

Complexity vs. Selection: Retuning Kauffman's "Tunable" Adaptive Landscape

Accepted at *Organization Science*, but withdrawn in favor of a substitute—currently under revision
All rights reserved. Not to be quoted, paraphrased, copied, or distributed in any fashion.

Bill McKelvey

The Anderson School at UCLA, 110 Westwood Plaza, Los Angeles, CA 90095-1481
Phone 310/825-7796 Fax 310/206-2002 mckelvey@anderson.ucla.edu

June 1997

In his recent book, *The Origins of Order*, Stewart Kauffman argues that complexity may intervene in special circumstances to arrest Darwinian natural selection processes thought to create order in the biological world. In the past two decades selectionist theory has been applied as a means of explaining emergent order in firms and populations of firms. It is possible that complexity theory poses an alternative explanation of organizational evolution as well, since organizations have been accepted as complex for many years. As microevolutionary theory developed in biology, adaptive landscapes or sequence spaces emerged as a way of modeling microevolutionary decision choices by metaphorical adaptive agents. One of these is the *NK* model developed by Kauffman. This model appears particularly relevant for modeling complexity effects on microevolutionary organizational processes. Its defining characteristic is the idea that suboptimal peaks can be "tuned" or reshaped by the effects of complex interdependencies, by buckling and heaving the landscape to make it more rugged—to the point of causing complexity catastrophes, that is, conditions when complexity dominates selection. These occur when: 1) selection forces are not strong enough to create a new population with novel properties; and 2) adaptive peaks are so flattened that even though strong selection forces exist, the selected population remains undifferentiated from the general population. Should Kauffman's *NK* model be imported into the study of firms? This article presents a critical analysis of Kauffman's method of operationalizing his key parameters: "number of agents," number of interdependencies," and distribution of interdependency effects among the agents. Results of the analysis indicate that 1) Kauffman's allowing of infinitesimal steps in the adaptive move rule works to negate the effects of interdependencies in creating a complexity catastrophe; 2) If a lower bound threshold to infinitesimal steps is introduced, the catastrophe occurs inevitably and more suddenly because of the threshold effect as complexity is increased; 3) Kauffman's definitions of correlation and ruggedness, requisite to the tunable landscape concept, fail; 4) Landscapes, by Kauffman's operationalization become more correlated than rugged as complexity increases—opposite of what he thinks happens; and 5) The use of a proper distribution parameter and the consequent protection of some agents from interdependencies appears to create the correlation of increased ruggedness with increased complexity effects—what Kauffman needs if the *NK* model is to properly represent his theory. The analysis also raises questions about some of Kauffman's computational simulation tables. Several alternative approaches to operationalizing the *NK* model are suggested, all of which serve to position the model to represent Kauffman's tunable landscape version of complexity theory more accurately. An improved *NK* model could play a significant role in better understanding organizational microevolution.

I. INTRODUCTION

Complexity has been associated with organizations for decades (Etzioni 1961, Blau and Scott 1962, Simon, 1962, Haas and Drabek 1973, Hall, Haas, and Johnson 1967, Miller and Friesen 1984, Donaldson 1985, Von Glinow and Mohrman 1990). A current trend is to call for increased levels of cross-functional integration and dynamic capability as the effects of global competition and change increase (Dimancescu 1992, Galbraith, Lawler, and Associates 1993, Graham 1994, Ashkenas et al. 1995, Leonard-Barton 1995, Johann, 1995). Firms have as much or more complexity now as they ever did (Donaldson 1995) and more complexity appears in sight, though simplicity may be under appreciated (Rommel et al. 1995).

Evolution toward complexity is taken as a given in biology (Bonner 1988, Depew and Weber 1995). Evolutionary explanations of complexity have also emerged in organization science (Kaufman 1975, Hannan and Freeman 1977, 1989; Aldrich 1979, Weick 1979, McKelvey 1982, Singh 1990, Hannan and Carroll 1992, Baum and Singh 1994, Baum, 1996). This trend is rooted in Darwinian theory wherein the "order" which scientists are trying to explain, defined as the extant patterned structure of economic and social organization, is attributed to selectionist effects. Selectionist theory also appears in economics (Alchian, 1950, Winter 1964, Hirshleifer 1977, Nelson and Winter 1982, Foster 1987, Goudy 1994). Friedman (1953), used a selectionist argument to defend the rational actor assumption, saying that all existing economic actors must be rational since less rational ones would have been selected out.

Consider the possibility that the organizational "order" to be explained could be as much the result of complexity as selection. Stewart Kauffman faces biologists with this possibility, recognizing that modern biologists universally adhere to the selectionist paradigm. The work of Kauffman and his coworkers over the past couple of decades is summarized in his book, titled, *The Origins of Order* (1993). In this book Kauffman draws on elements of evolutionary modeling as practiced by modern biologists and complexity theorists. His principle analytical tool is the *NK model*. By now this model has appeared in the literature for nearly a decade and is generally well known among analytical biologists and complexity theorists, since Kauffman has been a long-time editor of the *Journal of Theoretical Biology* and is a founding member of the Santa Fe Institute, the center of U. S. complexity theory. Kauffman is specifically credited for introducing the idea of a 'tunable adaptive landscape' (see Macken and Stadler 1995) for the purpose of connecting complexity analysis with adaptive landscape microstate modeling (Kauffman and Weinberger 1989). Inasmuch as firms are both complex and subject to selectionist forces, Kauffman's paradigm may apply to firms, and indeed, perhaps more so than to genes and organisms.

Biologists study evolutionary dynamics at multiple levels of analysis, ranging from very micro genetic mutations toward improved fitness to macro ecological dynamics affecting speciation and population regulation. They draw on *microstate analysis* in the context of macro competitive dynamics, adaptive landscape metaphors, statistical mechanics, and with Kauffman, complexity theory. Stochastic microstates, such as the Brownian

motion of atomic particles, genetic mutations, kinetic gas molecules, neuron firings, behavioral events, and specific process events in firms, characterize modern scientific assumptions about the bottom level phenomena of a specific discipline such as physics, biology, behavioral science, or organization science (McKelvey 1997). In developing his argument, Kauffman draws on modern biological methods for modeling mutational microstates, following Fisher (1930) and using Wright's (1931, 1932) notion of the adaptive landscape and various so-called adaptive learning models, particularly *spin glass* (Fischer and Hertz 1993) and *cellular automata* (Weisbuch 1993).

Organization scientists could continue to study evolutionary dynamics using the natural history or case method approach advocated by Eisenhardt (1989, 1991), Van Maanen (1989), and Porter (1991). This method focuses on what Darwin called "*contingent detail*" in his famous debate with Asa Grey (Gould 1989, p. 290). Alternatively we could reemphasize a search for what Darwin termed "*background laws*" as advocated by McKelvey (1997). This direction calls for adopting some methodological ideas from the biologists along with selectionist theory. A logical starting point would be the tunable landscape approach of Kauffman, along with his spin glass derived analytical and computational *NK* model. With the tunable landscape, Kauffman combines complexity effects with selection effects in one dynamic analysis.

The *NK* model appears to have compelling relevance to the study of adaptive competitive processes affecting the evolutionary dynamics of firms and industries, since it merges adaptive landscape modeling with microstate analysis and complexity theory. Consequently this model could easily be taken as a convenient starting point for applying microstate, adaptive landscape, and complexity analysis to firms (see for example, Levinthal 1995, Choi 1996, Maguire 1996, McKelvey and Kiousis 1996, Sorenson 1996, Westhoff et al. 1996, McKelvey forthcoming). But, before a microstate analytical procedure such as the *NK* model is imported widely into organization science, three concerns need to be addressed. **First**, is microstate analysis relevant to organization science? Elsewhere (McKelvey 1997, forthcoming) I argue that firms, as complex adaptive systems, are especially amenable to selectionist and complexity theories and microstate analysis, so won't repeat that discussion here. **Second**, does the *NK* model actually do what Kauffman says it does? I identify various descriptive statements by Kauffman and several consequent inadequacies in the model. **Third**, is it possible to "fix" his model and assure that it could represent firms. Several alterations to the *NK* model are suggested.

In Section 2, I introduce the adaptive landscape modeling approach and Kauffman's tunable variant along with a discussion of its relevance to organization science. Section 3 shifts to my critique of what the *NK* model does or does not do, based on Kauffman's various descriptions

of it. Based on this critique I conclude, in Section 4, by suggesting some alterations that make the model a truer representation of Kauffman's theory and more relevant to firms.

2. TUNABLE ADAPTIVE LANDSCAPES

I begin this Section by briefly discussing the development of the adaptive fitness landscape metaphor in biological microstate modeling. The presentation then turns to a consideration of the relevance of microstate landscapes to organizational evolution. The section ends with an "initializing" presentation of the *NK* model as a tunable landscape as received in the literature.

2.1 EVOLUTION ON FITNESS LANDSCAPES

The notion of an adaptive landscape is attributed to Sewall Wright (1932). This metaphor has subsequently retained considerable popularity among biologists, though in other disciplines *sequence*, *configuration*, or *search space* is preferred. An adaptive landscape has three elements: 1) A configuration space; 2) fitness functions; and 3) move rules which define the steps of the *adaptive walk*—modeling definitions of which are the subject of much of this article. Wright's landscape is an element of his overall contribution to the Modern Synthesis in biology beginning circa 1930, in which theories of evolution, taxonomy, and genetics are merged. One aftermath of this synthesis has been a prolonged debate between population geneticists and paleontologists about who really should sit at the High Table of evolutionary biology, and thus who best can explain the evolution of species (Eldredge 1995). This is in fact an argument between micro biologists and macro biologists. I mention this because I argue elsewhere (McKelvey 1997) that microevolutionary theory at the process event level could very well have considerable explanatory power and consequently set up a similar kind of debate in organization science—between explanations based on idiosyncratic process microstates and complexity driven emergent structures vs. macro contextualist explanations based on ecological analysis.

As one approaches explanation from a "micro" level, the landscape or search space becomes central. In biology there is no question that the doyen's of microevolutionary biology all draw on the landscape concept. As Macken and Stadler (1995) observe, Maynard Smith (1970) uses it to study protein evolution; Eigen (1971), Spiegelman (1971), and coworkers (Kramer et al. 1974) use it to investigate the *in vitro* evolution of RNA molecules; the Vienna group also uses it to study RNA adaptation (Fontana and Schuster 1987, Fontana, Schnabl, and Schuster 1989); and Kauffman and Weinberger (1989) introduce the idea of a tunable landscape in which complex interdependencies are allowed to affect fitness yields. Two key messages relevant to organization scientists come from this work, in addition to the fact that

adaptive landscape modeling is a cornerstone of modern theoretical biology. **First**, a number of methods of evolutionary computational modeling have been developed in biology and other disciplines; and **second**, microevolution in fact takes place at and may be studied at multiple levels. As McKelvey (1997) observes, the stage is set for multiple microstate levels of analysis in organization science.

Given a tunable landscape, Kauffman (1993) identifies two conditions when complexity effects may thwart selectionist effects as the root cause of order in biology:

1. In a “correlated” landscape containing some clearly advantageous fitness peaks, if selection forces are weak and thus fail to hold members of a population high up on the peaks, the apparent order in the population is due to the typical properties of the majority of the population still spread around the valley. That is, “adapting systems exhibit order not *because* of selection but *despite* it” (1993, p. 35; his italics).
2. In a “rugged” landscape, given that (a) as peaks proliferate they become less differentiated from the general landscape; (b) in precipitous rugged landscapes adaptive progression is trapped on the many suboptimal “local” peaks; and (c) even in the face of strong selection forces, the fittest members of the population exhibit characteristics little different from the entire population.

Kauffman labels these conditions ‘*complexity catastrophes*’ because one or the other inevitably happens if the “complexity of the entities under selection increases.” Thus complexity imposes an upper bound on adaptive progression via selection “when the number of parts exceeds a critical value” (1993, p. 36). The “catastrophe” is designated as such because complexity acts to thwart the selectionist process, thereby stopping progression toward improved fitness.

The bullets defining the catastrophe conditions introduce two concepts central to Kauffman’s thesis, ‘*correlated*’ and ‘*rugged*’ landscapes, which are also key determinants of his notion of tunable landscapes, and which form the key elements of the *NK* model. *N* represents the number of significant components comprising an adapting entity, such as a gene, chromosome, trait, or species, or in our case, number of process events, units, actors, firms, or generally the number of agents attempting to achieve higher fitness. *K* stands for the number of interdependencies among the agents; *K* can range from 0 to *N*–1. Thus, *K* is a measure of the complexity of interdependencies. Kauffman argues that *K* causes the landscape to buckle and deform with the result that it changes from a single dominant fitness peak at the *K* = 0 extreme to many low level peaks at the *K* = *N*–1 extreme. Kauffman defines the *K* = 0 landscape as highly correlated whereas the *K* = *N*–1 landscape is highly rugged. Rugged landscapes contain many peaks and valleys, steep slopes, many suboptimal peaks, and offer a greater chance of an agent being trapped on a suboptimal peak. Note that suboptimal in a land of many low peaks may not be much less than a considerably flattened global optimum.

Kauffman uses the *NK* model, a spin glass variant, to investigate the following kinds of questions: 1) How high

are the fitness levels of local optima? 2) How many and how similar are local optima? 3) How long are the walks to local optima? 4) What is the rate at which the number of more fit variants along a walk diminishes? 5) How long a wait before an agent discovers a fitter variant? 6) What sizes are the basins of attraction? Which is to ask, how many walks toward a particular peak from different starting points are possible? In total, these questions focus on the *rate of adaptation* and *level of success* likely on a particular landscape. With tunable landscapes one may ask how levels of complexity affect rates and levels of adaptive success by altering the ruggedness of the landscape. Kauffman’s applies the *NK* model in studies of fundamental biological questions pertaining to adaptive evolutionary rates in protein evolution, the crystallization theory of the origin of life, the origin of a connected metabolism, the formation of autocatalytic sets of RNA catalysts, and the evolution of genetic regulatory circuits. With the *NK*[*C*] model he uses cellular automata models to explore the distortion of landscapes due to micro level complexity effects on the coevolutionary dynamics between opponents, the complexity induced percolation of emergent ecological structures, and complexity induced alterations of the landscape affecting the relative height of Nash equilibrium levels.

McKelvey (forthcoming) applies the *NK*[*C*] model, where *C* represents interdependencies between opponents at the “parts” level, to value chains by translating results from models used by Kauffman for studying biological phenomena into the world of firms and their strategies. Thus coevolutionary value chain microstate activities give rise to complexity effects which in turn cause other outcomes. Some key strategic choices for firms suggested by the existing *NK*[*C*] model are: 1) Firms should keep internal and external levels of complexity at levels less than those of direct competitors; 2) Firms should keep internal complexity at about the same level as external complexity; 3) Firms should consider simultaneous advances in the adaptive fitness of only a few sites at any given time; and 4) Even if *N* grows, firms should limit the size of *K* and *C* to avoid complexity catastrophe.

2.2 RELEVANCE TO FIRMS AND ORGANIZATION SCIENCE

The relevance of the *NK* model to firms rests on two main questions: 1) Are complexity catastrophes likely alternatives to selection as explanations of organizational order? and 2) Is the *NK* model’s focus on issues related to the rate and level of adaptive success, that is, number and height of locally optimal peaks, length of walks, etc., relevant for organization scientists to attend to? As to the first question, is it possible for the two kinds of catastrophe to occur in firms?

First, a possible condition one: According to Pfeffer (1994, p. 241–242), a few organization design “peaks” offering adaptive advantage were laid down by the Americans in Japan after the Second World War—

specifically ideas by Deming and the “Training Within Industries” programs set up by the American Occupation authorities. The Japanese were motivated to scale these peaks with “green field” firms by virtue of having to jump start their postwar domestic economy. Consequently they had no initial existing complexity to overcome and they experienced strong adaptive tension later due to intense domestic competition and eventual entry into the U. S. markets. In contrast, in the U. S. postwar selection forces were weak, given the strong postwar domestic demand and ease of entry into foreign postwar markets. Consequently U. S. firms making the transition from a wartime to a peacetime economy, and burdened by existing complex webs of interdependencies, were not faced with a high enough adaptive tension and so failed to scale the peaks. This left the emergent “order” in American industries more a function of postwar complexity than postwar selection—a condition one complexity catastrophe.

Second, a possible condition two: American business schools may be trapped on a landscape having many possible peaks of adaptive success—for example, academic excellence, placement of Ph.D. graduates, teaching excellence, executive education excellence, quality of entering students, high ratings, Fortune 500 recruiter preference, Wall Street’s preference, salary levels, number of offers, analytical strength, high tech strength, global expertise, fund raising ability, and so forth. No single peak offers much differentiable advantage and any school achieving any peak does not gain much in selective advantage, even though competition among schools is quite stiff. Following Kauffman’s tunable landscape argument, the many peaks are caused by complex interdependencies, as follows. The parts of the business school value chain necessary for achieving fitness yields useful in climbing toward any of these peaks are interdependent in most schools because of a common faculty (often graduates from the same Ph.D. programs) and common core, common path dependencies, and common ties to outside stakeholders such as recruiters and executives. These interdependencies are also increasing because of attempts to install “more integrated” cores, or hire faculty “with both teaching and research” values. Therefore, the height of the fitness yield from, say, building up faculty research, is diminished because of interdependent lowered fitness yields from other parts of the value chain because less money is available for improving them. Thus, while research skills might increase overall fitness, specific fitness yields from research might decrease due to limitations imposed by the other interdependent agents or reputational elements, such as: spending on teaching, computer technology in the classroom, global relevance, interest and time for case writing. Consequently, succeeding with respect to a particular peak offers little advantage over other schools achieving success on one of the many other peaks—a consequence of the landscape consisting of multiple, considerably flattened peaks. Thus, even though strong competition and selectionist forces are present, schools

successful in achieving their highest feasible adaptive peak are little differentiated from other schools. This happens because “order” is more the result of a complexity induced “flattened peak” landscape rather than selectionist effects—a condition two complexity catastrophe.

As to the second question, a model is supposed to offer an idealized and reasonably valid representation of organizational phenomena. The *NK* model is aimed at evolutionary events most likely following the kind of environmental shock that shifts a population of firms from “convergent” to a “punctuated” form of adaptive change (Tushman and Romanelli 1985), Tushman and Anderson 1986), and Rosenkopf and Tushman 1994). Take the rate of development of digital technology over the past decade and the current shift of U. S. television from analog to digital as a result. The *NK[C]* part represents the microcoevolution of the various players as opponents in the competition attempt to steer the system toward the global optimum (dominant design) most beneficial to themselves. The *NK* part represents the rates at which the system adapts, the length of adaptive walks from firms’ starting points to the optimal peak, the likelihood of the dominant design (Anderson and Tushman 1990) getting trapped on localized suboptimal peaks, the rate and likelihood that the system will adopt and thus reduce the number of remaining technological fitness yields, how long before a firm is likely to discover a more preferable or fitter variant, and finally how many different “technological approach walks” toward a “technological peak” or dominant design are likely in a population.

The *NK* part also allows the representation of varying complexity effects. Given that complexity is defined in terms of the various technological or other “parts” comprising the value chain, which have diverse levels of interdependency, to what extent do the technological and other value chain interdependencies—of the cross-functional kind—buckle the landscape so as to flatten the technological adaptive peaks? And thus, does complexity thwart selectionist effects or prevent a coevolutionary pocket from reaching the global optimum? Models such as the *NK* also allow representation of the relative impact of micro and macro level effects on the movement of a system toward technological peaks, as well as micro and macro effects on fitness, survival, attaining competitive advantage, and ultimately rents. The *NK* model offers one platform for micro level coevolutionary analysis. Although its ostensible purpose is to study the impact of complexity, it also supports the investigation of the optimum balance between the autonomous specialization of value chain “parts” and cross-functional integration.

For example, firms producing notebook computers constitute an industry forming a coevolutionary pocket (Porter, 1990) of competitive advantage for the U. S. These firms compete against each other at the firm level on cost and quality. But they also compete at the “parts” levels of their computers—they compete over chip speed, mother board design, battery life, weight, mouse

technology, docking stations, expandability, and so forth. In this example, microstate analysis could be at the level of these parts, adaptive landscapes could be defined in terms of coevolutionary dynamics at the parts level, and there could easily be interdependencies among the organizational units producing the various parts. Given this, the coevolutionary pocket acts as a *complex adaptive system*, as do each of the firms within it. Such systems “contain many relatively independent parts which are highly interconnected and interactive and that a large number of such parts are required to reproduce the functions of truly complex, self-organizing, replicating, learning, and adaptive systems” (Cowan 1994, p. 2). The *NK* model is one member of the set of adaptive learning models that represent such systems computationally.

2.3 PARTICULARS OF THE *NK* MODEL

This section presents some definitional details of the *NK* model as received from the literature, particularly descriptions in Kauffman (1988, 1993), Kauffman and Weinberger (1989), Stauffer and Jan (1994), and Westhoff, Yarbrough, and Yarbrough (1996). I treat the received definitional material as the basic “data” of the analysis, presenting it in the various tables. This is not necessarily straightforward because of some vagueness and inconsistency in the descriptions.

In developing the *NK* model, Kauffman (1993: Ch. 2) assumes that the events by which chromosomes, genes, or organisms progress toward improved evolutionary fitness are myriad, obscure, complicated, and so poorly understood that to all intents and purposes their adaptive movements may be modeled most appropriately as randomly occurring microstates. McKelvey (1997, forthcoming) makes a parallel argument that individual agents’ decision events within firms appear stochastically idiosyncratic, drawing on the idiosyncrasy assumption that has increasingly formed the basis of the subjectivist, postpositivist, postmodernist movement in organization science (Silverman 1971, Burrell and Morgan 1979, Lincoln 1985, Reed and Hughes 1992, Weick 1995, Alvesson and Deetz 1996, Burrell 1996, Chia 1996), and thus also may be modeled as probabilistic microstates. McKelvey (1997) argues that organizational microstates can be similarly captured in terms of probabilistic master equations of the *Fokker-Planck* kind and nonlinear applications of statistical mechanics as a precursory platform for aggregate organizational analysis of Darwin’s background laws.

In the *NK* model,¹ N (agents or sites) are randomly assigned fitness contributions of 0.0 to 1.0, drawing from a uniform distribution. K stands for the *average* number of other agents that are interdependent with a given agent. Randomly assigned fitness contributions for the K effects

on fitness contributions are added to the fitness contribution by each agent. The final fitness contribution is an *average* of all fitness contributions pertaining to a given agent—the initial N -site contribution plus modifications due to fitness yield limitations imposed by each of the K interdependent agents (see Figure 1, which is a reproduction of Kauffman’s Figure 2.2, p. 42). Kauffman (1993: 52-54) makes a special point of noting that as K increases the average of the randomly assigned contributions recedes toward the mean of 0.5. This effect is how he models complexity catastrophe. That is, as the web of interdependencies (complexity) grows in size, the likelihood that a particular agent will achieve higher than average fitness tends toward zero. Thus, the basic Kauffman argument that under some conditions complexity effects dominate selection effects is represented by a simple averaging process and the effect of the Central Limit Theorem (1993: 53). One final point: Though Kauffman’s presentation of the basic *NK* model is analytical (mathematical) in Ch. 2, the “results” shown in the various tables in Chapter 2 come from computational simulations; the multiperiod *NK[C]* models (Ch. 6) are also computational simulations. At this “initializing” point I outline the basic definitions and parameters, solely based on Kauffman’s description.

>>>Insert Figure 1 about here<<<

To avoid confusion, I expand terminology somewhat. In Kauffman’s descriptions of the *NK* model he refers to three kinds of agents, loci, or sites, which I term ‘*N-sites*’, ‘*C-sites*’ and ‘*K-sites*’. He also defines three kinds of spaces, which I term ‘*N-strings*’, ‘*C-space*’ and ‘*K-space*’. Table 1 presents Kauffman’s definitions of these spaces. By way of illustration, consider the notebook firm example. To fit Kauffman’s typical sizing of the *NK* model, a notebook firm consists of $N = 24$ “parts,” that is, an N -string composed of various subunits responsible for various parts of the notebook primary and secondary value chains—one could also imagine this as a production value chain consisting of 24 workstations, each adding a component to the notebook. Each subunit is an N -site, with one agent presumed at each site.

>>>Insert Table 1 about here<<<

For “**nearest-neighbors**,” think of each agent as having $D = (A-1)N$ nearest-neighbors—or the 23 other workstations. Now imagine that it is possible that agents at one or another N -sites develop diverse value chain relevant managerial or other technologies, some of which might offer advantage to another site if copied. For example, suppose various subunits experiment with one or more of the 16 human resource practices discussed by Pfeffer (1994). We could add quality control improvements, computer-based design, control, and scheduling improvements, personnel selection approaches, equipment redesign approaches, and so forth, to the list of possible improvements at one subunit (site) that might be copied to advantage by other subunits.

¹ Hereinafter this paper focuses only on the *NK* model, eschewing the coevolutionary *NK[C]* aspects of the model.

In the NK model the “move rule” requires that for any given time period of the model, an agent can only copy one improvement idea from another agent. This is the so-called “**one-change neighbor rule**” (for Kauffman it is the “one-mutant neighbor” rule). Even with the limitation of the one-change rule, each agent has $N-1 = 23$ one-change neighbors. This leads to a C -space containing $A^N = 16,777,216$ C -sites, or what we might think of as imaginary nanoagents, each of which may choose to stay at its existing fitness level or adopt a higher level for each iteration of the model. For a small $N = 4$ N -string, such a space (or Boolean hypercube) is shown in Figure 2 (a reproduction of Figure 3c from Westhoff et al. 1996, p. 8)—it is Boolean in this instance because $A = 2$, meaning that all agents are limited to a binary 0 or 1 choice.

Kauffman takes the standard A^N combinatorial space, which is quite large even for $N = 24$, and further complicates it by allowing the possibility of a K -space at each site in the C -space. As Kauffman uses it, K , the complexity parameter, is allowed to affect the fitness yield at each C -space site with a uniform number of interdependencies, as depicted in Figure 3). Thus, a hypercube of fitness yields of the K interdependent agents appears at all of the N -sites, as shown in Figure 3. These depict the K -spaces. As Kauffman uses the K -space, the yields from the K interdependent agents are averaged in with the N -site fitness yield, to produce a modified N -site fitness yield (described in Figure 1—his Figure 2.2).

>>>Insert Figures 2 and 3 about here<<<

3. QUESTIONS ABOUT THE NK MODEL

As I have argued elsewhere (McKelvey 1997), philosophers of science have abandoned *logical positivism*, *logical empiricism*, Kuhnian *historical relativism*, and instead, based on more recent *historical realist* reanalysis of the evolution of science, have adopted a *scientific realist* approach along with the “*semantic conception of theories*” (Beth 1961, Suppe 1977, 1989; Lloyd 1988, Thompson 1989). The semantic conception philosophy divides science into two model centered activities: 1) testing whether the idealized model represents empirical phenomena; and 2) testing whether the theory accurately explains the behavior of the model. Accordingly, a model such as Kauffman’s NK model: 1) more or less accurately behaves according to his theory; and 2) more or less represents biological, or in my case, organizational microevolutionary phenomena. Thus, the NK model could: 1) be true to Kauffman but falsely represent reality; 2) accurately represent reality but show Kauffman’s theory to be false; 3) lie somewhere in between; or 4) fail with regard to both Kauffman and reality.

I start the critique with two givens: 1) Kauffman’s theory could be an accurate explanation of order and deserves to be tested with a model that accurately and fully represents the essence of the theory; and 2) The family,

NK models, is capable of accurately and fully representing *complex adaptive system* phenomena—biological and organizational—as generally described by various authors (including Kauffman) in Cowan, Pines, and Meltzer (1994) and Belew and Mitchell (1996). Thus I am a *friend of the court* regarding these two givens. They may not in fact be correct, but that is not my concern in this article. My question is: Does the model, as described in the various sources in the literature, actually do what it needs to do to accurately and fully represent the two givens? I will show in what follows that the answer to this question seems to be NO. In this regard it is possible that 1) the model, as programmed, behaves in a fashion true to the givens and is simply inaccurately described; or 2) the model is accurately described, behaves as described, and therefore is not true to the two givens.

In trying to unravel this process, key questions about the model are:

1. How the K interdependent neighbors are defined?
2. How complexity catastrophe is operationalized?
3. What happens if N is used as a measure of complexity?
4. What happens if fitness values are normalized by N ?
5. What happens if K effects are distributed uniformly across all sites?
6. What is the effect of allowing *infinitesimal* differences?
7. What determines the *correlation* and *ruggedness* of a landscape?

The first five are grouped in Section 3.1, labeled “Shaping the Landscape.” The last two are in Section 3.2, labeled “Making the Walk.” In Section 3.3 I draw on my analysis to explain some apparent anomalies in a couple of Kauffman’s computational findings tables.

3.1 SHAPING THE LANDSCAPE

This Section focuses on definitions of the adaptive landscape’s *fitness function* and how they speed up the onset of complexity catastrophe. Complexity catastrophe should be an emergent property stemming from K effects, as the NK model is conceived. But as Kauffman describes the model in various sources it appears that: 1) his operationalization of K and N arbitrarily speeds up the onset of complexity catastrophe; and 2) fitness peaks are arbitrarily flattened because the parameter affecting the distribution of K effects is not used, which also speeds up catastrophe.

3.1.1 HOW N AND K TUNE LANDSCAPES—SOME ISSUES

The first subsection works toward a more realistic selection of the random w_j fitness values. The next subsection suggests that subunit fitness is more a function of weakest links or bottle-necks than averaging as a means of separating N effects from K effects. The third subsection raises the possibility that N is not a realistic cause of complexity catastrophe. The final subsection focuses on issues pertaining to normalizing by N .

3.1.1.1 Problems in Operationalizing the K Interdependent Neighbors²

Several questions pertain to the operationalization of nearest-neighbors. **First**, do interdependent parts remain so throughout a genotype's or firm's adaptive walk? Presumably in rabbits or firms, subsystems do not randomly become independent or dependent—heart capacity and leg length fitness yields stay linked, as do linked parts of a value chain. And even if they did link or unlink, it probably would not be random or frequent. Yet in Kauffman's operationalization of the NK model the selection of which agents in the C -space should be included as interdependent one-change neighbors in the K -space randomly changes for each iteration, that is, each time fitness values are drawn from the K -space fitness table.

Second, Kauffman's description of how K fitness values w_j are selected and combined with the fitness level s_c of an agent at some combinatorial C -site are shown in Table 2. For each site s_c subject to K effects, imagine a table of fitness values. By Kauffman's theory, the table should appear as what I term the K -space, in which the fitness value of each of the interdependent agents is the same as that agent's initial C -site fitness value for the same iteration in the C -space—after all, for each iteration an agent can have only one initial fitness level.

>>>Insert Table 2 about here<<<

Third, Kauffman, in fact, doesn't model it this way. In Figure 1 one can see that Kauffman simply draws random fitness levels from a table of random numbers between 0.0 and 1.0—this being the simplest “random” approach. Other discussions of the NK model take alternative approaches to assigning the w_j values. Macken and Stadler (1995, p. 66) suggest a “circular” approach, which Kauffman also uses, that takes the w_j 's from alternating sides of the i^{th} position (thinking of the N -string as a circle). Their preferred approach is to choose the K contributions at random and not limited to the A^K K -space table—what Kauffman does in Figure 2. Westhoff, Yarbrough, and Yarbrough (1996, p. 7) draw their K values randomly from anywhere within the K -space table. Stauffer and Jan (1994, p. 212) base their function f_i on the elements of the A^K table relevant to the interdependent N -sites along the N -string for the given s_i choices for a given time period, which translates into taking as K values the fitnesses of all appropriate agents in the K -space. Ignoring the distribution parameter for the moment, this latter approach is illustrated in Figure 3, where uniform K -spaces are shown attached to each C -site vertex.

K -Space Selection. The restriction to the actual K nearest-neighbors of an N -site is critical, especially as K becomes small (Macken and Stadler 1995)—discussed further in Section 3.3—the results can be quite different. There is little positive argument for random drawing from the K -space or some other table given that the difference between the randomization and Stauffer and Jan approaches could be significant at small values of K —which in Kauffman's findings are usually the more telling values. Since there is no reason to expect that interdependencies with adjacent N -sites along the N -string are any more likely than those with more distant sites, I suggest picking K contributions by using the already random fitness values of the appropriate agents in a K -space defined by the interdependent N -sites—neither random nor circular.

3.1.1.2 Problems in Operationalizing Complexity Catastrophe

Some issues concern the operationalization of complexity catastrophe. Is the total fitness of a system the *average* of its interdependent parts or is it limited by the *weakest link* in the system of interdependent parts? Kauffman presents formula 1 to calculate fitness at each vertex of the C -space, shown in Figure 1:

$$W = \frac{1}{N} \sum_{i=1}^N f_i \quad 1$$

In fact, formula 2 better represents how Kauffman combines K effects with the s_c 's at a given C -site. For each c^{th} C -site there is a function, f_c , which the agent at that site attempts to maximize. This function results from the sum of the fitness contribution, s_c , of the agent at the c^{th} C -site (the agent has two fitness values, one for each of its two choices, given that $A = 2$), and the K interdependencies with other N -sites, according to the formula:

$$f_c = \frac{1}{1+K} \left(s_c + \sum_{j=1}^K w_j \right) \quad 2$$

where $c = 1 \rightarrow A^N$. As noted above, the w_j 's are selected from a K -space table having A^K , in this case, 2^K entries. Kauffman models the complexity catastrophe simply as a movement toward central tendency as a result of averaging of K effects in with C -site values. A key question is, How do the fitness levels of the K interdependent agents affect the C -site agent?

First, why an averaging of K effects approach? Why not a *bottle-neck* or *weakest link* approach? Whether a rabbit is running from a fox, or a manager is trying to speed up or improve the quality of a set of workstations, is it average fitness or the weakest link or lowest fitness that tells the tale? Possibly it could go either way. But presumably the rabbit is constrained by the weakest of its subsystems—heart, lungs, metabolism, muscle contraction rates, motor control, eyes, and so forth, and the manager is also limited by the weakest part of the value chain. Given

² A question is, for an example 2^{24} C -space, are or should randomly drawn fitness contribution values be limited to the first decimal place? Use of three decimal places would reduce ties if there is no K effect averaging, since, if $A^N = 2^{24} = 16,777,216$, there could be over 750,000 ties for each of the $A = 2$ choice spaces.

this, once an agent becomes subject to interdependencies, it is limited by the weakest link of interdependent K effects, not the average.³ Absent empirical evidence to the contrary, the weakest link assumption appears more accurate than the averaging assumption—hence the lowest w_j should enter the formula not the average of the w_j 's.

Second, should the K effect be averaged in with the s_c fitness value at a given C -site? 1) If an s_c fitness value is higher than the w_j value, it is arguably more likely that once an agent is subject to interdependencies, its fitness achievement is limited more by the weakest link rather than an average of the weak link and its own higher fitness. Thus, if a fast machine is suddenly made interdependent with a slow machine the result is that the joint production line is limited by the bottleneck machine, or say, if the rabbit breaks its leg, it is limited by its speed on three legs, not an average of its new “three-leg” subsystem and its other subsystems. 2) If the s_c fitness value is lower than the w_j 's, then surely an agent would be able to improve its fitness as fast as possible—as fast as the model allows—until the fitness limit imposed by the K effects is reached. 3) Can higher K effects actually speed up fitness improvement at a particular C -site? Does a superior heart system enhance the fitness yield from, say, short legs on a rabbit? Do faster machines on a production line enhance the speed of the slowest machine? Probably not. On balance it seems more realistic to posit that low K -site fitness yields decrease the yield at a high fitness C -site, but high fitness yields at K -sites do not increase yields at a low fitness C -site—it is asymmetric. Consequently, the following formula appears to make more sense:

$$f_c = \min_{j=1 \rightarrow K} w_j \quad 3$$

Weakest Link. If s_c is the weakest link, it may improve its fitness yield up to the next weakest link, that is the minimum w_j . If s_c is higher than the weakest link, the interdependency means that its yield is lowered to that of the w_j bottleneck. The bottom line is that once s_c becomes interdependent with other agents—the K effect—its yield moves to that of the weakest link, which is the lowest w_j . One may expect that in this framing K effects take hold quite gradually with the result that when $K = N-1$ all⁴ s_c yields are defined by the lowest w_j —in contrast to Kauffman's method where the largest K impact is when $K = 2$, which happens because on a probabilistic basis as soon as a higher and low yield are averaged the result is a quick move toward the mean. Clearly, Kauffman's approach speeds up the onset of complexity catastrophe.

³ An interesting analysis of why the bottleneck approach is to be preferred over the averaging approach is given, in novel form, in a book titled, *The Goal* (E. M. Goldratt and J. Cox, 1992, 2nd rev. ed., North River Press, Great Barrington, MA).

⁴ My use of “all” here stays within Kauffman's use of a uniform application of K effects to all C -sites. In Section 3.1.2 I take up the average vs. uniform distribution of K effects.

And then he adds in an additional speed-up by treating N as a measure of complexity as well, discussed next.

3.1.1.3 Problems If N Is Also a Measure of Complexity

Several questions relate to N as a measure of complexity. **First**, in Table 3 one may see that Kauffman makes a number of statements suggesting that in his view N acts as a measure of complexity independently of K . Bonner (1988, p. 98), a biologist, draws on Herb Simon (1962) to define complex systems as “a large number of parts that interact.” Cowan (1994, p. 2) describes complex adaptive systems as containing “many relatively independent parts which are highly interconnected and interactive.” Burton and Obel (1995, p. 70) define organizational complexity as “the degree of horizontal, vertical, and spatial differentiation.” Finally, Maynard Smith and Szathmáry (1995) define complexity in terms of the length of relevant DNA⁵ instructions for making an organism—the more instructions, the more complex. The foregoing definitions consistently connect “a large number of parts” with “interdependency.” The issue is whether a large number of parts, absent interdependency, is sufficient to cause complexity. Bonner (1988, p. 98) also says that, “the most interesting thing is that parts are often not only numerous, but frequently they are different in their structure and function,” a point also implicit in the Burton and Obel definition—differentiation plus number of differentiated units.

>>>Insert Table 3 about here<<<

Second, Kauffman holds that a sufficient number of (independent) parts (leaving degree of differentiation undefined) will cause a complexity catastrophe. He argues that the n^{th} additional part, after many, offers a fitness advantage diluted by $1/N$ —with the result that above a certain critical value, the fitness improvement of the n^{th} part offers no selective advantage, hence a complexity catastrophe. This seems a questionable argument in both organisms and firms. In recent years various genes have been discovered that have clearly unfavorable survival impact—for example some genes cause sickle cell anemia, immune deficiency, or cancer. In this case $1/N = 1/100,000+$ genes and yet the n^{th} gene offers a significant survival disadvantage. For notebook computer firms with $N = 24+$ parts, it seems quite possible that any one of a number of additional n^{th} subunits in their value chains (with $n > 24$) could assure a significant survival disadvantage, that is, not be diluted by the $1/N$ factor. On this basis, both real world examples and theoretical definitions of complexity do not readily favor Kauffman's approach.

⁵ Only a “varyingly small proportion of the DNA codes for anything” (Maynard Smith and Szathmáry 1995, p. 5) that biologists are presently aware of—estimated at as little as 1% (Calladine and Drew (1992, p. 14), which is still some 100,000+ genes in the human genome (Davis 1990).

Third, the questions then emerge, 1) Does an increase in N cause increased complexity? and 2) Is there a dilution of fitness advantage or disadvantage stemming from the n^{th} increment of an increasing N ? If (2) doesn't hold there is no catastrophe based on an increase of independent parts, and if (1) doesn't hold the catastrophe, if it happens, is not complexity caused. The foregoing definitions of complexity all suggest complexity is independent of mere size—it is the linking of differentiated parts that counts. The most prudent approach would be to limit “complexity” catastrophe to K effects and take pains to assure that increasing N does not operationally create catastrophe in the model. A fall-back position could be to operationally define N caused complexity in terms of the averaging effect of the Central Limit Theorem—as Kauffman does—while operationally defining K caused catastrophe in terms of the weakest link approach, as I propose. Thus, the effects remain separate.

N not complexity. Is complexity—as it might act to subvert the effects of selection—the result of the number of subunits N or whether it is the result of K interdependencies. By Kauffman's approach complexity is both N and K . By my analysis a better operationalization of Kauffman's theory suggests that 1) N , by itself, is not a measure of complexity; and 2) that the overall fitness of a genotype or firm, that is, an N -string, also should not be based on normalization by N , the subject of the next section.

3.1.1.4 Problems in Normalizing by N ?

Should N be normalized? Kauffman's Figure 2.2 (Figure 1 here) is actually ambiguous as to the use of formula 1. He says the formula is to “...define the *fitness of an entire genotype*...” (1993, p. 42; his italics). From Figure 1 it is clear that W is the row sum of $K+1$ agent fitnesses, with one row per C -site or vertex on the hypercube. But “fitness of an entire genotype” could also easily refer to fitness across all the N -sites, that is, across the $N = 3$ sites, shown as Kauffman's 2.2a. Thus, there are really two kinds of “*entire genotype*,” 1) the one-change neighbor defined genotype at each vertex (C -site) of the C -space; and 2) the genotype defined by summing across the N -sites. Though Kauffman's formula confuses the two, it would appear that he accepts normalizing by N for both kinds and—in fact, N , represented by the s_i fitness values at each N -site, gets averaged twice: 1) As shown in formula 2, Kauffman's sums the s_i value in with the w_j 's and then divides by $1+K$ —since for Kauffman $K = N-1$ at all vertices, formulas 1 and 2 are the same and the denominator is $1+K = N$. Thus each s_i takes the first move toward the mean on this step. 2) As noted above, Kauffman says complexity catastrophe results directly from an increase in N . This is what leads to formula 1 in which all s_i 's along the N -string are also divided by N —the second move toward the mean. Consequently the s_i 's are in effect divided by N twice.

There are three problems with Kauffman's approach: 1) As discussed earlier, normalizing by N should be

avoided because it is less appropriate than a “weakest link” approach, and because his averaging confounds N with K ; 2) normalizing by N doesn't make sense when $K+1$ is less than N ; and 3) normalizing is unnecessary at the C -site vertices anyway.

Referring to point 2, what happens when $K+1$ is smaller than N ? Obviously the denominator cannot remain as N —it has to become $1+K$, with the 1 representing the s_i fitness value and then the K , as in formula 2 above. In order for normalization to work, K fitness contributions cannot be *in addition* to the N i^{th} element fitness contributions—otherwise the denominator would be $N+K$ for normalization. Further, fitness cannot be a function of s_i plus the K w_j 's, as Kauffman (1988) implies on p. 133, and shows in my Figure 1, because if K is smaller than $N-1$ then normalization also would not work unless the denominator were $1+K$ instead of N . For Kauffman's described model to work and stay normalized, the fitness of an N -string is always the sum of N elements. This means that K effects have to be added to some or all of the i^{th} elements comprising the string so that the total number of “things” summed does not exceed N , as in formula 4, where i goes from 1 to N :

$$f_i = \frac{1}{2} \left(s_i + \frac{\sum_{j=1}^K w_j}{K} \right) \quad 4$$

Referring to point 3, there is the question of what actually happens at a particular C -site in C -space. This is important because when $N = 24$, C -space = $A^N = 2^{24} = 16,777,216$ combinations or C -sites. And for Kauffman, N typically equals 24. A “combination” at a C -site is by definition different from a neighboring combination by only one element of a binary choice (1 or 0), assuming $A = 2$, which is almost always the case. This means that the fitness contribution, going from one C -site to another is a reflection of a change at only a single i^{th} N -site. Thus, the fitness contribution from all the other N -sites stays the same. And it means that there is no reason for “normalizing” by $1/N$ at a given C -site because fitness changes from one C -site to another (from one combination to another), are never the result of a change at more than a single i^{th} N -site—they are by definition already of equal weight.

Misapplied Normalization. I conclude that formula 1 (2.1 in Kauffman's book) is misapplied by Kauffman: 1) It is a confusion of both the w_j effects at an i^{th} N -site (in terms of the N -string or c^{th} C -site in terms of the C -space) and the effects of function f_i across the i ($i = 1 \rightarrow N$) N -sites. In my view, formula 2 best describes Kauffman's definition of the function f_i and formula 1 describes how Kauffman calculates the *total* fitness, W , of the entire N -string of a whole entity—genotype, value chain, or firm. Section 3.1.1.2 suggests that a weakest link approach is

better than a normalizing by N approach in general. This holds for calculating s_i at N -sites, and for the calculation of W for the entire N -string. The present analysis suggests that normalizing is irrelevant at C -sites. It also shows that when $K < N-1$ normalization by N is mistaken. In fact Kauffman is vague on what happens if $K < N-1$, since his Figure 2.2 (Figure 1 here) illustration is for $K = N-1$, and he uses the same K effect at all vertices. I now turn to what happens if $K < N-1$ and if the K effects are distributed around an average of K , as Kauffman's theory calls for.

3.1.2 HOW UNIFORM DISTRIBUTION OF K EFFECTS PANCAKES THE LANDSCAPE

Table 4 quotes Kauffman on the *distribution* of K effects. If the first quote is taken as the ruling description of what he thinks should happen, then what happens in Figure 1 (Kauffman's Figure 2.2) is misleading. The figure shows K effects to be equal at $K = 2$ for all N -sites. If it had shown $K = 0$ or 1 or 2 at various i^{th} N -sites, it could have shown evidence of a distribution effect with K averaging $K = 1$ —some sites with zero K effects, some with one, and some with two. Thus Kauffman's textual description is inconsistent with what is shown in Figure 1.

>>>Insert Table 4 about here<<<

First, the most immediate observation about Kauffman's approach is that the "independent" fitness values of the s_c agents are very quickly subordinated to the K effects. For example, with $N = 24$ and $K = 4$, even though 83% of the agents are independent the fitness values of all agents are moved substantially toward the mean since at each site the s_c and the four w_j 's are summed and normalized by $1/5$. Consequently the K effects, the central tendency effect, and ultimately the complexity catastrophe as Kauffman operationalizes it, dominate the landscape very rapidly. Furthermore, as shown in Table 4, Kauffman also says that only the N and K parameters are important, the distribution (of K effects) parameter being unimportant. This is inevitable—if the K effects are applied uniformly how could a "distribution" parameter have any effect?

Second, consider fitness peaks as scattered throughout the C -space. By Kauffman's analysis we may expect the proliferation of local optima as follows:

$$\#LOpt = \frac{2^N}{N+1} \quad 5$$

Thus, the number of local optima becomes very large as N increases. Needless to say, local optima, as the highest s_c 's in C -space, become the adaptive targets of the c^{th} agents. It follows that if an agent keeps adopting the higher fitnesses of nearest-neighbors it might get trapped if, at a particular s_c^{th} site, all nearest-neighbors showed lower fitness yields. The problem with Kauffman's approach is that by allowing the K effects to apply with equal weight across all vertices of the C -space, as shown in Figure 1, all local optima or fitness peaks are systematically flattened.

If this happens, and the "move rule" remains unchanged (which is to say that the steps taken at each move cover the same amount of adaptive ground), then the flattening of all fitness peaks vastly speeds up the rate at which peaks are attained. Further, the attained local optima are far below the theoretically achievable global optimum fitness—in a typical Kauffman case, with a large A^N , one might expect the highest unflattened global optimum to approach 1.0.

Third, since Kauffman's mean is 0.5, the foregoing logic also indicates that all valleys are systematically raised while peaks are flattened. Thus, rather than K effects tuning the landscape toward an increasingly rugged or jagged landscape of many steep-sided peaks, the picture is, instead, one of a pancaked geography where peaks and valleys disappear into a broad flat plain. As one can also imagine, supposing we paired off the highest and lowest peaks and valleys across the landscape, taking the averages could instantly produce a complexity catastrophe in one step, given Kauffman's approach. Of course, it is unlikely that high and low s_c 's would be systematically averaged, but you get the idea that the tendency is in this direction. In contrast, if K effects are not uniformly applied, then some peaks will be flattened, but other peaks will remain at their initial higher fitness levels, possibly even approaching a maximum global optimum near 1.0. Clearly the "tuning," toward buckling, heaving, or ruggedness of the landscape is vastly different depending on whether all or some fitness peaks are flattened.

Fourth, if the distribution parameter is properly installed in the model, as per Kauffman's description, it surely would have a significant effect. The distribution is supposed to characterize "how the K genes are distributed among the N genes" (1993, p. 42). I conclude that a formula such as the following more accurately does what Kauffman calls for in his textual description:

$$f_c = \frac{1}{2} \left(s_c + \frac{\sum_{j=1}^{k_d} w_j}{k_d} \right) \quad 6$$

where s_c is the fitness contribution of the c^{th} agent's choice, and w_j is the k_d fitness effects bearing on the c^{th} agent, with k_d a number drawn from a uniform distribution and with a mean of K , but ranging within the limits of 0 and $N-1$. By this formula, K effects are inserted only at the specific c^{th} C -site where an agent's fitness yield is moderated by the K effects its site is dependent upon when making an adaptive choice, since all other C -sites in C -space do not change, as illustrated in Figure 4. Further, the number of K effects *averages* around whatever the value of K is, say for example $K = 7$ with $N = 24$. Thus, at each c^{th} C -site, the number of K effects may range from 0 to 23, depending on a randomly drawn number, k_d , subject to the constraint that the average across all c^{th} C -sites equals 7. With this

approach some C -sites could have no K effects and some peaks might never flatten, even though $K = N-1$, because k_d could be zero here and there. Needless to say, formula 6 follows Kauffman's *averaging* approach rather than the *weakest link* approach I have suggested. If the latter were followed, the formula would be:

$$f_c = \min_{j=1 \rightarrow k_d} w_j \quad 7$$

when $s_c > w_j$. If $s_c < w_j$ then s_c increases via the normal process of adopting a higher fitness from a nearest-neighbor—till a local optima or the minimum w_j is reached.

>>>Insert Figure 4 about here<<<

Distribution is significant. I conclude that if the distribution parameter is properly designed 1) some valleys probably will remain very low, some peaks probably will remain very high, and some adaptive walks could remain quite long—from a low valley to a high peak, even though most other peaks and valleys are pancaked; 2) the onset of complexity catastrophe will occur much slower than in Kauffman's present application of the model; and 3) the distribution parameter is a critically important parameter to the operation of the model and whether it is present or not has a significant effect on how the model behaves.

Overall, my discussion in Section 3.1 shows that Kauffman's application of the NK model arbitrarily speeds up the onset of the complexity catastrophe because of:

1. Adoption of an averaging of K effects approach rather than a weakest link approach;
2. Use of N as a complexity parameter which is also allowed to induce complexity catastrophe;
3. Inappropriate normalization of N .
4. Failure to use a K effects distribution parameter.

3.2 MAKING THE WALK

Section 3.1 focused on how Kauffman's definition of the fitness function arbitrarily speeds up the onset of catastrophe. Ironically, this Section focuses on approaches Kauffman takes in defining a *move rule* that undermine the tuning effects of the complexity driven fitness function. I first consider what happens if the move rule is defined to allow *infinitesimal steps* in the adaptive walk. Then I consider the effects of this on whether and how a landscape might become more or less *rugged* due to complexity effects.

3.2.1 WHAT IS THE EFFECT OF INFINITESIMAL DIFFERENCES

In physics and biology "nature" defines the move rule. Thus, "the flipping of a single spin might be a reasonable model for the elementary dynamic step. In molecular biology, the move rule is defined by the elementary mutational events: point mutations, insertions and deletions, and transpositions" (Macken and Stadler 1995, p. 59). In nature "steps" in an adaptive walk come with a "naturally" defined length—a discrete change in energy or

fitness greater than some threshold amount. This contrasts with Kauffman's operationalization of steps. He says, "I shall count a neighboring genotype fitter than another even if the fitness difference is infinitesimal. Ignoring the actual fitness differences amounts to a kind of *rank-ordering* of the fitness of all possible genotypes" (1993 p. 44–45; italics added). Absent any threshold limit, and given a binary choice of "copy the nearest neighbor or not," an agent's choice is essentially the same, whether the difference in fitness contribution between two C -sites is 1.0 (the maximum) or 10^{-15} —and possibly beyond the accuracy limitation of the computer. Since Kauffman treats steps as ordinal ranks, steps are all the same in the NK model whether they are near 1.0 or 0.0.

First, consider the effect of no threshold limiting the "smallness" of the infinitesimal steps. There is an obvious problem in that without a threshold there is the risk that random effects in the model could be as much or more due to rounding or other computational error effects than designed effects. This is a chronic issue in computational modeling (Morrison, 1991).

Second, if there is no threshold, and infinitesimal differences are allowed in the move rule, then choice rankings are never curtailed, no matter how small the averaged fitness differences are. That is, if the length of steps in the walk are allowed to shrink down to infinitesimal increments as the peaks and valleys are pancaked, then the number of steps in a walk from a filled in valley to a flattened peak could stay the same as the number of steps from a low valley to a tall peak. And, if steps are ordinalized into ranks, as Kauffman calls for, the number of rank increments also stays the same as peaks and valleys are pancaked toward the flat plain of the mean. Furthermore, if steps are changes in rank as is shown in Figure 5, then the number of ranks between the start and finish of a walk stays the same. Since the purpose of the spin glass part of the NK model is to focus on questions pertaining to rates of adaptation, or the rate at which the walks in the model converge toward adaptive peaks as the landscape is tuned by complexity effects, the consequence of infinitesimal steps is that tuning the landscape by making it rugged or pancaked may not have any effect on the rates of adaptive walks reaching peaks. Most importantly, given the foregoing points, the effects of flattening and the Central Limit Theorem are negated—meaning that in the model the number of steps, or rates of reaching peaks remains unchanged. The consequence of this is that allowing infinitesimal steps means that adaptive walks are not slowed down or shortened and, thus, **complexity catastrophe cannot affect rate of adaptation.**

>>>Insert Figure 5 about here<<<

Third, suppose Kauffman's NK model does have a threshold to curtail infinitesimal steps. What then? Remembering that he says the Central Limit Theorem is the basis of the complexity catastrophe (p. 52), the averaging effect of increasing K clearly pancakes the

valleys and peaks of fitness contributions toward the mean—0.5 in his case. Now, if a threshold exists and if K becomes large enough, then at some point virtually all differences between fitness contributions (1) are squeezed into the vicinity of the mean; (2) are closer to the mean than the threshold step size; and (3) fitness improvements are impossible since none would be larger than the threshold step size. By this logic, complexity catastrophe, in the model, could be more a function of step size and reaching the (hidden) threshold designation than it is a function of some particular size of N or K . Thus, the fact that a particular size of N or K appears to cause complexity catastrophe happens only because the system reaches the point where step increments fall below the threshold. And finally, since the threshold designation is of fixed step size, and step size invariably shrinks as N or K increase, complexity catastrophe is, thus, inevitable as the model is described and would happen immediately once the threshold was reached.

Fourth, Kauffman says his results are robust against uniform, Gaussian, or U-shaped distributions (1993, p. 45). This is inevitable. Because of averaging, the uniform, Gaussian, or U-shaped random selections will all end up sharply peaked around the central tendency any way—the greater is K the more sharply peaked the distributions will appear. More importantly, because of allowing metric infinitesimal differences with nonmetric binary—and thus distribution free—fitness improvement rankings (as in Figure 5), how can cardinal distributions have any effect, no matter what shape they have? They will all behave as if they are uniform—the nonmetric ranking “distribution free” effect. Because of allowing infinitesimal steps, the number of choices and number of suboptimal peaks is never less than what is probabilistically allowed by the size of the C -space. The consequence of mixing cardinal steps with ordinal fitness choices is that the various adaptive walk rates to be studied by the NK model are invariant no matter what the value of N or K or what shape are the fitness distributions. What appear to be the effects of tuning are really due to metric and nonmetric mixing, infinitesimal steps, or a possible threshold effect.

Yet his models show differences—the models do behave differently when parameters are changed—suggesting that widespread random effects from infinitesimal steps and computational errors are in fact avoided—presumably by some kind of threshold designation. I conclude that: (1) A threshold level is in the model and that infinitesimal differences are cut off at some point. Presumably the step threshold is higher than the error threshold of the computer being used. (2) If a threshold designation is included, as just suggested, it follows that adaptive walks in the model could be as much stopped by threshold effects as by hill-climbing to the tops of suboptimal peaks. (3) This effect could undermine Kauffman’s claims about the apparent number of peaks and the rapidity at which peaks are approached. (4)

Adaptive moves could be as much due to threshold effects as landscape effects, a significant problem since landscape features may then not have the correlation and ruggedness effects Kauffman claims, my next topic of concern.

3.2.2 WHAT ABOUT THE CORRELATION AND RUGGEDNESS OF LANDSCAPES?

Kauffman’s definitions of *correlation* and *ruggedness* are shown in Table 5. Thus, on p. 45 he says correlation is due to similarity of contributions and ruggedness is due to uncorrelation, but on p. 47 he says uncorrelation is due to new random fitness contributions.

>>>Insert Table 5 about here<<<

First, the definitions of correlation and ruggedness are problematic, as follows. The “new random fitness contributions” Kauffman refers to above, come from the imposition of the K effects drawn from the K -space table of fitness values on each iteration of the model, and in his approach (shown in Figure 1) applied uniformly to all C -space agents. When $K = N-1$ and the averaging process is used (Kauffman’s approach), the fitness values from the K -space are squeezed very close to the mean—since they are essentially divided by N twice, as discussed in Section 3.1.1.4. This means they are very similar in value, being almost all very near the mean of 0.5. In this situation the model produces new random fitness contributions that are very similar. Consequently, the fitness values are at the same time both 1) very similar and thus correlated and 2) newly random and thus uncorrelated. Since they can’t be both correlated and uncorrelated at the same time, Kauffman’s definition fails.

Second, if infinitesimal differences are allowed, I do not see how the $K = 0$ model is any more correlated than the $K = N-1$ model, since a ranking is always possible, unless Kauffman’s model stops at some threshold limiting the smallness of infinitesimal differences. He says $K = 0$ models are more correlated because “neighboring points have nearly the same fitness contribution” (p. 45). But if infinitesimal differences are allowed, and as I argue in Section 3.2.1, the metric differences are ordinalized into a nonmetric ranking of fitness choices (as shown in Figure 5), then it makes no difference in a binary choice ranking space whether the differences are large or small—any improvement an agent might adopt is the same—a one rank move. It follows that, since complexity tuning is based on shifting from a correlated to uncorrelated landscape, and given that the step size shrinks along with the pancaking due to increasing N or K , the landscape never becomes more or less correlated—since a rank is a rank. In fact, because of averaging, the similarity of fitness values actually increases as N and K are increased. Thus, in terms of Kauffman’s definition of correlation (Table 5), $K = N-1$ landscapes are more correlated than $K = 0$ landscapes—opposite of what Kauffman concludes. This is corroborated by Macken and

Stadler (1995: 66). Kauffman's tuning mechanism fails at best and at worst is opposite of what he thinks.

Third, does the ruggedness of landscapes really change as K goes from 0 to $N-1$? Kauffman says $K = N-1$ landscapes are "extremely rugged" (1993, p. 46). While averaged fitnesses obviously recede toward the global mean of 0.5 as K increases, I do not see how ruggedness—defined as an increase in number of suboptimal peaks, and steeper sides—increases. I agree that the number of local optima would increase—if K were infinite the fitness value of every C -site would equal 0.5 and there would be A^{infinite} number of local optima, each with a value of 0.5—a perfectly flat landscape. But rather than increasing ruggedness, as Kauffman claims (1993: 47), K , by making the landscape quite flat actually reduces ruggedness—the f_c values are increasingly similar (Kauffman's definition of correlation) even though they are still random (Kauffman's definition of ruggedness).

Fourth, a related comment here is that for $K = 0$ landscapes Kauffman says that "all other genotypes are suboptimal and can climb to the global optimum via fitter neighbors and that all one-mutant neighbors have nearly the same fitness" (1993, p. 45). This seems false on the face of it. It is clear from Kauffman's own Figure 2.1 (shown here as Figure 1) and also from a similar $N = 4$ N -string in Westhoff et al. (1996), that even a small hypercube may have trapping suboptima. You can confirm this by drawing your own set of 16 random "fitness" numbers and randomly placing them on the hypercube vertices. By this view it appears that ruggedness is highest when $K = 0$ —a few unflattened suboptima are likely and, absent K effects the range of probable fitness values is maximized at 0.0 to 1.0.

Fifth, do the foregoing points affect the approach taken by Macken and Stadler (1995)? They propose that "correlation length" is the most rigorous measure of ruggedness, where correlation length is a function of the size of K ; as K increases, correlation length decreases, thus tuning the landscape toward more ruggedness. A landscape is rugged "...if it has many local optima, if the length of adaptive and gradient walks are short compared to the maximal distance in the landscape, and if the correlation [length] between nearest neighbors is small" (1995, p. 69). Macken and Stadler define correlation length as "...how far neighbors must be separated in sequence space in order that they be essentially uncorrelated" (1995: 47). The questions are, how does K decrease correlation length and does reduced length mean less fitness value separation or less randomness? If infinitesimal differences and binary choices are allowed, correlation length, defined as level of similarity from a "walk" perspective, will not change as K increases. Peaks and valleys are pancaked but the steps shrink concurrently, hence still the same number of steps—the nonmetric distance measured as number of ranks does not change. And, since step length will not change, all fitness values will be one rank step apart no matter what K is and,

therefore, equally dissimilar and unpredictable from a "walk" perspective. Hence, for Macken and Stadler's approach to correlation and ruggedness to hold, steps cannot be infinitesimal or ranks. This puts their approach at odds with Kauffman's—a point confirmed by the fact that their move rule presumes fixed "natural" step size.

Distribution and Ruggedness. The NK model has the steepness effect Kauffman needs only if the K contributions are averaged in with some N -site contributions, depending on k_d , leaving some contributions at other C -sites at their initial higher or lower levels. This happens only if C -site fitnesses are not the result of averaging across all i^{th} N -sites, as Kauffman does in Figure 1. Steepness, then, is really a function of chance, depending on how many peaks and valleys remain separate from K effects. This in turn depends on how many C -sites have $k_d = 0$ —meaning no K effects. Here is one example of where the distribution parameter could have a significant effect by allowing some number of C -sites to remain unaffected by K effects while many other C -sites are pancaked. Since hill climbs and number of steps counted typically start at the lowest fitness C -sites (valleys), the chance availability of peaks and valleys unpancaked by K effects could significantly affect the "emergent ordering" properties of landscapes, a possibility not available to Kauffman's operationalization. If my proposed distribution parameter is used, there is always a probability of some peaks and valleys not being pancaked, the number of possible unflattened suboptima would grow along with N , and they will increasingly stand out relative to the peaks flattened by the K effects. Therefore the ruggedness of a landscape, defined as the number of unpancaked peaks and valleys relative to the number of pancaked peaks and valleys, will increase as K increases. This sets up a basis for expecting a tuning effect from K , which what the NK model needs.

Overall my discussion in Section 3.2 suggests:

1. That Kauffman's allowing of infinitesimal steps in the move rule works to negate the effects of K in creating a complexity catastrophe;
2. If a lower bound threshold to infinitesimal steps is introduced, the catastrophe occurs inevitably and more suddenly because of the threshold effect as K is increased;
3. Kauffman's definitions of correlation and ruggedness, requisite to the tunable landscape concept, fail;
4. Landscapes, by Kauffman's operationalization become more correlated than rugged as K increases—opposite of what he thinks happens;
5. The use of a proper distribution parameter and the consequent protection of some C -sites from K effects appears to create the correlation of increased ruggedness with increased K effects—what Kauffman needs if the NK model is to properly represent his theory.

3.3 CONFUSION IN KAUFFMAN'S TABLES

Table 6 reproduces Kauffman's Figure 2.4 and Tables 2.1 and 2.2.⁶ Here is where the definitional and operational confusion hits home. **First**, in his Figure 2.4 he shows that N produces complexity catastrophe—to be expected since he normalizes by N (1993, p. 52). But in Table 2.1 his simulation shows that when $K = 0$ an increasing N does not produce catastrophe. Both figure and table focus on “expected fitness of local optima. If N is normalized the results of Table 2.1 can't be correct. If N is not normalized then Table 2.1 is at odds with his description on pages 52–54 of how N is treated.

>>>Insert Table 6 about here<<<

Second, one of Kauffman's primary results is the computational discovery, shown in Table 2.1, that mean fitnesses of local optima are higher when $K = 2, 4,$ or 8 than when $K = 0$. This is critical since this is his main evidence that ruggedness increases with an increasing K . One might reasonably ask how this can happen? By Kauffman's descriptions and as illustrated in Figure 1 (his Figure 2.2), averaging of the s_i 's in with the w_j 's begins as soon as $K > 0$. Probabilistically, as soon as averaging begins the fitnesses of all s_i 's, including local optima, should shift toward the mean of 0.5. How could those above the mean possibly increase? Why should the Central Limit Theorem take hold only when $K > 8$? Since the table reports mean results from 100 simulated landscapes, the chances of some local optima having miraculously escaped an averaging process seem minuscule. These questions lead to two additional inquiries stemming from the two different approaches taken in Tables 2.1 and 2.1.

Third, it is bad enough that the w_j 's don't immediately progress toward the mean, following the Central Limit Theorem. Much more problematic, however, is how to explain why local optima increase when $K = 2, 4,$ or 8 ? How can the process of averaging in 2, 4, or 8 w_j 's with a local optimum s_i increase the value of such a local optimum? It appears impossible, given Kauffman's described methods (Figure 1). It gets worse. Table 2.1 is based on the so-called “circular” method⁷ in which the $K/2$ nearest-neighbors on each side of the local optimum s_i are used to define the w_j fitness values. Given this, the reality is that the $K/2$ nearest-neighbor s_i 's on either side—using the circular method—of the local optima s_i 's have to be lower—otherwise they wouldn't be on either side of a local optima. Thus, when averaged in with an s_i the latter has to decrease. For Table 2.1 to be correct the averaged w_j 's would collectively have to be higher than the local optimum s_i —an impossibility. But since the s_i 's studied in

Table 2.1 are all local optima, and the “circular” method actually forces the local optima toward the mean faster than the random method, Table 2.1 should show local optima values moving toward the mean even faster than in Table 2.2. Given that Kauffman's operationalization in the text, and as illustrated in Figure 1, shows that the w_j 's have to decrease, the fact that they go up in Table 2.1 has to be based on some kind of error or a programming approach inconsistent with the definitions and operationalizations. I don't see any other explanation. Consequently, his conclusion that small values of K increase the fitness values of local optima and therefore ruggedness, surely is questionable.

Even if a more correct distribution parameter, such as the k_d approach I discuss in Section 3.1.2, is used, the results don't figure. With the k_d approach some s_i 's would remain unaffected by the averaging approach. Thus, most s_i 's would be shifted toward the mean, but some local optima s_i 's would escape averaging, therefore relatively they would be increasingly higher than those flattened. It is unlikely that this effect would be more pronounced at lower values of K because both unflattened and flattened local optima would increase geometrically as K increases. Thus the ruggedness idea is preserved. But I see nothing in the k_d distribution approach that would work to actually increase values of local optima s_i 's.

Third, if Kauffman were drawing randomly from the fitness table, as he does in Figure 1, the random drawings might produce 2 or 4 or 8 w_j 's whose average fitness is higher than the local optima s_i . Possible, but remote! What are the odds that an average of the randomly drawn nearest-neighbor w_j 's would be higher than a local optima s_i ? In fact, he presents simulations by this approach in Table 2.2. And guess what—the $K = 2, 4,$ or 8 increase in local optima fitness disappears (except for a 0.01 increase at $K = 4$), a fact Kauffman fails to note in his discussion—he says “Table 2.2 shows similar results...” (1993, p. 56), when in fact they are fundamentally different. **This table indicates that landscapes do not increase in ruggedness as K increases**—undermining the main contention of the “tunable landscape” adherents, Kauffman, Weinberger, and Perelson. Instead landscapes linearly become more similar (correlated) as K increases—what my analysis would predict! This is confirmed in Kauffman's Tables 2.3 and 2.4 (not reproduced here), which show no difference between the circular and random draw approaches.

4. CONCLUSION: COUNTERACTING EFFECTS

Could a head-on collision be in sight between selectionist and complexity theories in organization science? Kauffman (1993) suggests two conditions when the creation of “order” via Darwinian natural selection processes may be thwarted—an event he terms *complexity catastrophe*: 1) selection forces are not strong enough to separate a novel, presumably more fit, population from the

⁶ The simulations also draw on Kauffman, Weinberger, and Perelson 1988.

⁷ In this drawing of K values, Kauffman draws from the table of K fitness values by successively taking a value from one side and then the other side of the local optima s_i —extending around the “circular” N -string as K approaches $N-1$.

general population long enough for it to develop distinct novel properties—the U. S. postwar example; and 2) competitive selection is strong but the many suboptimal fitness peaks are flattened such that even if a population achieves a peak it does not develop properties distinct from the general population—the business school example. He develops the *NK* model as a means of tuning adaptive landscapes so as to explore the effects of increasing amounts of complexity on the adaptive search for improved fitness.

I have presented a fairly detailed critique of the methods Kauffman uses to operationalize the functioning of the model. The review focuses on six questions: 1) How do subunit interdependencies (*K* effects) affect agents responsible for various parts of a firm? 2) Is a weakest link approach better than averaging *K* effects? 3) Is complexity simply a function of the number, *N*, of subunits in a firm? 4) Since averaging is the means of operationalizing complexity catastrophe, should the parameter *N* be normalized? 5) What is the impact of allowing move steps to reach infinitesimal sizes? And 6) What if ruggedness decreases as *K* increases instead of increasing as Kauffman expects? In addition I conclude from an analysis of a couple of key tables that his computational finding that ruggedness increases for small values of *K* must be an error, given his operational approach.

The bottom line of the Section 3.1 critique states that Kauffman's operationalizations of *N* and *K* act to **arbitrarily speed up the onset of complexity catastrophe**. In opposition, the bottom line of Section 3.2 suggests that his allowing of infinitesimal steps **undermines the tuning effect of *K***. Given that both his fitness function speed-up and his infinitesimal step move rule operate in direct opposition to each other, interpreting published results of Kauffman's *NK* model is surely problematic. It would seem that all of his conclusions that in one way or another start from the *NK* model platform should be reevaluated. This affects most of the chapters in his book. Most specifically Kauffman's operationalization of *K* is such that *K* produces more correlated landscapes and less rugged ones—opposite to what Kauffman thinks happens and what his theory calls for—thus fundamentally altering the role of *K* effects as modeled in his book.

Note that my critique focuses on the methods Kauffman uses to operationalize the *NK* model. I do NOT critique Kauffman's fundamental contribution, which is the insight that complexity effects may thwart selection effects. It is clear that many leading scholars have taken his work as is a major contribution to biology and complexity theory, not the least of which are Nobel Laureates Manfred Eigen, Murray Gell-Mann, and Philip Anderson along with Brian Arthur, Stephen Jay Gould, Richard Lewontin, John Maynard Smith, and Lewis Wolpert. I have argued that Kauffman's theory fits firms much better than organisms.

The *NK* model may be operationally rescued by introducing the following modifications:

1. Basing the selection of *K* effects on the proper interdependent sites in the *K*-space table.
2. Shifting from an averaging of the w_i 's in with the s_i 's to the use of the *weakest link* approach.
3. Decoupling the increase in *N* from the creation of complexity catastrophe.
4. Avoiding the normalization of the s_i 's by *N*.
5. Using the distribution parameter such that the *K* effects are not uniformly applied to all s_i 's but instead are designed to *average* around the *K* value—which Kauffman actually calls for in his text.
6. Stopping the use of infinitesimal steps in the move rule. Steps are better based on “naturally” defined step sizes. And this is more consistent with the use of nonmetric rankings.
7. The correct installation of the distribution parameter would properly operationalize the notion of ruggedness and result in ruggedness increasing with *K*, which is what Kauffman calls for and needs for the *NK* model to result in a *tunable* landscape.

Given these modifications, organization scientists are in a position to use the *NK* model to better understand 1) the role of complexity in the fitness-based selection of firms, 2) rates at which firms might move toward improved fitness, and 3) the likelihood of their being trapped on suboptimal peaks. The *NK* model also allows organization scientists to 4) explore the relative importance of micro and macro evolutionary effects, as has recently been shown by Carley and Svoboda (1996), using a simulated annealing modeling approach. Given a properly working tunable landscape method, of particular interest is 5) the chance to study the role of varying amounts of subunit interdependency in the evolutionary movement of firms and their ultimate adaptive health. Finally, 6) the *NK* model offers promise of studying the approach firms should take toward cross-functional integration—should they have more or less?

Given postmodernist assumptions that behavioral and process events in firms appear as stochastic idiosyncratic Brownian motion (Chia 1996, McKelvey 1997), organizations may very well follow paths taken by other sciences in shifting their investigations down to microlevel analyses and studying how explainable order emerges from apparently random events. Adaptive learning models, such as Kauffman's *NK* model, become important tools for this kind of organization science. As with any other complicated method, computational models can be as much of a black art as factor analysis ever was. They are no better than the operational care that goes into them. And what comes out is not “magic”—it is simply a result of how parameters are operationalized. As organization scientists move toward a more model centered concept of their science, modeling skills and mistakes will become increasingly important. This article points to some advantages of adaptive learning models as well as some of the dangers of computational modeling.

REFERENCES

- Alchian, A. A. (1950), “Uncertainty, Evolution and Economic Theory,” *Journal of Political Economy*, 58, 211–222.

- Aldrich, H. (1979), *Organizations and Environments*, Englewood Cliffs, NJ: Prentice Hall.
- Alvesson, M. and S. Deetz (1996), "Critical Theory and Postmodernism Approaches to Organizational Studies," in S. R. Clegg, C. Hardy and W. R. Nord (Eds.), *Handbook of Organization Studies*, Thousand Oaks, CA: Sage, 191–217.
- Anderson, P. and M. L. Tushman (1990), "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly*, 35, 604–633.
- Ashkenas, R., D. Ulrich, T. Jick and S. Kerr (1995), *The Boundaryless Organization*, San Francisco: Jossey-Bass.
- Baum, J. A. C. (1996), "Organizational Ecology," in S. R. Clegg, C. Hardy and W. R. Nord (Eds.), *Handbook of Organization Studies*, Thousand Oaks, CA: Sage, 77–114.
- Baum, J. A. C. and J. V. Singh (Eds.) (1994), *Evolutionary Dynamics of Organizations*, New York: Oxford University Press.
- Belew, R. K. and M. Mitchell (Eds.) (1996), *Adaptive Individuals in Evolving Populations: Models and Algorithms*, Proceedings Vol. XXVI, Santa Fe Institute, Reading, MA: Addison-Wesley.
- Beth, E. (1961), "Semantics of Physical Theories," in H. Freudenthal (Ed.), *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*, Dordrecht, Netherlands: Reidel, 48–51.
- Blau, P. M. and W. R. Scott (1962), *Formal Organizations*, San Francisco: Chandler.
- Bonner, J. T. (1988), *The Evolution of Complexity*, Princeton, NJ: Princeton University Press.
- Burrell, G. (1996), "Normal Science, Paradigms, Metaphors, Discourses and Genealogies of Analysis," in S. R. Clegg, C. Hardy and W. R. Nord (Eds.), *Handbook of Organization Studies*, Thousand Oaks, CA: Sage, 642–658.
- Burrell, G. and G. Morgan (1979), *Sociological Paradigms and Organizational Analysis*, London: Heinemann.
- Burton, R. M. and B. Obel (Eds.) (1995), *Design Models for Hierarchical Organizations: Computation, Information, and Decentralization*, Boston, MA: Kluwer.
- Calladine, C. R. and H. R. Drew (1992), *Understanding DNA: The Molecule and How it Works*, London: Academic Press.
- Carley, K. M. and D. M. Svoboda (1996), "Modeling Organizational Adaptation as a Simulated Annealing Process," *Sociological Methods and Research*, 25, 138–168.
- Chia, R. (1996), *Organizational Analysis as Deconstructive Practice*, Berlin: Walter de Gruyter.
- Choi, C. (1996), "Detecting Chaotic Behavior in an Organizational Population," presented at INFORMS College of Organization Science, November, Atlanta, GA.
- Cowan, G. A. (1994), "Conference Opening Remarks," in G. A. Cowan, D. Pines and D. Meltzer (Eds.) (1994), *Complexity: Metaphors, Models, and Reality*, Proceedings Vol. XIX, Santa Fe Institute, Reading, MA: Addison-Wesley, 1–4.
- Cowan, G. A., Pines and D. Meltzer (Eds.) (1994), *Complexity: Metaphors, Models, and Reality*, Proceedings Vol. XIX, Santa Fe Institute, Reading, MA: Addison-Wesley.
- Depew, D. J. and B. H. Weber (1995), *Darwinism Evolving: Systems Dynamics and the Genealogy of Natural Selection*, Cambridge, MA: Bradford/MIT Press.
- Dimancecu, D. (1992), *The Seamless Enterprise: Making Cross Functional Management Work*, New York: Harper Business.
- Donaldson, L. (1985), *In Defence of Organization Theory*, Cambridge, UK: Cambridge University Press.
- Donaldson, L. (1995), *American Anti-Management Theories of Organization*, Cambridge, UK: Cambridge University Press.
- Eigen, M. (1971), "Self-Organization of Matter and the Evolution of Biological Macro-molecules," *Naturwissenschaften*, 58, 465–523.
- Eisenhardt, K. M. (1989), "Making Fast Strategic Decisions in High-Velocity Environments," *Academy of Management Journal*, 32, 543–576.
- Eisenhardt, K. M. (1991), "Better Stories and Better Constructs: The Case for Rigor and Comparative Logic," *Academy of Management Journal*, 16, 620–627.
- Eldredge, N. (1995), *Reinventing Darwin*, New York: Wiley.
- Etzioni, A. (1961), *A Comparative Analysis of Complex Organizations*, New York: Free Press.
- Fischer, K. H. and J. A. Hertz (1993), *Spin Glasses*, New York: Cambridge University Press.
- Fisher, R. A. (1930), *The Genetical Theory of Natural Selection*, Oxford, UK: Clarendon Press.
- Fontana, W. and P. Schuster (1987), "A Computer Model of Evolutionary Optimization," *Biophysical Chemistry*, 26, 123–147.
- Fontana, W., W. Schnabl and P. Schuster (1989), "Physical Aspects of Evolutionary Optimization and Adaptation," *Physical Review*, A40, 3301–3321.
- Foster, J. (1987), *Evolutionary Macroeconomics*, London: Allen & Unwin.
- Friedman, M. (1953), *Essays in Positive Economics*, Chicago: University of Chicago Press.
- Galbraith, J. R., E. E. Lawler III and Associates (1993), *Organizing for the Future: The New Logic for Managing Complex Organizations*, San Francisco: Jossey-Bass.
- Glinow, M. A. Von and S. A. Mohrman (1990), *Managing Complexity in High Technology Organizations*, New York: Oxford University Press.
- Goldratt, E. M. and J. Cox (1992), *The Goal* (2nd rev. ed.), Great Barrington, MA: North River Press.
- Gould, S. J. (1989), *Wonderful Life: The Burgess Shale and the Nature of History*, New York: Norton.
- Gowdy, J. M. (1994), *Coevolutionary Economics*, Boston, MA: Kluwer.
- Graham, M. A. and M. J. LeBaron (1994), *The Horizontal Revolution: Reengineering Your Organization Through Teams*, San Francisco: Jossey-Bass.
- Haas, J. E. and T. E. Drabek (1973), *Complex Organizations*, New York: Macmillan.
- Hall, R. H., J. E. Haas and N. J. Johnson (1967), "Organizational Size, Complexity, and Formalization," *American Sociological Review*, 32, 903–912.
- Hannan, M. T. and G. R. Carroll (1992), *Dynamics of Organizational Populations*, New York: Oxford University Press.
- Hannan, M. T. and J. Freeman (1977), "The Population Ecology of Organizations," *American Journal of Sociology*, 83, 929–984.
- Hannan, M. T. and J. Freeman (1989), *Organizational Ecology*, Cambridge, MA: Harvard University Press.
- Hirshleifer, J. (1977), "Economics from a Biological Viewpoint," *Journal of Law and Economics*, 20, 1–52.
- Johann, B. (1995), *Designing Cross-Functional Business Processes*, San Francisco: Jossey-Bass.
- Kauffman, S. A. (1988), "The Evolution of Economic Webs," in P. W. Anderson, K. J. Arrow and D. Pines (eds.), *The Economy as an Evolving Complex System*, Proceedings Vol. V, Santa Fe Institute, Reading, MA: Addison-Wesley, 125–146.
- Kauffman, S. A. (1993), *The Origins of Order: Self-Organization and Selection in Evolution*, New York: Oxford University Press.
- Kauffman, S. A. and E. D. Weinberger (1989), "The N-K Model of Rugged Fitness Landscapes and its Application to Maturation of the Immune Response," *Journal of Theoretical Biology*, 141, 211–245.
- Kauffman, S. A., E. D. Weinberger and A. S. Perelson (1988), "Maturation of the Immune Response via Adaptive Walks on Affinity Landscapes," in A. S. Perelson (Ed.), *Theoretical*

- Immunology I*, Proceedings Vol. II, Santa Fe Institute, Reading, MA: Addison-Wesley.
- Kaufman, H. (1975), "The Natural History of Organizations," *Administration and Society*, 7, 131–149.
- Kramer, F. R., D. R. Mills, P. E. Cole, T. Nishihara and S. Spiegelman (1974), "Evolution *in vitro*: Sequence and Phenotype of a Mutant RNA Resistant to Ethidium Bromide," *Journal of Molecular Biology*, 89, 719–736.
- Leonard-Barton, D. (1995), *Wellsprings of Knowledge*, Boston, MA: Harvard Business School Press.
- Levinthal, D. (1995), "Adaptation in Rugged Landscapes," unpublished manuscript, Wharton School, University of Pennsylvania, Philadelphia, PA.
- Lincoln, Y. S. (Ed.) (1985), *Organizational Theory and Inquiry*, Newbury Park, CA: Sage.
- Lloyd, E. A. (1988), *The Structure and Confirmation of Evolutionary Theory*, Princeton, NJ: Princeton University Press.
- Macken, C. A. and P. F. Stadler (1995), "Evolution on Fitness Landscapes," in L. Nadel and D. L. Stein (Eds.) *1993 Lectures in Complex Systems*, Lectures Vol. VI, Santa Fe Institute, Reading, MA: Addison-Wesley, 43–86.
- Maguire, S. (1996), "A Rugged Landscape Framework for Understanding Configurations," presented at INFORMS College of Organization Science, November, Atlanta, GA.
- Maynard Smith, J. (1970), "Natural Selection and the Concept of a Protein Space," *Nature*, 225, 563.
- Maynard Smith, J. and E. Szathmáry (1995), *The Major Transitions in Evolution*, Oxford, UK: Freeman/Spektrum.
- McKelvey, B. (1982), *Organizational Systematics: Taxonomy, Evolution and Classification*, Berkeley, CA: University of California Press.
- McKelvey, B. (1997), "Quasi-Natural Organization Science," *Organization Science*, forthcoming.
- McKelvey, B. (forthcoming), "Complexity vs. Selection Among Coevolutionary Microstates in Firms: Complexity Effects on Strategic Organizing," *Comportamento Organizacional E Gestão*.
- McKelvey, B. and P. Konstantina Kiouisis (1996), "Complexity vs. Selection Among Coevolutionary Firms: A Field Effect," presented at INFORMS College of Organization Science, November. Atlanta, GA.
- Miller, D. and P. H. Friesen (1984), *Organizations: A Quantum View*, Englewood Cliffs, NJ: Prentice-Hall.
- Nelson, R. R. and S. G. Winter (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Belknap/Harvard.
- Pfeffer, J. (1994), *Competitive Advantage Through People*, Boston, MA: Harvard Business School Press.
- Porter, M. E. (1990), *Competitive Advantage of Nations*, New York: Free Press.
- Porter, M. E. (1991), "Towards a Dynamic Theory of Strategy," *Strategic Management Journal*, 12, 95–117.
- Reed, M. and M. Hughes (Eds.) (1992), *Rethinking Organization: New Directions in Organization Theory and Analysis*, London: Sage.
- Rommel, G., J. Kluge, R. Kempis, R. Diederichs and F. Brück (1995), *Simplicity Wins*, Cambridge, MA: Harvard Business School Press.
- Rosenkopf, L. and M. L. Tushman (1994), "The Coevolution of Technology and Organization," in J. A. C. Baum and J. V. Singh (Eds.), *Evolutionary Dynamics of Organizations*. New York: Oxford University Press, 403–424.
- Silverman, D. (1971), *The Theory of Organisations*, New York: Basic Books.
- Simon, H. A. (1962), "The Architecture of Complexity," *Proceedings of the American Philosophical Society*, 106, 467–482.
- Singh, J. V. (Ed.) (1990), *Organizational Evolution*, Menlo Park, CA: Sage.
- Sorenson, O. (1996), "The Complexity Catastrophe and the Evolution of the Computer Workstation Industry," presented at INFORMS College of Organization Science, November. Atlanta, GA.
- Spiegelman, S. (1971), "An Approach to the Experimental Analysis of Precellular Evolution," *Quarterly Review of Biophysics*, 4, 36.
- Stauffer, D. and N. Jan (1994), "Size Effects in Kauffman Type Evolution for Rugged Fitness Landscapes," *Journal of Theoretical Biology*, 168, 211–218.
- Suppe, F. (1977), *The Structure of Scientific Theories* (2nd ed.), Urbana, IL: University of Illinois Press.
- Suppe, F. (1989), *The Semantic Conception of Theories and Scientific Realism*, Urbana, IL: University of Illinois Press.
- Thompson, P. (1989), *The Structure of Biological Theories*, Albany, NY: State University of New York Press.
- Tushman, M. L. and E. Romanelli (1985), "Organizational Evolution: A Metamorphosis Model of Convergence and Reorientation," in B. M. Staw and L. L. Cummings (Eds.), *Research in Organizational Behavior*, 7, 171–222.
- Tushman, M. L. and P. Anderson (1986), "Technological Discontinuities and Organizational Environments," *Administrative Science Quarterly*, 31, 439–465.
- Van Maanen, J. (1989), "Some Notes on the Importance of Writing in Organization Studies," in Harvard Business School Research Colloquium, 27–33, Boston: Harvard Business School.
- Weick, K. E. (1979), *The Social Psychology of Organizing* (2nd ed.), Reading, MA: Addison-Wesley.
- Weick, K. E. (1995), *Sensemaking in Organizations*, Thousand Oaks, CA: Sage.
- Weisbuch, G. (1993), *Complex Systems Dynamics: An Introduction to Automata Networks*, (trans. S. Ryckebusch), Lecture Notes Vol. II, Santa Fe Institute, Reading MA: Addison-Wesley.
- Westhoff, F. H., B. V. Yarbrough and R. M. Yarbrough (1996), "Modelling Complexity, Adaptation and Evolution," *Journal of Economic and Organizational Behavior*, 29, 1–25.
- Winter, S. G. (1964), "Economic 'Natural Selection' and the Theory of the Firm," *Yale Economic Essays*, 4, 225–272.
- Wright, S. (1931), "Evolution in Mendelian Populations," *Genetics*, 16, 97–159.
- Wright, S. (1932), "The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution," *Proceedings of the Sixth International Congress on Genetics*, 1, 356–366.

dFigure 1 Kauffman's Figure 2.2

Figure 2 Westhoff, Yarbrough and Yarbrough's Figure 3

Figure 3 Kauffman's Boolean Hypercube: Uniform K -Spaces with $K = 3$

Figure 4 Boolean Hypercube: Distributed K -Spaces with $K = 2$

Figure 5 Kauffman's Boolean Hypercube with Ranks

Table 1
(a) Basic NK Model Definitions— N -strings and C -space

-
1. Typically, for Kauffman, $N = 24$, mostly, though N ranges up to 96 in some models, K ranges from 0 to $N-1$, and $A = 2$ though it could range higher.
 2. Agents choosing between two states, that is $A = 2$ —choose to copy the state of a nearest “one-change neighbor” only if the neighbor has a higher fitness contribution—the “nearest-neighbor” modeling approach.
 3. An *adaptive landscape* is termed a ‘**C-space**’, which has $D = (A-1)N$ as the number of dimensions, and also the number of one-change neighbors of each agent.
 4. The combinatorial C -space contains A^N sites (loci or agents), termed ‘**C-sites**’ s_c , in this case, 2^{24} C -sites, where c goes from $1 \rightarrow A^N$. That is, the space includes all possible differences among A^N one-change neighbors.
 5. Each “ C -site” imaginary *nonoagent* attempts to maximize fitness in terms of an ‘**N-string**’ of $i = 1 \rightarrow N$ nearest-neighbors, termed ‘**N-sites**’, each showing a choice state with an attached fitness value. Thus, the N -string (at a C -site) may differ from a nearest-neighbor on any one, but only one, of its i^{th} N -sites. Kauffman terms this the one-mutant neighbor difference between two genotypes—here a one-change neighbor difference between two subunits. In the entire C -space there are thus c one-change neighbor N -sites (referred to as C -sites), with c going from $1 \rightarrow A^N$.
 6. A very small C -space is depicted in Westhoff et al.’s Figure 3 (1996), reproduced here as Figure 2. It is an $N = 4$ N -string creating a space of A^N vertices or C -sites with c going from $1 \rightarrow 16$.
 7. Fitness values (contributions) s_c of each C -site are drawn randomly from a uniform distribution, with contributions from 0.0 to 1.0, and mean of 0.5, suggesting only the 1st decimal place is used. According to Figure 1 also, contributions are only to the first decimal place.
-

(b) Basic NK Model Definitions— K -Space

-
1. At each hypercube vertex, i.e., C -site, the fitness yield of the c^{th} C -site in question may be a function of the fitness of some number, K , of other interdependent N -sites. This is also a combinatorial space, termed ‘**K-space**’, having A^K ‘**K-sites**’. Kauffman shows this space in Figure 2.2 (1993: 42), reproduced below as Figure 1.
 2. Contributions for K are also randomly assigned and range between 0.0 and 1.0, also uniformly distributed around a mean of 0.5.
 3. There are several methods of determining which fitness values in the K -space are actually used, (as will be discussed). Kauffman uses both the “random” and “circular” methods.
 4. The actual fitness values in the K -space, for my purposes, will not be drawn via the random or circular methods. Instead, K -space fitness values are based on the values of the interdependent “one-change neighbors of the C -space as it stands at the beginning of an iteration.
-

Table 2. Creation of the K Fitness Values

-
1. “The fitness contribution of the allele at the i^{th} locus depends upon itself (in other words, whether it is a 1 or 0) and the alleles, 1 or 0, at K other loci, hence upon $K+1$ alleles.”
 2. “Therefore the fitness contribution w_i of the i^{th} locus is specified by a list of random decimals between 0.0 and 1.0, with 2^{K+1} entries (Figure 2.2)” (I have labeled this the K -space).
 3. “The fitness contribution of each allele at each gene in the context of the K other genes which impinge on that gene must be specified.”
 4. “For each gene, its fitness contribution w_j is generated by random assignment on the 2^{K+1} allele combinations of the $K+1$ genes which impinge upon it.”
 5. “Having assigned the fitness contributions, we may now define the *fitness of an entire genotype* as the average of the contributions of all the loci (all from 1993: 42):

$$W = \frac{1}{N} \sum_{i=1}^N w_i \quad 1$$

7. “In the $K = N-1$ limit, the fitness vector w_i for each gene i , $i = 1, 2, \dots, N$, is a function of all $K+1 = N$ genes. Consider any initial genotype among the $2N$ genotypes with two alternative alleles at each locus. Alteration of the allele at any single locus affects each of the N genes, since that alteration alters the combination of the $K+1 = N$ alleles which bear on the fitness of each gene. In turn, this second alteration alters the fitness contribution of each gene to a different randomly chosen value between 0.0 and 1.0. The fitness of the new one-mutant neighbor genotype is therefore a new sum of N random decimals between 0.0 and 1.0” (1993: 46).
 8. “I can find out the fitness of the entire entity by summing the fitness contribution of each part in the context of itself and the K which bear upon it. For simplicity, I shall then normalize fitness...by dividing this sum by N ” (1988: 133).
-

Table 3. N as a measure of Complexity

-
1. “[T]he fitness of any genotype is the average of the contributions from all N loci” (p. 52).
 2. “[A]s N increases, the fitness values of attainable local optima decrease toward 0.5” (p. 53).
 3. “[A]s the number of genes N increases, the accessible optima dwindle in height toward the average unselected fitness in the space of genotypes.... Thus inexorably in these landscapes, adaptive walks terminate on poorer ‘solutions’ as N increases” (p. 53).
 4. “In the limiting case where the parts are independent, each part typically makes, to the overall function of the system, a contribution which *decreases* in relative importance as the total number of parts in the system increases” (p. 36).
 5. “This contention of mutational and selective forces leads...to a complexity catastrophe when the number of parts exceeds a critical value” (p. 36).
-

Table 4. Distribution of K Effects a c^{th} C-site*

-
1. “Let K stand for the *average* number of other loci which epistatically affect the fitness contribution of *each* locus. Thus the two main parameters of the NK model are the number of genes N and the average number of other genes K which epistatically influence the fitness contributions of *each* gene” (my italics).
 2. A third ‘*distribution*’ parameter “*characterizes how the K genes are distributed among the N genes*” (my italics).
 3. “...only the N and K parameters matter.
-

* (All from 1993: 41-42)

Table 5. Definitions of “Correlation” and “Ruggedness”

Kauffman (1993), in fact, offers two definitions of correlation:

-
1. He says that when $K = 0$, highly correlated fitness contributions result—“A *smooth* landscape is one in which neighboring points in the space have nearly the same fitness value” (p.45). And at the bottom of the page (45-46) he says, “This is clear because the fitness of one-mutant neighbors cannot differ by more than $1/N$. Therefore, for large N , the fitness of one-mutant neighbors is very similar”
 2. “At the opposite extreme, a *maximally rugged* landscape is one in which the fitness values are entirely uncorrelated” (p. 45; his italics).
 3. Alternatively, on page 47 he says, “The fitness of the new one-mutant neighbor genotype is therefore a new sum of N random decimals between 0.0 and 1.0. Therefore, the new fitness contribution is *entirely uncorrelated* with the old fitness contribution” (his italics)
-

Table 6**Kauffman’s Local Optima Computational Findings**

- a) Kauffman’s Figure 2.4
- b) Kauffman’s Table 2.1
- c) Kauffman’s Table 2.2