pools.

As Durham (1991) and Richerson and Boyd (2004) explore at some length, humans are endowed with two distinct but interacting heritage systems: genes and culture. Durham (1991), for example, writes that genes and culture constitute two distinct but interacting systems of information inheritance within human populations and information of both kinds has influence, actual or potential, over behaviors, which creates a real and unambiguous symmetry between genes and phenotypes on the one hand, and culture and phenotypes on the other. Genes and culture, in his view, are best represented as two parallel lines

or tracks of hereditary influence on phenotypes.

Both genes and culture can be envisioned as generalized languages in that they have recognizable 'grammar' and 'syntax', in the sense of Wallace (2005) and Wallace and Wallace (2008, 2009).

More recent work has identified epigenetic heritage mechanisms involving such processes as environmentally-induced gene methylation, that can have strong influence across several generations (e.g., Jablonka and Lamb, 1995, 1998; Jabolonka, 2004), and are the subject of intense current research.

There are, it seems, two powerful heritage mechanisms in addition to the genetic where one may perhaps find the 'missing heritability of complex diseases' that Manolio et al. seek.

Here, however, we are particularly interested in the phenotypes of madness, and in their relations to genes, culture, and environment.

1.2 Two case histories

1.2.1 Gene-environment interaction

Much recent work in American biological psychiatry has emphasized the search for gene-environment interactions. Caspi and Moffitt (2006), for example, claim that such interactions occur when the effect of exposure to an environmental pathogen on a person's health is conditional on his or her genotype. The first evidence that genotype moderates the capacity of an environmental risk to bring about mental disorders was, according to them, reported in 2002, (Caspi et al., 2002), in a study of the role of genotype in the cycle of violence in maltreated children. Caspi and Moffitt (2006) claim that the gene-environment interaction approach brings opportunities for extending the range and power of neuroscience by introducing opportunities for collaboration between experimental neuroscience and research on gene-environment interactions. Successful collaboration can, in their view, solve the biggest mystery of human psychopathology: how does an environmental factor, external to the person, get inside the nervous system and alter its elements to generate the symptoms of a disordered mind? Concentrating the considerable resources of neuroscience and gene-environment interaction on this question will, they claim, bring discoveries that advance the understanding of mental disorders, and increase the potential to control and prevent them.

One of the most cited of recent studies of gene-environment interactions is, indeed, the work of Caspi et al. (2003), who found that genetic variation in the promoter region of the serotonin transporter gene (5-HTTLPR;[OMIM182138]), in interaction with stressful life events, contributes to a predisposition to major depression. As Risch et al. (2009) put it, this result was striking and potentially paradigm shifting because numerous previous studies of this same polymorphism, without examining environmental risk factors or life events, had not consistently shown either a strong or replicated association with depression. A subsequent meta-analysis was conducted by Risch et al. (2009) that combined data from some 14 studies having a total of 14,250 participants, some 1769 of

whom met criteria for depression.

Risch et al. state that most of the participants were white, except for a multiethnic sample in one study, and an Asian sample in another. Contrary to the results of Caspi et al. (2003), they found no evidence that the serotonin transporter genotype alone, or in interaction with stressful life events, is associated with an elevated risk of depression.

The Asian study, by J. Kim et al. (2006), involved 732 Korean community residents ages 65+, a fair number indeed. Some 88 percent at baseline did not meet criteria for depression. Kim et al., in contrast with Risch et al., in spite of using 'standard' instruments for both measures of depression and life events (translated into Korean), found a strong statistical trend suggesting that environmental risk of depression is indeed modified by at least two genes, and that gene-environment interactions are found even into old age.

Given the scathing analyses by Arnett, Heine, and Henrich et al., the bitter conflict between the results of Caspi et al. (2003) and Risch et al. (2009) is in serious danger of becoming simply a culture-bound tempest in a distinctly American teapot.

1.2.2 Gene-culture interaction

The necessity for the inclusion of culture in the operation of fundamental psychological phenomena is emphasized by the observations of Nisbett et al. (2001), and others, following the tradition of Markus and Kitayama (1991), regarding profound differences in basic 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 built into them, and that users accept these theories, albeit unknowingly, when they use these tools.

More recently, Masuda and Nisbett (2006) examined cultural variations in change blindness, a phenomenon related to inattentional blindness, and found striking differences between Western and East Asian subjects. They presented participants with still photos and with animated vignettes having changes in

focal object information and contextual information. Compared to Americans, East Asians were more sensitive to contextual changes than to focal object changes. These results, they conclude, suggest that there can be cultural variation in what may seem to be basic perceptual processes.

H. Kim et al. (2010) have extended this line of work to examine the interaction between genes and culture as determinants of individuals' locus of attention. As the serotonin (5-HT) system has been associated with attentional focus and the ability to adapt to changes in reinforcement, they examined the serotonin 1A receptor polymorphism (5-HTR1A). Koreans and European Americans were geneotyped and reported their chronic locus of attention. They found a significant interaction between 5-HTR1A and culture in the locus of attention. Koreans reported attending to the field more than European Americans, and this cultural difference was moderated by 5-HTR1A. There was a linear pattern such that those homozygous for the G allele, which is associated with reduced ability to adapt to changes in reinforcement, more strongly endorsed the culturally reinforced mode of thinking than those homozygous for the C allele, with those heterozygous in the middle. Kim et al. claim that their findings suggest that the same genetic predisopsition can result in divergent psychological outcomes, depending on an individual's cultural context.

The sample used in this study included 149 Korean and 140 European subjects. Given the problems with the Caspi et al. work, it is clear that replication across larger samples will be needed.

That being said, the results of H. Kim et al. do indeed underline the necessity of expanding work on psychiatric disorders to gene-culture-environment interactions. It seems likely, however, that, overall, culture-environment interaction effects will predominate. Nonetheless, the effects of genetic structure on that interaction might well provide important insights as to etiology and possible treatment.

1.3 Global broadcast models

Recent research on schizophrenia, dyslexia, and autism, broadly supports a 'brain connectivity' model for these disorders that is of considerable interest from the perspective of global workspace/global broadcast models of consciousness that are the foundation of our work (e.g., Baars, 1989; Wallace, 2005), since large-scale brain connectivity is essential for the operation of consciousness, a principal, and very old, evolutionary adaptation (e.g., Wallace and Wallace, 2009).

For example, Burns et al. (2003), on the basis of sophisticated diffusion tensor magnetic resonance imaging studies, argue 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 Global Workspace/Broadcast models of individual mental function including consciousness (e.g., Baars, 1989; Wallace, 2005), 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. For humans, of course, one of the principal regulatory mechanisms for mental function is the embedding culture and culturally-mediated social interaction, in addition to culture's role as the second system of human heritage. As the evolutionary anthropologist Robert Boyd has put it, culture is as much a part of human biology as the enamel on our teeth (e.g., Richerson and Boyd, 2004).

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. Yet the experience of individual consciousness cannot be disentangled from social interaction and culture (Wallace, 2005). Distortion of a culturally-mediated consciousness in mental disorders inhibits both rou-

tine social interchange 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 causing decoupling from social and cultural context is usually a threat to long-term individual survival, and those with mental disorders significantly affecting consciousness typically experience severely shortened lifespans.

Here we will expand recent explorations of a cognitive paradigm for gene expression (Wallace and Wallace, 2008, 2009) that incorporates the effects of surrounding epigenetic regulatory machinery as a kind of catalyst to include the effects of the embedding information source of human culture on the ontology of the human mind. The essential feature is that a cognitive process, including gene expression, can instantiate a dual information source that can interact with the generalized language of culture in which, for example, social interplay and the interpretation of socioeconomic and environmental stressors, involve complicated matters of symbolism and its grammar and syntax. These information sources interact by a crosstalk that, over the life course, determines the ontology of mind, including its manifold dysfunctions.

That is, contemporary American work on gene-environment interactions in psychiatry must be extended to the study of gene-culture-environment interactions. This is no small matter, and the probability models we present here are at the borders of current applied mathematics.

1.4 A cognitive paradigm for gene expression

A cognitive paradigm for gene expression has emerged in which contextual factors determine the behavior of what Cohen characterizes as a 'reactive system', not at all a deterministic, or even stochastic, mechanical process (e.g., Cohen, 2006; Cohen and Harel, 2007; Wallace and Wallace, 2008, 2009). The very different formal approaches are, however, all in the spirit of Maturana and Varela (1980, 1992) who understood the central role that cognitive process must play across a vast array of biological phenomena.

O'Nuallain (2008) has placed gene expression firmly in the realm of complex linguistic behavior, for which context imposes meaning, claiming that the analogy between gene expression and language production is useful both as a fruitful research paradigm and also, given the relative lack of success of natural language processing (nlp) by computer, as a cautionary tale for molecular biology. A relatively simple model of cognitive process as an information source permits use of Dretske's (1994) insight that any cognitive phenomenon must be constrained by the limit theorems of information theory, in the same sense that sums of stochastic variables are constrained by the Central Limit Theorem. This perspective permits a new formal approach to gene expression and its dysfunctions, in particular suggesting new and powerful statistical tools for data analysis that could contribute to exploring both ontology and its pathologies. Wallace and Wallace (2009, 2010) apply the perspective, respectively, to infec-

tious and chronic disease. Here we extend the mathematical foundations of that work to examine the topological structures of development and developmental disorder, in the context of an embedding information source representing the compelling varieties of human culture.

This approach is consistent with the broad context of epigenetics and epigenetic epidemiology. In particular, Jablonka and Lamb (1995, 1998) argue that information can be transmitted from one generation to the next in ways other than through the base sequence of DNA. It can be transmitted through cultural and behavioral means in higher animals, and by epigenetic means in cell lineages. All of these transmission systems allow the inheritance of environmentally induced variation. Such Epigenetic Inheritance Systems are the memory systems that enable somatic cells of different phenotypes but identical genotypes to transmit their phenotypes to their descendants, even when the stimuli that originally induced these phenotypes are no longer present.

After a decade of research and debate, the epigenetic perspective has received much empirical confirmation (e.g., Backdahl et al. 2009; Turner, 2000; Jaenish and Bird, 2003; Jablonka, 2004).

Foley et al. (2009) argue that epimutation is estimated to be 100 times more frequent than genetic mutation and may occur randomly or in response to the environment. Periods of rapid cell division and epigenetic remodeling are likely to be most sensitive to stochastic or environmentally mediated epimutation. Disruption of epigenetic profile is a feature of most cancers and is speculated to play a role in the etiology of other complex diseases including asthma allergy, obesity, type 2 diabetes, coronary heart disease, autism spectrum disorders and bipolar disorders and schizophrenia.

Important work by Scherrer and Jost (2007a, b) that is similar to the approach of this paper explicitly invokes information theory in their extension of the definition of the gene to include the local epigenetic machinery, a construct they term the 'genon'. Their central point is that coding information is not simply contained in the coded sequence, but is, in their terms, *provided by* the genon that accompanies it on the expression pathway and controls in which peptide it will end up. In their view the information that counts is not about the identity of a nucleotide or an amino acid derived from it, but about the relative frequency of the transcription and generation of a particular type of coding sequence that then contributes to the determination of the types and numbers of functional products derived from the DNA coding region under consideration.

The proper formal tools for understanding phenomena that 'provide' information – that are information sources – are the Rate Distortion Theorem and its zero error limit, the Shannon-McMillan Theorem.

2 Models of development

The currently popular spinglass model of development (e.g., Ciliberti et al., 2007a, b) assumes that N transcriptional regulators, are represented by their expression patterns

$$\mathbf{S}(t) = [S_1(t), ..., S_N(t)] \tag{1}$$

at some time t during a developmental or cell-biological process and in one cell or domain of an embryo. The transcriptional regulators influence each other's expression through cross-regulatory and autoregulatory interactions described by a matrix $w = (w_{ij})$. For nonzero elements, if $w_{ij} > 0$ the interaction is activating, if $w_{ij} < 0$ it is repressing. w represents, in this model, the regulatory genotype of the system, while the expression state $\mathbf{S}(t)$ is the phenotype. These regulatory interactions change the expression of the network $\mathbf{S}(t)$ as time progresses according to a difference equation

$$S_i(t + \Delta t) = \sigma[\textstyle\sum_{j=1}^{N} w_{ij} S_j(t)], \tag{2}$$

where Δt is a constant and σ a sigmodial function whose value lies in the interval $(-1, 1)$. In the spinglass limit σ is the sign function, taking only the values ± 1.

The regulatory networks of interest here are those whose expression state begins from a prespecified initial state $\mathbf{S}(0)$ at time $t = 0$ and converge to a prespecified stable equilibrium state $\mathbf{S}_\infty$. Such networks are termed *viable* and must necessarily be a very small fraction of the total possible number of networks, since most do not begin and end on the specified states. This 'simple' observation is not at all simple in our reformulation, although other results become far more accessible, as we can then invoke the asymptotic limit theorems of information theory.

The spinglass approach to development is formally similar to spinglass neural network models of learning by selection, e.g., as proposed by Toulouse et al. (1986) nearly a generation ago. Much subsequent work, summarized by Dehaene and Naccache (2001), suggests that such models are simply not sufficient to the task of understanding high level cognitive function, and these have been largely supplanted by complicated 'global workspace' concepts whose mathematical characterization is highly nontrivial (Atmanspacher, 2006).

Wallace and Wallace (2008, 2009) shift the perspective on development by invoking a cognitive paradigm for gene expression, following the example of the Atlan/Cohen model of immune cognition.

Atlan and Cohen (1998), in the context of a study of the immune system, argue that the essence of cognition is the comparison of a perceived signal with

an internal, learned picture of the world, and then choice of a single response from a large repertoire of possible responses.

Such choice inherently involves information and information transmission since it always generates a reduction in uncertainty, as explained by Ash (1990, p. 21).

More formally, a pattern of incoming input – like the $\mathbf{S}(t)$ above – is mixed in a systematic algorithmic manner with a pattern of internal ongoing activity – like the (w_{ij}) above – to create a path of combined signals $x = (a_0, a_1, ..., a_n, ...)$ – analogous to the sequence of $\mathbf{S}(t + \Delta t)$ above, with, say, $n = t/\Delta t$. Each a_k thus represents some functional composition of internal and external signals.

This path is fed into a highly nonlinear decision oscillator, h, a 'sudden threshold machine', in a sense, that 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 us define the sets B_k as

$$B_0 \equiv \{b_0, ..., b_k\},$$
$$B_1 \equiv \{b_{k+1}, ..., b_m\}. \tag{3}$$

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 has been recognized, and some action $b_j, k + 1 \leq j \leq m$ takes place.

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

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

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

$$H \equiv \lim_{n\to\infty} \frac{\log[N(n)]}{n} \tag{4}$$

both exists and is independent of the path x.

Define such a pattern recognition-and-response cognitive process as *ergodic*. Not all cognitive processes are likely to be ergodic in this sense, implying that H, if it indeed exists at all, is path dependent, although extension to nearly ergodic processes seems possible (Wallace and Fullilove, 2008).

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

$$\begin{array}{l} H[\mathbf{X}] = \lim_{n\to\infty} \frac{\log[N(n)]}{n} \\ = \lim_{n\to\infty} H(X_n|X_0, ..., X_{n-1}) \\ = \lim_{n\to\infty} \frac{H(X_0,...,X_n)}{n+1}. \end{array} \tag{5}$$

This information source is defined as *dual* to the underlying ergodic cognitive process.

Adiabatic means that the source has been parametrized according to some scheme, and that, over a certain range, along a particular piece, as the parameters vary, the source remains as close to stationary and ergodic as needed for information theory's central theorems to apply. *Stationary* means that the system's probabilities do not change in time, and *ergodic*, roughly, that the cross sectional means approximate long-time averages. Between pieces it is necessary to invoke various kinds of phase transition formalisms, as described more fully in Wallace (2005) or Wallace and Wallace (2008).

In the developmental vernacular of Ciliberti et al., we now examine paths in phenotype space that begin at some $\mathbf{S}_0$ and converge $n = t/\Delta t \to \infty$ to some other $\mathbf{S}_\infty$. Suppose the system is conceived at $\mathbf{S}_0$, and h represents (for example) reproduction when phenotype $\mathbf{S}_\infty$ is reached. Thus $h(x)$ can have two values, i.e., B_0 not able to reproduce, and B_1, mature enough to reproduce. Then $x = (\mathbf{S}_0, \mathbf{S}_{\Delta t}, ..., \mathbf{S}_{n\Delta t}, ...)$ until $h(x) = B_1$.

Structure is now subsumed *within the sequential grammar and syntax of the dual information source* rather than within the cross sectional internals of (w_{ij})-space, a simplifying shift in perspective.

This transformation carries computational burdens, as well as providing mathematical insight.

First, the fact that viable networks comprise a tiny fraction of all those possible emerges easily from the spinglass formulation simply because of the 'mechanical' limit that the number of paths from $\mathbf{S}_0$ to $\mathbf{S}_\infty$ will always be far smaller than the total number of possible paths, most of which simply do not end on the target configuration.

From the information source perspective, which inherently subsumes a far larger set of dynamical structures than possible in a spinglass model – not simply those of symbolic dynamics – the result is what Khinchin (1957) characterizes as the 'E-property' of a stationary, ergodic information source. This property allows, in the limit of infinitely long output, the classification of output strings into two sets:

[1] a very large collection of gibberish which does not conform to underlying (sequential) rules of grammar and syntax, in a large sense, and which has near-zero probability, and

[2] a relatively small 'meaningful' set, in conformity with underlying structural rules, having very high probability.

The essential content of the Shannon-McMillan Theorem is that, if $N(n)$ is the number of meaningful strings of length n, then the uncertainty of an information source X can be defined as

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

that can be expressed in terms of joint and conditional probabilities. Proving these results for general stationary, ergodic information sources requires considerable mathematical machinery (e.g., Khinchin, 1957; Cover and Thomas, 1991; Dembo and Zeitouni, 1998).

Second, according to Ash (1990) information source uncertainty has an important heuristic interpretation in that we may regard a portion of text in a particular language as being produced by an information source. A large uncertainty means, by the Shannon-McMillan Theorem, a large number of 'meaningful' sequences. Thus given two languages with uncertainties H_1 and H_2 respectively, if $H_1 > H_2$, then in the absence of noise it is easier to communicate in the first language; more can be said in the same amount of time. On the other hand, it will be easier to reconstruct a scrambled portion of text in the second language, since fewer of the possible sequences of length n are meaningful.

Third, information source uncertainty is homologous with free energy density in a physical system, a matter having implications across a broad class of dynamical behaviors.

The free energy density of a physical system having volume V and partition function $Z(K)$ derived from the system's Hamiltonian – the energy function – at inverse temperature K is (e.g., Landau and Lifshitz 2007)

$$F[K] = \lim_{V\to\infty} -\frac{1}{K}\frac{\log[Z(K,V)]}{V}$$
$$= \lim_{V\to\infty} \frac{\log[\hat{Z}(K,V)]}{V}, \tag{6}$$

where $\hat{Z} = Z^{-1/K}$.

The partition function for a physical system is the normalizing sum in an equation having the form

$$P[E_i] = \frac{\exp[-E_i/kT]}{\sum_j \exp[-E_j/kT]} \tag{7}$$

where E_i is the energy of state i, k a constant, and T the system temperature.

Feynman (2000), following the classic approach by Bennett (1988), who examined idealized machines using information to do work, concludes that *the information contained in a message is most simply measured by the free energy needed to erase it.*

Thus, according to this argument, source uncertainty is homologous to free energy density as defined above, i.e., from the similarity with the relation $H = \lim_{n\to\infty} \log[N(n)]/n$.

Ash's perspective then has an important corollary: If, for a biological system, $H_1 > H_2$, source 1 will require more metabolic free energy than source 2.

3 Tunable epigenetic catalysis

Following the direction of Wallace and Wallace (2009), incorporating the influence of embedding contexts – generalized epigenetic effects – is most elegantly done by invoking the Joint Asymptotic Equipartition Theorem (JAEPT) (Cover and Thomas, 1991). For example, given an embedding epigenetic information source, say Y, that affects development, then the dual cognitive source uncertainty $H[X]$ is replaced by a joint uncertainty $H(X,Y)$. The objects of interest then become the jointly typical dual sequences $z^n = (x^n, y^n)$, where x is associated with cognitive gene expression and y with the embedding epigenetic regulatory context. Restricting consideration of x and y to those sequences that are in fact jointly typical allows use of the information transmitted from Y to X as the splitting criterion.

One important inference is that, from the information theory 'chain rule' (Cover and Thomas, 1991), $H(X,Y) = H(X)+H(Y|X) \leq H(X)+H(Y)$, while there are approximately $\exp[nH(X)]$ typical X sequences, and $\exp[nH(Y)]$ typical Y sequences, and hence $\exp[n(H(x)+H(Y))]$ independent joint sequences, there are only

$$\exp[nH(X,Y)] \leq \exp[n(H(X)+H(Y))]$$

jointly typical sequences, so that the effect of the embedding context is to lower the relative free energy of a particular developmental channel. Equality occurs only for stochastically independent processes.

Thus the effect of epigenetic regulation is to channel development into pathways that might otherwise be inhibited by an energy barrier. Hence the epigenetic information source Y acts as a *tunable catalyst*, a kind of second order

cognitive enzyme, to enable and direct developmental pathways. This result permits hierarchical models similar to those of higher order cognitive neural function that incorporate contexts in a natural way (e.g., Wallace and Wallace, 2008; Wallace and Fullilove, 2008). The cost of this ability to channel is the metabolic necessity of supporting two information sources, X and Y, rather than just Y itself.

This elaboration allows a spectrum of possible 'final' phenotypes, what S. Gilbert (2001) calls developmental or phenotype plasticity. Thus gene expression is seen as, in part, responding to environmental or other, internal, developmental signals.

Including the effects of embedding culture in the development of the human mind is, according to this formalism, quite straightforward: Consider culture as another embedding information source, Z, having source uncertainty $H(Z)$. Then the information chain rule becomes

$$H(X,Y,Z) \leq H(X) + H(Y) + H(Z) \tag{8}$$

and

$$\exp[nH(X,Y,Z)] \leq \exp[n(H(X) + H(Y) + H(Z))], \tag{9}$$

where, again, equality occurs only under stochastic independence.

In this model, following explicitly the direction indicated by Boyd, Kleinman and their colleagues, culture is seen here as an essential component of the catalytic epigenetic machinery that regulates the development of the human mind. This is not to say that the development of mind in other animals, particularly those that are highly social, does not undergo analogous regulation by larger scale structures of interaction. For human populations, however, social relations are themselves very highly regulated through an often strictly formalized cultural grammar and syntax.

4 Spontaneous symmetry breaking

A formal equivalence class algebra can now be constructed by choosing different origin and end points $\mathbf{S}_0, \mathbf{S}_\infty$ and defining equivalence of two states by the existence of a high probability meaningful path connecting them with the same origin and end. Disjoint partition by equivalence class, analogous to orbit

equivalence classes for dynamical systems, defines the vertices of the proposed network of cognitive dual languages, much enlarged beyond the spinglass example. We thus envision a network of metanetworks. Each vertex then represents a different equivalence class of information sources dual to a cognitive process. This is an abstract set of metanetwork 'languages' dual to the cognitive processes of gene expression and development.

This structure generates 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 to the same base and end points, and tuning across the various possible ways in which that can happen – the different cognitive languages – parameterizes the set of equivalence relations and creates the (very large) groupoid. See the mathematical appendix for a summary of standard material on groupoids.

There is a hierarchy in groupoid structures. First, there is structure *within the system having the same base and end points*, as in Ciliberti et al. Second, there is a complicated groupoid structure defined by sets of dual information sources surrounding the variation of base and end points. We do not need to know what that structure is in any detail, but can show that its existence has profound implications.

First we examine the simple case, the set of dual information sources associated with a fixed pair of beginning and end states.

Taking the serial grammar/syntax model above, we find that not all high probability meaningful paths from $\mathbf{S}_0$ to $\mathbf{S}_\infty$ are the same. They are structured by the uncertainty of the associated dual information source, and that has a homological relation with free energy density.

Let us index possible dual information sources connecting base and end points by some set $A = \cup\alpha$. Argument by abduction from statistical physics is direct: Given metabolic energy density available at a rate M, and an allowed development time τ, let $K = 1/\kappa M\tau$ for some appropriate scaling constant κ, so that $M\tau$ is total developmental free energy. Then the probability of a particular H_α will be determined by the standard expression (e.g., Landau and Lifshitz, 2007),

$$P[H_\beta] = \frac{\exp[-H_\beta K]}{\sum_\alpha \exp[-H_\alpha K]}, \tag{10}$$

where the sum may, in fact, be a complicated abstract integral.

This is just a version of the fundamental probability relation from statistical mechanics, as above. The sum in the denominator, the partition function in statistical physics, is a crucial normalizing factor that allows the definition of of $P[H_\beta]$ as a probability.

A basic requirement, then, is that the sum/integral always converges. K is the inverse product of a scaling factor, a metabolic energy density rate term, and

a characteristic development time τ. The developmental energy might be raised to some power, e.g., $K = 1/(\kappa(M\tau)^b)$, suggesting the possibility of allometric scaling.

Some dual information sources will be 'richer'/smarter than others, but, conversely, must use more metabolic energy for their completion.

While we might simply impose an equivalence class structure based on equal levels of energy/source uncertainty, producing a groupoid, we can do more by now allowing both source and end points to vary, as well as by imposing energy-level equivalence. This produces a far more highly structured groupoid that we now investigate.

Equivalence classes define groupoids, by standard mechanisms (e.g., Weinstein, 1996; Brown, 1987; Golubitsky and Stewart, 2006). The basic equivalence classes – here involving both information source uncertainty level and the variation of $\mathbf{S}_0$ and $\mathbf{S}_\infty$, will define transitive groupoids, and higher order systems can be constructed by the union of transitive groupoids, having larger alphabets that allow more complicated statements in the sense of Ash above.

Again, given an appropriately scaled, dimensionless, fixed, inverse available metabolic energy density rate and development time, so that $K = 1/\kappa M\tau$, we propose that the metabolic-energy-constrained probability of an information source representing equivalence class D_i, H_{D_i}, will again be given by the classic relation

$$P[H_{D_i}] = \frac{\exp[-H_{D_i}K]}{\sum_j \exp[-H_{D_j}K]},$$

where the sum/integral is over all possible elements of the largest available symmetry groupoid. By the arguments of Ash above, compound sources, formed by the union of underlying transitive groupoids, being more complex, generally having richer alphabets, as it were, will all have higher free-energy-density-equivalents than those of the base (transitive) groupoids.

Let $Z_D \equiv \sum_j \exp[-H_{D_j}K]$. We now define the *Groupoid free energy* of the system, F_D, at inverse normalized metabolic energy density K, as

$$F_D[K] \equiv -\tfrac{1}{K}\log[Z_D[K]], \tag{11}$$

again following the standard arguments from statistical physics (again, Landau and Lifshitz, 2007, or Feynman, 2000).

The groupoid free energy construct permits introduction of important ideas from statistical physics.

We have expressed the probability of an information source in terms of its relation to a fixed, scaled, available (inverse) metabolic free energy density, seen as a kind of equivalent (inverse) system temperature. This gives a statistical

thermodynamic path leading to definition of a 'higher' free energy construct – $F_D[K]$ – to which we now apply Landau's fundamental heuristic phase transition argument (Landau and Lifshitz 2007; Skierski et al. 1989; Pettini 2007). See, in particular, Pettini (2007) for details.

Landau's insight was that second order phase transitions were usually in the context of a significant symmetry change in the physical states of a system, with one phase being far more symmetric than the other. A symmetry is lost in the transition, a phenomenon called spontaneous symmetry breaking, and symmetry changes are inherently punctuated. The greatest possible set of symmetries in a physical system is that of the Hamiltonian describing its energy states. Usually states accessible at lower temperatures will lack the symmetries available at higher temperatures, so that the lower temperature phase is less symmetric: The randomization of higher temperatures – in this case limited by available metabolic free energy densities – ensures that higher symmetry/energy states – mixed transitive groupoid structures – will then be accessible to the system. Absent high metabolic free energy rates and densities, however, only the simplest transitive groupoid structures can be manifest. A full treatment from this perspective seems to require invocation of groupoid representations, no small matter (e.g., Buneci, 2003; Bos 2006).

Something like Pettini's (2007) Morse-Theory-based topological hypothesis can now be invoked, i.e., that changes in underlying groupoid structure are a necessary (but not sufficient) consequence of phase changes in $F_D[K]$. Necessity, but not sufficiency, is important, as it, in theory, allows mixed groupoid symmetries.

Using this formulation, the mechanisms of epigenetic catalysis are accomplished by allowing the set B_1 above to span a distribution of possible 'final' states $\mathbf{S}_\infty$. Then the groupoid arguments merely expand to permit traverse of both initial states and possible final sets, recognizing that there can now be a possible overlap in the latter, and the epigenetic effects are realized through the joint uncertainties $H(X_{D_i}, Z)$, so that the epigenetic information source Z serves to direct as well the possible final states of X_{D_i}. Again, Scherrer and Jost (2007a, b) use information theory arguments to suggest something similar.

5 'Phase change' and the developmental holonomy groupoid in phenotype space

There is a more direct way to look at phase transitions in cognitive, and here culturally-driven, gene expression, adapting the topological perspectives of homotopy and holonomy directly within phenotype space.

We begin with ideas of directed homotopy.

In conventional topology one constructs equivalence classes of loops that can be continuously transformed into one another on a surface. The prospect of interest is to attempt to collapse such a family of loops to a point while remaining within the surface. If this cannot be done, there is a hole. Here we are concerned,

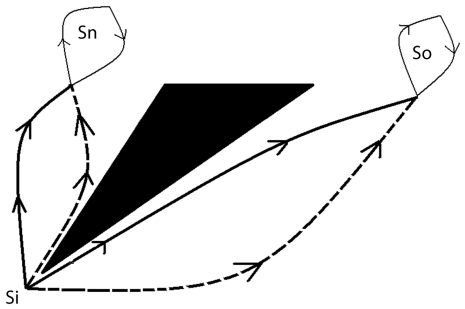

Figure 1: Developmental homotopy equivalence classes in phenotype space. The set on one-way paths from $\mathbf{S_i}$ to $\mathbf{S_n}$ represents an equivalence class of developmental trajectories converging on a particular phenotype, here representing a highly dynamic normal mind structure. In the presence of a noxious external epigenetic catalyst, developmental trajectories can converge on an abnormal mind structure, represented by the dynamic phenotype $\mathbf{S_o}$.

as in figure 1, with sets of one-way developmental trajectories, beginning with an initial phenotype $\mathbf{S_i}$, and converging on some final phenotype, here characteristic (highly dynamic) brain phenotypes labeled, respectively, $\mathbf{S_n}$ and $\mathbf{S_o}$. One might view them as, respectively, 'normal' and 'other', and the developmental pathways as representing convergence on the two different configurations. The filled triangle represents the effect of a composite external epigenetic catalyst – including the effects of culture and culturally-structured social interaction – acting at a critical developmental period represented by the initial phenotype $\mathbf{S_i}$.

We assume phenotype space to be directly measurable and to have a simple 'natural' metric defining the difference between developmental paths.

Developmental paths continuously transformable into each other without crossing the filled triangle define equivalence classes characteristic of different information sources dual to cognitive gene expression, as above.

Given a metric on phenotype space, and given equivalence classes of developmental trajectories having more than one path each, we can *pair one-way developmental trajectories* to make loop structures. In figure 1 the solid and dotted lines above and below the filled triangle can be pasted together to make loops characteristic of the different developmental equivalence classes. Although figure 1 is represented as topologically flat, there is no inherent reason for the phenotype manifold itself to be flat. The existence of a metric in phenotype space permits determining the degree of curvature, using standard methods. Figure 2 shows a loop in phenotype space. Using the metric definition it is possible to *parallel transport* a tangent vector starting at point s around the loop, and to measure the angle between the initial and final vectors, as indicated. A central result from elementary metric geometry is that the angle α will be given by the integral of the curvature tensor of the metric over the interior of the loop

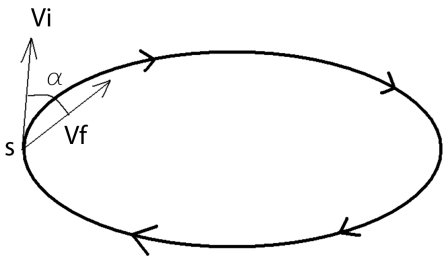

Figure 2: Parallel transport of a tangent vector $V_i \rightarrow V_f$ around a loop on a manifold. Only for a geometrically flat object will the angle between the initial and final vectors be zero. By a fundamental theorem the path integral around the loop by parallel displacement is the surface integral of the curvature over the loop.

(e.g., Frankel, 2006, Section 9.6).

The *holonomy group* is defined as follows (e.g., Helgason, 1962):

If s is a point on a manifold M having a metric, then the holonomy group of M is the group of all linear transformations of the tangent space M_s obtained by parallel translation along closed curves starting at s.

For figure 1 the *phenotype holonomy groupoid* is the disjoint union of the different holonomy groups corresponding to the different branches separated by 'developmental shadows' induced by epigenetic information sources acting as developmental catalysts.

The relation between the phenotype groupoid as defined here and the phase transitions in $F_D[K]$ as defined above is an open question, and is a central focus of ongoing work.

6 Holonomy on the manifold of dual information sources

6.1 Basic structure

Glazebrook and Wallace (2009a) examined holonomy groupoid phase transition arguments for networks of interacting information sources dual to cognitive phenomena. A more elementary form of this arises directly through extending holonomy groupoid arguments to a manifold of different information source dual to cognitive phenomena as follows.

Different cognitive phenomena will have different dual information sources, and we 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 an open set U of closely similar languages $\hat{A}$, such that $A, \hat{A} \subset U$. It may be necessary to coarse-grain the 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 elsewhere, 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 (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,

$$\mathcal{M}(A, \hat{A}) = |\lim \frac{\int_{A,\hat{A}} d(Ax, \hat{A}x)}{\int_{A,A} d(Ax, A\hat{x})} - 1|, \tag{12}$$

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 $\hat{A}$. Other metric constructions on U seem possible. Structures weaker than a conventional metric would be of more general utility, but the mathematical complications are formidable.

Note that these conditions can be used to define equivalence classes of *languages* dual to cognitive processes, where previously we defined equivalence classes of *states* that could be linked by high probability, grammatical and syntactical paths connecting two phenotypes. This led to the characterization of different information sources. Here we construct an entity, formally a topological manifold, *that is an equivalence class of information sources.* This is, provided $\mathcal{M}$ is a conventional metric, a classic differentiable manifold. The set of such equivalence classes generates the *dynamical groupoid*, and questions arise regarding mechanisms, internal or external, which can break that groupoid symmetry.

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

$$dH/d\mathcal{M} = \lim_{\hat{A} \to A} \frac{H(A) - H(\hat{A})}{\mathcal{M}(A, \hat{A})}, \tag{13}$$

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

$$S \equiv H - \mathcal{M} dH/d\mathcal{M}. \tag{14}$$

The simplest possible generalized Onsager relation – here seen as an empirical, fitted, equation like a regression model – is

$$d\mathcal{M}/dt = L dS/d\mathcal{M}, \tag{15}$$

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.

Explicit parameterization of $\mathcal{M}$ introduces standard – and quite considerable – notational complications (Burago et al., 2001): Imposing a metric for different cognitive dual languages parameterized by $\mathbf{K}$ leads to Riemannian, or even Finsler, geometries, including the usual geodesics (e.g., Wallace and Fullilove, 2008; Glazebrook and Wallace, 2009a, b).

The dynamics, as we have presented them so far, have been noiseless. The simplest generalized Onsager relation in the presence of noise might be rewritten as

$$d\mathcal{M}/dt = L dS/d\mathcal{M} + \sigma W(t),$$

where σ is a constant and $W(t)$ represents white noise. Again, S is seen as a function of the parameter $\mathcal{M}$. This leads directly to a family of classic stochastic differential equations of the form

$$d\mathcal{M}_t = L(t, \mathcal{M})dt + \sigma(t, \mathcal{M})dB_t, \tag{16}$$

where L and σ are appropriately regular functions of t and $\mathcal{M}$, and dB_t represents the noise structure, characterized by its quadratic variation. In the sense of Emery (1989), this leads into complicated realms of stochastic differential geometry and related topics.

The natural generalization is to a system of developmental processes that interact via mutual information crosstalk, as described by Wallace and Wallace (2009).

6.2 'Coevolutionary' development

The model can be applied to multiple interacting information sources representing simultaneous gene expression processes. This is, in a broad sense, a 'coevolutionary' phenomenon in that the development of one process may affect that of others.

Most generally we assume that different cognitive developmental subprocesses of gene expression characterized by information sources H_m interact through chemical or other signals and assume that *different processes become each other's principal environments*, a broadly coevolutionary phenomenon.

We write

$$H_m = H_m(K_1 ... K_s, ... H_j ...), \tag{17}$$

where the K_s represent other relevant parameters and $j \neq m$.

The dynamics of such a system is driven by a recursive network of stochastic differential equations, similar to those used to study many other highly parallel dynamic structures (e.g., Zhu et al., 2007).

Letting the K_j and H_m all be represented as parameters Q_j, (with the caveat that H_m not depend on itself), one can define, according to the generalized Onsager development of Wallace and Wallace (2009),

$$S^m = H_m - \sum_i Q_i \partial H_m / \partial Q_i$$

to obtain a complicated recursive system of phenomenological 'Onsager relations' stochastic differential equations,

$$dQ_t^j = \sum_i [L_{j,i}(t, ... \partial S^m / \partial Q^i ...)dt + \sigma_{j,i}(t, ... \partial S^m / \partial Q^i ...)dB_t^i], \tag{18}$$

where, again for notational simplicity only, we have expressed both the H_j and the external K's in terms of the same symbols Q_j.

m ranges over the H_m and we could allow different kinds of 'noise' dB_t^i, having particular forms of quadratic variation that may, in fact, represent a

projection of environmental factors under something like a rate distortion manifold (Glazebrook and Wallace, 2009).

As usual for such systems, there will be multiple quasi-stable points within a given system's H_m, representing a class of generalized resilience modes accessible via holonomy punctuation.

There are other possible patterns:

1. Setting equation (18) equal to zero and solving for stationary points gives attractor states since the noise terms preclude unstable equilibria.

2. This system may, however, converge to limit cycle or 'strange attractors' that are very highly dynamic.

3. What is converged to in both cases is not a simple state or limit cycle of states. Rather it is an equivalence class, or set of them, of generalized language information sources coupled by mutual interaction through crosstalk. Thus 'stability' in this extended model represents particular patterns of ongoing dynamics rather than some identifiable 'state', although such dynamics may be indexed by a 'stable' set of phenotypes.

Here we become enmeshed in a system of highly recursive phenomenological stochastic differential equations, but at a deeper level than the standard stochastic chemical reaction model (e.g., Zhu et al., 2007), and in a dynamic rather than static manner: the objects of this system are equivalence classes of information sources and their crosstalk, rather than simple final states of a chemical system.

We have defined a groupoid for the system based on a particular set of equivalence classes of information sources dual to cognitive processes. That groupoid parsimoniously characterizes the available dynamical manifolds, and breaking of the groupoid symmetry by epigenetic crosstalk creates more complex objects of considerable interest. This leads to the possibility, indeed, the necessity of epigenetic *Deus ex Machina* mechanisms – analogous to programming, stochastic resonance, etc. – to force transitions between the different possible modes within and across dynamic manifolds. In one model the external 'programmer' creates the manifold structure, and the system hunts within that structure for the 'solution' to the problem according to equivalence classes of paths on the manifold. Noise, as with random mutation in evolutionary algorithms, precludes simple unstable equilibria, but not other possible structures.

Equivalence classes of *states* gave dual information sources. Equivalence classes of *information sources* give different characteristic dynamical manifolds. Equivalence classes of one-way developmental *paths* produce different directed homotopy topologies characterizing those dynamical manifolds. This introduces the possibility of having different quasi-stable modes *within* individual manifolds, and leads to ideas of holonomy and the holonomy groupoid of the set of quasi-stable developmental modes.

7 Extending the mathematical formalism

We have, in the context of the tunable epigenetic catalysis of Wallace and Wallace (2009), developed three separate phase transition/branching models of cognitive gene expression based on groupoid structures that may be applied to the development of the human mind and its dysfunctions, as known to be particularly influenced by embedding culture. The first used Landau's spontaneous symmetry breaking to explore phase transitions in a groupoid free energy $F_D[K]$. The second examined a holonomy groupoid in phenotype space generated by disjoint developmental homotopy equivalence classes, and 'loops' constructed by pairing one-way development paths. The third introduced a metric on a manifold of different information sources dual to cognitive gene expression, leading to a more conventional picture of parallel transport around a loop leading to holonomy. The dynamical groupoid of Wallace and Fullilove (2008, Sec. 3.8) is seen as involving a disjoint union across underlying manifolds that produces a holonomy groupoid in a natural manner.

There are a number of evident mathematical questions.

The first is the relation between the Landau formalism and the structures of phenotype space S and those of the associated manifold of dual information sources, the manifold M having metric $\mathcal{M}$. How does epigenetic catalysis in M-space imposes structure on S-space? How is this related to spontaneous symmetry breaking?

What would a stochastic version of the theory, in the sense of Emery (1989), look like? It is quite possible, using appropriate averages of the stochastic differential equations that arise naturally, to define parallel transport, holonomy, and the like for these structures. In particular a stochastic extension of the results of the first question would seem both fairly direct and interesting from a real-world perspective, as development is always 'noisy'.

The construction of loops from directed homotopy arcs in figure 1 is complicated by the necessity of imposing a consistent piecewise patching rule for parallel translation at the end of each arc, say from $\mathbf{S}_i$ to $\mathbf{S}_n$. This can probably be done by some standard fiat, but the details will likely be messy.

One extension of theory would be of great interest: We have imposed metrics on S and M space, making possible a fairly standard manifold analysis of complex cognitive processes of gene expression and development. While this is no small thing, an important 'natural' generalization, given the ubiquity of groupoids across our formalism, would be to a *groupoid atlas*, in the sense of Bak et al., (2006) and Glazebrook and Wallace (2009b, Section 7.4). The groupoid atlas permits a weaker structure compared with that of a conventional manifold since no condition of compatibility between arbitrary overlaps of the patches is necessary. It is possible that the groupoid atlas will become, to complicated problems in biological cognitive process, something of what the Riemannian manifold has been to physics. The groupoid atlas, unlike the Riemannian manifold, is quite new and under active study.

With regard to questions of 'smoothness', we are assuming that the cognitive landscape of gene expression is sufficiently rich that discrete paths can be well

approximated as continuous where necessary, the usual physicist's hack.

Finally, sections 5 and 6 are based on existence of more-or-less conventional metrics, and this may not be a good approximation to many real systems. Extending topological phase transition theory to 'weaker' topologies, e.g., Finsler geometries and the like, is not a trivial task.

8 Discussion

Culturally structured psychosocial stress, and similar noxious exposures, can write distorted images of themselves onto human ontology – both child growth, and, if sufficiently powerful, adult development as well – by a variety of mechanisms, initiating a punctuated trajectory to characteristic forms of comorbid mind/body dysfunction. This occurs in a manner recognizably analogous to resilience domain shifts affecting stressed ecosystems (e.g., Wallace, 2008; Holling, 1973; Gunderson, 2000). Consequently, like ecosystem restoration, reversal or palliation may often be exceedingly difficult once a generalized domain shift has taken place. Thus a public health approach to the prevention of mental disorders may be paramount: rather than seeking to understand why half a population does not respond to the LD50 of a teratogenic environmental exposure, one seeks policies and social reforms that limit the exposure.

Both socio-cultural and epigenetic environmental influences – like gene methylation – are heritable, in addition to genetic mechanisms. The missing heritability of complex diseases that Manolio et al. (2009) seek to find in more and better gene studies is most likely dispersed within the 'dark matter' of these two other systems of heritage that together constitute the larger, and likely highly synergistic, regulatory machinery for gene expression. More and more purely genetic studies would, under such circumstances, be akin to using increasingly powerful microscopes to look for cosmic membranes of strewn galaxies.

A crucial matter for future work is the conversion of the probability models we present here into statistical tools suitable for analyzing real data. This requires not only programming the models for use, but identifying appropriate real-world problems, assembling available data sets, transforming the data as needed for the models, and actually applying the programs. Indeed, the environmental health literature contains numerous examples of developmental deviations due to either chemical exposures or interaction between chemical and socioeconomic exposures, and these could serve as sources of data for direct analysis (e.g., Needleman et al., 1996; Fullilove, 2004; Dietrich et al., 2001; Miranda et al., 2007; Glass et al., 2009; Jacobson and Jacobson, 2003; Shankardass et al., 2009; Clougherty et al., 2007; Ben Jonathan et al., 2009; Karp et al., 2005; Sarlio-Lahteenkorva and Lahelma, 2001; Wallace and Wallace, 2005; Wallace, Wallace and Rauh, 2003). Thus, quite a number of data sets exist in the environmental health and socioeconomic epidemiological literature that could be subjected to meta-analysis and other review for model verification and fitting. Our topological models, when converted to statistical tools for data analysis, hold great potential for understanding developmental trajectories and interfer-

ing factors (teratogens) through the life course. Sets of cross cultural variants of these data focusing specifically on mental disorders, would be needed to address the particular concerns of this paper.

Nonetheless, what we have done is of no small interest for understanding the ontology of the human mind and its pathologies. West-Eberhard (2003, 2005) argues that any new input, whether it comes from the genome, like a mutation, or from the external environment, like a temperature change, a pathogen, or a parental opinion, has a developmental effect only if the preexisting phenotype is responsive to it. A new input causes a reorganization of the phenotype, or 'developmental recombination'. In developmental recombination, phenotypic traits are expressed in new or distinctive combinations during ontogeny, or undergo correlated quantitative change in dimensions. Developmental recombination can result in evolutionary divergence at all levels of organization.

According to West-Eberhard, individual development can be visualized as a series of branching pathways. Each branch point is a developmental decision, or switch point, governed by some regulatory apparatus, and each switch point defines a modular trait. Developmental recombination implies the origin or deletion of a branch and a new or lost modular trait. The novel regulatory response and the novel trait originate simultaneously. Their origins are, in fact, inseparable events: There cannot, West-Eberhard concludes, be a change in the phenotype, a novel phenotypic state, without an altered developmental pathway.

Our analysis provides a new formal picture of this process as it applies to the human mind: The normal branching of developmental trajectories, and the disruptive impacts of teratogeneic events of various kinds, can be described in terms of a growing sequence of holonomy groupoids, each associated with a set of dual information sources representing patterns of cognitive gene expression catalyzed by epigenetic information sources that, for humans, must include culture and culturally-modulated social interaction as well as more direct mechanisms like gene methylation. This is a novel way of looking at human mental development and its disorders that may prove to be of some use. The most important innovation of this work, however, seems to be the natural incorporation of embedding culture as an essential component of the epigenetic regulation of human mental development, and in the effects of environment on the expression of mental disorders, bringing what is perhaps the central reality of human biology into the center of contemporary biological psychiatry.

In sum, we have outlined a broad class of probability models of gene-culture-environment interaction that might help current studies of gene-environment interaction in American psychiatry avoid Heine's (2001) trap of developing an understanding of the self, and its disorders, that is peculiar in the context of the world's cultures.

9 Mathematical Appendix: Groupoids

9.1 Basic Ideas

Following Weinstein (1996) closely, a groupoid, G, is defined by a base set A upon which some mapping – a morphism – can be defined. Note that not all possible pairs of states (a_j, a_k) in the base set A can be connected by such a morphism. 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, 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, $(g_1 g_2)g_3 = g_1(g_2 g_3)$.

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

An orbit of the groupoid G over A is an equivalence class for the relation $a_j \sim Ga_k$ if and only if there is a groupoid element g with $\alpha(g) = a_j$ and $\beta(g) = a_k$. Following Cannas da Silva and Weinstein (1999), we note that a groupoid is called transitive if it has just one orbit. The transitive groupoids are the building blocks of groupoids in that there is a natural decomposition of the base space of a general groupoid into orbits. Over each orbit there is a transitive groupoid, and the disjoint union of these transitive groupoids is the original groupoid. Conversely, the disjoint union of groupoids is itself a groupoid.

The isotropy group of $a \in X$ consists of those g in G with $\alpha(g) = a = \beta(g)$. These groups prove fundamental to classifying groupoids.

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.

Groupoids may have additional structure. As Weinstein (1996) explains, a groupoid G is a topological groupoid over a base space X if G and X are topological spaces and α, β and multiplication are continuous maps. A criticism sometimes applied to groupoid theory is that their classification up to isomorphism is nothing other than the classification of equivalence relations via the orbit equivalence relation and groups via the isotropy groups. The imposition of a compatible topological structure produces a nontrivial interaction between the two structures.

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

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.

Brown (1987) describes the fundamental structure as follows:

> A groupoid should be thought of as a group with many objects, or with many identities... A groupoid with one object is essentially just a group. So the notion of groupoid is an extension of that of groups. It gives an additional convenience, flexibility and range of applications...
>
> EXAMPLE 1. A disjoint union [of groups] $G = \cup_\lambda G_\lambda, \lambda \in \Lambda$, is a groupoid: the product ab is defined if and only if a, b belong to the same G_λ, and ab is then just the product in the group G_λ. There is an identity 1_λ for each $\lambda \in \Lambda$. The maps α, β coincide and map G_λ to λ, $\lambda \in \Lambda$.
>
> EXAMPLE 2. An equivalence relation R on [a set] X becomes a groupoid with $\alpha, \beta : R \to X$ the two projections, and product $(x, y)(y, z) = (x, z)$ whenever $(x, y), (y, z) \in R$. There is an identity, namely (x, x), for each $x \in X$...

Weinstein (1996) makes the following fundamental point:

> Almost every interesting equivalence relation on a space B arises in a natural way as the orbit equivalence relation of some groupoid G over B. Instead of dealing directly with the orbit space B/G as an object in the category S_{map} of sets and mappings, one should consider instead the groupoid G itself as an object in the category G_{htp} of groupoids and homotopy classes of morphisms.

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).

9.2 Global and Local Symmetry Groupoids

Here we follow Weinstein (1996) fairly closely, using his example of a finite tiling.

Consider a tiling of the euclidean plane R^2 by identical 2 by 1 rectangles, specified by the set X (one dimensional) where the grout between tiles is $X = H \cup V$, having $H = R \times Z$ and $V = 2Z \times R$, where R is the set of real numbers and Z the integers. Call each connected component of $R^2 \backslash X$, that is, the complement of the two dimensional real plane intersecting X, a tile.

Let Γ be the group of those rigid motions of R^2 which leave X invariant, i.e., the normal subgroup of translations by elements of the lattice $\Lambda = H \cap V = 2Z \times Z$ (corresponding to corner points of the tiles), together with reflections through each of the points $1/2\Lambda = Z \times 1/2Z$, and across the horizontal and

vertical lines through those points. As noted by Weinstein (1996), much is lost in this coarse-graining, in particular the same symmetry group would arise if we replaced X entirely by the lattice Λ of corner points. Γ retains no information about the local structure of the tiled plane. In the case of a real tiling, restricted to the finite set $B = [0, 2m] \times [0, n]$ the symmetry group shrinks drastically: The subgroup leaving $X \cap B$ invariant contains just four elements even though a repetitive pattern is clearly visible. A two-stage groupoid approach recovers the lost structure.

We define the transformation groupoid of the action of Γ on R^2 to be the set

$$G(\Gamma, R^2) = \{(x, \gamma, y | x \in R^2, y \in R^2, \gamma \in \Gamma, x = \gamma y\},$$

with the partially defined binary operation

$$(x, \gamma, y)(y, \nu, z) = (x, \gamma\nu, z).$$

Here $\alpha(x, \gamma, y) = x$, and $\beta(x, \gamma, y) = y$, and the inverses are natural.

We can form the restriction of G to B (or any other subset of R^2) by defining

$$G(\Gamma, R^2)|_B = \{g \in G(\Gamma, R^2) | \alpha(g), \beta(g) \in B\}$$

[1]. An orbit of the groupoid G over B is an equivalence class for the relation $x \sim_G y$ if and only if there is a groupoid element g with $\alpha(g) = x$ and $\beta(g) = y$.

Two points are in the same orbit if they are similarly placed within their tiles or within the grout pattern.

[2]. The isotropy group of $x \in B$ consists of those g in G with $\alpha(g) = x = \beta(g)$. It is trivial for every point except those in $1/2\Lambda \cap B$, for which it is $Z_2 \times Z_2$, the direct product of integers modulo two with itself.

By contrast, embedding the tiled structure within a larger context permits definition of a much richer structure, i.e., the identification of local symmetries.

We construct a second groupoid as follows. Consider the plane R^2 as being decomposed as the disjoint union of $P_1 = B \cap X$ (the grout), $P_2 = B \backslash P_1$ (the complement of P_1 in B, which is the tiles), and $P_3 = R^2 \backslash B$ (the exterior of the tiled room). Let E be the group of all euclidean motions of the plane, and define the local symmetry groupoid G_{loc} as the set of triples (x, γ, y) in $B \times E \times B$ for which $x = \gamma y$, and for which y has a neighborhood $\mathcal{U}$ in R^2 such that $\gamma(\mathcal{U} \cap P_i) \subseteq P_i$ for $i = 1, 2, 3$. The composition is given by the same formula as for $G(\Gamma, R^2)$.

For this groupoid-in-context there are only a finite number of orbits:

$\mathcal{O}_1$ = interior points of the tiles.

$\mathcal{O}_2$ = interior edges of the tiles.

$\mathcal{O}_3$ = interior crossing points of the grout.

$\mathcal{O}_4$ = exterior boundary edge points of the tile grout.

$\mathcal{O}_5$ = boundary 'T' points.

$\mathcal{O}_6$ = boundary corner points.

The isotropy group structure is, however, now very rich indeed:

The isotropy group of a point in $\mathcal{O}_1$ is now isomorphic to the entire rotation group O_2.

It is $Z_2 \times Z_2$ for $\mathcal{O}_2$.

For $\mathcal{O}_3$ it is the eight-element dihedral group D_4.

For $\mathcal{O}_4, \mathcal{O}_5$ and $\mathcal{O}_6$ it is simply Z_2.

These are the 'local symmetries' of the tile-in-context.

10 Acknowledgments

The author thanks M. Weissman for critical comments useful in revision, and S. Heine for access to the Henrich et al. and H. Kim preprints. The opinions expressed in the paper, however, remain distinctly those of the author.

11 References

Arnett, J., 2008, The neglected 95 %, *The American Psychologist*, 63:602-614.

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, Toward an information theoretical implementation of contextual conditions for consciousness, *Acta Biotheoretica*, 54:157-160.

Baars, B., 1989, *A Cognitive Theory of Consciousness*, Cambridge University Press, New York.

Backdahl, L., A. Bushell, and S. Beck, 2009, Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation, *The International Journal of Biochemistry and Cell Biology*,

doi:10.1016/j.biocel.2008.08.023.

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.

Bebbington, P., 1993, Transcultural aspects of affective disorders, *International Review of Psychiatry*, 5:145-156.

Ben-Jonathan, N., E. Hugo, T. Brandenbourg, 2009, Effects of bisphenol A on adipokine release from human adipose tissue: implications for the metabolic syndrome, *Molecular Cell Endocrinology*, 304:49-54.

Bennett, C., 1988, Logical depth and physical complexity. In *The Universal Turing Machine: A Half-Century Survey*, R. Herkin (ed.), pp. 227-257, Oxford University Press.

Bos, R., 2007, Continuous representations of groupoids, arXiv:math/0612639.

Bossdorf, O., C. Richards, and M. Pigliucci, 2008, Epigenetics for ecologists, *Ecology Letters*, 11:106-115.

Brown, G., T. Harris, and J. Peto, 1973, Life events and psychiatric disorders, II: nature of causal link, *Psychological Medicine*, 3:159-176.

Brown, R., 1987, From groups to groupoids: a brief survey, *Bulletin of the London Mathematical Society*, 19:113-134.

Buneci, M., 2003, *Representare de Groupoizi*, Editura Mirton, Timisoara.

Burago, D., Y. Burago, and S. Ivanov, 2001, *A Course in Metric Geometry*, Graduate Studies in Mathematics 33, American Mathematical Society.

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.

Cannas Da Silva, A., and A. Weinstein, 1999, *Geometric Models for Noncommutative Algebras*, American Mathematical Society, RI.

Caspi, A., et al., 2002, Role of geneotype in the cycle of violence in maltreated children, *Science*, 297:851-854.

Caspi, A., K. Sugden, T. Moffitt, et al., 2003, Influence of life stress on depression : moderation by a polymorphism in the 5-HTT gene, *Science*, 301:386-389.

Caspi, A., and T. Moffitt, 2006, Gene-environment interactions in psychiatry: joining forces with neuroscience, *Nature Reviews Neuroscience* 7:583-590.

Ciliberti, S., O. Martin, and A. Wagner, 2007a, Robustness can evolve gradually in complex regulatory networks with varying topology, *PLoS Computational Biology*, 3(2):e15.

Ciliberti, S., O. Martin, and A. Wagner, 2007b, Innovation and robustness in complex regulatory gene networks, *Proceeding of the National Academy of Sciences*, 104:13591-13596.

Clougherty, J., J. Levy, L. Kubzansky, P. Ryan, S. Suglia, M. Canner, and R. Wright, 2007, Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology, *Environmental Health Perspectives*, 115:1140-1146.

Cohen, I., 2006, Immune system computation and the immunological homunculus. In Nierstrasz, O., J. Whittle, D. Harel, and G. Reggio (eds.), MoDELS 2006, LNCS, vol. 4199, pp. 499-512, Springer, Heidelberg.

Cohen, I., and D. Harel, 2007, Explaining a complex living system: dynamics, multi-scaling, and emergence. *Journal of the Royal Society: Interface*, 4:175-182.

Coplan, J., 2005, personal communication.

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.

Dembo, A., and O. Zeitouni, 1998, *Large Deviations: Techniques and Applications*, 2nd edition, Springer, New York.

Dietrich, K., R. Douglas, P. Succop, O. Berger, R. Bornschein, 2001, Early exposure to lead and juvenile delinquency, *Neurotoxicology and Teratology*, 23:511-518.

Dohrenwend B.P., and B.S. Dohrenwend, 1974, Social and cultural influences on psychopathology, *Annual Review of Psychology*, 25:417-452.

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

Durham, W., 1991, *Coevolution: Genes, Culture and Human Diversity*, Stanford University Press, Palo Alto, CA.

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

Eaton, W., 1978, Life events, social supports, and psychiatric symptoms: a re-analysis of the New Haven data, *Journal of Health and Social Behavior*, 19:230-234.

Emery, M., 1989, *Stochastic Calculus on Manifolds*, Springer, New York.

Feynman, R., 2000, *Lectures on Computation*, Westview Press, New York.

Foley, D., J. Craid, R. Morley, C. Olsson, T. Dwyer, K. Smith, and R. Saffery, 2009, Prospects for epigenetic epidemiology, *American Journal of Epidemiology*, 169:389-400.

Frankel, T., 2006, The Geometry of Physics: An Introduction, Second Edition, Cambridge University Press.

Fullilove, M., 2004, *Root Shock: How Tearing Up City Neighborhoods Hurts America and What we can do about it*, Balantine 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.

Gilbert, S., 2001, Mechanisms for the environmental regulation of gene expression: ecological aspects of animal development, *Journal of Bioscience*, 30:65-74.

Glass, T., K. Bandeen-Roche, M. McAtee, K. Bolla, A. Todd, B. Schwartz, 2009, Neighborhood psychosocial hazards and the association of cumulative lead dose with cognitive function in older adults, *American Journal of Epidemiology*, 169:683-692.

Glazebrook, J.F., and R. Wallace, 2009a, Small worlds and red queens in the global workspace: an information-theoretic approach, *Cognitive Systems Reserch*, 10:333-365.

Glazebrook, J.F., and R. Wallace, 2009b, Rate distortion manifolds as model spaces for cognitive information. In press, *Informatica* Overview Article.

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

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.

Helgason, S., 1962, *Differential Geometry and Symmetric Spaces*, Academic Press, New York.

Henrich J., S. Heine, and A. Norenzayan, 2010, The Weirdest people in the world? In press, *Behavioral and Brain Sciences*.

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

Jablonka, E., and M. Lamb, 1995, *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*, Oxford University Press, Oxford, UK.

Jablonka, E., and M. Lamb, 1998, Epigenetic inheritance in evolution, *Journal of Evolutionary Biology*, 11:159-183.

Jablonka, E., 2004, Epigenetic epidemiology, *International Journal of Epidemiology*, 33:929-935.

Jaenisch, R., and A. Bird, 2003, Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals, *Nature Genetics Supplement*, 33:245-254.

Jacobson J., and Jacobson S. 2002, Breast-feeding and gender as moderators of teratogenic effects on cognitive development, *Neurotoxicological Teratology*, 24:349-358.

Jenkins, J., A. Kleinman, and B. Good, 1990, Cross-cultural studies of depression. In J. Becker and A. Kleinman, *Advances in mood disorders: Theory and Research*, (pp. 67-99), L. Erlbaum, Los Angeles, CA.

Johnson-Laird, P., F. Mancini, and A. Gangemi, 2006, A hyper-emotion theory of psychological illnesses, *Psychological Reviews*, 113:822-841.

Karp, R., C. Chen, A. Meyers, 2005, The appearance of discretionary income: influence on the prevalence of under- and over-nutrition, *International Journal of Equity in Health*, 4:10.

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

Kim, J., R. Stewart, S. Kim, et al., 2007, Interactions between life stressors and susceptibility genes (5-HTTLPR and BDNF) on depression in Korean elders, *Biological Psychiatry*, 62:423-428.

Kim, H., D. Sherman, S. Taylor, J. Sasaki, C. Ryu, and J. Xu, 2010, Culture, serotonin receptor polymorphism (5-HTR1A) and locus of attention. In press, *Social, Cognitive, and Affective Neurosciences*.

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*, 276(3):86-89.

Landau, L., and E. Lifshitz, 2007, *Statistical Physics, 3rd Edition*, Part I, Elsevier, New York.

Manolio, T., F. Collins, N. Cox, et al., 2009, Finding the missing heritability of complex diseases, *Nature*, 461:747-753.

Manson, S., 1995, Culture and major depression: Current challenges in the diagnosis of mood disorders, *Psychiatric Clinics of North America*, 18:487-501.

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.

Maturana, H., and F. Varela, 1980, *Autopoiesis and Cognition*, Reidel Publishing Company, Dordrecht, Holland.

Maturana, H., and F. Varela, 1992, *The Tree of Knowledge*, Shambhala Publications, Boston, MA.

Marsella, A., 2003, Cultural aspects of depressive experience and disorders. In W. Lonner, D. Dinnel, S, Hays, and D. Sattler (Eds.), *Online Readings in Psychology and Culture* (Unit 9, Chapter 4) (http://www.wwu.edu/ culture), Center for Cross-Cultural Research, Western Washington University, Bellingham, WA.

Miranda, M., D. Kim, M. Overstreet Galeano, C. Paul, A. Hull, S. Morgan. 2007. The relationship between early childhood blood lead levels and performance on end-of-grade tests, *Environmental Health Perspectives*, 115:1242-1247.

Needleman, H., Riess J., Tobin M., Biesecker G., and Greenhouse J., 1996, Bone lead levels and delinquent behavior. *Journal of the American Medical Association*, 275:363-369.

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

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

Norenzayan, A., and S. Heine, 2005, Psychological universals: what are they and how can we know? *Psychological Bulletin*, 131:763-784.

O'Nuallain, S., 2008, Code and context in gene expression, cognition, and consciousness. Chapter 15 in Barbiere, M., (ed.), *The Codes of Life: The Rules of Macroevolution*, Springer, New York, pp. 347-356.

Pettini, M., 2007, *Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics*, Springer, New York.

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

Risch, N., R. Herrell, T. Lehner, K. Liang, L. Eaves, J. Hoh, A. Griem, M. Kovacs, J. Ott, and K.R. Merikangas, 2009, Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression, *Journal of the American Medical Association*, 301:2462-2472.

Sarlio-Lahteenkorva, S., E. Lahelma, 2001, Food insecurity is associated with past and present economic disadvantage and body mass index, *Journal of Nutrition*, 131:2880-2884.

Scherrer, K., and J. Jost, 2007a, The gene and the genon concept: a functional and information-theoretic analysis, *Molecular Systems Biology* 3:87-93.

Scherrer, K., and J. Jost, 2007b, Gene and genon concept: coding versus regulation, *Theory in Bioscience* 126:65-113.

Shankardass, K., McConnell, R., Jerrett, M., Milam, J., Richardson, J., and Berhane, K., 2009, Parental stress increases the effect of traffic-related air pollution on childhood asthma incidence, *Proceedings of the National Academy of Sciences*, 106:12406-12411.

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 vowel based morphometry study, *Brain*, 128(Pt.10):2453-2461.

Skierski, M., A. Grundland, and J. Tuszynski, 1989, Analysis of the three-dimensional time-dependent Landau-Ginzburg equation and its solutions, *Journal of Physics A* (Math. Gen.), 22:3789-3808.

Toulouse, G., S. Dehaene, and J. Changeux, 1986, Spin glass model of learning by selection, *Proceedings of the National Academy of Sciences*, 83:1695-1698.

Turner, B., 2000, Histone acetylation and an epigeneticv code, *Bioessays*, 22:836-845.

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

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

Wallace, D., and R. Wallace, 2000, Life and death in Upper Manhattan and the Bronx: Toward evolutionary perspectives on catastrophic social change, *Environment and Planning A*, 32:1245-1266.

Wallace, D., R. Wallace, V. Rauh, 2003, Community stress, demoralization, and body mass index: evidence for social signal transduction, *Social Science and Medicine*, 56:2467-2478.

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

Wallace, R., 2007, Culture and inattentional blindness, *Journal of Theoretical Biology*, 245:378-390.

Wallace, R., 2008, Developmental disorders as pathological resilience domains, *Ecology and Society*, 13:29 (online).

Wallace, R., and M. Fullilove, 2008, *Collective Consciousness and its Discontents: Institutional Distributed Cognition, Racial Policy, and Public Health in the United States*, Springer, New York.

Wallace, R. and D. Wallace, 2005, Structured psychosocial stress and the US obesity epidemic, *Journal of Biological Systems*, 13:363-384.

Wallace, R., and D. Wallace, 2008, Punctuated equilibrium in statistical models of generalized coevolutionary resilience: how sudden ecosystem transitions can entrain both phenotype expression and Darwinian selection, *Transactions on Computational Systems Biology IX*, LNBI 5121:23-85.

Wallace, R., and D. Wallace, 2009, Code, context, and epigenetic catalysis in gene expression, *Transactions on Computational Systems Biology XI*, LNBI 5750:283-334.

Wallace R.G., and R. Wallace, 2009, Evolutionary radiation and the spectrum of consciousness, *Consciousness and Cognition*, 18:160-167.

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

West-Eberhard, M., 2003, *Developmental Placisticity and Evolution*, Oxford University Press, New York.

West-Eberhard, M., 2005, Developmental plasticity and the origin of species differences, *Proceedings of the National Academy of Sciences*, 102:6543-6549.

Zhu, R., Rebirio, A., Salahub, D., Kaufmann, S., 2007, Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models, *Journal of Theoretical Biology*, 246:725-745.