Multiscale Variety in Complex Systems

YANEER BAR-YAM
New England Complex Systems Institute, 24 Mt. Auburn Street, Cambridge, Massachusetts 02138

Received April 15, 2002; revised October 27, 2003; accepted November 25, 2003

The Law of Requisite Variety is a mathematical theorem relating the number of control states of a system to the number of variations in control that is necessary for effective response. The Law of Requisite Variety does not consider the components of a system and how they must act together to respond effectively. Here we consider the additional requirement of scale of response and the effect of coordinated versus uncoordinated response as a key attribute of complex systems. The components of a system perform a task, with a number of such components needed to act in concert to perform subtasks. We apply the resulting generalization—a Multiscale Law of Requisite Variety—to understanding effective function of complex biological and social systems. This allows us to formalize an understanding of the limitations of hierarchical control structures and the inadequacy of central control and planning in the solution of many complex social problems and the functioning of complex social organizations, e.g., the military, healthcare, and education systems. © 2004 Wiley Periodicals, Inc. Complexity 9: 37–45, 2004

Key Words: complexity; multiscale; variety; hierarchy; organization management

If we make the assumption that there is a system subject to an environment that has a number, W, conditions each of which requires a different response for some specific measure of success to be achieved, then it is a tautology that the system must have at least W distinct states in order to respond effectively. Although this is not a sufficient condition for effective response (the states of the system must be valid responses and matched to each of the conditions), it is a necessary condition. More generally, the variety in actions of a system can be used to reduce the variety of outcomes (next time conditions) when the environment changes in time. This is a statement of the Law of Requisite Variety formulated by Ashby [1]. Its utility arises in the context of design evaluation. Even if a system is ideally designed, if it does not have enough actions that it can take, its overall effectiveness is limited. In this article we make the further assumption that the system is composed of a number of components and that these components can be combined to perform specific tasks that might require more than a single component to perform. We make no assumption about the physical nature of these components, allowing statements that can apply to biological or social systems and to many different types of systems in each of these broad categories.

Specifically, we assume that the responding system is composed of a number of subsystems, N, that are variously

Correspondence to: Yaneer Bar-Yam, E-mail: yaneer@necsi.org

coordinated to respond to external contexts. The number of possible actions that the system can take, M, is not more than, m^N, the product of the possible actions of each part, m. We could directly apply the Law of Requisite Variety for the total number of actions, but we further constrain the problem of effective function by assuming that effective actions require some number of components to perform: If the components are people, then a task may require a certain (minimum) number of people acting together; if the components are biological cells or biological molecules, similar considerations apply. We consider the number of entities that must act together as a measure of the "scale" of the task, which is a distinct property from the necessary variety to perform the task.

It is conventional to measure variety, like information, in logarithmic units so that the total variety of a set of independent components $V = \log(M)$ is the sum of the variety of the components, $V = N\nu$, where $V = \log(M)$.

When components act together on a task, we can assume for simplicity that the maximal variety of the coordinated group of components is the variety of one of the components. This is true as long as the actions of one of the components require certain actions of the other components regardless of the specifics of coordination. For example, if two people must lift at the same moment in order to lift something heavy, they are coordinated, but this is also the case if two individuals must push and pull alternately, as in the operation of a cross-cut saw. If the components are only coordinated to some degree, we can consider a generalization of the treatment using coordination probabilities, but there is a difference between complete coordination that occurs only under some conditions (at some times or with some probability) and partial coordination. We will start by allowing the former and not the latter.

To clarify this point, we assume the existence of multiple behaviors, with each behavior being a response to a set of external conditions. A behavior does not correspond to a particular state of the system, but rather to a particular scheme of coordination of the parts. For example, under certain conditions two people lift at the same time, under other conditions one of them pushes while the other one pulls; under yet other conditions they act (push or pull) independently. Each of these is a separate behavior. We could consider the information necessary to describe the actions directly; however, this would not necessarily reveal the underlying coordination. For example, if there are two possible actions (R, L) of two individuals, and if half of the time according to the external conditions the individuals are coordinated to act jointly with equal probability (RR or LL), and half they are coordinated to act in opposition with equal probability (RL or LR), then the set of actions would include four actions with equal likelihood (RR, LL, RL, LR). Without recognizing the effect of the coordination, we would assume that the individuals are not coordinated at all. Instead, we consider each coordination scheme and the (conditional) information necessary to specify the behavior of the system under that scheme, then we average over the coordination mechanisms. For each type of coordination we assume that the system is divided into subsets that are independent of each other, but the components that make up each subset are fully coordinated [2, 3].

With these assumptions, we can see that the attributes of scale and variety constrain each other and what a system of a certain number of components can do. Because coordination reduces variety, the same number of components cannot both have a large variety and a large scale, though various tradeoffs are possible. A generalized Law of Requisite Variety expresses both the scale and variety necessary, and specifically states that at every scale the variety necessary to meet the tasks, at that scale, must be larger for the system than the task requirements, self-consistently defined as a necessary variety at each scale. For a particular behavior the variety at scale k is $D(k) = \nu n(k)$, where $n(k)$ is the number of different k-member fully coordinated groups needed to perform the entire task, which therefore at a minimum requires $N = \Sigma k n(k)$ components to perform. The total variety of the task is proportional to the total number of subsets of any scale $V = \Sigma D(k)$. The same number, N, of components can, in the extreme, perform a task of scale N, with variety equal to that of one component, or a task of scale one with variety N times as great. More generally, the equation [obtained from $N = \Sigma k n(k)$],

$$N\nu = \sum kD(k) \tag{1}$$

can be considered a constraint on the possible behavior patterns (sum rule) of a system due to different mechanisms of organization. It is often convenient to think about the variety of a system, $V(k)$, that has a scale k or larger, because this is the set of possible actions that can have at least that scale,

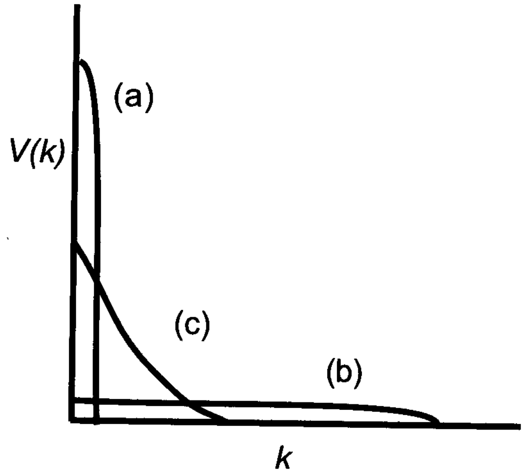
$$V(k) = \sum_{k'=k}^{N} D(k'). \tag{2}$$

Then the total variety of the system is $V(1)$, and the sum rule can be written as

$$\sum_{k=1}^{N} V(k) = N\nu. \tag{3}$$

The sum rule given by Equation (1) or (3) describes the existence of a tradeoff between variety at different scales. Increasing the variety at one scale, by changing the organizational form, must come at the expense of variety at other

FIGURE 1



Schematic illustration of the constraint on variety $V(k)$ (vertical axis) as a function of scale, k (horizontal axis, increasing to the right). A system with the highest possible fine-scale variety corresponds to a system with independent degrees of freedom (curve a). When all degrees of freedom are coherent the system has the largest-scale behavior, but the same low value of variety at all scales (curve b). Intuitively, complex nonequilibrium physical, biological, and social systems have an increasing variety as the scale decreases due to coordination at various scales (curve c). The scale-dependent variety characterizes the system organization and its functional capabilities. The area under all curves is the same for systems organized from the same components.

scales. A schematic illustration of several systems having the same number of components but different coordination mechanisms is shown in Figure 1.

We can also extend the discussion of coordination between parts to coordination between times. Imagine a set of photographs of the system taken at intervals of time; for example, images of the individual frames that make up a movie. Each of the components has a certain time interval between independent actions. For simplicity, we take the frames at this interval of time. Then the organization of the system can impose that the parts of the system persist in a specific state or change that state over time possibly in response to changes in the environment. In the first case we have fewer possibilities; in the second we have more possibilities. These possibilities, like the possibilities at a particular time, affect the set of possible reactions over the course of the sequence of frames (movie) that the system can execute. Repetition over time (more generally, the coupling of states over time) causes an increase in scale similar to the coordination of components at a particular time. Some actions may require persistent behavior, whereas other do not. Often we consider cases where the action of one component over time may be substituted for the action of many at a particular time, as, for example, in applying a large force at one time or applying a smaller force over time in achieving the same mechanical effect. Thus it is sometimes reasonable, at a first order of approximation, to summarize the entire system in number of components and time of action as a single measure of scale. In other cases this may not be so simply done.

In order to understand the physical implications of the scale-based decomposition of a system behavior, we consider the relationship between the behavior of a system and the impact on the environment. The Law of Requisite Variety in and of itself does not tell us what affects the impact on the environment. Intuitively, however, the larger the number of components that are coordinated, the larger the potential impact. To clarify this point, we should recognize that additive superposition of subtasks to achieve a desired task has a simple linear relationship between coordination and impact. More interesting is the case of tasks that are characterized by nonlinearity. Because we have assumed that certain tasks require a certain scale of coordination to achieve, we need to characterize these tasks. Such tasks requiring a certain scale are characteristic of nonlinear threshold processes. These arise in cases of metastability where in response to a small push the system will return to its original state, but a larger push will move the system to a different state. They are also characteristic of competitive processes, especially winner takes all cases, such as those that are considered in evolutionary biology. A predator chasing a prey typically either gets it or it does not. Additional speed on the part of one or the other can result in one case or the other. The implication of the existence of tasks that require a certain scale to accomplish is also that the impact of the system on the environment can be larger for larger scale actions. This suggests that scale can also be viewed as a measurement of the system's impact on the environment.

One way to make this more precise is to consider the environment as a system that itself has coordinated components in groups. Any transitions of the system are changes of state that change the states of all of the components in a particular group. This implies that in order for such a transition to occur the system performing the change must have a number of coordinated components as large as that of the environment. We could also generalize this by considering a "coupling constant," reflecting the degree to which a single component is able to change a single component of the environment. Still, the insight is clear that the scale necessary for a task is related to the number of coordinated components in the environment that must be changed in order for the task, understood as a change in the environment, to be achieved.

We can consider the relationship of scale of the system and scale of the environment behavior as a manifestation of coupling of system with environment when there are no specific assumptions about that coupling. This means that change in one is coupled to change in the other. When there are no specific assumptions, we should assume that on average the number of components of the system and the number of components of the environment that are coupled to each other are equal. Clearly this is not necessarily true; still it is a reasonable first assumption. In essence, the idea is that the boundary between system and environment does not necessarily characterize the coupling of components, so that coupled components are equally likely to be in and out of the system.

It is significant that the discussion of the level of a system's impact on the environment is not the same as the Law of Requisite Variety, because the latter does not describe the nature of the task to be performed. The level of impact is also not a mathematical identity or theorem but rather a statement that relies upon observations about the world. This would be a "Law" (if verified/not contradicted by observation) rather than a theorem, whereas the Law of Requisite Variety is a theorem rather than a Law. A reasonable concern about the relevance of scale of behavior to scale of response might be raised because of the existence of amplification and nonlinear dynamics/chaotic behavior of systems. This suggests that we consider the possibility that effects that are delayed over time (and possibly displaced in space) allow small actions to have large effects. Still, the relevance of scale to the interaction of systems is clearly present and should be accounted for in formal discussions.

Our generalization of the Law of Requisite Variety is directly relevant to the analysis of coordination mechanisms of an organization, biological or social. Specifically, how such coordination mechanisms are well or ill suited to the tasks being performed. Given the constraint imposed by the number of components, a successful organization has a coordination mechanism that ensures that the groups are coordinated at the relevant scale of tasks to be performed. This simple and intuitive statement is captured by the Multiscale version of the Law of Requisite Variety.

These ideas can be illustrated using examples of human organizations and the tasks they perform [5–7]. Ancient empires were constructed based on armies organized on repetitive actions in time and coordinated actions of individuals in groups like the Greek Phalanx and the Roman Legion, that were capable of marching large distances [8]. The task they performed is intuitively large scale as measured by the size of empires they created or the number of people they controlled. More recently, the Model-T Ford factory was based on a concept of repetitive actions of individuals, with distinct individuals performing distinct tasks, combined together to produce large quantities of relatively complex objects [9]. Given different environmental conditions, the production line can create a large variety through different actions of different individuals, but the large scale measured by the number produced in an interval of time arises from the simplified but large-scale repetition of the task of each individual.

Similarly, for biological systems there are manifest examples in the behavior of various molecular, cellular, physiological, or population systems. Conventional understanding of coordination and independence is sufficient for illustration [10]. At the molecular level, the contrast between structural and energy storing polymers, including carbohydrates on one hand and enzymatic or information storing proteins and DNA on the other, illustrates the role of variety and scale in polymer composition. At the cellular level, the barrier function of membranes is consistent with a homogenous aggregate and is different from the role of the variety of sensors embedded in it. At the physiological level, the coherent behavior of muscle cells can be contrasted with the variety of neural system responses. The combination of the two in the neuromuscular system provides large-scale responses at a particular time, with high variety over time. Considering populations of animals evokes images of large-scale herds and smaller groups of predators. In general, social coordination varies widely between species, as has recently been revealed in studies of various primates [11]. Analysis of the functional (selective) context of the environment and its relevance to social organization is an ongoing endeavor.

The discussion of the organization of social systems and their capabilities has become more interesting in a context where networks of dependency are increasingly being relied on rather than conventional hierarchies for coordination. The importance of complex systems concepts in this context has been recognized [12]. There is general agreement that uncertain (complex, dynamic, hostile) environments lead to a need for decentralization [13–17]. A key issue is the concept of a hierarchical organization, which is considered a mechanism of decentralized control because decisions can be made at many levels of hierarchy and can progressively be decentralized through delegation within the hierarchy. Limitations of the individual in obtaining information or making decisions ("bounded rationality") have been the core of discussions of the human limitations in central control. The idea that information collection limits the effectiveness of central control has led to the perspective that such problems might be solved through information technology (specifically vertical information systems) [16]. Below, this notion is directly challenged. It is also well recognized [15–17] that coordination between divisions of an organization may be addressed by lateral communication. Possible architectures of organizations to reduce the need for coordination have been discussed (divisions for product, process, geographic, and market segment) and the increase coordination through, "matrix organizations," and looser structures such as "adhocracies."

We will use the multiscale generalization of variety to consider the topic of centralization/decentralization of control and coordination. In considering variety—the limits of the number of possible actions that can be taken, we also gain a useful practical advantage over discussions of information and decision making. In discussing action, we have a better opportunity to observe and apply our conclusions, because decision making and control is often difficult to observe directly.

From the point of view of variety, the limitations of central control can be understood if we consider that a strictly centrally controlled system, where action is determined by one (or a few) individuals, has a variety that is not more than the variety of one (or a few) individuals, the individuals making the decisions, $V \leq v$. Distributing this control increases the variety. This is true for various forms of decentralization including hierarchies when delegation of authority is used.

In considering the requirements of multiscale variety more generally, we can state that for a system to be effective, it must be able to coordinate the right number of components to serve each task, while allowing the independence of other sets of components to perform their respective tasks without binding the actions of one such set to another. This now serves as a key characterization of system organization. Specifically, the Multiscale Law of Requisite Variety implies that in order for a system to be successful its coordination mechanisms must allow independence and dependence between components so as to allow the right number of sets of components at each scale.

When we consider modern organizational structures, we see that the military has different types of organization for different types of tasks [18, 19].

In the Gulf War in 1991, the military used a conventional military buildup to a force consisting of hundreds of thousands of soldiers, many in tank divisions. The coordination of these forces was distinctly large scale in the sense that maneuvers could be characterized by the movement of large numbers of individuals at the same time [20]. It is readily understood that this task could not be decomposed into actions of individuals acting independently, i.e., independently choosing when to move forward to attack. Thus the coordination/control mechanisms of the military provided for tight coordination of the movement of all the forces.

In contrast, in the War in Afghanistan, the military force used consisted of "special forces": teams of individuals each of which was largely independent of the others [21]. In some cases these teams were tightly coordinated amongst themselves; in others, individual members of these teams performed actions that were quite different from each other and loosely coordinated in the sense that acts of one individual did not depend on their timing or effectiveness on acts of other individuals.

The essential problem in each case is determining the number of individuals necessary to perform specific tasks, and therefore the coordination necessary between the individuals involved. Although this seems a simple statement, a key to improved understanding arises from the principles we are discussing of requisite variety and scale. Communication is often considered to increase capability; here we see that its effect is to increase coordination, which reduces variety of the whole system and particularly at a fine scale. When individuals are independent, they may perform many different possible tasks. When they are coordinated, they only perform a few possible tasks. Therefore the issue of independence versus dependence results in a tradeoff of variety versus scale.

A direct analysis of the problem of coordination provides insight into the organizational structures that are needed to perform high-variety tasks. A natural place to begin is the consideration of control hierarchies, the traditional form of human organizations. In such a system, the communication only proceeds up and down the hierarchy, not laterally. Coordination between workers of different parts of the organization must be directed by (or at least communicated through) the manager, whose domain of authority includes all the participants in that task.

Coordination corresponds to communication that is responsible for the mutual information between two or more workers. The information communicated by a part of a system does not exceed that part's variety. Each of these (information, variety) is measured per unit time, where a certain amount of time is required to switch to the next state. Assuming that a manager has a limited variety, this bound on the communication capacity (bandwidth) limits the coordination of workers under the supervision of the manager. Specifically, if we measure the conditional information, which assumes that we know when the coordination is taking place and when it is not, then the information in the manager's actions is equal to the mutual information of the workers. (If we do not assume the conditional information, then the mutual information is always less than the information in the actions of the manager, because the information describing the manager includes the coordination state as well as what is being coordinated.)

How do we describe a manager? A manager specifies the state of the subordinates and a coordination mechanism. We assume that at any particular time the manager can only coordinate a particular subset, indexed by w, of the subordinates, and at that time these subordinates are fully coordinated, whereas the others act independently (one cannot be in two places at the same time). $q(w)$ is the number of subordinates that are being coordinated, which, for values of zero or one corresponds to no coordination. A specification of the manager at a particular time can thus be written (s_m, w), where the state of s_m specifies the states of all the coordinated subordinates, whereas w specifies which subordinates are coordinated. For simplicity we do not count the redundancy provided by the manager (who we assume does not do the action only specifies it), and therefore s_m is not needed in the description of the system because it is redundant to the actions of the subordinates. We also neglect the information in specifying w by treating the information as conditional on the coordination mechanism. These assumptions can be relaxed without changing the conclusions. Then we have the multiscale variety for a particular coordination state given by

$$D(k|w) = \nu(N - q(w))\delta_{k,1} + \nu\delta_{k,q(w)}. \quad (4)$$

Combining coordination states, each with a probability $P(w)$, we have

$$D(k) = \sum_{w} P(w)\delta_{q(w),k}\nu + \delta_{k,1} \sum_{w} P(w)(N - q(w))\nu. \quad (5)$$

This gives the expected bound on the total coordination:

$$V(2) = \sum_{k=2}^{N} D(k) = \sum_{k=2}^{N} \sum_{w} P(w)\delta_{q(w),k}\nu \leq \nu. \quad (6)$$

where the inequality is the quite reasonable statement that the variety of the system for scales larger than one individual cannot be greater than the variety of the manager.

This coordination limitation is recursively applied to each level of managers for the set of individuals under their supervision so that the mutual information between individuals (workers or managers) at one level of organization is limited by the manager that supervises them. This implies, for example, that the combined mutual information between all workers is no more than the variety of the first level supervisors. Assuming that the variety of a manager is typically no more than the variety of a worker, we would expect that the limit of mutual information to be N/B, where B is the branching ratio, i.e., the number of workers supervised by a single manager. Higher level managers are similarly restricted in their ability to coordinate the managers at the lower level. We note that in a conventional hierarchy, when an upper level manager coordinates parts of the organization, this information must be communicated through the lower level managers. This also reduces the degree to which their own inter-worker coordination can be performed (i.e:, to the extent that the higher level manager performs coordination, this reduces the capacity of the lower level managers to coordinate).

We can make a more direct connection to multiscale variety if we consider a somewhat generalized version of hierarchical control. In the generalized version of the hierarchy, managers exist at a certain level of authority supervising a certain fraction of the organization, but do not have a particular set of subordinates that they supervise (the "matrix organization" [16] is an intermediate case). By not including the constraint of a strict hierarchy—that a manager has a particular subset of the individuals and cannot coordinate others outside of this subset, we obtain an upper bound on the coordination of a more conventional hierarchy. If we were to include this additional constraint, then the coordination of the system is further limited because then even only two individuals who are in different divisions of the organization require coordination by the CEO. For the generalized hierarchical model, we can generalize the equations above and reach a conclusion that

$$V(2) = \sum_{k=2}^{N} D(k) \leq C\nu. \quad (7)$$

where C is the number of managers. This states quite reasonably that the total variety of actions greater than the scale of one individual is not greater than the total variety of the managers. For managers having a certain limit on how many subordinates they can control, so that managers at level l can coordinate up to B^l subordinates, we further limit the number of those coordinated at larger scales by

$$V(B^{l-1} + 1) = \sum_{k=B^{l-1}+1}^{N} D(k) \leq \sum_{l' \geq l} C_{l'}\nu, \quad (8)$$

which reasonably states that the variety of behaviors associated with a number of individuals is only as great as the variety of the managers that can coordinate that number of individuals.

For example, we consider the role of the CEO and assign him/her the obligation of determining those issues that are of relevance to the actions of a large proportion of individuals that are part of the organization. If we consider 10% to be the threshold fraction, then all decisions involving 10% of the individuals of the organization are coordinated by the CEO. The maximal possible variety of such portions (at this scale of action) is 10 times the variety of a single individual. However, this cannot be done when coordinated by a single individual; the maximum is the CEO's variety. More generally, we can categorically state that to the extent that a single individual is coordinating the behavior of an organization, to that extent the coordination (as defined by mutual information) cannot have a higher variety than an individual.

We see that for a hierarchically coordinated system, the combined conditional mutual information of subunits of a manager cannot be greater than the variety of that manager. This is not a problem for either of two cases (dictated by environmental conditions): if the system has a simple coherent behavior or if the manager exercises very little control so that the workers are almost totally independent of each other. It is a problem, however, when the behaviors of subunits themselves have a high variety (greater than that of an individual) and these must be coordinated. Thus, a hierarchical control system is well designed for relatively simple large-scale behaviors or for systems with very distributed control, but not for highly coordinated behaviors, i.e., when the coordination of these behaviors is more complex than a human being can communicate [22, 23].

We can consider these concepts from a phenomenological point of view. Centralized coordination of components was characteristic of scientific management as applied to the economy of the USSR that specified the coordination of industrial enterprises. Failures of this system in providing agricultural products of appropriate quantity but possibly more importantly of sufficient variety [7, 24] led Gorbachev, First Deputy Prime Minster in charge of agriculture before becoming General Secretary of the Central Committee of the Communist Party, to institute reforms that preceded the collapse of the Soviet Union.

The general discussion of hierarchical coordination applies in many contexts. For example, a context where the same limitation applies is in traditional "systems engineering." In the conventional systems engineering approach the project is recursively broken into subparts. The parts are then put together, with the task of selecting and coordinating the subprojects, the domain of the systems engineer. The failure rate of such engineering projects in recent years has been remarkably high costing many billions of dollars [25–27].

The characteristics of the environment that require high variety, largely independent individuals doing diverse tasks, rather than large scale, many individuals performing the same task, has been most clearly learned by the U.S. military from difficulties in Vietnam and from the Soviet experience in Afghanistan. Unlike more conventional warfare, complex terrains and enemy guerrilla tactics require high-variety rather than large-scale forces. This can be readily understood when we consider the nature of the task of firing at a dispersed and hidden enemy, possibly among civilians. Highly coordinated large-scale actions are manifestly incapable of the variety of actions necessary, as measured by the large number of possible specific locations of the enemy forces. This shows how the variety of the environment at different scales is directly related to the variety of a successful system.

The even more highly dispersed Al-Queda and the asymmetric war against terrorists is generally understood to be highly complex, and the large number of possible actions that might have to be taken is a clear indicator of the variety necessary of the system that might effectively oppose them.

We can consider this same lesson for other societal problems [28].

Among these problems is the poverty in developed nations and third world development. Despite many efforts involving substantial resources these problems have proven to be highly intractable to efforts by governments and charitable organizations. The main observation that we make here is that many of these efforts, such as the "War on Poverty" and the actions of the World Bank in the Third World, involved actions based on highly coordinated efforts using large amounts of financial capital and human beings following a script, corresponding to a centrally defined set of actions to be taken. The dominant "holistic" but still centrally planned approach [29] is insufficient. It has become increasingly apparent that the problem of poverty involves many different acts to respond to many different kinds of local, often individual, needs. And the task of Third World development, with objectives of creating functional societies that are themselves organizations capable of performing tasks of high variety, is unlikely to be addressed by large-scale actions.

Similarly, we might consider the problems of healthcare in the United States, where cost containment is a uniform large-scale force being applied to a system that involves high variety in individual treatment by individual doctors. The effort to apply effective controls over the expenditure of resources for individual care can be directly related to the Law of Requisite Variety, which suggests the need for a high-variety control system is incompatible with a centrally directed system with comparatively low variety. A direct symptom of application of a uniform approach to a high-variety problem is a high rate of failure. The documented problems with medical errors and quality of care appear to confirm the diagnosis presented here [30].

We can also consider the education system in the United States. In recent years the concept of standardized testing has become the primary mechanism for evaluation as well as application of force for change toward solving problems of the education system. Again, the fundamental question to be asked is whether the task can be effectively met by such a high-scale, low-variety strategy. The answer might be framed in terms of the necessary variety of the society that is to be formed out of the next generation of children, because this is an essential outcome of the educational process [31].

To further formalize the Multiscale Law of Requisite Variety to arbitrary coordination, we use an information theory framework that can represent any order of correlations between components [5, 6, 32–39]. Conventionally in studies of physical, biological, and social systems, the focus is on correlations between pairs of components. A formalism that can consider the coupling of arbitrary numbers of elements is necessary for a complete multiscale treatment. The coordination between elements of a system is captured in the probability over all possible states of the system. We label the states of the components of the system $\{s\} = \{s_i: i = 1,\ldots N\}$ and the probability distribution of states $P(\{s\})$. We are interested in characterizing the variety of responses of the system, where each action involves a certain number, k, of elements. An explicit expression can be found in terms of the information in subsets of the system. The derivation requires ensuring that a proper count is made of the different levels of coupling. When parts of the system are coupled, the variety in a particular subset of the system cannot be counted separately from the variety of other parts, i.e. it may overlap the variety of other subsets of the system. The degree to which this overlap occurs depends on the multiplicity of the underlying dependencies of the variables. A careful treatment of the combinatorics of k-fold couplings for all k is necessary. The result is

$$D(k) = \sum_{j=0}^{k} (-1)^{k-j+1} \binom{N-j}{k-j} Q(N-j), \qquad (9)$$

where

$$Q(N-j) = -\sum_{\{i_1,\ldots,i_j\}} \sum_{s} P(s_1,\ldots,s_N)\log_2 \sum_{\{s_{i_1},\ldots,s_{i_j}\}} P(s_1,\ldots,s_N) \qquad (10)$$

is the sum over the information in all groups of $N - j$ components, i.e., the information excluding sets of j components from the N total number of components. The sum over $\{i_1,\ldots,i_j\}$ implies a sum over all distinct subsets of j variables out of the entire set of N variables. To our knowledge this expression has not appeared in the literature and provides an explicit form for the information in k-fold dependencies for any system described by a probability distribution over the ensemble. Using this expression, we can prove by evaluation that, quite generally, a system satisfies the sum rule given by Equation (1) or (3). This means that organizational structures, i.e., coordination mechanisms, always have a tradeoff in behaviors at different scales and particularly that large-scale behaviors must come at the expense of high variety at a fine scale. For the case where the system partitions into independent groups of fully interdependent components, Equation (9) reduces to a count of the number of subgroups that have k-components and thus to the discussion above. More generally, however, Equation (9) does not reduce to a count of subgroups and quite different behaviors can be found [38].

In summary, a generalization of the Law of Requisite Variety suggests that the effectiveness of a system organization can be evaluated by its variety at each scale of tasks to be performed. In its simplest form, when a system has a high degree of coordination, then it is large scale. When it is not coordinated, allowing for independent component action, then it has high variety. The tradeoff of large-scale action, as compared to the variety possible when actions of components are independent provides a direct analysis of system organization. Although it does not specify that a particular system is capable of performing a task, it can provide a necessary condition for such effectiveness. In considering biological and social systems, such analysis provides a way of classifying their behavior and considering the functional role they play in survival and societal function. The U.S. military became successful in complex warfare situations by adopting appropriate organizational structures to these tasks. In contrast, in many other responses to difficult social problems, an approach analogous to conventional large-scale military confrontation has been adopted. When considering the details of action needed and the lack of success of traditional approaches, it is reasonable to suggest that high-variety organizational structures that are appropriate to these tasks may be more successful [28].

ACKNOWLEDGMENTS

I thank Michel Baranger, Irving Epstein, Mehran Kardar, David Meyer, Larry Rudolph, and Hiroki Sayama for helpful discussions. This work was supported in part by the National Science Foundation under Grant No. 0083885.

REFERENCES

1. Ashby, W.R. An Introduction to Cybernetics; Chapman & Hall, 1956.
2. The total information describing the system would also include the information needed to specify the coordination mechanism. The approximation used here is valid if the amount of this information is small compared with the information in the behavior. If the information in the coordination mechanism is not small compared with the information in the behavior, we can treat this case using the more general formalism discussed later.
3. Although the discussion is eventually quite general, the immediate description can be considered as a strong version of a "nearly decomposable," modular/hierarchic system; see [4], Chapter 8, and [5], Chapters 2 and 3.
4. Simon, H.A. Sciences of the Artificial; 3rd Ed., MIT Press, 1996; Chapter 8.
5. Bar-Yam, Y. Dynamics of Complex Systems; Perseus Press, 1997.
6. Bar-Yam, Y. General Features of Complex Systems, UNESCO Encyclopedia of Life Support Systems, in press.
7. Bar-Yam, Y. Complexity Rising, From Human Beings to Human Civilization, a Complexity Profile; UNESCO Encyclopedia of Life Support Systems, in press.
8. Webster, G. The Roman Imperial Army: Of the First and Second Centuries A.D.; 3rd Ed.; University of Oklahoma Press, 1998.
9. Weitzman, D. Model T: How Henry Ford Built a Legend; Crown, 2002.
10. For example, Campbell, N.A.; Reece, J.B. Biology; Benjamin Cummings, 2002.
11. Smuts, B.B.; Dorothy, L.C.; Seyfarth, R.M.; Wrangham, R.W.; Struhsaker, T. T.; Eds. Primate Societies; University of Chicago Press, 1987.
12. Anderson, P. Complexity theory and organization science. Organization Sci 1999, 10, 216.

13. Statements about the inadequacy of central planning and central control appear to have originated in the context of the effectiveness of free markets as opposed to planned economies [14]. In this context free markets were described as mechanisms of coordination. We feel it is necessary, however, to distinguish markets from organizational structures since markets are not specific coordination structures but rather contexts in which coordination may take place.
14. Hayek, F.A. The Road to Serfdom; Routledge, 1944.
15. March, J.G.; Simon, H.A. Organizations; 2nd Ed.; John Wiley & Sons: New York, 1958.
16. Galbraith, J. Designing Complex Organizations; Addison-Wesley, 1973.
17. Mintzberg, H. The Structuring of Organizations; Prentice-Hall, 1979.
18. Bar-Yam, Y. Multiscale Representation Phase I, Final Report to Chief of Naval Operations Strategic Studies Group, 2001.
19. Bar-Yam, Y. Complexity of Military Conflict: Multiscale Complex Systems Analysis of Littoral Warfare Report to Chief of Naval Operations Strategic Studies Group, 2003.
20. Hutchison, K.D. Operation Desert Shield/Desert Storm; Greenwood, 1995.
21. A new breed of soldier. Newsweek Dec. 10, 2001, 24.
22. Introducing lateral communication [16] into a hierarchical system, e.g., by using a liaison between divisions of an organization or an integrating manager, introduces additional coordination at the level of the variety of one individual, v. This is still a very limited degree of coordination. A separate discussion of matrix structures that connect divisions of organization through dual reporting to a few managers is not necessary because they are already described in the generalized model of hierarchy discussed in the text. If these managers talk to each other (which is not assumed) as opposed to their subordinates, then there is actually a reduction in coordination capability.
23. There is one form of hierarchical control that is not ruled out by our discussions. When the set of possibilities is only a few, even if they are radically different from each other (involving changes in the action of many individuals), then the coordination/decision can be made by a single individual. This implies that that aspect of the organization is coherent, i.e., large scale and not of high variety. For example, the choice of whether or not to go to war can be made by an individual with only two possible decision states. However, this reflects the assumption that all aspects of the internal coordination necessary for the two states are made by others. Although this aspect of central control is not limited by our discussion, it is important to recognize the applicability of limitations by other arguments: The availability of the necessary information [14] and information processing to make the decision [15]. This information is related to the structure of the decision making process. The process must be able to contain prototypes of conditions and pair them with actions (or conditions and actions with effects).
24. Birman, I. Personal Consumption in the USSR and the USA; Palgrave Macmillan, 1989.
25. Bar-Yam, Y. Enlightened Evolutionary Engineering/Implementation of Innovation in FORCEnet, Report to Chief of Naval Operations Strategic Studies Group, May 1, 2002.
26. Bar-Yam, Y. Unifying Principles in Complex Systems. In: Converging Technology (NBIC) for Improving Human Performance; Roco, M.C., Bainbridge, W.S., Eds.; Kluver, 2003.
27. Bar-Yam, Y. When Systems Engineering Fails—Toward Complex Systems Engineering; IEEE International Conference on Systems, Man & Cybernetics, in press.
28. Bar-Yam, Y. The Complex World, http://www.necsi.org.
29. Wolfensohn, J.D. The challenge of inclusion. World Bank, 1997.
30. Kohn, L.T.; Corrigan, J.M.; Donaldson, M.S., Eds. To Err Is Human: Building a Safer Health System; Institute of Medicine (IOM), National Academy Press, 2000.
31. Senge, P.M.; Cambron McCabe, N.H.; Lucas, T.; Kleiner, A.; Dutton, J.; Smith, B. Schools That Learn: A Fifth Discipline Fieldbook for Educators, Parents, and Everyone Who Cares About Education; Doubleday, 2000.
32. Shannon, C.E. A mathematical theory of communication. Bell Systems Technical Journal July and October 1948; reprinted in Shannon, C.E.; Weaver, W. The Mathematical Theory of Communication; University of Illinois Press: Urbana, 1963.
33. Wolf, D.R. Information and Correlation in Statistical Mechanical Systems, Ph.D. Dissertation, University of Texas, Austin, TX 1996; A quantity discussed here appears to be the same as $D(k)$, though no explicit form was given.
34. Sporns, O.; Tononi, G. Classes of network connectivity and dynamics. Complexity 2001, 7, 28. Information theory (including multicomponent correlations) was used here to define a single measure of complexity as a measure of universal quality of a neural system and was applied to interdependence, subdivision, redundancy, and consciousness in neural systems (see also [4], Chapters 2 and 3). In their discussion of universality used to characterize the nervous system, the contrast of, e.g., muscles and the nervous system, that violates their notion of universal effective response, has not been included.
35. For related discussions and applications, see [5], Chapters 8–9, and [36–39].
36. Bar-Yam, Y. Sum rule for multiscale representations of kinematic systems. Advances Complex Systems 2002, 5, 409.
37. Davidson, A.; Teicher, M.H.; Bar-Yam, Y. The role of environmental complexity in the well being of the elderly. Complexity Chaos Nursing 1997, 3, 5.
38. Gheorghiu-Svirschevski, S.; Bar-Yam, Y. K-fold mutual information in the long range Ising model. manuscript submitted.
39. Bar-Yam, Y. Multiscale Complexity/Variety, Advances in Complex Systems (in press).