Chapter 2
Cultural Epigenetics: On the Missing Heritability of Complex Diseases

Written in collaboration with Dr. D.N. Wallace.

Summary We extend a cognitive paradigm for gene expression based on the asymptotic limit theorems of information theory to the epigenetic epidemiology of complex developmental disorders in humans. 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 development and its dysfunctions. The cultural and epigenetic systems of heritage may well provide the "missing" heritability of complex diseases now under so much intense discussion.

2.1 Introduction

Mental Disorders and Culture

We begin with a discussion of human mental disorders, that, while increasingly recognized as quintessentially developmental in nature, are still not well understood. The classic scientific task will then be to infer the general from the particular, extending focus on the central role of culture in human mental dysfunction to a vastly larger spectrum of developmental pathologies. This is a task that requires cutting-edge mathematical methods for its formal expression.

The understanding of mental disorders is in considerable disarray. Official classifications of mental illness such as the *Diagnostic and Statistical Manual of Mental Disorders - fourth edition*, (DSM-IV 1994, and now its replacement as "DSM-V"), the standard descriptive nosology in the USA, 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 nineteenth 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,

R. Wallace, *Computational Psychiatry*, DOI 10.1007/978-3-319-53910-2_2

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 association 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 birth weight, or other functional teratogens.

P. Gilbert (2001) 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 downregulates 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 the publication of that landmark study, a number of comprehensive overviews have been published that support its conclusions, for example, Bebbington (1993), Jenkins et al. (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 the late twentieth century American research not only stands the risk of developing models that are particular to that culture, but also 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 century, 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.

Many other observers have made similar points over the years (e.g., Arnett 2008; Henrich et al. 2010; Markus and Kitayama 1991; Matsuda and Nisbett 2006; Nisbett et al. 2001; Wallace 2007).

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; Jablonka 2004; Wallace and Wallace 2009), and are the subject of intense current research, a matter to which we will return below.

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 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 human ontology. 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 human ontology and its manifold dysfunctions.

2.2 A Cognitive Paradigm for Gene Expression

Recapitulating something of the previous chapter, a cognitive paradigm for gene expression is under active study, a model in which contextual factors determine the behavior of what Cohen characterizes as a "reactive system," not at all a deterministic—or even simple stochastic—mechanical process (e.g., Cohen 2006; Cohen and Harel 2007; Wallace and Wallace 2008, 2009). The various 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) puts gene expression directly in the realm of complex linguistic behavior, for which context imposes meaning. He claims 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 (1998, 2009) apply the perspective, respectively, to infectious 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. Jablonka and Lamb (1995, 1998), for example, 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.

This epigenetic perspective has received much empirical confirmation (e.g., Backdahl et al. 2009; Turner 2000; Jaenisch 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.

Similarly, Scherrer and Jost (2007a,b) explicitly invoke information theory in their extension of the definition of the gene to include the local epigenetic machinery, a construct they term the "genon." Their 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.

Again, 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.

2.3 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

$$\mathbf{S}(t) = [S_1(t), \ldots, S_N(t)] \tag{2.1}$$

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

$$S_i(t + \Delta t) = \sigma \left[\sum_{j=1}^{N} w_{ij} S_j(t) \right] \tag{2.2}$$

where Δt is a constant and σ a sigmoidal 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 converges 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).

It is possible to 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 or inherited) picture of the world, and then, upon that comparison, the choice of a single response from a larger 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 $x = (a_0, a_1, \ldots, a_n, \ldots)$—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.

For a cognitive process, this path is supposed to be fed into a "highly nonlinear decision oscillator," h, a sudden threshold machine whose canonical model could well be taken as the famous integrate-and-fire neuron (e.g., Hoppensteadt and Izhikevich 1997, Proposition 8.12). $h(x)$, otherwise seen as a "black box," thus generates an output 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

$$\begin{aligned} B_0 &\equiv \{b_0, \ldots, b_k\} \\ B_1 &\equiv \{b_{k+1}, \ldots, b_m\} \end{aligned} \tag{2.3}$$

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.

Rather than focusing on the properties of h, *we shift the perspective*: The principal objects of formal interest become *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 $h(a_0, \ldots, a_j) \in B_0$ for all $0 < j < m$, but $h(a_0, \ldots, a_m) \in B_1$.

Several points are central to the shift in perspective we are making:

1. It is very important to understand that the fundamental core of the argument does not regard the exact internal details of the inferred (but perhaps not easily observable) function $h(x)$, but rather in the "grammar" and "syntax" of the strings $x = a_0, a_1, \ldots$ leading to action of that function, and which are more likely to be observable.
2. 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 $h(x) \in B_1$.
3. While the combining algorithm, the form of the "nonlinear oscillator" h, and the details of grammar and syntax are all unspecified in this model, *the critical assumption* that permits inference of the necessary conditions constrained by the asymptotic limit theorems of information theory is that the finite limit

$$H \equiv \lim_{n\to\infty} \frac{\log[N(n)]}{n} \tag{2.4}$$

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 $P(a_0, \ldots, a_n)$ and $P(a_n|a_0, \ldots, a_{n-1})$ such that appropriate conditional and joint Shannon uncertainties satisfy the classic relations

$$\begin{aligned} H[\mathbf{X}] &= \lim_{n\to\infty} \frac{\log[N(n)]}{n} \\ &= \lim_{n\to\infty} H(X_n|X_0, \ldots, X_{n-1}) \\ &= \lim_{n\to\infty} \frac{H(X_0, \ldots, X_n)}{n+1} \end{aligned} \tag{2.5}$$

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

Adiabatic means that the source has been parameterized 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 $n = t/\Delta t \rightarrow \infty$ 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 $x = (\mathbf{S}_0, \mathbf{S}_{\Delta t}, \ldots, \mathbf{S}_{n\Delta t}, \ldots)$ 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 considerable computational burdens, as well as, and perhaps in consequence of, providing deep 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

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

that can be expressed in terms of joint and conditional probabilities. Proving these results for general stationary, ergodic information sources require considerable mathematical machinery (e.g., Khinchin 1957; Cover and Thomas 2006; 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)

$$\begin{aligned} F[K] &= \lim_{V\to\infty} -\frac{1}{K}\frac{\log[Z(K,V)]}{V} \\ &= \lim_{V\to\infty} \frac{\log[\hat{Z}(K,V)]}{V} \end{aligned} \tag{2.6}$$

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

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

$$P[E_i] = \frac{\exp[-E_i/kT]}{\sum_j \exp[-E_j/kT]} \tag{2.7}$$

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.

2.4 Tunable Epigenetic Catalysis

Incorporating the influence of embedding contexts—generalized epigenetic effects—is most elegantly done by invoking the Joint Asymptotic Equipartition Theorem (JAEPT) (Cover and Thomas 2006). 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.

From the information theory "chain rule" (Cover and Thomas 2006),

$$H(X, Y) = H(X) + H(Y|X) \leq H(X) + H(Y)$$

Equality occurs only for stochastically independent processes.

Interpreting the homology between information and free energy rather broadly, the embedding context, in effect *lowers an analog to the relative activation energy of a particular developmental channel*, at the expense of raising the total free energy needed, since the system must now support two information sources instead of one, i.e., that regulated, and that providing the regulation.

Thus the effect of epigenetic regulation is to change the probability of developmental pathways, while requiring more total energy for development. 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 X 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 human ontology is, according to this formalism, straightforward: Consider culture as another embedding information source, Z, having source uncertainty $H(Z)$. Then the information chain rule becomes

$$H(X, Y, Z) \leq H(X) + H(Y) + H(Z) \tag{2.8}$$

and the numbers of "typical" sequences of length n are then approximately

$$\exp[nH(X, Y, Z)] \leq \exp[n(H(X) + H(Y) + H(Z))] \tag{2.9}$$

where, again, equality occurs only under stochastic independence.

A cultural regulatory apparatus, however, has very considerable free energy requirements, to grossly understate the matter.

In this model, following explicitly the direction indicated by Boyd, Kleinman, and their colleagues, culture is seen as an essential component of the catalytic epigenetic machinery that regulates human ontology, including development of the human mind. This is not to say that the development 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.

2.5 The Groupoid Free Energy

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 (fixed) characteristic 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 we take the probability of a particular H_α as determined by a standard expression (e.g., Landau and Lifshitz 2007),

$$P[H_\beta] = \frac{\exp[-H_\beta K]}{\sum_\alpha \exp[-H_\alpha K]} \tag{2.10}$$

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 $P[H_\beta]$ 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 (presumed fixed) 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, will require 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 G_i, H_{G_i}, will be given by the classic relation

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

where, now, we have shifted perspective, and *the sum/integral is over all possible elements of the largest available symmetry groupoid representing the equivalence class structure*. 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 $Z_G \equiv \sum_\alpha \exp[-H_{G_\alpha}K]$. We now define the *Groupoid free energy* of the system, F_G, at inverse normalized metabolic energy density K, as

$$F_G[K] \equiv -\frac{1}{K}\log[Z_G[K]] \tag{2.11}$$

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

Spontaneous Symmetry Breaking

The groupoid free energy permits introduction 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—$F_G[K]$—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 2007).

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 $F_G[K]$. Necessity, but not sufficiency, is important, as it, in theory, allows mixed groupoid symmetries, leading to comorbidity in "condensation" dysfunctions.

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 $H(X_{G_\alpha}, Z)$, so that the epigenetic information source Z serves to direct as well the possible final states of X_{G_α}. Again, Scherrer and Jost (2007a,b) use information theory arguments to suggest something similar.

The Groupoid Atlas

The groupoid free energy inherently defines a groupoid atlas in the sense of Bak et al. (2006). Following closely Glazebrook and Wallace (2009a,b), the set of groupoids G_α comprise a groupoid atlas $\mathcal{A}$ as follows.

A family of local groupoids $(G_\mathcal{A})$ is defined with respective object sets $(X_\mathcal{A})_\alpha$, and a *coordinate system* $\Phi_\mathcal{A}$ of $\mathcal{A}$ equipped with a reflexive relation $\leq$. These satisfy the following conditions:

1. If $\alpha \leq \beta$ in $\Phi_\mathcal{A}$, then $(X_\mathcal{A})_\alpha \cap (X_\mathcal{A})_\beta$ is a union of components of $(G_\mathcal{A})$, that is, if $x \in (X_\mathcal{A})_\alpha \cap (X_\mathcal{A})_\beta$ and $g \in (G_\mathcal{A})_\alpha$ acts as $G : x \to y$, then $y \in (X_\mathcal{A})_\alpha \cap (X_\mathcal{A})_\beta$.
2. If $\alpha \leq \beta$ in $\Phi_\mathcal{A}$, then there is a groupoid morphism defined between the restrictions of the local groupoids to intersections

$$(G_\mathcal{A})_\alpha | (X_\mathcal{A})_\alpha \cap (X_\mathcal{A})_\beta \to (G_\mathcal{A})_\beta | (X_\mathcal{A})_\alpha \cap (X_\mathcal{A})_\beta,$$

and which is the identity morphism on objects.

Thus each of the G_α with its associated dual information source H_{G_α} constitutes a component of an atlas that incorporates the dynamics of an interactive system by means of the intrinsic groupoid actions.

These are matters currently under very active study (e.g., del Hoyo and Minian 2008).

2.6 "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, as in Fig. 2.1, 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 $\mathbf{S_i}$.

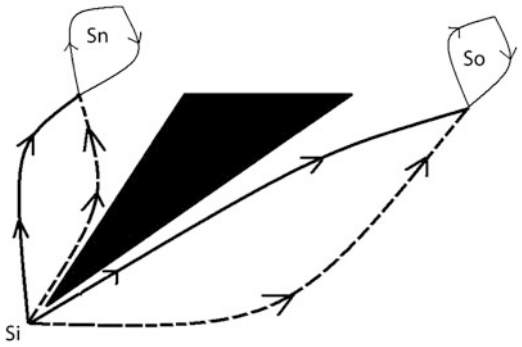

Fig. 2.1 Developmental homotopy equivalence classes in phenotype space. The set on one-way paths from $\mathbf{S_i}$ to $\mathbf{S_n}$ represents an equivalence class of developmental trajectories converging on a particular phenotype, here representing a highly dynamic normal mind structure. In the presence of a noxious external epigenetic catalyst, developmental trajectories can converge on an abnormal mind structure, represented by the dynamic phenotype $\mathbf{S_o}$

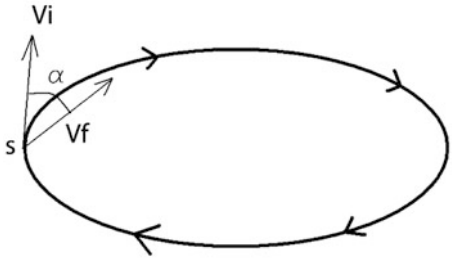

Fig. 2.2 Parallel transport of a tangent vector $V_i \rightarrow V_f$ around a loop on a manifold. Only for a geometrically flat object will the angle between the initial and final vectors be zero. By a fundamental theorem the path integral around the loop by parallel displacement is the surface integral of the curvature over the loop

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 Fig. 2.1 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 Fig. 2.1 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 2.2 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 (e.g., Frankel 2006, Sect. 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 Fig. 2.1 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.

2.7 Holonomy on the Manifold of Dual Information Sources

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 Chap. 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 2006). More detail on distortion measures is given in the section below on the Rate Distortion Theorem. Note that these languages do not interact, in this approximation.

3. Define a metric on U, for example,

$$\mathcal{M}(A,\hat{A}) = \left| \lim \frac{\int_{A,\hat{A}} d(Ax,\hat{A}x)}{\int_{A,A} d(Ax,A\hat{x})} - 1 \right| \tag{2.12}$$

integrating over the sets of 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, leading to similar results, just as different definitions of distortion lead to the same end in the Rate Distortion Theorem.

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

$$dH/d\mathcal{M} = \lim_{\hat{A}\to A} \frac{H(A) - H(\hat{A})}{\mathcal{M}(A,\hat{A})} \tag{2.13}$$

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

The essential point of a "covariant" derivative is that it is *independent of coordinate systems*, a condition this definition fulfills. As we will show below, this leads directly to ideas of a derivative along a tangent vector and to ideas of parallel transport leading to deep topological concepts such as holonomy. Introduction of a coordinate system in the definition of $\mathcal{M}$ quickly leads to the usual Christoffel symbols and the familiar geodesic equations (e.g., Wallace and Fullilove 2008, Sect. 8.3).

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, an "entropy" analog, as

$$S \equiv H - \mathcal{M}dH/d\mathcal{M} \tag{2.14}$$

The simplest possible generalized Onsager relation—here seen as an empirical, fitted, equation like a regression model—is in terms of the gradient of S,

$$d\mathcal{M}/dt = LdS/d\mathcal{M} \tag{2.15}$$

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.

Again, explicit parameterization of $\mathcal{M}$—that is, imposing a coordinate system—introduces standard, and quite considerable, notational complications (Burago et al. 2001). Defining 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

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

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

$$d\mathcal{M}_t = L(t, \mathcal{M})dt + \sigma(t, \mathcal{M})dB_t \tag{2.16}$$

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).

"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

$$H_m = H_m(K_1 \ldots K_s, \ldots H_j \ldots) \tag{2.17}$$

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), "entropies" via the Legendre transforms as

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

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

$$dQ_t^j = \sum_i [L_{j,i}(t, \ldots \partial S^m / \partial Q^i \ldots)dt + \sigma_{j,i}(t, \ldots \partial S^m / \partial Q^i \ldots)dB_t^i] \tag{2.18}$$

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).

It is important to realize that, for this formulation, one does not necessarily have the equivalent of "Onsager's fourth law" of thermodynamics, i.e., the symmetry relation $L_{i,j} = L_{j,i}$. This is because such a symmetry, at base, is a statement of local time reversal invariance (e.g., de Groot and Mazur 1984, pp. 35–41). But information sources are notoriously one-way in time, for example, someone speaking or writing in English is much more likely to utter the five-character string "the" than its reverse. Or, to put it another way, long palindromes, such as "Able was I ere I saw Elba," or "A man, a plan, a canal: Panama," are quite rare, and always relatively short, while information theory is based on asymptotic limit theorems most often involving very long strings of symbols.

As usual, for a system of equations like (2.18), there will be multiple quasi-stable nonequilibrium steady states, representing a class of generalized resilience modes accessible via holonomy punctuation.

There are, indeed, many possible patterns:

1. Setting the expectation of Eq. (2.18) 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.

2.8 Rate Distortion Models

The Rate Distortion Theorem

The interaction between cognitive structures can be restated from a highly formal, but far more restricted, Rate Distortion Theorem perspective. Suppose a sequence of signals is generated by an information source dual to a cognitive process, Y having output $y^n = y_1, y_2, \ldots$. This is "digitized" in terms of the observed behavior of the system with which it communicates, say a sequence of observed behaviors $b^n = b_1, b_2, \ldots$. Often the b_i will happen in a characteristic "real time" τ. Assume each b^n is then deterministically retranslated back into a reproduction of the original biological signal,

$$b^n \rightarrow \hat{y}^n = \hat{y}_1, \hat{y}_2, \ldots.$$

Define a distortion measure $d(y, \hat{y})$ that compares the original to the retranslated path. Many such measures are possible. The Hamming distortion, for example, is

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

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

For continuous variates the squared error distortion is

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

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

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

A remarkable fact of the Rate Distortion Theorem is that *the basic result is independent of the exact distortion measure chosen* (Cover and Thomas 2006; Dembo and Zeitouni 1998).

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

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

The average distortion is defined as

$$D \equiv \sum_{y^n} p(y^n) d(y^n, \hat{y}^n) \tag{2.19}$$

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

$$I(Y, \hat{Y}) \equiv H(Y) - H(Y|\hat{Y}) = H(Y) + H(\hat{Y}) - H(Y, \hat{Y}) \tag{2.20}$$

where $H(\ldots, \ldots)$ is the joint and $H(\ldots|\ldots)$ the conditional uncertainty (Cover and Thomas 2006; Ash 1990).

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

In general, of course, this will not be true.

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

$$R(D) = \min_{p(y,\hat{y}); \sum_{(y,\hat{y})} p(y) p(y|\hat{y}) d(y,\hat{y}) \leq D} I(Y, \hat{Y}) \tag{2.21}$$

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

The *Rate Distortion Theorem* states that $R(D)$ is the minimum necessary rate of information transmission which ensures the communication between the modules does not exceed average distortion D. Thus $R(D)$ defines a minimum necessary channel capacity. Cover and Thomas (2006) or Dembo and Zeitouni (1998) provide details. The Rate Distortion Function has been calculated for a number of systems using Lagrange multiplier and related methods.

There is an absolutely central fact characterizing the Rate Distortion Function: Cover and Thomas (2006) show that $R(D)$ *is necessarily a decreasing convex function of* D for any reasonable definition of distortion.

That is, $R(D)$ *is always* a reverse J-shaped curve. This will prove crucial for the overall argument. Indeed, convexity is an exceedingly powerful mathematical condition, and permits deep inference (e.g., Rockafellar 1970). Ellis (1985, Chap. VI) applies convexity theory to conventional statistical mechanics.

For a Gaussian channel having noise with zero mean and variance σ^2 under the squared distortion measure (Cover and Thomas 2006),

$$\begin{aligned} R(D) &= 1/2 \log[\sigma^2/D], 0 \leq D \leq \sigma^2 \\ R(D) &= 0, D \geq \sigma^2 \end{aligned} \tag{2.22}$$

Recall, now, the relation between information source uncertainty and channel capacity (e.g., Ash 1990):

$$H[\mathbf{X}] \leq C \tag{2.23}$$

where H is the uncertainty of the source X and C the channel capacity, defined according to the relation (Ash 1990)

$$C \equiv \max_{P(X)} I(X|Y) \tag{2.24}$$

where $P(X)$ is chosen so as to maximize the rate of information transmission along a channel Y.

Finally, recall the analogous definition of the Rate Distortion Function from Eq. (2.21), again an extremum over a probability distribution.

Rate Distortion Dynamics

$R(D)$ defines the minimum channel capacity necessary for the system to have average distortion less than or equal to D, placing a limit on information source uncertainty. Thus, we suggest distortion measures can drive information system

dynamics. That is, the Rate Distortion Function also has a homological relation to free energy density, similar to the relation between free energy density and information source uncertainty.

We are led to propose, as a heuristic, that the dynamics of cognitive modules interacting in a characteristic "real time" τ will be constrained by the system as described in terms of a parameterized Rate Distortion Function. To do this, take R as parameterized, not only by the distortion D, but also by some vector of variates $\mathbf{Q} = (Q_1, \ldots, Q_k)$, for which the first component is the average distortion. The assumed dynamics are, as in Wallace and Wallace (2008), then driven by gradients in the rate distortion disorder defined as the Legendre transform of the Rate Distortion Function

$$S_R \equiv R(\mathbf{Q}) - \sum_{i=1}^{k} Q_i \partial R / \partial Q_i \tag{2.25}$$

This leads to the deterministic and stochastic systems of equations analogous to the Onsager relations of nonequilibrium thermodynamics:

$$dQ_j/dt = \sum_i L_{j,i} \partial S_R / \partial Q_i \tag{2.26}$$

and

$$dQ_t^j = L^j(Q_1, \ldots, Q_k, t)dt + \sum_i \sigma^{j,i}(Q_1, \ldots, Q_k, t)dB_t^i \tag{2.27}$$

where the dB_t^i represent added, often highly structured, stochastic "noise" whose properties are characterized by the quadratic variation (e.g., Protter 1995).

Even for this simplified structure, it is not clear under what circumstances "Onsager reciprocal relations" are possible. Since average distortion is a scalar, however, some systems may indeed display the kind of time reversal invariance required for those symmetries.

A central focus of this chapter, however, is to generalize these equations in the face of richer structures, for example, interactions between cognitive modules that may not be time-reversible, the existence of characteristic time constants within nested processes, and the influence of an embedding source of free energy.

For a simple Gaussian channel with noise having zero mean and variance σ^2, an entropy can be defined as the Legendre transform

$$S_R(D) = R(D) - DdR(D)/dD = 1/2 \log(\sigma^2/D) + 1/2 \tag{2.28}$$

The simplest possible Onsager relation becomes

$$dD/dt = -\mu dS_R/dD = \frac{\mu}{2D} \tag{2.29}$$

where $-dS_R/dD$ represents the force of an entropic wind, a kind of internal dissipation inevitably driving the real time, system of interacting (cognitive) information sources toward greater distortion.

This has the solution

$$D = \sqrt{\mu t} \tag{2.30}$$

so that the average distortion increases monotonically with time, for this model.

A central observation is that *similar results must necessarily apply to any of the reverse-J-shaped relations that inevitably characterize* $R(D)$, since the Rate Distortion Function is necessarily a convex decreasing function of the average distortion D, whatever distortion measure is chosen. Again, see Cover and Thomas (2006) for details.

The explicit implication is that a system of cognitive modules interacting in real time will inevitably be subject to a relentless entropic force, requiring a constant free energy expenditure for maintenance of some fixed average distortion in the communication between them: The distortion in the communication between two interacting modules will, without free energy input, have time dependence

$$D = f(t) \tag{2.31}$$

with $f(t)$ monotonic increasing in t.

This necessarily leads to the punctuated failure of the system.

Note that Eq. (2.30) is similar to classical Brownian motion as treated by Einstein: Let $p(x,t)dx$ be the probability a particle located at the origin at time zero and undergoing Brownian motion is found at locations $x \to x + dx$ at time t. Then, p satisfies the diffusion equation $\partial p(x,t)/\partial t = \mu \partial^2 p(x,t)/\partial x^2$. Einstein's solution is that

$$p(x,t) = \frac{1}{\sqrt{4\pi\mu t}} \exp[-x^2/4\mu t].$$

It is easy to show that the standard deviation of the particle position increases in proportion to $\sqrt{\mu t}$, just as above.

Some comment is appropriate. Following, e.g., Chung and Williams (1990), a process $B = B_t, t \in \mathcal{R}_+$ is called a Brownian motion in $\mathcal{R}_+$ iff:

1. for $0 \leq s < t < \infty, B_t - B_s$ is a normally distributed random variate with mean zero and variance $|t - s|$.
2. for $0 \leq t_0 < t_1 < \cdots < t_k < \infty$,

$$\{B_{t_0}; B_{t_j} - B_{t_{j-1}}, j = 1, \ldots, k\}$$

is a set of independent random variates.

An information source, of course, generates a *highly correlated sequence* that grossly violates these simple assumptions. What we have shown is that the *distortion* in the communication between interacting cognitive modules, under appropriate empirical Onsager relations, can behave as if it were undergoing Brownian motion.

This is a simple, but far from trivial, result.

Prandolini and Moody (1995) have, in fact, observed something much like this in the time base error of recorded signals. Wow and flutter are the instantaneous speed error between recording and reproduction epochs. The time base error (TBE) in the reproduced signal is a function of the wow and flutter. They show, empirically, that the nonperiodic TBE is a *fractional Brownian motion*. The implication is that the nonperiodic flutter is fractional Gaussian, and thus what they call a "blind" TBE system is impractical for the design of a TBE compensation system.

Normalized fractional Brownian motion on $(0, t), t \in \mathcal{R}_+$ is a continuous time Gaussian process starting at zero, with mean zero, and having the covariance function (Beran 1994)

$$E[B^H(t)B^H(s)] = (1/2)[|t|^{2H} + |s|^{2H} - |t-s|^{2H}].$$

If $H = 1/2$ the process is a regular Brownian motion. Otherwise, for $H > 1/2$, the increments are positively correlated, and for $H < 1/2$, negatively correlated.

We will explore this kind of relation in more detail below.

Rate Distortion Coevolutionary Dynamics

A simplified version of Eq. (2.18) can be constructed using the Rate Distortion Functions for mutual crosstalk between a set of interacting cognitive modules, using the homology of the Rate Distortion Function itself with free energy, as driven by the inherent convexity of the Rate Distortion Function $R(D)$. That convexity is, in fact, why we invoke the Rate Distortion Function.

Given different cognitive processes $1 \ldots s$, the quantities of special interest thus become the mutual Rate Distortion Functions $R_{i,j}$ characterizing communication (and the distortion $D_{i,j}$) between them, while the essential parameters remain the characteristic time constants of each process, $\tau_j, j = 1 \ldots s$, and an overall, embedding, available free energy density, F.

Taking the Q^α to run over all the relevant parameters and mutual Rate Distortion Functions (including distortion measures $D_{i,j}$), Eq. (2.14) becomes

$$S_R^{i,j} \equiv R_{i,j} - \sum_k Q^k \partial R_{i,j} / \partial Q^k \tag{2.32}$$

Equation (2.18) accordingly becomes

$$dQ_t^\alpha = \sum_{\beta=(i,j)} [L_\beta(t, \ldots \partial S_R^\beta / \partial Q^\alpha \ldots)dt + \sigma_\beta(t, \ldots \partial S_R^\beta / \partial Q^\alpha \ldots)dB_t^\beta] \tag{2.33}$$

and this generalizes the treatment in terms of crosstalk, its distortion, the inherent time constants of the different cognitive modules, and the overall available free energy density.

This is a very complicated structure indeed, but its general dynamical behaviors will obviously be analogous to those described just above. For example, setting the expectation of Eq. (2.33) to zero gives the "coevolutionary stable states" of a system of interacting cognitive modules. Again, limit cycles and strange attractors seem possible as well. And again, what is converged to is a dynamic behavior pattern, not some fixed "state." And again, such a system will display highly punctuated dynamics almost exactly akin to resilience domain shifts in ecosystems (e.g., Holling 1973, 1992; Gunderson 2000). Indeed, the formalism seems directly applicable to ecosystem studies.

And again, because these are highly self-dynamic cognitive phenomena and not simple crystals or other physical objects, it may not often be possible to invoke time reversal invariance to give Onsager-like reciprocal symmetries to Eq. (2.33).

2.9 Expanding the Mathematical Approach

We have, in the context of a tunable epigenetic catalysis, developed several 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, Sect. 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 outstanding 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 Fig. 2.1 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.

On another matter, 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, the "natural" generalization, given the ubiquity of groupoids across our formalism, would be to a more complete groupoid atlas treatment in the spirit of section "The Groupoid Atlas". 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.

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, Sects. 2.6 and 2.7 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.

2.10 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 sociocultural 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 is the conversion of the probability models we present here into statistical tools suitable for analyzing real data, and hence actually testing the theoretical models we present here. Some work in this direction has been done in Sect. 2.8, but the problem involves not just programming such 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 statistical models. 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 (2002); 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 et al. 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 interfering 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 chapter.

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 human development: The normal branching of developmental trajectories, and the disruptive impacts of teratogenic 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 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 ontology, and in the effects of environment on the expression of a broad spectrum of developmental disorders: the missing heritability of complex diseases found.

Acknowledgements We thank 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, however, remain distinctly those of the authors.

References

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

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

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.

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, and 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*, ed. R. Herkin, 227–257. Oxford: Oxford University Press.

Beran, J. 1994. *Statistics for Long-Memory Processes*. New York: Chapman and Hall.

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

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

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

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

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

Chung, K., and R. Williams. 1990. *Introduction to Stochastic Integration*, 2nd ed. Boston, MA: Birkhauser.

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 *MoDELS 2006*, ed. O. Nierstrasz, J. Whittle, D. Harel, and G. Reggio. Lecture Notes in Computer Science, vol. 4199, 499–512. Heidelberg: Springer.

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.

Cover, T., and J. Thomas. 2006. *Elements of Information Theory*, 2nd ed. New York: Wiley.

de Groot, S., and R. Mazur. 1984. *Non-equilibrium Thermodynamics*. New York: Dover.

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

del Hoyo, M., and E. Minian. 2008. Classical invariants for global actions and groupoid atlases. *Applied Categorical Structures* 18: 689–721.

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

Dietrich, K., R. Douglas, P. Succop, O. Berger, and 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*. Palo Alto, CA: Stanford University Press.

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.

Ellis, R. 1985. *Entropy, Large Deviations, and Statistical Mechanics*. New York: Springer.

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

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

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*, 2nd ed. Cambridge: Cambridge University Press.

Fullilove, M. 2004. *Root Shock: How Tearing Up City Neighborhoods Hurts America and What We Can Do About It*. New York: Ballantine Books.

Gilbert, P. 2001. Evolutionary approaches to psychopathology: The role of natural defenses. *Australian and New Zealand Journal of Psychiatry* 35: 17–27.

Glass, T., K. Bandeen-Roche, M. McAtee, K. Bolla, A. Todd, and 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, 2009, Rate distortion manifolds as model spaces for cognitive information. *Informatica* 33: 309–346.

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

Glazebrook, J.F., and R. Wallace. 2009b. Rate distortion manifolds as model spaces for cognitive information. *Informatica* 33: 309–346.

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*. New York: Academic.
Henrich, J., S. Heine, and A. Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33: 61–83.
Holling, C. 1973. Resilience and stability of ecological systems. *Annual Reviews of Ecological Systematics* 4: 1–23.
Holling, C. 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs* 62: 447–502.
Hoppensteadt, F., and E. Izhikevich. 1997. *Weakly Connected Neural Networks*. New York: Springer.
Jablonka, E. 2004. Epigenetic epidemiology. *International Journal of Epidemiology* 33: 929–935.
Jablonka, E., and M. Lamb. 1995. *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford: Oxford University Press.
Jablonka, E., and M. Lamb. 1998. Epigenetic inheritance in evolution. *Journal of Evolutionary Biology* 11: 159–183.
Jacobson, J., and S. Jacobson. 2002. Breast-feeding and gender as moderators of teratogenic effects on cognitive development. *Neurotoxicological Teratology* 24: 349–358.
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.
Jenkins, J., A. Kleinman, and B. Good. 1990. Cross-cultural studies of depression. In *Advances in Mood Disorders: Theory and Research*, ed. J. Becker and A. Kleinman, 67–99. Los Angeles, CA: L. Erlbaum.
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, and 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*. New York: Dover.
Kleinman, A., and A. Cohen. 1997. Psychiatry's global challenge. *Scientific American* 276(3): 86–89.
Kleinman, A., and B. Good. 1985. *Culture and Depression: Studies in the Anthropology of Cross-Cultural Psychiatry of Affect and Depression*. Berkeley, CA: University of California Press.
Landau, L., and E. Lifshitz. 2007. *Statistical Physics*, 3rd ed., Part I. New York: Elsevier.
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.
Marsella, A. 2003. Cultural aspects of depressive experience and disorders. In *Online Readings in Psychology and Culture (Unit 9, Chapter 4)*, ed. W. Lonner, D. Dinnel, S. Hays, and D. Sattler. Bellingham, WA: Center for Cross-Cultural Research, Western Washington University. http://www.wwu.edu/~culture
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*. Dordrecht: Reidel.
Maturana, H., and F. Varela. 1992. *The Tree of Knowledge*. Boston, MA: Shambhala Publications.
Miranda, M., D. Kim, M. Overstreet Galeano, C. Paul, A. Hull, and 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., J. Riess, M. Tobin, G. Biesecker, and J. Greenhouse. 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. In *The Codes of Life: The Rules of Macroevolution*, ed. M. Barbiere, Chapter 15, 347–356. New York: Springer.
Pettini, M. 2007. *Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics*. New York: Springer.
Prandolini, R., and M. Moody. 1995. Brownian nature of the Time-Base Error in tape recordings. *Journal of the Audio Engineering Society* 43: 241–247.
Protter, P. 1995. *Stochastic Integration and Differential Equations: A New Approach*. New York: Springer.
Richerson, P., and R. Boyd. 2004. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: Chicago University Press.
Risch, N., R. Herrell, T. Lehner, K. Liang, L. Eaes, J. Hoh, A. Griem, M. Kovacs, J. Ott, and K. 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.
Rockafellar, R. 1970. *Complex Analysis*. Princeton, NJ: Princeton University Press.
Sarlio-Lahteenkorva, S., and 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 K. Berhane. 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.
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 epigenetics code. *Bioessays* 22: 836–845.
Wallace, R. 2005. *Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model*. New York: Springer.
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*. New York: Springer.
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, 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, D., R. Wallace, and V. Rauh. 2003. Community stress, demoralization, and body mass index: Evidence for social signal transduction. *Social Science and Medicine* 56: 2467–2478.

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

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

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

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

http://www.springer.com/978-3-319-53909-6

Computational Psychiatry
A Systems Biology Approach to the Epigenetics of Mental Disorders
Wallace, R.
2017, XII, 236 p. 31 illus., Hardcover
ISBN: 978-3-319-53909-6