

VERSION 7.0

**PREDICTING HEALTH IMPACTS
OF THE WORLD TRADE CENTER
DISASTER: 1. Halogenated
hydrocarbons, symptom syndromes,
secondary victimization, and the
burdens of history**

Rodrick Wallace, PhD
The New York State Psychiatric Institute
Deborah N Wallace, Ph D
Joseph L Mailman School of Public Health
Columbia University*

November 14, 2001

Abstract

The recent attack on the World Trade Center, in addition to direct injury and psychological trauma, has exposed a vast population to dioxins, dibenzofurans, related endocrine disruptors, and a multitude of other physiologically active chemicals arising from the decomposition of the massive quantities of halogenated hydrocarbons and other plastics within the affected buildings. The impacts of these chemical species have been compounded by exposure to asbestos, fiberglass,

*Correspondence: R Wallace, PISCS Inc., 549 W 123 St., Suite 16F, New York, NY, 10027. Telephone (212) 865-4766, email rdwall@ix.netcom.com.

crushed glass, concrete, plastic, and other irritating dusts. To address the manifold complexities of this incident we combine recent theoretical perspectives on immune, CNS, and sociocultural cognition with empirical studies on survivors of past large toxic fires, other community-scale chemical exposure incidents, and the aftereffects of war. Our analysis suggests the appearance of complex, but distinct and characteristic, spectra of synergistically linked social, psychosocial, psychological and physical symptoms among the 100,000 or so persons most directly affected by the WTC attack. The different ‘eigenpatterns’ should become increasingly comorbid as a function of exposure. The expected outcome greatly transcends a simple ‘Post Traumatic Stress Disorder’ model, and may resemble a particularly acute form of Gulf War Syndrome. We explore the role of external social factors in subsequent exacerbation of the syndrome – secondary victimization – and study the path-dependent influence of individual and community-level historical patterns of stress. We suggest that workplace and other organizations can act as ameliorating intermediaries. Those without access to such buffering structures appear to face a particularly bleak future.

Key Words: chemical exposure, disaster, ecological resilience, Gulf War Syndrome, historical burdens, immune cognition, racism, secondary victimization, terrorism, traumatic perturbation, weathering, World Trade Center

Introduction

The fiery collapse of New York City’s World Trade Center has, at this writing, killed thousands and driven tens of thousands to seek medical attention. It appears to have directly and profoundly affected several hundred thousand people. Impacts range from expected extremities of psychological and psychosocial stress, to physical trauma, burns, and exposure to irritating dusts containing concrete, fiberglass, asbestos, ground plastics, and pulverized human remains. Less well understood is the potential for significant inhalation and skin-dose exposure to the highly toxic and carcinogenic combustion and other decomposition products of the thousand or so tons of halogenated and nitrogenated plastics and plastics additives in the affected buildings. Great quantities of these highly unstable and acutely combustion-toxic materials continue to smolder and burst into flame almost two months

after the event, covering lower Manhattan with a persistent, eye-burning cloud of stinking, choking, smoke.

Combinations of smoke inhalation, burns and trauma are reported to be the most common factors in the immediate hospitalizations. Very significantly, in our view, large numbers of persons with allergic skin and anaphylactic reactions were seen during the initial stages of the incident (Larkin, 2001). If recent case histories of multi-fatal and other toxic plastics fires and chemical releases provide guidance, psychological and psychosocial trauma, burns and other physical injuries, and toxic exposure, will interact synergistically to produce complex but characteristic patterns of symptoms on several timescales in a path-dependent manner at the individual and group levels – what has already come to be called ‘World Trade Center Syndrome’ by the popular press (NY Post, 2001; Newsweek, 2001). The long-term path of this syndrome will be profoundly affected by the response of larger, embedding social structures, in likely synergism with individual and community-level historical burdens carried by those exposed. Community response is most likely to occur as a virulent secondary victimization (e.g. Tarrier et al., 1999; Campbell and Raja, 1999), but can, in a best case scenario, serve as an ameliorating influence.

We begin with a review of recent theoretical developments which cross link processes of central nervous system (CNS), immune, and sociocultural cognition. This allows analysis of the effects of a sudden, traumatic, perturbation – characteristic combinations of physical, social and chemical assault – on ‘condensed’ cognitive systems at different timescales, and in the context of individual and community experiences of historical burden.

Next, we introduce the ‘new’ combustion toxicology, and summarize observed effects of large toxic fire and other chemical release events on injured survivors from this perspective. We then reinterpret studies of Gulf War Syndrome, and ultimately combine the analyses to try and predict the developing spectra of symptoms to be expected from the WTC disaster at different ecosystem levels, in the context of couplings across them.

We are particularly concerned with the likely interaction of larger, embedding crosssectional and longitudinal social structures and historical experiences with symptom patterns.

The problems are subtle and complex, and we are driven to employ cutting-edge methods. These require some initial development.

Genes, cognition, and culture

Recently, interactions between the central nervous system (CNS) and the immune system, and between the genetic heritage and the immune system have been recognized and even codified through journals such as *Neuroimmunology* and *Immunogenetics*. We argue that a cognitive socioculture – a social network embodying culture – in which individuals are embedded, and through which they are both acculturated and function to meet collective challenges of threat and opportunity, may interact strongly with individual immune function to produce a composite entity which might well be labeled an *Immunocultural Condensation* (ICC). We propose that ultimately CNS, immune system, and embedding socioculture become jointly convoluted to form a composite entity which will respond coherently, but in a complex manner, to sudden ‘delta function’ external perturbation. Elsewhere we have explored the response of such systems to long-term ‘structured’ systems of perturbation, for example embedding within an Apartheid system (R Wallace, 2001a, b). We will apply that work to understand how embedding social structures may either exacerbate or ameliorate the long-term effects of a sudden perturbation.

Increasingly, biologists are roundly excoriating simple genetic reductionism which neglects the role of environment. Lewontin (2000), for example, explains that genomes are not ‘blueprints,’ as genes do not ‘encode’ for phenotypes. Organisms are instead outgrowths of fluid, conditional interactions between genes and their environments, as well as developmental ‘noise.’ Organisms, in turn, shape their environments, generating what Lewontin terms a triple helix of cause and effect. Such interpenetration of causal factors may be embodied by an array of organismal phenomena, including, as we shall discuss, culture’s relationships with the brain and the immune system.

The current vision of human biology among evolutionary anthropologists is consistent with Lewontin’s analysis. That work is summarized by Durham (1991) to the effect that genes and culture constitute two distinct but interacting systems of 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 are best represented as two parallel lines or ‘tracks’ of hereditary influence on phenotypes.

Regarding the interaction of these tracks, evolutionary anthropologists have concluded that, over hominid evolution, genes came to encode for increasing hypersociality, learning, and language skills, so the complex cul-

tural structures which better aid in buffering the local environment became widespread in successful populations (Bonner, 1980).

Every successful human population seems to have a core of tool usage, sophisticated language, oral tradition, mythology and music, focused on relatively small family/extended family groupings of various forms. More complex social structures are build on the periphery of this basic genetic/cultural object (Richerson and Boyd, 1995, 1998).

At the level of the individual human, the genetic-cultural object appears to be mediated by what evolutionary psychologists postulate are cognitive modules within the human mind. Each module was shaped by natural selection in response to specific environmental and social conundrums Pleistocene hunter-gatherers faced. One set of such domain-specific cognitive adaptations addresses problems of social interchange (Barkow et al., 1992; Cosmides and Tooby, 1992). The human species' very identity may rest, in part, on its unique evolved capacities for social mediation and cultural transmission. Anthropologist Robert Boyd has remarked that culture is as much a part of human biology as the enamel on our teeth.

Indeed, a brain-and-culture condensation has been adopted as a kind of new orthodoxy in recent studies of human cognition. For example Nisbett et al. (2001) review an extensive literature on empirical studies of basic cognitive differences between individuals raised in what they call 'East Asian' and 'Western' cultural heritages. They view Western-based pattern cognition as 'analytic' and East-Asian as 'holistic.' Nisbett et al. (2001) find that

1. Social organization directs attention to some aspects of the perceptual field at the expense of others.
2. What is attended to influences metaphysics.
3. Metaphysics guides tacit epistemology, that is, beliefs about the nature of the world and causality.
4. Epistemology dictates the development and application of some cognitive processes at the expense of others.
5. Social organization can directly affect the plausibility of metaphysical assumptions, such as whether causality should be regarded as residing in the field vs. in the object.

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

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

We may assume, then, the existence of both gene-culture and brain-culture condensations.

Immune cognition

Atlan and Cohen (1998) have proposed an information-theoretic adaptation of IR Cohen's (1992) 'cognitive principle' model of immune function and process, a paradigm incorporating pattern recognition behaviors analogous to those of the central nervous system.

We paraphrase Atlan and Cohen's description of immune system cognitive pattern recognition-and-response behavior as follows:

The meaning of an antigen can be reduced to the type of response the antigen generates. That is, the meaning of an antigen is functionally defined by the response of the immune system. The meaning of an antigen to the system is discernible in the type of immune response produced, not merely whether or not the antigen is perceived by the receptor repertoire. Because the meaning is defined by the type of response there is indeed a response repertoire and not only a receptor repertoire.

To account for immune interpretation IR Cohen has proposed a cognitive paradigm for the immune system. The immune system can respond to a given antigen in various ways, it has 'options.' Thus the particular response we observe is the outcome of internal processes of weighing and integrating information about the antigen.

In contrast to Burnet's view of the immune response as a simple reflex, it is seen to exercise cognition by the interpolation of a level of information processing between the antigen stimulus and the immune response. A cognitive immune system organizes the information borne by the antigen stimulus within a given context and creates a format suitable for internal processing; the antigen and its context are transcribed internally into the 'chemical language' of the immune system.

IR Cohen's (2000) cognitive paradigm suggests a language metaphor to describe immune communication by a string of chemical signals. This metaphor is apt because the human and immune languages can be seen to manifest several similarities such as syntax and abstraction. Syntax, for example, enhances both linguistic and immune meaning.

Although individual words and even letters can have their own meanings, an unconnected subject or an unconnected predicate will tend to mean less than does the sentence generated by their connection.

The immune system, in Atlan and Cohen’s view, creates a ‘language’ by linking two ontogenetically different classes of molecules in a syntactical fashion. One class of molecules are the T and B cell receptors for antigens. These molecules are not inherited, but are somatically generated in each individual. The other class of molecules responsible for internal information processing is encoded in the individual’s germline.

Meaning, the chosen type of immune response, is the outcome of the concrete connection between the antigen subject and the germline predicate signals.

The transcription of the antigens into processed peptides embedded in a context of germline ancillary signals constitutes the functional ‘language’ of the immune system. Despite the logic of clonal selection, the immune system does not respond to antigens as they are, but to abstractions of antigens-in-context.

Immune and sociocultural cognition

As we show at length in the Appendix, it is possible to give Atlan and Cohen’s language metaphor of meaning-from-response a precise information-theoretic characterization, and to place that characterization within a context of recent developments which propose the ‘coevolutionary’ mutual entrainment – in a large sense – of different information sources to create larger meta-languages containing the original as subdialects. This work, a highly natural extension of formalism based on the Large Deviations Program of applied probability, also permits treating gene-culture and brain-culture condensations using a similar, unified, conceptual framework of information source ‘coevolutionary condensation’.

Cohen’s immune cognition model suggests the possibility that culture, in the sense of a local cognitive sociocultural network by which individuals are acculturated and within which they participate in collective response to patterns of threat and opportunity, and the individual immune system may be jointly convoluted. That is, there would appear to be, in precisely the sense of the gene-culture and brain-culture condensations of *immunogenetics* and *neuroimmunology*, an ‘immunoculture condensation’ as well.

Ultimately, however, these arguments suggest that CNS, immune and sociocultural cognition are strongly linked into a single composite entity by various kinds of ‘crosstalk’ – hormones and cytokines are neurologically active, while neurotransmitters are well known to have impact on the immune system. Contact with our fellows affects both.

Sudden perturbation

The WTC disaster constitutes significant and extremely broad perturbation of individual and collective life for a very large population, including both commuters and local residents of lower Manhattan. The modeling exercise of the Appendix suggests that the response to perturbation of an ‘information source’ made up of the convolution of CNS, immune and sociocultural cognition will be in terms of characteristic but ‘nonorthogonal’ eigenmodes of mixed patterns of pathology, so that a variable degree of mixing of different identifiable patterns of pathology – comorbidity of syndromes – will occur in proportion to the magnitude of the imposed perturbation. Equation (42) in particular suggests that, depending on the ‘time constants’ and ‘lag times’ of the elements of the syndrome, different aspects of pathology will ‘kick in’ at different times following the disaster. That is, although the ‘incident’ occurred at a single discrete time and place, the social, physiological, and psychological sequelae will not all appear simultaneously, but rather will develop according to their own internal ‘time clocks’ after delays which may themselves be highly, if systematically, variable. The perturbation itself will, further, interact with individual and group histories so that different individuals or groups may embark on different ‘paths’ of response.

The response may be multiply comorbid: At low levels of applied stress, be it physical, psychosocial or chemical trauma or their mix, a relatively small number of distinct syndromes of behavioral, psychological, social, or immune pathology should be observed, according to the model of the Appendix. As the ‘magnitude’ of perturbation experienced increases, however, these syndromes should become mixed, i.e. increasingly comorbid. Again, this effect should be complicated, as the ‘time of integration’ increases after perturbation, by the sequential appearance of symptoms and symptom patterns with differing time constants or lag times, in a broad sense.

Equation (43) suggests that ‘eigenpatterns’ of mixed pathology appear as synergistically amplified internal representations of external perturbation. If $\sigma(y)$ represents the standard deviation of pathology, and $\sigma(w)$ that of perturbation, then they will be related as

$$\sigma(y) \propto \frac{\sigma(w)}{|1 - \lambda|},$$

where λ is a kind of multiple correlation coefficient, so that, for ‘tightly coupled’ systems, $|\lambda| \rightarrow 1$, and this expression implies that applied perturbations will be greatly amplified by mixed pathologies of physiological,

psychological and social response. Ives (1995) has characterized natural systems for which $\lambda \rightarrow 0$ as ‘resilient’ in the sense that applied perturbations will not be amplified by internal system structures. See the Appendix for details.

Note particularly that, in this formulation, symptom patterns may express themselves within embedding social networks and associated social cognition as well as simply within individuals, that is, collective, as well as individual, ‘eigenpatterns’ of pathology. Thus patterns of pathology will be functions of ecosystem scale as well as of applied stress. Individual and collective ecologies are separate, but linked.

Further consideration, leading to equations (47-49) in the Appendix, suggests that the multiple correlation coefficient λ *may itself be a function of the maximum applied stress*, again depending on ecosystem scale. If $|w|_M$ represents that maximum applied perturbation, then

$$\lambda \rightarrow \lambda(|w|_M)$$

and the equation above becomes somewhat complicated. For example, if $\lambda(|w|_M)$ has a typical S-shaped dose-response to the maximum applied perturbation – rising toward 1 – then the ‘filter’ of the transformation $\lambda \rightarrow 1/|1 - \lambda|$ produces a sharply rising J-shaped effect.

Similarly, if $\lambda(|w|_M)$ displays an inverted-U ‘generalized signal transduction’, then that transformation produces a much sharper peak on the same scale. Signal transduction effects are to be expected where the applied perturbation carries physiological or social information, be it chemical, sensory, or whatever. Beyond a certain maximum amplitude, in a large sense, the signal ceases to be meaningful, and becomes only noise. The question of which pattern might predominate at individual and group levels of scale is an important empirical one.

Generalized signal transduction, as we have characterized it, is consistent with what Lifton terms ‘psychic numbing,’ a condition resulting from extreme or prolonged stress in which a person or group is so battered as to become more-or-less permanently unable to respond appropriately to further stress (e.g. Lifton and Markusen, 1990). Such social signal transduction, which perhaps might better be termed ‘pathological resilience’, has been observed in the South Bronx section of New York City, which, after suffering an official policy of ‘planned shrinkage’, saw vast areas lose between 50 and 80 % of their housing and population to contagious urban decay in a period

of a few years (D Wallace and R Wallace, 1998, 2000). Patterns of relation among low-weight birth, homicide rate, unemployment rate, percentage foreign-born, and percent on welfare were similar for Upper Manhattan and the South Bronx – the same ‘eigenmodes’ of exication – but the amplification factor for the former was 5.3 times as large for the latter, 12.5 vs. 2.4. We interpreted this to mean that the internal social structures of the Bronx had been permanently changed, i.e. ‘selected’ by the extreme experiences of forced population displacement and massive urban decay, so as to limit the magnitude of response to perturbation. Recognizably analogous effects, we propose, may be seen at the individual level, where a person is driven beyond the ability to respond.

Social factors can exacerbate or ameliorate symptom eigenmodes

Those exposed to the ‘perturbation’ of the WTC collapse will, according to our model, have a raised probability of developing a characteristic system of symptoms, a ‘World Trade Center Syndrome’, to quote the popular press. That syndrome will, in all likelihood, involve serious exacerbation of preexisting low level multifactorial ‘eigenpatterns’ already present within the affected population. This will be complicated, according to our model, by the emergence of a stratified ‘comormidity’ driven by increasing exposure to the disaster itself.

Once such a syndrome – or mixture of them – emerges, it or they will be affected by embedding social structures. According to the arguments of the later sections of the Appendix, the pattern of larger social response will, itself, likely be highly coherent, and indeed may constitute a kind of ‘higher language’ interacting with the perturbed ‘condensed information source’ of those within the affected group. The arguments of the Appendix suggest that the ‘language’ of that social response is very powerful indeed and can, as an externally-imposed ‘selection pressure’, literally write itself upon developing symptom patterns of embedded individuals suffering exposure to the WTC disaster.

Typical collective social responses might range from public displays of expressed emotion repeatedly triggered by public officials for political reasons of ‘national solidarity’, to grossly unrealistic expectations of ‘closure’ or ‘getting back to normal’. These may be compounded by continued exposure to the effects of the WTC disaster, repeated alarms, false or otherwise, for example relatively small ‘bioterrorist’ outbreaks have already been exaggerated for

political or other purposes, or simply from induced hypersensitivity. Indeed, other large multifatal attacks are highly likely, in the context of economic downturn, job-loss, and other stressors.

The latter part of the Appendix suggests that the onset of these effects – the ‘writing’ of systematic external stressors on symptom patterns – may itself be highly punctuated. That is, beyond a critical point, ‘collective expressed emotion’ will suddenly begin to exacerbate symptom patterns of those suffering ‘WTC syndrome’.

An alternative model suggests that intervening social structures – community groups, church groups, labor unions, artificial ‘created therapeutic communities’, or even workplace initiatives – might serve to buffer affected individuals from the larger embedding social pathology. That is, having one’s employer, union, or church take appropriate measures might well serve to dampen down patterns of symptoms rather than amplify them.

These speculations can be made quite formal by invoking the ‘higher order’ pattern recognition argument of the Appendix, assuming that the ‘WTC syndrome’ response constitutes, in fact, a phenotypic ‘choice’ made by the condensation of local cognitive sociocultural, CNS, and immune systems. That cognitive ‘choice’ of response might be expected to involve four possible outcomes:

- (1) no exacerbation of preexisting patterns,
- (2) short-term exacerbation,
- (3) persistent ‘sensitization’ to periodic re-excitation of pathological eigenmodes, i.e. generalized ‘flashbacks’ given appropriately conditioned triggers, or
- (4) permanently elevated pathological eigenpatterns.

The external selection pressures of the larger society will then, through the usual rate distortion arguments, ‘write themselves’ on the ‘language’ associated with and defined by this ‘decision oscillator.’ Thus the outputs of that oscillator will reflect those selection pressures. Supportive embedding social environments would be expected to produce outcomes (1) or (2), while ‘secondary victimization’ would induce the latter two.

Path dependence and symptom patterns: the burdens of history

The effects of both sudden perturbation and subsequent secondary victimization will be greatly affected by individual and community-level histories of stress, if recent studies can be generalized. Geronimus (1992, 1996) has,

for example, invoked individual-level ‘weathering’ in the context of social deprivation to account for differences in US Black/White birth outcomes. As described above, D Wallace and R Wallace (2000) found a draconian policy-driven ‘selection pressure’ acting on social network structures accounted for differences in ecosystem resilience between Upper Manhattan and the South Bronx. More recently Fischer et al. (2001) have applied a recognizably similar analysis to argue that ecological history is essential in predicting community responses to environmental perturbation. They conclude that if historical conditions strongly influence community and ecosystem responses to sudden change, then explicit consideration of ecological history may improve forecasting of ecological responses to environmental change and guide identification of highly sensitive systems. Fischer et al. (2001) find that past stressful environmental conditions can indeed increase sensitivity to subsequent perturbation, but that natural selection caused by perturbations may, on the other hand, alter responses to a repeated perturbation by increasing the tolerance of populations that succeed in persisting, a conclusion which seems consistent with the community-scale Bronx/Manhattan results of D Wallace and R Wallace (2000).

These considerations suggest, for the World Trade Center disaster, a scale-determined path dependence in the effects of both the initial perturbation and secondary victimization. Individuals with histories of chemical, income, psychosocial, or other stress would be expected to have exacerbated responses to perturbation and secondary victimization, unless they have been driven into the realm of generalized psychic numbing and pathological resilience. Similarly, communities which have traditionally borne burdens of discrimination or racism should be expected to have different, and possibly exacerbated, responses compared with those which have not.

At the individual level, we may expect stratification in symptom pattern and response to secondary victimization according to age, income, occupational history, and social status, while at the population level, we may expect significant differences according to ethnicity and ‘race’. Under the highly segregated US system of de-facto Apartheid, Zip Code of residence should serve as a good surrogate for population-level histories of deprivation and racism.

After a review of combustion toxicity, we are prepared to reinterpret several case histories.

A brief introduction to the ‘new’ combustion toxicology

Fumes and smoke from modern office and construction materials are not

like those from natural materials. Although wood, paper, natural fabrics, etc., can emit irritating fumes and soot in a fire, the intensity and irritancy is much less than most synthetic materials now commonly used as substitutes for them. Rigid polyvinyl chloride (PVC) emits 60 % of its weight as hydrogen chloride, HCl, in the early stages of a fire, which coats soot particles. These are small, smaller than those given off by, say, wood or cotton. They get into the lung and deliver a huge dose in a short time, destroying lung tissue by sheer corrosion.

Because of the flame retardant properties of chlorine, much more mass of the PVC goes off as soot and the smoke is extremely dense. People cannot see to escape. Plasticized PVC, as is used in communication cable, also produces highly acidic soot and fumes of dense smoke, but also gives off masses of phthalate anhydride, very irritating and explosive. Both forms of PVC produce large quantities of benzene, xylene, and toluene which are narcotic, and may give rise to secondary explosions.

A variety of halogenated species are present both as gases and in the soot. Large quantities of dioxins and dibenzofurans – notoriously powerful endocrine disruptors – form during the cooling phase, as they do in incinerators or burning landfills. PVC fumes and soot may contain as many as 300 species, as reported in the literature, ranging from formaldehyde to the polyaromatics, with and without chlorine. Most PVCs also contains a variety of additives which may include lead stearate as a stabilizer, phthalate plasticizers, and metal-based colorants.

The brominated fire retardants can be present in other plastics and in fabrics at more than 10 % by weight and form similar pyrolysis and combustion products. Other halogenated plastics and highly brominated-flame retarded plastics and fabrics behave analogously to PVC.

Nitrogen-based polymers such as urethane foam, acrylonitrile-butadiene-styrene (ABS) and some of the acrylics will likely have played a considerable role. Hydrogen cyanide, HCN, is emitted massively from these materials shortly before ignition. Cyanide also delivers a skin dose, besides a lung dose. Nitrogenated organics also form, of course. These should contain large quantities of flame retardants of different kinds, including brominated ethers or metal salts.

The WTC fuel load, between the two planes and the several buildings which burned, collapsed, and then smoldered copiously for weeks, would make the 1975 New York Telephone Exchange fire look like a minor event. That fire involved over 200,000 pounds of plasticized PVC (D Wallace, 1981,

1982, 1990, 1992).

The 1975 NY Telephone firefighters, and survivors of the 1980 MGM Grand Hotel plastics fire, suffered a spectrum of subsequent symptoms which we examine in some detail below.

Two toxic fire case histories

The New York Telephone Company and MGM Grand Hotel fires, and their effects on survivors, are described at great length in D Wallace (1981, 1982, 1990, 1992). We summarize that work briefly here.

As we indicated above, the Telephone Company fire involved about 100 tons of heavily plasticized PVC cable burning over a 14 hour period. Some 700 firefighters responded to the blaze, 239 of them reporting themselves as injured. 113 of these latter subsequently participated in a 1980 survey sponsored by the Uniformed Firefighters Association, the firefighters union.

The MGM Grand Hotel fire involved a very large, mixed load of various common plastics burning at ground level – PVC, ABS, etc. Rising smoke trapped hotel patrons in their rooms, killing 85 of them at a considerable distance from the fire itself. 54 injured survivors participated in a subsequent detailed study sponsored by the Plaintiffs Legal Committee, which oversaw the resulting lawsuit.

Firefighters and civilians constitute separately socialized populations – different and quite distinct ‘sociocultures’ – in their response to the fire environment. To the former, it is the ‘normal’ working condition, for which they are rigorously trained. To the latter, it is a horror and a gross and traumatic aberration from normal life. Very large or multifatal fires, or persistent episodes of contagious urban decay which devastate communities, however, will breach even the paramilitary cultural discipline of the fire service.

Tables 1-3, below, describe the pattern of reported firefighter symptoms from the NY Telephone Company fire on short, intermediate and long time scales, respectively at the fire, two weeks to three months after, and six months or more after the fire. 64 % of the respondents complained of persistent symptoms. It is of some note that the pattern differed significantly between those with long and short work histories. Those over thirty-five years of age and those who had served more than ten years as firefighters tended to have permanent damage. Indeed, of those who had served more than 15 years, 80 % reported permanent injury.

Tables 4 and 5, by contrast, examine the respiratory and non-respiratory complaints of the sample of injured survivors of the MGM Grand Hotel fire.

Reproductive, neurological, and psychological dysfunction, in contrast to the firefighter sample, dominate.

Besides the group pattern shown in Tables 4 and 5, certain individuals experienced special disabilities from the smoke, for example muscle spasms which show that certain parts of the nervous system that cause involuntary movement of the normally voluntary skeletal muscles were affected. Both Parkinsonism and epilepsy are based on these parts of the nervous system. In addition, sufferers of Parkinsonism and epilepsy have low levels of one or more catecholamines, as do depressed individuals. One injured survivor had Parkinsonism even before the fire, but required a much greater amount of medication to control his condition after the fire. A second person became epileptic after the fire and suffered grand mal seizures. These two individuals are index cases of the neurotoxic nature of the smoke.

Besides the physical injuries and individual psychological distress, social interactions appear to have been affected by the experience of the MGM fire, possibly synergistic with exposure to neurotoxins. Marriages became troubled and were dissolved. The divorce rate among the survivors was very high. This high rate of divorce is also seen in other disasters besides fires, and reflects personality changes and relational changes that often result from having risked death and knowing that others died in the same event.

In addition to divorce, another socioeconomic consequence of the MGM fire was the decline of several businesses that had sent key people to conventions or seminars that were held at the hotel. These people stayed overnight in the hotel, and several within the same office may have suffered personality changes. Individual productivity typically declined, but group interactions also suffered so that decisions could not be made, consensus could not be reached, and actions could not be taken. At least two mid-sized companies suffered greatly and became much smaller when several key managers could not perform.

The MGM Grand Hotel fire in particular shows precisely the inextricably mixed pattern of physical, neurological, psychological, and social pathology to be expected from the ‘perturbation’ of a traumatic toxic exposure incident.

Traumatic toxic incident of sufficient magnitude should be expected to produce a similar pattern among affected emergency services personnel, in spite of both discipline and experience. ‘Post Traumatic Stress Disorder’ is a grossly inadequate characterization.

These two examples appear to be fully consistent with the theoretical development we have presented as an introductory context, showing a distinct

and characteristic pattern of symptoms. The next case histories explore in more detail the relation between preexisting 'eigenmodes' and exposure.

Table 1: Immediate symptoms of Firefighters

INJURY	PERCENT AFFECTED
Sore throat, irritated eyes, dizziness, aching nostrils, confusion, weakness, and exhaustion	Over 50
Chest pains, nausea, chest congestion, and headache	35-50
Irritated skin and faintness	20-30
Loss of control of arms and/or legs	10-20

Table 2: Intermediate time symptoms of Firefighters

SYMPTOMS	PERCENT AFFECTED
RESPIRATORY:	
Chest congestion	51.2
Chronic cough	22.1
Sore throat	18.5
Sore chest	9.3
Hoarseness, wheezing, allergy to smoke, difficulty breathing, irritated nasal	

membranes, shortness of breath	Less than 5

NEUROLOGICAL:	
Muscular weakness	19.8
Impaired smell/taste	16.3
Increased irritability	10.5
Headaches	10.5
Perception difficulty	7.0
Confusion, anxiety, numbness of extremities	Less than 5

MISCELLANEOUS SYMPTOMS:	
Heart trouble, irritated eyes, irritated skin	2 complaints each
Chills, sinus trouble, weight loss, bowel problem, nausea, head congestion	1 complaint each
GENERAL WELL-BEING	
Fatigue	5.8
Impaired endurance	3.5

Table 3: Long-Term Effects among Firefighters

INJURY	PERCENT AFFECTED

RESPIRATORY:	
Impaired disease resistance	37.5
Coughing	33.3
Hoarseness	23.6
Shortness of breath	9.7
'Lung function' or pain	15.8
Chest congestion	9.7
Sensitivity to smoke	11.1
Sinus or nasal drip	6.9
Repetitive bronchitis	8.3
Sore throat	8.3

Asthma	6.9
Allergy, unspecified upper respiratory problem	Less than 5%

GROWTHS (epidermal or membrane lining)	13.9

HEART:	
Palpitations, acute myocardial infarction, prolapsed mitral valve, enlarged heart, unspecified damage	8.3

Headaches, perception difficulty	4 individual complaints
Fatigue, kidney-urinary tract	3 individual complaints
Weakness, pancreatitis/diabetes	2 individual complaints
Elevated blood count, elevated bilirubin, high pressure, gall bladder deterioration, irritation of hemorrhoid, irritated eyes, convulsive seizures	1 individual complaint

Table 4: Respiratory Complaints of 58 MGM Survivors

COMPLAINT	NUMBER	PERCENT

Frequent sore throat	30	51.7
Hoarseness	26	44.8
Sinusitis	23	39.7
Sensitive to dust/smoke	44	75.9
Coughs	27	46.6
Wheezing	26	44.8
Phlegm production	29	50.0
Shortness of breath	38	65.5
Bronchitis	21	36.2
Frequent upper respiratory		

infection	25	43.1
-----------	----	------

Table 5: Non-Respiratory Symptoms of 57 MGM Survivors

SYMPTOMS	NUMBER	PERCENT
<hr/>		
CIRCULATORY SYSTEM:		
Developed heart problem	12	21.1
Hands/feet fall asleep	39	68.4
Hands/feet easily cold	22	38.6
Mottling or blue hands/feet	8	14.0
Blood pressure change	20	35.1
<hr/>		
NEUROLOGICAL/PSYCHONEUROLOGICAL:		
Headaches	34	59.6
Change in sleep pattern	42	73.7
Memory lapses	32	56.1
Irritable	43	75.4
Difficulty learning	14	24.6
Change in perception abilities	31	54.4
Coordination decline	15	26.3
Dizziness	27	47.4
Disorientation	20	35.1
<hr/>		
KIDNEY/BLADDER:		
Frequent urination	15	26.3
Pain/burning during urination	12	21.1
Lower back pains	27	47.4
<hr/>		
SKIN CHANGES:		
Acne-like breakouts	13	22.8
Rashes	12	21.1
Burns from the smoke	4	7.0
Infections	6	10.5
Skin growths (warts, moles)	3	5.3
Skin dry, sensitive	14	24.6

REPRODUCTIVE:*		
New menstrual difficulties	11	84.6*
Miscarriage	2	15.4*
Hysterectomy	2	15.4*
Dilation and curettage	1	7.7*

PSYCHOLOGICAL:		
Nightmares	42	73.7
Depression	40	70.2
Guilt	15	26.3
Anger	17	29.8
Change in relationships	24	42.1

GENERAL WELL-BEING:		
Change in appetite	16	28.1
Less endurance	31	54.4
General weakness	24	42.1
Ceased/decreased activity	17	29.8
Lack of sexual desire	14	24.6

OTHER COMPLAINTS:		
Sensitized eyes	11	19.3
Dizziness with position change	8	14.0
Heavy sweating	7	12.3
Drink more water	3	5.3
Tremors, spasms, clenching	6	10.5

* Percentage is that of the women of reproductive age in the data base, not of all 57 people.

The following affected less than 5 percent:
liver dysfunction, nausea, neck pains, impotence,
pigmentation of skin, loss of hair, ears hurt,
swollen glands, onset of epilepsy,
exacerbation of Parkinsonism,
return of menses to elderly women,

swollen testes, prostate infections.

Other chemical disasters

Bowler et al. (1994, 1996, 1997) have examined a spectrum of toxic disasters which spanned both ‘ethnicity’ and chemical species, finding unexpectedly analagous population-level patterns of symptom response – headaches, respiratory, visual, gastrointestinal and dermatologic. Exposure relationships were found with increasing symptoms and worsening of preexisting illnesses. The three groups were Whites exposed to metam sodium, Hispanics working in a microelectronics plant exposed to multiple organic solvents, and African-Americans exposed to sulphuric acid. Using a similar symptom instrument across exposures, Bowler et al. (1996) conclude that their results “suggest a robust symptom complex following chemical exposure regardless of the specific chemical”.

Bowler et al. (1997) state that the relationship of stress and illness has been well established since Hans Selye’s early work (1976) on the Stress Response Syndrome, and is further shown in the association of such reactive and serious disorders as PTSD. They cite the work of McFarlane et al. (1994), who found physical symptoms to be an ‘accompaniment’ of PTSD in a sample of firefighters who had higher rates of cardiovascular, respiratory, musculoskeletal, and neurological symptoms. McFarlane et al. suggest, as did Kinson and Rosser (1974), that much psychological work is needed to minimize the impact of chemical releases and disasters on subsequent physical and psychological illness.

While Bowler et al. (1997) did not conduct a multivariate analysis of their data, they found, in one case/control study of matched exposed/unexposed communities, that the same patterns were seen in both communities, with those of the exposed community much worse than in the unexposed, concluding that

“Although both towns were similar on reported illnesses prior to the [chemical] release and similar on other nonreported illnesses, it is notable that many of the exposed reported that their illnesses were worse six months after the release. Those who had prior acute or chronic bronchitis reported worsening of their condition by 14 and 7 times more (respectively). By a magnitude

of 5 times, those who had prior asthma or allergies also reported their condition to be worse.”

This result, again, suggests the amplification of similar underlying community ‘eigenmodes’ by chemical exposure.

The next studies, however, do use multivariate methods, and, in our view, successfully isolate eigenmodes.

Gulf War Syndrome

Recently a number of researchers have begun to critically examine characteristic patterns of ‘medically unexplained somatic symptoms’, including multiple chemical sensitivity (MCS), chronic fatigue syndrome (CFS), fibromyalgia (FM), irritable bowel syndrome (IBS) and Gulf War Syndrome (GWS). Kipen and Fiedler (1999) find that psychiatric explanations and nomenclature have less than a perfect fit for many cases of unexplained symptoms, although there is often very great comorbidity among them, often as much as 50 % between MCS and CFS, for example. In the case of Gulf War symptoms, careful epidemiology has shown that, compared with soldiers who did not deploy to the Persian Gulf, those who deployed had two- to threefold increases in symptoms, without apparent medical explanations (Kipen and Fieldler, 1999; The Iowa Group, 1997; Fukuda et al., 1998).

Haley et al. (1997) used multivariate methods to examine symptom patterns among about 250 Gulf War veterans in a construction battalion, and found three particularly distinct and characteristic clusters, roughly characterized as ‘impaired cognition’, ‘confusion-ataxia’, and ‘arthro-myoneuropathy’. Veterans with the second pattern were found to be 12.5 times more likely to be unemployed than those with no health problems. Haley et al. conclude that clusters of symptoms of many Gulf War veterans represent discrete factor-analysis derived syndromes that appear to reflect a spectrum of neurologic injury involving the central, peripheral, and autonomic nervous systems. Subsequent work by Haley et al. (2001) extends these studies.

Knoke et al. (2000) responded directly to the report of Haley et al. (1997) with their own multivariate study which compared Gulf War-era veterans who had and had not been deployed to the Persian Gulf. Closely similar clusters of syndromes were identified within both groups, although the deployed veterans showed these clusters with greater prevalence than the non-deployed, and the strongest clusters matched fairly well with the observations of Haley et al. (1997).

Doebbeling et al. (2000) conducted a similar deployed/non-deployed veterans study, and again found similar clusters of symptoms within both groups, with the deployed again having very much greater prevalence than the non-deployed.

Doebbeling et al. (2000) conclude, however, that the markedly increased prevalence of nearly every symptom assessed from all bodily organs among the Gulf War veterans is difficult to explain pathophysiologically as a single condition. Identification of the same patterns of symptoms among the deployed veterans and the nondeployed controls suggests that the health complaints of Gulf War veterans are similar to those of the general military population and are not, in their view, consistent with the existence of a unique Gulf War syndrome.

The modeling exercise of the Appendix, leading to equation (43), however, suggests a more consistent explanation of these results. Deployed and non-deployed veterans were embedded within very similar – and highly rigid and nonresilient – sociocultural structures, and seemed to represent very similar cross-sections of the US population. This suggests that the ‘multiple correlation coefficient’ representing the linking of internal mechanisms would satisfy the condition $\lambda \rightarrow 1$, giving a very large amplification pattern for perturbations, $\sigma(w)$. Combat experience is indeed a significant perturbation at both individual and social levels, and amplified ‘eigenmodes’ of characteristic patterns of mixed pathology are, from this model, the expected result at both scales. Some degree of comorbidity is inherent, since the response ‘eigenmodes’ are nonorthogonal.

Predicting the spectrum of WTC pathologies

On the basis of our theoretical development and these case histories, it is not difficult to construct an expected pattern of symptoms for both civilians and emergency workers having significant exposures to the WTC disaster, and this is left as an exercise for the reader. The great prevalence of uterine dysfunction among women exposed in the MGM Grand Hotel fire suggests, however, particularly nasty consequences for those in the WTC disaster: the presence of massive quantities of endocrine disruptors and other physiologically active chemical species in the aftermath of the WTC attack takes the matter far beyond either Post Traumatic Stress Disorder or Gulf War Syndrome.

To reiterate, responses will be scale-dependent, with different characteristic patterns seen the individual and group levels – the usual ‘ecosystem’

effect. These responses will also be dependent on individual and group histories of stress, deprivation, and the effects of racism. As described above, dose-response and signal transduction may be observed at both individual and group levels, for a number of systems, depending on the magnitude of perturbation. At the purely psychological level, signal transduction might well be described by Lifton's term 'psychic numbing'.

These patterns will be exacerbated or ameliorated by the imposed 'selection pressure' of influence by embedding social structures, in a manner determined in no small part by past individual and group experiences of stress, deprivation, and racism. External social pressures will drive, according to our model, much of the 'phenotype choice' of syndromic response to applied perturbation by affected individuals and groups. Path-dependent 'secondary victimization' by the larger society seems the most likely outcome, but intermediate structures such as labor unions, work-place initiatives, church groups, or artificial therapeutic communities, may serve to buffer such effects and perhaps ameliorate symptoms.

Note that this is a true 'selection pressure' effect, since those suffering syndrome exacerbation face shortened lifespan.

Less obvious, however, are the inevitable consequences of the couplings of cognitive process *across scale*. As one example of what are likely to be many subtle but critical cross-system impacts we examine in more detail the possible convolution of immune function into the more general pathology.

As is well known, dioxins can cause severe immunosuppression (e.g. Tonn et al., 1996), an effect possibly complicated by an inverted-U dose-response effect (e.g. Fang et al., 1996). This is to be contrasted to the standard dose-response effect on the mucociliary tract which is expected from exposure to irritants such as HF, HCl and HBr. As is increasingly understood, however, psychosocial stress may itself have an immunosuppressive effect which is indeed sufficient to interfere with vaccine efficacy.

Recent work by Kiecolt-Glaser and Glaser (1996, 1998, 2000), for example, has examined the effect of 'chronic stress' on the efficacy of influenza, hepatitis-B, and pneumococcal pneumonia vaccine among elderly caregivers of dementia patients, and among medical students.

They found, for influenza, that the caregivers showed a poorer antibody response following vaccination relative to control subjects, as assessed by ELISA and hemagglutination inhibition. Caregivers also had lower levels of *in vivo* virus-specific-induced interleukin 2 levels and interleukin 1β . The data demonstrate that down-regulation of the immune response to influenza

virus vaccination is associated with a chronic stressor in the elderly.

Similar effects were found among the elderly caregivers for response to pneumococcal pneumonia vaccination, leading to the conclusion that chronic stress can inhibit the stability of the IgG antibody response to a bacterial vaccine.

Medical students who reported greater social support and lower anxiety and stress demonstrated a higher antibody response to HEP-B surface antigen at the end of the study period.

Glaser et al. (2000) conclude that the differences in antibody and T-cell responses to HEP-B and influenza virus vaccinations provide a demonstration of how stress may be able to alter both the cellular and humoral immune responses to vaccines and novel pathogens in both younger and older adults.

In addition a vast body of animal model studies involving socially structured populations shows clear impacts of acute and chronic social and other stressors on immune competence (e.g. deGroot et al., 2001; Gryazeva et al., 2001). Elenkov and Chrousos (1999) in particular suggest that glucocorticoids and catecholamines, the end-products of the stress system at the individual level, might selectively suppress cellular immunity, Th1 phenotype, in favor of humoral response – again at the individual level.

We suggest, however, that the particular role of cognitive socioculture in human biology takes matters considerably beyond such individual-level stress models, and into realms for which, to paraphrase Robert Boyd’s aphorism, culture is as much a part of the human immune system as are T-cells.

It seems likely that those exposed to both stress and toxic combustion products at the WTC disaster will enter the coming influenza season during the peak 3-9 month period of ‘normal’ grieving, and may be both markedly more susceptible to infection and markedly less responsive to the available vaccine than unexposed populations. This effect may be complicated by persisting mucociliary damage consequent on exposure to smoke irritants. Vaccine strategy for the coming season, and probably for many seasons to come, must be significantly altered in the New York metropolitan region if a large number of excess flu/pneumonia deaths are to be averted.

Analogous coupling mechanisms at both individual and larger social scales will become clear as matters unfold. A particularly important effect might well involve carcinogenesis, since dioxins are among the most potent of cancer-inducing substances. Affected individuals or groups may, however, be damaged by pathological group processes, so that ‘second victimization’, by further suppressing immune function, may indirectly promote cancer induction

as well.

‘Post Traumatic Stress Disorder’ does not seem an adequate description of the likely consequences of this attack for individuals, families, social networks, enterprises, communities, and larger organizational structures. In particular, the massive quantity of acutely combustion-toxic halogenated hydrocarbons within the affected buildings constitutes a serious, and continuing, exacerbation of the disaster.

Medical practitioners treating victims of the WTC disaster and their families over the next several years, and administrators or others dealing with larger affected groups, should expect characteristic but complex patterns of respiratory, immune, reproductive, neurological, psychological, behavioral, social and other pathology to emerge, persist, and interact for the foreseeable future. Further, this is not the first, nor is it likely the last, such incident: repeated exposures seem increasingly likely, and will be synergistic.

Given the current political climate, considerable secondary victimization – through neglect, or the result of deliberately cultivated collective expressed emotion – seems inevitable. This will likely much exacerbate symptom patterns. To reiterate, it may prove necessary to invoke labor union, workplace, religious, or artificial therapeutic communities as intermediate structures to ameliorate or even reverse these effects. Those without access to such structures face a bleak future indeed.

References

- Ash R, 1990, *Information Theory*, Dover Publications, New York.
- Atlan H, 1983, “Information theory,” pp. 9-41 in R Trappl (ed.), *Cybernetics: Theory and Application*, Springer-Verlag, New York.
- Atlan H, 1987, “Self creation of meaning,” *Physica Scripta*, **36**, 563-576.
- Atlan H, 1998, “Intentional self-organization, emergence and reduction: towards a physical theory of intentionality,” *Theses Eleven*, **52**, 6-34.
- Atlan H and IR Cohen, 1998, Immune information, self-organization and meaning, *International Immunology*, **10**, 711-717.
- Barkow J, L Cosmides and J Tooby (eds.), 1992, *The Adapted Mind: Biological approaches to mind and culture*, University of Toronto Press.
- Beck C and F Schlogl, 1993, *Thermodynamics of Chaotic Systems*, Cambridge University Press, New York.
- Bennett C, 1988, “Logical depth and physical complexity”. In: Herkin R (Ed) *The Universal Turing Machine: A Half-Century Survey*, Oxford University Press, pp. 227-257.

- Bonner J, 1980, *The Evolution of Culture in Animals*, Princeton University Press, Princeton, NJ.
- Bowler R, D Mergler, G Huel and J Cone, 1994, "Psychological, psychosocial, and psychophysiological sequelae in a community affected by a railroad chemical disaster", *Journal of Traumatic Stress*, **7**, 601-624.
- Bowler R, G Huel, D Mergler, J Cone, S Rauch and C Hartney, 1996, "Symptom base rates after chemical exposure for white, Hispanic and African-Americans", *Neurotoxicology*, **17**, 793-802.
- Bowler R, L Ngo, C Hartney, K Lloyd, I Tager, J Midtling and G Huel, 1997, "Epidemiological health study of a town exposed to chemicals", *Environmental Research*, **72**, 93-108.
- Campbell R and S Raja, 1999, "Secondary victimization of rape victims: insights from mental health professionals who treat survivors of violence", *Violence and Victimization*, **14**, 261-275.
- Cohen IR, 1992, The cognitive principle challenges clonal selection, *Immunology Today*, **13**, 441-444.
- Cohen IR, 2000, *Tending Adam's Garden: evolving the cognitive immune self*, Academic Press, New York.
- Cohen IR, 2001, "Immunity, set points, reactive systems, and allograft rejection." To appear.
- Cosmides L and J Tooby, 1992, Cognitive adaptations for social exchange, in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford University Press, New York.
- Courant R and D Hilbert, 1989, *Methods of Mathematical Physics Volume II*, John Wiley and Sons, New York.
- Cover T and J Thomas, 1991, *Elements of Information Theory*, Wiley, New York.
- de Groot J, M Ruis, J Scholten, J Koolhaas and W Boersma, 2001, "Long-term effects of social stress on antiviral immunity in pigs", *Physiology and Behavior*, **73**, 145-158.
- Dembo A and O Zeitouni, 1998, *Large Deviations: Techniques and Applications, 2nd Ed.*, Springer-Verlag, New York.
- Doebbling B, W Clarke, D Watson, J Torner, R Woolson, M Voelker, D Barrett and D Schwartz, 2000, "Is there a Persian Gulf War syndrome? Evidence from a large population-based survey of veterans and nondeployed controls", *American Journal of Medicine*, **108**, 695-704.
- Durham W, 1991, *Coevolution: Genes, Culture and Human Diversity*, Stanford University Press, Palo Alto, CA.

Elenkov I and G Chrousos, 1999, Stress, cytokine patterns and susceptibility to disease, *Bailliere's Clinical Endocrinology and Metabolism*, **13**, 583-595.

Elitzur 1996, "Life's emergence is not an axiom: a reply to Yockey", *Journal of Theoretical Biology*, **180**, 175-180.

Fang 1996, *Toxicology*, **106**, 221-228.

Feigenbaum M, 1988, "Presentation functions, fixed points and a theory of scaling function dynamics", *Journal of Statistical Physics*, **52**, 527-569.

Fischer J, J Klug, A Ives and T Frost, 2001, "Ecological history affects zooplankton community responses to acidification", *Ecology*, **82**, 2984-3000.

Fukuda K, R Nisenbaum, G Stewart, et al., 1998, "Chronic multisymptom illness affecting Air Force veterans of the Gulf War", *Journal of the American Medical Association*, **280**, 981-988.

Geronimus A, 1992, "The weathering hypothesis and the health of African-American women and infants: evidence and speculations", *Ethnic Disparities*, **2**, 207-221.

Geronimus A, 1996, "Black/white differences in the relationship of maternal age to birthweight: a population-based test of the weathering hypothesis", *Social Science and Medicine*, **42**, 589-597.

Glaser R, J Kiecolt-Glaser, W Malarkey and J Sheridan, 1998, "The influence of psychological stress on the immune system response to vaccines," *Annals of the New York Academy of Sciences*, **840**, 649-655.

Glaser R, J Sheridan, W Malarkey, R MacCallum and K Kiecolt-Glaser, 2000, "Chronic stress modulates the immune response to a pneumococcal pneumonia vaccine," *Psychosomatic Medicine*, **62**, 804-807.

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

Gryazeva N, A Shurlygina, L Verbitskaya, E Mel'nikova, N Kudryavtseva and V Trufakin, 2001, "Changes in various measures of immune status in mice subject to chronic social conflict", *Neuroscience and Behavioral Physiology*, **31**, 75-81.

Haley R, T Kurt and J Hom, 1997, "Is there a Gulf War Syndrome? Searching for syndromes by factor analysis", *Journal of the American Medical Association*, **211**, 215-222.

Haley R, G Luk and F Petty, 2001, "Use of structural equation modeling to test the construct validity of a case definition of Gulf War syndrome: Invariance over developmental and validation samples, service branches and publicity", *Psychiatry Research*, **102**, 175-200.

- Higgins K, A Hastings, J Sarvaela and L Botsford, 1997, "Stochastic dynamics and deterministic skeletons: population behavior of Dungeness Crab", *Science*, **276**, 1431-1435.
- Holling C, 1973, "Resilience and stability of ecological systems", *Annual Reviews of Ecological Systematics*, **4**, 1-25.
- Holling C, 1992, Cross-scale morphology, geometry, and dynamics of ecosystems, *Ecological Monographs*, **62**, 447-502.
- Ives A, 1995, "Measuring resilience in stochastic systems", *Ecological Monographs*, **65**, 217-233.
- Khinchine A, 1957, *The Mathematical Foundations of Information Theory*, Dover Publications, New York.
- Kiecolt-Glaser K, R Glaser, S Gravenstein, W Malarkey and J Sheridan, 1996, "Chronic stress alters the immune response to influenza virus vaccine in older adults", *Proceedings of the National Academy of Sciences*, **93**, 3043-3047.
- Kinson W and R Rosser, 1974, "Disaster: effects on mental and physical state", *Journal of Psychosomatic Research*, **18**, 437-456.
- Kipen H and N Fiedler, 1999, "Invited commentary: sensitivities to chemicals – context and implications", *American Journal of Epidemiology*, **150**, 13-16.
- Knoke J, T Smith, G Gray, K Kaiser and A Hawksworth, 2000, "Factor analysis of self-reported symptoms: Does it identify a Gulf War syndrome?", *American Journal of Epidemiology*, **152**, 379-388.
- Landau L and E Lifshitz, 1959, *Classical Mechanics*, Addison-Wesley, Reading, MA.
- Larkin M, 2001, "New York physicians respond to the terror, tragedy, and trauma," *The Lancet*, **358**, 940.
- Lewontin R, 2000, *The Triple Helix: Gene, Organism and Environment*, Harvard University Press, Cambridge, MA.
- Lifton R and E Markusen, 1990, *The Genocidal Mentality: Nazi Holocaust and Nuclear Threat*, Basic Books, New York.
- Luce R, 1977, "Several unresolved conceptual problems of mathematical psychology", *Journal of Mathematical Psychology*, **41**, 79-87.
- McCauley L, 1993, *Chaos, Dynamics and Fractals: an algorithmic approach to deterministic chaos*, Cambridge University Press, Cambridge, UK.
- McFarlane A, M Atchison, E Rafalowicz, and P Papay, 1994, "Physical symptoms in post-traumatic stress disorder", *Journal of Psychosomatic Research*, **38**, 715-726.

Newsweek Magazine, November 5, 2001, “Now ‘WTC Syndrome’ ”. Article by David France, p. 10.

New York Post, October 29, 2001, “Docs: ‘WTC Syndrome’ afflicting thousands,” *New York Post*. Story by Rod Dreher.

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

Park H, S Amari and K Fukumizu, 2000, “Adaptive natural gradient learning algorithms for various stochastic models,” *Neural Networks*, **13**, 755-764.

Richerson P and R Boyd, 1995, The evolution of human hypersociality. Paper for Ringberg Castle Symposium on Ideology, Warfare and Indoctrinability (January, 1995), and for HBES meeting, 1995.

Richerson P and R Boyd, 1998, “Complex societies: the evolutionary origins of a crude superorganism,” to appear.

Rojdestvenski I and M Cottam, 2000, “Mapping of statistical physics to information theory with applications to biological systems,” *Journal of Theoretical Biology*, **202**, 43-54.

Seyle H, 1976, *The Stress of Life*, McGraw-Hill, New York.

Stefanski V, G Knopf and S Schulz, 2001, “Long-term colony housing of Long Evans rats: immunological, hormonal, and behavioral consequences”, *Journal of Neuroimmunology*, **114**, 122-130.

Spelberg B and J Edwards, 2001, Type 1/type 2 immunity in infectious diseases, *Clinical Infectious Diseases*, **32**, 76-102.

Tarrier N, C Sommerfeld and H Pilgrim, 1999, “Relatives’ expressed emotion (EE) and PTSD treatment outcome”, *Psychological Medicine*, **29**, 801-811.

The Iowa Persian Gulf Study Group, 1997, “Self-reported illness and health status among Gulf War veterans – a population-based study”, *Journal of the American Medical Association*, **277**, 238-245.

Tonn et al. 1996, *Environmental Health Perspectives*, **104**, 422-426.

Wallace D, 1981, “Dangers of polyvinyl chloride wire insulation decomposition. I. Long term health impairments: studies of firefighters of the 1975 New York Telephone fire and of survivors of the 1977 Beverly Hills Supper Club fire,” *Combustion Toxicology*, **8**, 205-232.

Wallace D, N Nelson and T Gates, 1982, “Polyvinyl chloride wire insulation decomposition II. Consideration of long term health effects from chlorinated hydrocarbons, *Combustion Toxicology*, **9**, 105-112.

Wallace D, 1990, *In the Mouth of the Dragon: Toxic fires in the age of plastics*, Avery, Garden City, NY.

Wallace D, 1992, "Analysis and interpretation of the plumbing materials file of the NYS combustion toxicity data bank," *Fire and Materials*, **16**, 77-94.

Wallace D and Wallace R, 1998, *A Plague On Your Houses*, Verso Publications, New York,

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

Wallace R and RG Wallace, 1998, Information theory, scaling laws and the thermodynamics of evolution, *Journal of Theoretical Biology*, **192**, 545-559.

Wallace R and RG Wallace, 1999, Organisms, organizations and interactions: an information theory approach to biocultural evolution, *BioSystems*, **51**, 101-119.

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

Wallace R, 2000b, Information resonance and pattern recognition in classical and quantum systems: toward a 'language model' of hierarchical neural structure and process,

www.ma.utexas.edu/mp_arc-bin/mpa?yn=00-190.

Wallace R, 2001a, Adaptation, punctuation and rate distortion: non-cognitive 'learning plateaus' in evolutionary process,

www.ma.utexas.edu/mp_arc-bin/mpa?yn=01-163

Wallace R, 2001b, Selection pressure and organizational cognition: implications for the social determinants of health, cogprints.soton.ac.uk, Manuscript number cog00001526.

Wallace R, Y Huang, P Gould and D Wallace, 1997, "The hierarchical diffusion of AIDS and violent crime among US metropolitan regions: inner-city decay, stochastic resonance and reversal of the mortality transition," *Social Science and Medicine*, **44**, 935-947.

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

Wilson K, 1971, “Renormalization group and critical phenomena. I Renormalization group and the Kadanoff scaling picture”, *Physical Review B*, **4**, 3174-3183.

MATHEMATICAL APPENDIX

The essence of the modeling approach is to, first, express the cognitive pattern recognition-and-response described as characterizing immune cognition by Atlan and Cohen (1998) in terms of a ‘language,’ in a broad sense, and then to show how that language can interact and coalesce with similar cognitive languages at larger scales – central nervous system (CNS) and the embedding local sociocultural network.

The next step is to model the way such a coalesced system may, in turn, interact with sudden external ‘perturbation.’ The ‘nonorthogonal eigenmode’ solution emerges from this discussion after a certain amount of development. We then proceed to model the effect of embedding structures as a kind of ‘selection pressure’ which writes itself on embedded information sources.

We begin with a summary of relevant information theory formalism.

Information theory preliminaries

Suppose we have an ordered set of random variables, X_k , at ‘times’ $k = 1, 2, \dots$ – which we call \mathbf{X} – that emits sequences taken from some fixed alphabet of possible outcomes. Thus an output sequence of length n , x_n , termed a path, will have the form

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

where α_k is the value at step k of the stochastic variate X_k ,

$$X_k = \alpha_k.$$

A particular sequence x_n will have the probability

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

(1)

with associated conditional probabilities

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

(2)

Thus substrings of x_n are not, in general, stochastically independent. That is, there may be powerful serial correlations along the x_n . We call \mathbf{X} an information source, and are particularly interested in sources for which the long run frequencies of strings converge stochastically to their time-independent probabilities, generalizing the law of large numbers. These we call *ergodic* (Ash, 1990, Cover and Thomas, 1992, Khinchine, 1957. We will refer to these in the future as ACTK). If the probabilities of strings do not change in time, the source is called *memoryless*.

We shall be interested in sources which can be parametrized and that are, with respect to that parameter, *piecewise memoryless*, i.e. probabilities do not change markedly within a ‘piece,’ but may do so between pieces. This allows us to apply the simplest results from information theory, and to use renormalization methods to examine transitions between ‘pieces.’ Learning plateaus represent regions where, with respect to the parameter, the system is, to first approximation, memoryless in this sense. In what follows we use the term ‘ergodic,’ to mean ‘piecewise memoryless ergodic.’

For any ergodic information source it is possible to divide all possible sequences of output, in the limit of large n , into two sets, S_1 and S_2 , having, respectively, very high and very low probabilities of occurrence. Sequences in S_1 we call *meaningful*.

The content of information theory’s Shannon-McMillan Theorem is twofold:

First, if there are $N(n)$ meaningful sequences of length n , where $N(n) \ll$ than the number of all possible sequences of length n , then, for each ergodic information source \mathbf{X} , there is a unique, path-independent number $H[\mathbf{X}]$ such that

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

(3)

See ACTK for details.

Thus, for large n , the probability of *any* meaningful path of length $n \gg 1$ – independent of path – is approximately

$$P(x_n \in S_1) \propto \exp(-nH[\mathbf{X}]) \propto 1/N(n).$$

(3)

This is the *asymptotic equipartition property* and the Shannon-McMillan Theorem is often called the Asymptotic Equipartition Theorem (AEPT).

$H[\mathbf{X}]$ is the *splitting criterion* between the two sets S_1 and S_2 , and the second part of the Shannon-McMillan Theorem involves its calculation. This requires introduction of some nomenclature.

Suppose we have stochastic variables X and Y which take the values x_j and y_k with probability distributions

$$P(X = x_j) = P_j$$

$$P(Y = y_k) = P_k$$

Let the joint and conditional probability distributions of X and Y be given, respectively, as

$$P(X = x_j, Y = y_k) = P_{j,k}$$

$$P(Y = y_k | X = x_j) = P(y_k | x_j)$$

The *Shannon uncertainties* of X and of Y are, respectively

$$H(X) = - \sum_j P_j \log(P_j)$$

$$H(Y) = - \sum_k P_k \log(P_k)$$

(4)

The *joint uncertainty* of X and Y is defined as

$$H(X, Y) = - \sum_{j,k} P_{j,k} \log(P_{j,k}).$$

(5)

The *conditional uncertainty* of Y given X is defined as

$$H(Y|X) = - \sum_{j,k} P_{j,k} \log[P(y_k|x_j)].$$

(6)

Note that by expanding $P(y_k|x_j)$ we obtain

$$H(X|Y) = H(X, Y) - H(Y).$$

The second part of the Shannon-McMillan Theorem states that the – path independent – splitting criterion, $H[\mathbf{X}]$, of the ergodic information source \mathbf{X} , which divides high from low probability paths, is given in terms of the sequence probabilities of equations (1) and (2) as

$$H[\mathbf{X}] = \lim_{n \rightarrow \infty} H(X_n | X_0, X_1, \dots, X_{n-1}) =$$

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

(7)

The AEPT is one of the most unexpected and profound results of 20th Century applied mathematics.

Ash (1990) describes the uncertainty of an ergodic information source as follows;

“...[W]e may regard a portion of text in a particular language as being produced by an information source. the probabilities $P[X_n = \alpha_n | X_0 = \alpha_0, \dots, X_{n-1} = \alpha_{n-1}]$ may be estimated from the available data about the language. A large uncertainty means, by the AEPT, 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.”

Languages can affect each other, or, equivalently, systems can translate from one language to another, usually with error. The Rate Distortion Theorem, which is one generalization of the SMT, describes how this can take place. As IR Cohen (2001) has put it, in the context of the cognitive immune system,

“An immune response is like a key to a particular lock; each immune response amounts to a functional image of the stimulus that elicited the response. Just as a key encodes a functional image of its lock, an effective [immune] response encodes a functional image of its stimulus; the stimulus and the response fit each other. The immune system, for example, has to deploy different types of inflammation to heal a broken bone, repair an infarction, effect neuroprotection, cure hepatitis, or contain tuberculosis. Each aspect of the response is a functional representation of the challenge.

Self-organization allows a system to adapt, to update itself in the image of the world it must respond to... The immune system, like the brain... aim[s] at representing a part of the world.”

These considerations suggest that the degree of possible back-translation between the world and its image within a cognitive system represents the profound and systematic coupling between a biological system and its environment, a coupling which may particularly express the way in which the system has ‘learned’ the environment. We attempt a formal treatment, from which it will appear that both cognition and response to systematic patterns of selection pressure are – almost inevitably – highly punctuated by ‘learning plateaus’ in which the two processes can become inextricably intertwined.

Suppose we have a ergodic information source \mathbf{Y} , a generalized language having grammar and syntax, with a source uncertainty $H[\mathbf{Y}]$ that ‘perturbs’ a system of interest. A chain of length n , a path of perturbations, has the form

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

Suppose that chain elicits a corresponding chain of responses from the system of interest, producing another path $b^n = (b_1, \dots, b_n)$, which has some ‘natural’ translation into the language of the perturbations, although not, generally, in a one-to-one manner. The image is of a continuous analog audio signal which has been ‘digitized’ into a discrete set of voltage values. Thus, there may well be several different y^n corresponding to a given ‘digitized’ b^n . Consequently, in translating back from the b-language into the y-language, there will generally be information loss.

Suppose, however, that with each path b^n we specify an inverse code which identifies exactly one path \hat{y}^n . We assume further there is a measure

of distortion which compares the real path y^n with the inferred inverse \hat{y}^n . Below we follow the nomenclature of Cover and Thomas (1992).

The *Hamming distortion* is defined as

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

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

(8)

For continuous variates the *Squared error distortion* is defined as

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

(9)

Possibilities abound.

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

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

(10)

We suppose that with each path y^n and b^n -path translation into the y -language, denoted \hat{y}^n , there are associated individual, joint and conditional probability distributions $p(y^n), p(\hat{y}^n), p(y^n, \hat{y}^n)$ and $p(y^n|\hat{y}^n)$.

The *average distortion* is defined as

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

(11)

It is possible, using the distributions given above, to define the information transmitted from the incoming Y to the outgoing \hat{Y} process in the usual manner, using the appropriate Shannon uncertainties:

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

(12)

If there is no uncertainty in Y given \hat{Y} , then no information is lost. In general, this will not be true.

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

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

(13)

where 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.

The Rate Distortion Theorem states that $R(D)$, as we have defined it, is the maximum achievable rate of information transmission which does not exceed distortion D . Note that the result is *independent of the exact form of the distortion measure* $d(y, \hat{y})$.

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

$$p(\hat{y}^n) \geq p(\hat{y}^n | y^n) \exp[-nI(Y, \hat{Y})]. \quad (14)$$

Thus, roughly speaking, $I(Y, \hat{Y})$ embodies the splitting criterion between high and low probability pairs of paths. These pairs are, again, the input ‘training’ paths and corresponding output path.

Note that, in the absence of a distortion measure, this result remains true for two interacting information sources, the principal content of the *joint asymptotic equipartition theorem*, [Cover and Thomas, 1992, Theorem 8.6.1].

Thus the imposition of a distortion measure results in a limitation in the number of possible jointly typical sequences to those satisfying the distortion criterion.

For the theory we will explore later – of pairwise interacting information sources – $I(Y, \hat{Y})$ (or $I(Y_1, Y_2)$ without the distortion restriction), can play the role of H in the critical development of the next section.

The RDT is a generalization of the Shannon-McMillan Theorem which examines the interaction of two information sources under the constraint of a fixed average distortion. For our development we will require one more iteration, studying the interaction of three ‘languages’ under particular conditions, and require a similar generalization of the SMT in terms of the splitting criterion for triplets as opposed to single or double stranded patterns. The tool for this is at the core of what is termed *network information*

theory [Cover and Thomas, 1992, Theorem 14.2.3]. Suppose we have (piecewise memoryless) ergodic information sources Y_1, Y_2 and Y_3 . We assume Y_3 constitutes a critical embedding context for Y_1 and Y_2 so that, given three sequences of length n , the probability of a particular triplet of sequences is determined by *conditional probabilities with respect to Y_3* :

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) =$$

$$\prod_{i=1}^n p(y_{1i}|y_{3i})p(y_{2i}|y_{3i})p(y_{3i}).$$

(15)

That is, Y_1 and Y_2 are, in some measure, driven by their interaction with Y_3

Then, in analogy with the previous two cases, triplets of sequences can be divided by a splitting criterion into two sets, having high and low probabilities respectively. For large n the number of triplet sequences in the high probability set will be determined by the relation [Cover and Thomas, 1992, p. 387]

$$N(n) \propto \exp[nI(Y_1; Y_2|Y_3)],$$

(16)

where splitting criterion is given by

$$I(Y_1; Y_2|Y_3) \equiv$$

$$H(Y_3) + H(Y_1|Y_3) + H(Y_2|Y_3) - H(Y_1, Y_2, Y_3)$$

Below we examine phase transitions in the splitting criteria H , which we will generalize to both $I(Y_1, Y_2)$ and $I(Y_1, Y_2|Y_3)$. The former will produce punctuated cognitive and non-cognitive learning plateaus, while the latter characterizes the interaction between selection pressure and sociocultural cognition, leading to a model of the effects of social structure on ‘eigenmodes’ of pathological response to perturbation.

Phase transition and coevolutionary condensation

The essential homology relating information theory to statistical mechanics and nonlinear dynamics is twofold (Wallace and Wallace, 1998, 1999):

(1) A ‘linguistic’ equipartition of probable paths consistent with the Shannon-McMillan and Rate Distortion Theorems serves as the formal connection with nonlinear mechanics and fluctuation theory – a matter we will not fully explore here, and

(2) A correspondence between information source uncertainty and statistical mechanical free energy density, rather than entropy. See Wallace and Wallace (1998, 1999) and Rojdestvenski and Cottam, (2000) for a fuller discussion of the formal justification for this assumption, described by Bennett (1988) as follows:

“...[T]he value of a message is the amount of mathematical or other work plausibly done by the originator, which the receiver is saved from having to repeat.”

This is a central insight to which we will return at some length below. In sum, we will generally impose invariance under renormalization symmetry on the ‘splitting criterion’ between high and low probability states from the Large Deviations Program of applied probability (e.g. Dembo and Zeitouni, 1998). Free energy density (which can be reexpressed as an ‘entropy’ in microscopic systems) is the splitting criterion for statistical mechanics, and information source uncertainty is the criterion for ‘language’ systems. Imposition of renormalization on free energy density gives phase transition in a physical system. For information systems it gives interactive condensation.

This analogy is indeed a mathematical homology:

The definition of the free energy density for a parametrized physical system is

$$F(K_1, \dots, K_m) = \lim_{V \rightarrow \infty} \frac{\log[Z(K_1, \dots, K_m)]}{V} \quad (17)$$

where the K_j are parameters, V is the system volume and Z is the ‘partition function’ defined from the energy function, the Hamiltonian, of the system.

For an ergodic information source the equivalent relation associates source uncertainty with the number of ‘meaningful’ sequences $N(n)$ of length n , in the limit

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

We will *parametrize* the information source to obtain the crucial expression on which our version of information dynamics will be constructed;

$$H[K_1, \dots, K_m, \mathbf{X}] = \lim_{n \rightarrow \infty} \frac{\log[N(K_1, \dots, K_m)]}{n}. \quad (18)$$

The essential point is that while information systems do not have ‘Hamiltonians’ allowing definition of a ‘partition function’ and a free energy density, they may have a source uncertainty obeying a limiting relation like that of free energy density. Importing ‘renormalization’ symmetry gives phase transitions at critical points (or surfaces), and importing a Legendre transform in a ‘natural’ manner gives dynamic behavior far from criticality, a matter leading to our analysis of the effects of sudden perturbation on coalesced cognitive systems.

As neural networks demonstrate so well, it is possible to build larger pattern recognition systems from assemblages of smaller ones. We abstract

this process in terms of a generalized linked array of subcomponents which ‘talk’ to each other in two different ways. These we take to be ‘strong’ and ‘weak’ ties between subassemblies. ‘Strong’ ties are, following arguments from sociology (Granovetter, 1973), those which permit disjoint partition of the system into equivalence classes. Thus the strong ties are associated with some reflexive, symmetric, and transitive relation between components. ‘Weak’ ties do not permit such disjoint partition. In a physical system these might be viewed, respectively, as ‘local’ and ‘mean field’ coupling.

We fix the magnitude of strong ties, but vary the index of weak ties between components, which we call P , taking $K = 1/P$.

We assume the array, sensory activity and ongoing activity depend on three parameters, two explicit and one implicit. The explicit are K as above and an ‘external field strength’ analog J , which gives a ‘direction’ to the system. We may, in the limit, set $J = 0$.

The implicit parameter, which we call r , is an inherent generalized ‘length’ on which the phenomenon, including J and K , are defined. That is, we can write J and K as functions of averages of the parameter r , which may be quite complex, having nothing at all to do with conventional ideas of space, for example degree of niche partitioning in ecosystems.

Rather than specify complicated patterns of individual dependence or interaction for sensory activity, ongoing activity and array components, we follow the direction suggested above and instead work entirely within the domain of the uncertainty of the ergodic information source dual to the large-scale pattern recognition process, which we write as

$$H[K, J, \mathbf{X}]$$

Imposition of invariance of H under a renormalization transform in the implicit parameter r leads to expectation of both a critical point in K , which we call K_C , reflecting a phase transition to or from collective behavior across the entire array, and of power laws for system behavior near K_C . Addition of other parameters to the system, e.g. some Q , results in a ‘critical line’ or surface $K_C(Q)$.

Let $\kappa = (K_C - K)/K_C$ and take χ as the ‘correlation length’ defining the average domain in r -space for which the dual information source is primarily dominated by ‘strong’ ties. We begin by averaging across r -space in terms of ‘clumps’ of length R , defining J_R, K_R as J, K for $R = 1$. Then, following Wilson’s [53] physical analog, we choose the renormalization relations as

$$H[K_R, J_R, \mathbf{X}] = R^{\mathcal{D}} H[K, J, \mathbf{X}]$$

$$\chi(K_R, J_R) = \frac{\chi(K, J)}{R}$$

(19)

where \mathcal{D} is a non-negative real constant, possibly reflecting fractal network structure. The first of these equations states that ‘processing capacity,’ as indexed by the source uncertainty of the system which represents the ‘richness’ of the inherent language, grows as $R^{\mathcal{D}}$, while the second just states that the correlation length simply scales as R .

Other, very subtle, symmetry relations – not necessarily based on elementary physical analogs – may well be possible. For example McCauley, (1993, p.168) describes the counterintuitive renormalization relations needed to understand phase transition in simple ‘chaotic’ systems.

For K near K_C , if $J \rightarrow 0$, a simple series expansion and some clever algebra (e.g. Wilson, 1971, the argument leading from his eqs. 4 and 5 to his eqs. 22 and 23) gives

$$H = H_0 \kappa^{s\mathcal{D}}$$

$$\chi = \chi_0 \kappa^{-s}$$

(20)

where s is a positive constant. Some rearrangement produces, near K_C ,

$$H \propto \frac{1}{\chi^{\mathcal{D}}}$$

(21)

This suggests that the ‘richness’ of the pattern recognition language is inversely related to the domain dominated by disjointly partitioning strong ties near criticality. As the nondisjunctive weak ties coupling declines, the efficiency of the coupled system as an information channel declines precipitously near the transition point: see ACTK for discussion of the relation between channel capacity and information source uncertainty.

Far from the critical point matters are considerably more complicated, apparently driven by appropriate (and usually counterintuitive) generalizations of a physical system’s ‘Onsager relations,’ described below in terms of a ‘thermodynamics’ of information sources.

The essential insight is that *regardless of the particular renormalization symmetries involved, sudden critical point transition is possible in the opposite direction for this model*, that is, from a number of independent, isolated and fragmented pattern recognition systems operating individually and more or less at random, into a single large, interlocked, coherent pattern recognition system, once the parameter K , the inverse strength of weak ties, falls below threshold, or, conversely, once the strength of weak ties parameter $P = 1/K$ becomes large enough.

Thus, increasing weak ties between them can bind several different pattern recognition or other ‘language’ processes into a single, embedding hierarchical metalanguage which contains the different languages as linked sub-dialects.

This heuristic insight can be made exact using a rate distortion argument:

Suppose that two ergodic information sources \mathbf{Y} and \mathbf{B} begin to interact, to ‘talk’ to each other, i.e. to influence each other in some way so that it is possible, for example, to look at the output of \mathbf{B} – strings b – and infer something about the behavior of \mathbf{Y} from it – strings y . We suppose it possible to define a retranslation from the B-language into the Y-language through a deterministic code book, and call $\hat{\mathbf{Y}}$ the translated information source, as mirrored by \mathbf{B} .

Take some distortion measure d comparing paths y to paths \hat{y} , defining $d(y, \hat{y})$. We invoke the Rate Distortion Theorem's mutual information $I(Y, \hat{Y})$, which is a splitting criterion between high and low probability pairs of paths. Impose, now, a parametrization by an inverse coupling strength K , and a renormalization symmetry representing the global structure of the system coupling. This may be much different from the renormalization behavior of the individual components. If $K < K_C$, where K_C is a critical point (or surface), the two information sources will be closely coupled enough to be characterized as condensed.

We will make much of this below; cultural and genetic heritages are generalized languages, as are neural, immune, and sociocultural pattern recognition.

Pattern recognition as language

The task of this section is to express cognitive pattern recognition-and-response in terms of an ergodic information source constrained by the AEPT. This general approach would then apply to the immune system, the CNS and sociocultural networks. Pattern recognition, as we will characterize it here, proceeds by convoluting an incoming 'sensory' signal with an internal 'ongoing activity' and, at some point, triggering an appropriate action based on a decision that the pattern of the sensory input requires a response. For the purposes of this work we do not need to model in any particular detail the manner in which the pattern recognition system is 'trained,' and thus adopt a 'weak' model which may have considerable generality, regardless of the system's particular learning paradigm, which can be more formally described using the RDT.

We will, fulfilling Atlan and Cohen's (1998) criterion of meaning-from-response, define a language's contextual meaning entirely in terms of system output.

The model is as follows: A pattern of sensory input is convoluted with a pattern of internal 'ongoing activity' to create a path

$$x = (a_0, a_1, \dots, a_n, \dots).$$

This is fed into a (highly nonlinear) 'decision oscillator' which generates an output $h(x)$ that is an element of one of two (presumably) disjoint sets B_0 and B_1 .

We take

$$B_0 = b_0, \dots, b_k$$

$$B_1 = b_{k+1}, \dots, b_m.$$

Thus we permit a graded response, supposing that if

$$h(x) \in B_0$$

the pattern is not recognized, and

$$h(x) \in B_1$$

that the pattern is recognized and some action $b_j, k+1 \leq j \leq m$ takes place.

We are interested in paths which trigger pattern recognition exactly once. That is, given a fixed initial state a_0 such that $h(a_0) \in B_0$, we examine all possible subsequent paths x beginning with a_0 and leading exactly once to the event $h(x) \in B_1$. Thus $h(a_0, a_1, \dots, a_j) \in B_0$ for all $j < m$ but $h(a_0, \dots, a_m) \in B_1$.

For each positive integer n , let $N(n)$ be the number of paths of length n which begin with some particular a_0 having $h(a_0) \in B_0$, and lead to the condition $h(x) \in B_1$. We shall call such paths ‘meaningful’ and assume $N(n)$ to be considerably less than the number of all possible paths of length n – pattern recognition is comparatively rare – and in particular assume that the finite limit

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

exists and is independent of the path x . We will – not surprisingly – call such a pattern recognition process ergodic.

We may thus define a ergodic information source \mathbf{X} associated with stochastic variates X_j having joint and conditional probabilities $P(a_0, \dots, a_n)$ and

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

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

We say this ergodic information source is *dual* to the pattern recognition process.

Different ‘languages’ will, of course, be defined by different divisions of the total universe of possible responses into different pairs of sets B_0 and B_1 , or perhaps even by requiring more than one response in B_1 along a path. Like the use of different distortion measures in the RDT, however, it seems obvious that the underlying dynamics will all be qualitatively similar.

Meaningful paths – creating an inherent grammar and syntax – are defined entirely in terms of system response, as Atlan (1983, 1987, 1998) and Atlan and Cohen (1998) propose, quoting Atlan (1987)

“...[T]he perception of a pattern does not result from a two-step process with first perception of a pattern of signals and then processing by application of a rule of representation. Rather, a given pattern in the environment is perceived at the time when signals are received by a kind of resonance between a given structure of the environment – not necessarily obvious to the eyes of an observer – and an internal structure of the cognitive system. It is the latter which defines a possible functional meaning – for the system itself – of the environmental structure.”

Elsewhere (R Wallace, 2000b, 2001) we have termed this process an ‘information resonance.’

Although we do not pursue the matter here, the ‘space’ of the a_j can be partitioned into disjoint equivalence classes according to whether states can be connected by meaningful paths. This is analogous to a partition into domains of attraction for a nonlinear or chaotic system, and imposes a ‘natural’

algebraic structure which can, among other things, enable multitasking (R Wallace, 2000b, 2001).

We can apply this formalism to the stochastic neuron: A series of inputs $y_i^j, i = 1 \dots m$ from m nearby neurons at time j is convoluted with ‘weights’ $w_i^j, i = 1 \dots m$, using an inner product

$$a_j = \mathbf{y}^j \cdot \mathbf{w}^j = \sum_{i=1}^m y_i^j w_i^j$$

(22)

in the context of a ‘transfer function’ $f(\mathbf{y}^j \cdot \mathbf{w}^j)$ such that the probability of the neuron firing and having a discrete output $z^j = 1$ is $P(z^j = 1) = f(\mathbf{y}^j \cdot \mathbf{w}^j)$. Thus the probability that the neuron does not fire at time j is $1 - f(\mathbf{y}^j \cdot \mathbf{w}^j)$.

In the terminology of this section the m values y_i^j constitute ‘sensory activity’ and the m weights w_i^j the ‘ongoing activity’ at time j , with $a_j = \mathbf{y}^j \cdot \mathbf{w}^j$ and $x = a_0, a_1, \dots, a_n, \dots$

A little more work, described below, leads to a fairly standard neural network model in which the network is trained by appropriately varying the \mathbf{w} through least squares or other error minimization feedback. This can be shown to, essentially, replicate rate distortion arguments, as we can use the error definition to define a distortion function $d(y, \hat{y})$ which measures the difference between the training pattern y and the network output \hat{y} as a function of, for example, the inverse number of training cycles, K . As we will discuss in some detail, ‘learning plateau’ behavior follows as a phase transition on the parameter K in the mutual information $I(Y, \hat{Y})$.

Park et al. (2000) treat the stochastic neural network in terms of a space of related probability density functions $[p(\mathbf{x}, \mathbf{y}; \mathbf{w}) | \mathbf{w} \in \mathcal{R}^m]$, where \mathbf{x} is the input, \mathbf{y} the output and \mathbf{w} the parameter vector. The goal of learning is to find an optimum \mathbf{w}^* which maximizes the log likelihood function. They define a loss function of learning as

$$L(\mathbf{x}, \mathbf{y}; \mathbf{w}) \equiv -\log p(\mathbf{x}, \mathbf{y}; \mathbf{w}),$$

and one can take as a learning paradigm the gradient relation

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta_t \partial L(\mathbf{x}, \mathbf{y}; \mathbf{w}) / \partial \mathbf{w},$$

where η_t is a learning rate.

Park et al. (2000) attack this optimization problem by recognizing that the space of $p(\mathbf{x}, \mathbf{y}; \mathbf{w})$ is Riemannian with a metric given by the Fisher information matrix

$$G(\mathbf{w}) = \int \int \partial \log p / \partial \mathbf{w} [\partial \log p / \partial \mathbf{w}]^T p(\mathbf{x}, \mathbf{y}; \mathbf{w}) d\mathbf{y} d\mathbf{x}$$

where T is the transpose operation. A Fisher-efficient on-line estimator is then obtained by using the ‘natural’ gradient algorithm

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta_t G^{-1} \partial L(\mathbf{x}, \mathbf{y}; \mathbf{w}) / \partial \mathbf{w}.$$

Again, through the synergistic family of probability distributions $p(\mathbf{x}, \mathbf{y}; \mathbf{w})$, this can be viewed as a special case – a ‘representation’, to use physics jargon – of the general ‘convolution argument’ given above.

Again, it seems that a rate distortion argument between training language and network response language will nonetheless produce learning plateaus, even in this rather elegant special case.

The foundation of our mathematical modeling exercise is to assume that both the immune system and a sociocultural network’s pattern recognition behavior, like that of other pattern recognition systems, can also be represented by the language arguments given above, and is thus dual to a ergodic information source, a context-defining language in Atlan and Cohen’s sense (1998), having a grammar and syntax such that meaning is explicitly defined in terms of system response.

Sociogeographic or sociocultural networks – social networks embedded place and embodying culture – serve a number of functions, including acting as the local tools for teaching cultural norms and processes to individuals. Thus, for the purposes of this work, a person’s social network – family and friends, workgroup, church, etc. – becomes the immediate agency of cultural dynamics, and provides the foundation for analysis of the ‘brain/culture condensation’.

Sociocultural networks serve also, however, as instruments for collective decision-making, a cognitive phenomenon. Such networks serve as hosts to a political, in the large sense, process by which a community recognizes and

responds to patterns of threat and opportunity. To treat pattern recognition on sociocultural networks we impose a version of the structure and general formalism relating pattern recognition to a dual information source:

We envision problem recognition by a local sociocultural network as follows: A ‘real problem,’ in some sense, becomes convoluted with a community’s internal sociocultural ‘ongoing activity’ to create the path of a ‘perceived problem’ at times $0, 1, \dots$, producing a path of the usual form $x = a_0, a_1, \dots, a_n, \dots$. That serially correlated path is then subject to a decision process across the sociocultural network, designated $h(x)$ which produces output in two sets B_0 and B_1 , as before. The problem is officially recognized and resources committed to if and only if $h(x) \in B_1$, a rare event made even more rare if resources must then be diverted from previously recognized problems.

For the purposes of this work, then, we will view ‘culture’ as, in fact, a sociocultural cognitive process which can entrain individual CNS and immune cognition.

To reiterate, we have, following the earlier discussion of Atlan and Cohen’s work (1998), implicitly assumed that the immune cognition can likewise be expressed as a pattern recognition-and-response language characterized by an information source uncertainty.

Toward a ‘thermodynamics’ of information source uncertainty

We suppose a cognitive system – more generally a linked and broadly coevolutionary condensation of several such systems – is exposed to a sudden perturbation of sensory and other activity – trauma, toxic chemicals, and the like – and wish to estimate the response of that system. This requires some considerable development.

Since the source uncertainty of a (coevolutionarily condensed) behavioral language, $H[\mathbf{X}]$ is the average uncertainty per symbol, we must have $H[\mathbf{X}] \leq C$, where C is the capacity of the underlying geographically-focused social network as an information channel. We suppose that capacity is a function of system-wide average variables, K, J, Q which represent the *ensemble* indices – associated with the entire individual-and-group.

Thus we may write

$$C = C(K, J, Q)$$

$$H = H[K, J, Q, \mathbf{X}].$$

(23)

We assume that as K, J, Q increase, that $C(K, J, Q)$ and H decrease monotonically.

The essential trick – which is highly counterintuitive – is to recognize that, in its definition,

$$H[K, J, Q, \mathbf{X}] = \lim_{n \rightarrow \infty} \frac{\log[N(K, J, Q, n)]}{n}$$

(24)

is the analog of the *free energy density* of statistical mechanics: Take a physical system of volume V which can be characterized by an inverse temperature parameter $K = 1/T$. The ‘partition function’ for the system is (K. Wilson, 1971)

$$Z(K) = \sum_j \exp[-K E_j]$$

(25)

where E_j represents the energy of the individual state j . The probability of the state j is then

$$P_j = \frac{\exp[-K E_j]}{Z(K)}.$$

Then the ‘free energy density’ of the *entire* system is defined as

$$F(K) = \lim_{V \rightarrow \infty} \frac{\log[Z(K)]}{V}. \quad (26)$$

The relation between information and thermodynamic free energy has long been recognized. As Elitzur (1996, p. 179) puts it

“Recall... the lesson of Maxwell’s Demon: Information, when applied under appropriate circumstances, can save work.”

Again, Bennett (1988, p. 230) finds that the value of a message is the amount of mathematical or other work plausibly done by its originator, which the receiver is saved from having to repeat.

Equation (24) expresses one of the central theorems of information theory in a form similar to equation (26). It is this similarity which suggests that, for some systems under proper circumstances, there may be a ‘duality’ which maps Shannon’s ergodic source uncertainty onto thermodynamic free energy density, and perhaps vice versa.

The method we propose here, based entirely on equation (24), the Shannon-McMillan Theorem for ergodic information sources, may prove to be more generally applicable to information systems, not requiring the explicit identification a ‘duality’ in each and every case.

The formal analogy – the thermodynamic duality – between free energy density for a physical system and ergodic source uncertainty, based on equations (24) and (26), suggests that we impose a *thermodynamics* on source uncertainty.

By a thermodynamics we mean, in the sense of Feigenbaum (1988, p. 530),

“...[T]he deduction of variables and their relations to one another that in some well-defined sense are averages of exponential quantities defined microscopically on a set.”

In our context the relation between the number of meaningful sequences on length n , $N(n)$, for (fixed) large n and the source uncertainty, i.e.

$$N(n) \approx \exp(nH[\mathbf{X}])$$

for large n provides the exponential dependence exactly analogous to performing statistical mechanics. We are, to reiterate, going to express $H[\mathbf{X}]$ in terms of a number of parameters which characterize the underlying community which carries the behavioral language.

We suppose that H , representing the information source of our coevolutionary condensation of CNS, immune and sociocultural cognition, is allowed to depend on a number of observable parameters, which we will not fully specify here.

If source uncertainty H is the analog to free energy density in a physical system, K is the analog to inverse temperature, the next ‘natural’ step is to apply a *Legendre transformation* to H so as to define a generalized ‘entropy,’ S , and other (very) rough analogs to classical thermodynamic entities, depending on the parameters.

Courant and Hilbert (1989, p.32) characterize the Legendre transformation as defining a surface as the envelope of its tangent planes, rather than as the set of points satisfying a particular equation.

Their development shows the Legendre transformation of a well-behaved function $f(Z_1, Z_2, \dots, Z_w)$, denoted g , is

$$g = f - \sum_{i=1}^w Z_i \partial f / \partial Z_i \equiv f - \sum_{i=1}^w Z_i V_i. \quad (27)$$

with, clearly, $V_i \equiv \partial f / \partial Z_i$.

This expression is assumed to be invertible, hence the ‘duality:’

$$f = g - \sum_{i=1}^w V_i \partial g / \partial V_i.$$

Transformation from the ‘Lagrangian’ to the ‘Hamiltonian’ formulation of classical mechanics (Landau and Lifshitz, 1959) is via a Legendre transformation.

The generalization when f is not well-behaved is via a variational principle (Beck and Schlogl, 1993) rather than a tangent plane argument. Then

$$g(V) = \min_Z [f(Z) - VZ]$$

$$f(Z) = \min_V [g(V) - VZ].$$

(28)

In the first expression the variation is taken with respect to Z , in the second with respect to V .

For a badly behaved function it is usually possible to fix up a reasonable invertible structure since the singularities of f or g will usually belong to ‘a set of measure zero,’ for example a finite number of points on a line or lines on a surface where we may designate inverse values by fiat.

We first consider a very simple system in which the ergodic source uncertainty H depends *only* on the inverse strength of weak ties K , giving an analog to the ‘canonical ensemble’ of statistical mechanics which depends only on temperature. We define S , an entropy-analog which we term the ‘disorder,’ as the Legendre transform of the Shannon uncertainty $H[K, \mathbf{X}]$:

$$S = H - KdH/dK \equiv H - KU$$

(29)

where we take $U = dH/dK$ as an analog to the ‘internal energy’ of a system. Note that S and H have the same physical dimensionality.

Since

$$dS/dK = dH/dK - U - KdU/dK = -KdU/dK$$

we have

$$dS/dU = -K$$

and

$$dU \propto PdS.$$

This is the analog to the classic thermodynamic relation $dQ = TdS$ for physical systems. Thus what we have defined here as the disorder S is indeed a generalized entropy.

Note that since $dS/dU = -K$ we have

$$S = H - KU = H + U dS/dU$$

or

$$H = S - U dS/dU$$

which explicitly shows the dual relation between H and S .

Again let $N(n)$ represent the number of meaningful sequences of length n emitted by the source \mathbf{X} . Since

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

for large n , we have

$$U = dH/dK = \lim_{n \rightarrow \infty} \frac{1}{nN} dN/dK.$$

(30)

For fixed (large) n , U is thus the *proportionate* rate of change in number of meaningful sequences of length n with change in K . This is something like the rate of change of mass per unit mass for a person losing weight: A small value will not be much noticed, while a large one may represent a rigorous starvation causing considerable distress.

Some rearrangement gives

$$I \equiv (S - H) = -KdH/dK = UdS/dU \quad (31)$$

We define $I = S - H$ as the *instability* of the system.

If $-dH/dK$ is approximately constant – something like a heat capacity in a physical system – then we have the approximate linear relation

$$I \approx b_K K$$

with

$$b_K \equiv -\partial H/\partial K.$$

We generalize this as follows:

Allow H to depend on a number of parameters, for example average probability of weak ties, the inverse level of community resources, and or other factors which we call $Z_i, i = 1 \dots$. Then, taking $H = H[Z_i, \mathbf{X}]$, we obtain the equations of state

$$S = H - \sum_{j=1}^s Z_j \partial H / \partial Z_j = H - \sum_{j=1}^s Z_j V_j. \quad (32)$$

$$V_i \equiv \partial H / \partial Z_i$$

and the *instability relation*

$$\begin{aligned}
I = S - H &= - \sum_{j=1}^s Z_j \partial H / \partial Z_j = \sum_{j=1}^s V_j \partial S / \partial V_j, \\
&= - \sum_{j=1}^s V_j Z_j = -\mathbf{Z} \cdot \nabla_{|\mathbf{Z}} H,
\end{aligned}
\tag{33}$$

taking $\mathbf{Z} = (Z_1, Z_2, \dots, Z_s)$.

I represents the degree of disorder above and beyond that which is inherent to the ergodic information source itself.

We suggest that risk behaviors, patterns of anxiety, and of depression and demoralization, or other indicators of distress, are primary environmental indices of community instability.

Instability, as we have defined it, is driven by the declining capacity of behavioral language to convey messages across a community's sociogeographic networks.

We suppose that the capacity, C , of the underlying communication channel defined by the sociogeographic networks linked across the community declines with increasing K , so that $C = C(K)$ is monotonically decreasing in K . An ergodic information source can be transmitted without error by a channel only if $H[K, \mathbf{X}] < C(K)$ – again see ACTK – so that declining C will inevitably result in rising I .

I is, according to this development, driven by parameters characterizing public policy and underlying economic and social constraints – the Z_j and V_j .

For a social system, equation 33 is interpreted as stating that the rate of intentional violent deaths, depression or other indices of distress, is proportional to a community's experience of instability.

It may be possible to generalize the development to include temporal effects if we suppose that H depends on $d\mathbf{Z}/dt \equiv \dot{\mathbf{Z}}$ as well as on \mathbf{Z} . Note that terms of the form $\partial H/\partial t$ would violate ergodicity. Then we would take

$$I = -(\mathbf{Z}, \dot{\mathbf{Z}}) \cdot \nabla|_{(\mathbf{Z}, \dot{\mathbf{Z}})} H \equiv$$

$$\sum_j -Z_j \partial H / \partial Z_j - \dot{Z}_j \partial H / \partial \dot{Z}_j.$$

(34)

This suggests that both parameter gradients and their rates of change can be globally destabilizing.

In linear approximation, assuming $-\partial H/\partial Z_i = -V_i = b_{Z_i} \approx \text{constant}$, equation (22) can be rewritten as

$$I \approx b_K K + b_J Q + b_M M.$$

The use of environmental index variates for critical system parameters will generally result in a nonzero intercept, giving the final equation

$$I \approx b_K K + b_J Q + b_M M + b_0.$$

(35)

Note that the intercept b_0 may, in fact, be quite complex, perhaps incorporating other parameters not explicitly included in the model. But it may include as well an ‘error term’ representing stochastic fluctuations not entirely damped by large population effects, or even some ‘nonlinear’ structure when the b_{Z_i} are not quite constant.

Most importantly for our analysis here, if the ‘potentials’ $V_i = \partial H / \partial Z_i$ cannot be approximated as constants, then simple linear regression will fail entirely, and equations (32) and (35) will represent an appropriate generic model – possibly with ‘error terms’ – however *the system will be both nonlinear and nonmonotonic*, hence representing signal transduction in physiological systems.

In sum, we claim the instability relation derived from a fairly simple quasi-thermodynamic argument applied to an ergodic information source parametrized by various significant indices, (as well as, perhaps, their rates of change), explains the high degree to which simple regression models based on those parameters account for observed patterns of physiological, psychological, psychosocial, or immune response to perturbation. We further claim that extension of this model into nonlinear and nonmonotonic realms is quite direct, a matter we will not further treat here.

Ecological resilience of cognitive condensations

In reality, matters are significantly more complex than we have described so far: physiological and psychological responses are indeed produced by sudden deterioration of an individual’s underlying circumstances, but may, in turn, affect these as well, basically through nonlinear, feedback loops.

Thus the inherently nonlinear model for response as produced by increasing stimulation, $I = -\sum_j V_j Z_j$ is replaced by an even more nonlinear structure:

$$I(t) = -\sum_j V_j(I(t)) Z_j(I(t)).$$

(36)

In a first iteration using linear approximation, we can replace this equation with a series for which *each* of s variates – ‘independent’ as well as ‘dependent’ – is expressed in terms of s linear regressions on all the others:

$$x_i(t) = \sum_{j \neq i}^s b_{i,j} x_j(t) + b_{i,0} + \epsilon(t, x_1(t) \dots x_s(t)).$$

(37)

Here the $x_j, j = 1 \dots s$ are both ‘independent’ and ‘dependent’ variates involved in the feedback, $b_{i,0}$ is the intercept constant, and the ϵ terms are ‘error’ terms which may not be small, in this approximation.

In matrix notation this set of equations becomes

$$X(t) = \mathbf{B}X(t) + U(t)$$

(38)

where $X(t)$ is a s -dimensional vector, \mathbf{B} is an $s \times s$ matrix of regression coefficients having a zero diagonal and U is an s -dimensional vector containing the constant and ‘error’ terms. *‘Error’ terms are taken as including public and economic policy changes which are external to the internal feedback loops.*

We suggest that, on the timescale of applied perturbations and of initial responses, the \mathbf{B} -matrix remains relatively constant. Following the analysis of Ives (1995) this structure has a number of interesting properties which permit estimation of the effects of external perturbation on the likelihood of triggering phase change in marginalized communities.

We begin by rewriting the matrix equation as

$$[\mathbf{I} - \mathbf{B}]X(t) = U(t)$$

(39)

where \mathbf{I} is the $s \times s$ identity matrix and, to reiterate, \mathbf{B} has a zero diagonal.

We reexpress matter *in terms of the eigenstructure of \mathbf{B}* .

Let \mathbf{Q} be the matrix of eigenvectors which diagonalizes \mathbf{B} . Take $\mathbf{Q}Y(t) = X(t)$ and $\mathbf{Q}W(t) = U(t)$. Let \mathbf{J} be the diagonal matrix of eigenvalues of \mathbf{B} so that $\mathbf{B} = \mathbf{Q}\mathbf{J}\mathbf{Q}^{-1}$. In R. Wallace, Y Huang, P Gould and D Wallace (1997) we show the eigenvalues of \mathbf{B} are all real. Then, for the eigenvectors Y_k of \mathbf{B} , corresponding to the eigenvalues λ_k ,

$$Y_k(t) = \mathbf{J}Y_k(t) + W_k(t). \quad (40)$$

Using a term-by-term shorthand for the components of Y_k , this becomes

$$y_k(t) = \lambda_k y_k(t) + w_k(t). \quad (41)$$

Define the mean $E[f]$ of a time-dependent function $f(t)$ over the time interval $0 \rightarrow \Delta T$ as

$$E[f] = \frac{1}{\Delta T} \int_0^{\Delta T} f(t) dt. \quad (42)$$

We assume an appropriately ‘rational’ structure as $\Delta T \rightarrow \infty$, probably some kind of ‘ergodic’ hypothesis.

Note that this form of expectation does not include the effects of differing timescales or lag times. Under such circumstances, increasing ΔT will begin to ‘pick up’ new effects, in a path-dependent manner: The mathematics of equation (42) suddenly becomes extremely complicated.

The variance $V[f]$ over the same interval is defined as $E[f - E[f]]^2$.

Again taking matters term-by-term, we take the variance of the $y_k(t)$ from the development above to obtain

$$V[(1 - \lambda_k)y_k(t)] = V[w_k(t)]$$

so that

$$V[y_k] = \frac{V[w_k]}{(1 - \lambda_k)^2}$$

or

$$\sigma(y_k) = \frac{\sigma(w_k)}{|1 - \lambda_k|}.$$

(43)

The y_k are the components of the eigentransformed pathology, income, crowding, community size etc. variates x_i and the w_k are the similarly transformed variates of the driving externalities u_i .

The eigenvector components y_i are characteristic but non-orthogonal combinations of the original variates x_i whose standard deviation is that of the (transformed) externalities w_i , but *synergistically amplified* by the term $1/|1 - \lambda_i|$, a function of the eigenvalues of the matrix of regression coefficients \mathbf{B} . If $\lambda_i \rightarrow 1$ then any change in driving factors external to the community, like public policy or economic practice, will cause great instability within the affected community.

A simple example suffices. Suppose we have the two empirical regression equations

$$x_1 = b_1x_2 + b_{01}$$

and

$$x_2 = b_2 x_1 + b_{02}$$

where x_1 is, for example, an index of violent crime and x_2 is an index of the ‘strength of weak ties.’ These equations say that weak ties affect violence and violence affects weak ties. Then, after normalizing x_1 and x_2 to zero mean and unit variance, the \mathbf{B} -matrix becomes

$$\mathbf{B} = \begin{pmatrix} 0 & R \\ R & 0 \end{pmatrix}$$

where $R = b_1 = b_2$ is simply the correlation between x_1 and x_2 .

This matrix has eigenvalues $\pm|R|$ and eigenvectors $[\pm 1/\sqrt{2}, 1/\sqrt{2}]$. As the variates become more closely correlated, $R \rightarrow 1$ and the ratio of the standard deviation of the eigenvector with positive components and that of the external perturbations, $1/[1 - R]$, diverges.

There is a kind of physical picture for this model. Imagine a violin strung with limp, wet cotton twine. Then $R \approx 0$ and no amount of bowing – an external perturbation – will excite any sound from the instrument. Now restring that violin with finely tuned catgut (to be somewhat old fashioned). Then $R \approx 1$ and external perturbation – bowing – will now excite loud and brisant eigenresonances.

Ives (1995) defines an ecosystem for which $\lambda \approx 0$, so that $1/|1 - \lambda| \approx 1$, as *resilient* in the sense that applied perturbations will have no more effect than their own standard deviation.

As a brief digression, in consonance with our previous arguments on phase transition, we might suppose that there are domains of quasi-stability marked by different possible relations of the form

$$X(t) = \mathbf{B}_k X(t) + U_k(t)$$

(44)

and that a sufficiently perturbed system will suddenly change between them. Suppose there to be a threshold value for an eigenresonance Y , which

we call Y_C , such that if it is exceeded, then the system ‘falls off the table’ into a more pathological one. If Y has a normal distribution, then the probability that the system exceeds the critical limit Y_C , which we write Q_C , will be given by

$$\begin{aligned}
 Q_C &= \frac{1}{\sqrt{2\pi}\sigma(Y)} \int_{Y_C}^{\infty} \exp[-Y/2\sigma^2(Y)] dY \\
 &= \frac{1}{2} [1 - \text{ERF}(\frac{Y_C}{\sqrt{2}\sigma(Y)})]
 \end{aligned}$$

(45)

where ERF is the error function.

For large Y_C and $\lambda \rightarrow 1$ this becomes

$$Q_C \approx \frac{1}{Y_C|1-\lambda|} \sqrt{V[w]/2\pi} \propto \frac{1}{Y_C|1-\lambda|} \sigma(w).$$

(46)

A full nonlinear address of these matters is clearly necessary, but will be difficult. Such treatments are becoming routine in population and community ecology (e.g. Higgins et al., 1997), but are still rare in epidemiological or physiological studies. The essential point is that a nonlinear deterministic ‘backbone’ serves to amplify external perturbations (Higgins et al., 1997):

“...[R]elatively small environmental perturbations can markedly alter the dynamics of deterministic biological mechanisms, producing very large fluctuations...”

This, then, is our version of Holling’s (1973, 1992) mechanism for the loss of ecological resilience by which the small can affect the large.

In the next section we begin to examine one form of this mechanism which generalizes our linearized treatment.

‘Nonlinearizing’ the model by connecting domains of quasi-stability

Equation (44) suggests, in consonance with ecosystem theory (e.g. Holling, 1973, 1992), that different ‘domains of quasi-stability’ obeying equation (43) may emerge in a punctuated manner as a function of the magnitude of external perturbation. Indeed, D Wallace and R Wallace (2000) invoked precisely this mechanism to explain observation of different amplification factors associated with a similar excited eigenvector for a comparative study of large scale pathology in Upper Manhattan and the Bronx. These are two communities with greatly differing histories of perturbation in the 1970’s. While Upper Manhattan was greatly damaged by politically-induced outbreaks of contagious urban decay, the ‘South Bronx’ was literally devastated, losing between 50 and 80 % of its occupied housing in the same period, and became an urban proverb (e.g. D Wallace and R Wallace, 1998). The amplification factors for a mix of pathologies ranging from homicide to low birth weight at the scale of about 30 Health Areas in each Borough differed by a factor of five, with the Bronx showing the lower value, while the respective excited eigenvectors were virtually identical, a highly counterintuitive result. We attributed this pattern to the evolution of a qualitatively different social network structure in the Bronx, driven by the selection pressure of massive burnout and population displacement.

We are led to generalize equation (43) as

$$\sigma(y) = \frac{\sigma(w)}{|1 - \lambda(|w|_M)|} \quad (47)$$

where $|w|_M$ is the maximum absolute value of the ‘magnitude’ of perturbation in an appropriate sense over the time average integral of equation(42).

That is, domains of quasi-stability represented by equation (44) are, in this model, seen as smoothly connected, depending on both the ecosystem scale and the scale of perturbation.

At the individual level of scale we might expect the multiple correlation coefficient $\lambda(|w|_M)$ to take the usual S-shaped dose-response form, for example

$$\lambda(|w|_M) = \frac{1}{1 + \exp(-a|w|_M + b)}, \quad (48)$$

where a and b are appropriate scaling constants. When plugged into equation (47) this produces a sharply rising J-shaped curve near b/a for the amplification factor.

At the group level of scale, the individual survivors of the process leading to equation (48) might respond collectively to the massive perturbation with a signal transduction pattern: beyond a certain point of collective stress, a signal ceases to be collectively ‘meaningful,’ and the social group undergoes a kind of phase-transition-like withdrawal, so that $\lambda(|w|_M)$ takes on a broad ‘inverted-U-shaped’ peak. This kind of effect is often modeled in the signal transduction literature as a stochastic resonance, having the functional form

$$\lambda(|w|_M) = a \frac{1/x^2}{1 + b \exp[1/(2|w|_M)]} \quad (49)$$

where a and b again are appropriate scaling constants. Plugging equation (49) into equation (47) results in an acutely sharpened peak, with a considerable reduction in the full width at half height (FWHH). Signal-sharpening of

this nature is, for neural networks, often seen as a critical step in processes of perception. We are led to suggest that roughly analogous mechanisms may characterize the organizational and social levels of scale.

The behavior of equation (49), if seen at the individual level, might characterize what Lifton has called ‘psychic numbing’, in which a profoundly battered individual permanently loses the ability to respond emotionally, a condition which Lifton views as ‘not compatible with a definition of survival’ (Lifton and Markusen, 1990).

‘Higher order’ pattern recognition

The next step in the argument is to examine the way in which ‘structured’ social factors can affect individual or collective response to perturbation. This will require some development.

Suppose the pattern recognition language previously considered is, in some sense, iterated, so that its output sequence of the decision oscillator $h(x) \in B_0, B_1$ serves as the input to a higher order decision oscillator, h^b having, for example, a simple binary output: $h^b = Th1$ or $h^b = Th2$. Thus the sequence of $x^b \equiv b_0, b_1, \dots, b_k, \dots$ becomes the object of interest, leading to the binary output. In all likelihood, however, the output of h will be condensed by integration or renormalization, as discussed before. Thus the x^b will first be mapped or coded onto a simplified alphabet. In any event, the *original* paths in $x = a_0, a_1, \dots, a_k, \dots$ can be placed in equivalence classes, according to those which, ultimately, after iteration, lead to outputs which might be labeled ‘Th1’ or to ‘Th2’.

The process can, of course, be iterated to a higher level, so that, if the switch is indeed thrown, the Th2 paths will become input for another decision oscillator.

Obviously, at this point mathematical possibilities begin to grow exponentially, including the necessity of examining ‘meaningful’ binary sequences of internal and external signals, and it seems likely that recourse to empirical study will be required to prune the thicket.

Generalized cognitive condensations

We suppose a cognitive system – more generally a linked, hierarchically structured, and broadly coevolutionary condensation of several such systems – is exposed to a structured pattern of sensory activity – the training pattern – to which it must learn an appropriate matching response. From that

response we can infer, in a direct manner, something of the form of the excitatory sensory activity. We suppose the training pattern to have sufficient grammar and syntax so as to itself constitute a ergodic information source Y . The output of the cognitive system, B , is deterministically backtranslated into the ‘language’ of Y , and we call that translation \hat{Y} . The rate distortion behavior relating Y and \hat{Y} , is, according to the RDT, determined by the mutual information $I(Y, \hat{Y})$. We take the index of coupling between the sensory input and the cognitive system to be the number of training cycles – an exposure measure – having an inverse K , and write

$$I(Y, \hat{Y}) = I[K]$$

(50)

$I[K]$ defines the splitting criterion between high and low probability pairs of training and response paths for a specified average distortion D , and is analogous to the parametrized information source uncertainty upon which we imposed renormalization symmetry to obtain phase transition.

We thus interpret the sudden changes in the measured average distortion $D \equiv \sum p(y) d(y, \hat{y})$ which determines ‘mean square error’ between training pattern and output pattern, e.g. the *ending* of a learning plateau, as representing onset of a phase transition in $I[K]$ at some critical K_C , consonant with our earlier developments.

Note that $I[K]$ constitutes an interaction between the cognitive system and the impinging sensory activity, so that its properties may be quite different from those of the cognitive condensation itself.

From this viewpoint learning plateaus are an inherently ‘natural’ phase transition behavior of pattern recognition systems. While one may perhaps, in the sense of Park et al. (2000), find more efficient gradient learning algorithms, our development suggests learning plateaus will be both ubiquitous and highly characteristic of a cognitive system. Indeed, it seems likely that proper analysis of learning plateaus will give deep insight into the structures underlying that system.

This is not a new thought: Mathematical learning models of varying complexity have been under constant development since the late 1940’s (e.

g. Luce, 1997), and learning plateau behavior has always been a focus of such studies.

The particular contribution of our perspective to this debate is that the distinct coevolutionary condensation of immune, CNS, and local sociocultural network cognition which distinguishes human biology must respond as a composite in a coherent, unitary and coupled manner to sensory input. Thus the ‘learning curves’ of the immune system, the CNS and the embedding sociocultural network are inevitably coupled and must reflect each other. Such reflection or interaction will, of necessity, be complicated.

Our analysis, however, has a particular implication. Learned cultural behavior – sociocultural cognition – is, from our viewpoint, a nested hierarchy of phase transition learning plateaus which carries within it the history of an individual’s embedding socioculture. Through the cognitive condensation which distinguishes human biology, that punctuated history becomes part of individual cognitive and immune function. Simply removing ‘constraints’ which have deformed individual and collective past is unlikely to have the desired impact: one never, really, forgets how to ride a bicycle, and a social group, in the absence of affirmative redress, will not ‘forget’ the punctuated adaptations ‘learned’ from experiences of slavery or holocaust. Indeed, at the individual level, sufficiently traumatic events may become encoded within the CNS and immune systems to express themselves as Post Traumatic Stress Disorder.

Noncognitive condensation in response to selection pressure

As discussed above, sociocultural networks serve multiple functions and are not only decision making cognitive structures, but are cultural repositories which embody the history of a community. Sociocultural networks, like human biology in the large, and the immune system in the small, have a duality in that they make decisions based on recognizing patterns of opportunity and threat by comparison with an internalized picture of the world, and they respond to selection pressure in the sense that cultural patterns which cannot adapt to external selection pressures simply do not survive. This is not learning in the traditional sense of neural networks. Thus the immune system has both ‘innate’ genetically programmed and ‘learned’ components, and human biology in the large is a convolution of genetic and cultural systems of information transmission.

We suggest that sociocultural networks – the instrumentalities of culture – likewise contain both cognitive and selective systems of information transmission which are closely intertwined to create a composite whole.

We now examine processes of ‘punctuated evolution’ inherent to evolutionary systems of information transmission.

We suppose a self-reproducing cultural system – more specifically a linked, and in the large sense coevolutionary, condensation of several such systems – is exposed to a structured pattern of selective environmental pressures to which it must adapt if it is to survive. From that adaptive selection – changes in genotype and phenotype analogs – we can infer, in a direct manner, something, but not everything, of the form of the structured system of selection pressures. That is, the culture contains markers of past ‘selection events’.

We suppose the system of selection pressures to have sufficient internal structure – grammar and syntax – so as to itself constitute an ergodic information source Y whose probabilities are fixed on the timescale of analysis. The output of that system, B , is backtranslated into the ‘language’ of Y , and we call that translation \hat{Y} . The rate distortion behavior relating Y and \hat{Y} , is, according to the RDT, determined by the mutual information $I(Y, \hat{Y})$.

We take there to be a measure of the ‘strength’ of the selection pressure, P , which we use as an index of coupling with the culture of interest, having an inverse $K = 1/P$, and write

$$I(Y, \hat{Y}) = I[K]. \quad (51)$$

P might be measured by the rate of attack by predatory colonizers, or the response to extreme environmental perturbation, and so on.

$I[K]$ thus defines the splitting criterion between high and low probability pairs of input and output paths for a specified average distortion D , and is analogous to the parametrized information source uncertainty upon which we imposed renormalization symmetry to obtain phase transition. The result is robust in the absence of a distortion measure through the joint asymptotic equipartition theorem, as discussed above.

We thus interpret the sudden changes in the measured average distortion $D \equiv \sum p(y)d(y, \hat{y})$ which determines ‘mean error’ between pressure and response, i.e. the *ending* of a ‘learning plateau’, as representing onset of a phase transition in $I[K]$ at some critical K_C , consonant with our earlier developments. In the absence of a distortion measure, we may still expect phase transition in $I[K]$, according to the joint AEPT.

Note that $I[K]$ constitutes an interaction between the self-reproducing system of interest and the impinging ecosystem’s selection pressure, so that its properties may be quite different from those of the individual or conjoined subcomponents (R Wallace and RG Wallace, 1998, 1999).

From this viewpoint highly punctuated ‘non-cognitive condensations’ are an inherently ‘natural’ phase transition behavior of evolutionary systems, even in the absence of a distortion measure. Again, while there may exist, in the sense of Park et al. (2000), more efficient convergence algorithms, our development suggests plateaus will be both ubiquitous and highly characteristic of evolutionary process and path. Indeed, it seems likely that proper analysis of non-cognitive evolutionary ‘learning’ plateaus – to the extent they can be observed or reconstructed – will give deep insight into the mechanisms underlying that system.

Convolution between selection pressure and sociocultural cognition

Selection pressure acting on sociocultural networks can be expected to affect their cognitive function, their ability to recognize and respond to relatively immediate patterns of threat and opportunity. In fact, those patterns themselves may in no small part represent factors of that selection pressure, conditionally dependent on it. We assume, then, the linkage of *three* information sources, two of which are conditionally dependent on and may indeed be dominated by, a highly structured embedding system of externally imposed selection pressure which we call Y_3 . Y_2 we will characterize as the pattern recognition-and-response language of the sociocultural network itself.

In IR Cohen’s sense, this involves comparison of sensory information with an internalized picture of the world, and choice of a response from a repertory of possibilities. Y_1 we take to be a more rapidly changing, but nonetheless structured, pattern of immediate threat-and-opportunity which demands appropriate response and resource allocation – the ‘training pattern’. We reiterate that Y_1 is likely to be conditionally dependent on the embedding selection pressure, Y_3 , as is the hierarchically layered history expressed by Y_2 .

According to the triplet version of the SMT which we discussed above, then, for large n , triplets of paths in Y_1, Y_2 and Y_3 may be divided into two sets, a smaller ‘meaningful’ one of high probability – representing those paths consistent with the ‘grammar’ and ‘syntax’ of the interaction between the selection pressure, the cognitive sociocultural process, and the pattern of immediate ‘sensory challenge’ it faces – and a very large set of vanishingly small probability. The splitting criterion is the conditional mutual information:

$$I(Y_1, Y_2 | Y_3).$$

We parametrize this splitting criterion by a variate K representing the inverse of the strength of the coupling between the system of selection pressure and the linked complex of the sociocultural cognitive process and the structured system of day-to-day problems it must address. $I[K]$ will, according to the ‘phase transition’ developments above, be highly punctuated by ‘mixed’ plateau behavior representing the synergistic and inextricably intertwined action of both externally imposed ‘selection pressure’, in the most general sense, and internal sociocultural cognition.