Connectivity, Extremes, and Adaptation: A Power-Law Perspective of Organizational Effectiveness

Max Boisot and Bill McKelvey

Journal of Management Inquiry 2011 20: 119 originally published online 13 January 2011
DOI: 10.1177/1056492610385564

Journal of Management Inquiry
20(2) 119–133

DOI: 10.1177/1056492610385564
http://jmi.sagepub.com

Connectivity, Extremes, and Adaptation: A Power-Law Perspective of Organizational Effectiveness

Max Boisot[1] and Bill McKelvey[2]

Abstract

Managers are often required to respond in adaptive ways to the threats and opportunities presented by rare, extreme outcomes. Given these, management scholars frequently face a stark choice: say something useful to practitioners using narratives in which dramatic effects are often achieved at the expense of academic rigor or maintain the latter by sacrificing practitioner relevance. Recent developments in complexity science offer a new perspective. The article distinguishes between the simplicities achieved by reductionism (equilibrium, law-like equations, linearity, and predictability) and the complexity triggered by initiating "butterfly events"—nonlinearity, scale-free causes, and power laws (PLs). Schema formation and adaptation within Gaussian and PL ontologies are framed in terms of Ashby's law of requisite variety. Variety perceived to be requisite is sensitive to the type of ontological assumptions that are made. PL approaches to management inquiry focusing on rank/frequency distributions, fractal structures, and scale-free dynamics are outlined.

Keywords

power laws, adaptation, Ashby's Law, complexity, connectivity

In a world characterized by turbulence and uncertainty, managers are often required to respond in adaptive ways to the threats and opportunities presented by extreme, discontinuous, and hence rare events—not only negative ones such as the Asian financial meltdown, 9/11, the 07/08 liquidity crisis, and so on, but also positive ones such as Microsoft, the Internet, Google, and the hot economies of China and India. Management research, however, earns its spurs as a science by studying nicely behaved linear trends, normal distributions, and well-behaved variances—take a sample of 25 database studies published in *Academy of Management Journal* or *Strategic Management Journal* and you will get the point.

Although popular management books tend to attract their readers by focusing on rare extremes—of success, of failure, of leadership, and so on—academics attempt to analyze and interpret these in terms of averages and variances, publishing in journals little read by practitioners. Management scholars thus face a stark choice: (a) either say something that practitioners want to hear but do so through narratives in which rhetorically dramatic effects are achieved at the expense of academic rigor or (b) maintain academic integrity by sacrificing perceived practitioner relevance (Bennis & O'Toole, 2005; Pfeffer & Fong, 2002). In the latter case, management research strategies that build incrementally on testable law-like regularities—strategies that in science date back to Kaplan (1964) and Popper (1935)—often turn out to deliver little to the practitioner in need of guidance. Why is it so? Management researchers seek inspiration from intellectual traditions established in the disciplines they look up to—physics, economics, and sociology (Friedman, 1953; Lewin, 1951; White, 1963). These tend to share a Gaussian perspective of the world, one built on assumptions of independent (and normally) and identically distributed events (*i.i.d.*)—that is, on *atomism*. Where events can be made to conform to *i.i.d.* assumptions, the Gaussian perspective leads to robust and useful results—as for example in the case of controlled experiments in which selected variables can effectively be isolated and subjected to independent, repetitive testing. However, too often the circumstances in which practitioners might look to researchers for guidance do not warrant *i.i.d.* assumptions. The challenges they face typically involve anomalies and events that cannot be isolated from the context in which they are embedded, that cannot be repeated, and whose magnitude lies beyond the range

[1]Escuela Superior de Administración y Dirección de Empresas (ESADE), Barcelona, Spain
[2]UCLA Anderson School of Management, Los Angeles, CA

Corresponding Author:
Bill McKelvey, UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095-1481
Email: mckelvey@anderson.ucla.edu

of expectations that are consistent with a Gaussian perspective. In this perspective, they are then treated as "outliers." In this article, we focus on the conceptual challenges that anomalies and extreme events pose for management inquiry. At present, these remain undertheorized and are therefore dealt with for the most part descriptively—through case studies and other forms of narrative discourse. We believe that management inquiry into such anomalies and events now needs to go beyond case-study narratives. But how?

With their more biologically oriented concept of organization as it applies to living systems, the sciences of complexity can help us better understand the emergence of extreme outcomes (Bak, 1996; Gell-Mann, 2002). The complexity perspective has only recently captured the attention of management researchers. It takes emergent organization to be an attribute of the elements of an aggregation in interaction (Kauffman, 1993), that is, an emergent property of the aggregation's various *connectivities* (Holland, 1995). In addition to the *old simplicity* of reductionism, equations, linearity, and predictions of traditional physics, Gell-Mann (2002), it offers us the *new simplicity* of tiny initiating events (what we will call "*butterfly events*"),[1] nonlinearity, scale-free causes, and power laws (PLs).[2]

In what follows, in Section 1, we first introduce relevant complexity concepts and briefly compare Gaussian and PL ontologies. In Section 2, we frame the challenge of organizational adaptation within these ontologies in terms of Ashby's (1956) Law of Requisite Variety. It turns out that the variety perceived by agents to be requisite is sensitive to the ontological assumptions that they make and shapes their strategic adaptive schemas. In Section 3, we start outlining the basic elements of PL-based management inquiry. A conclusion follows this section.

From Gauss to Pareto: Basics of PL Science

A Gaussian world is populated by stable objects that are *i.i.d.* This world, inhabited by most of neoclassical economics (Lawson, 1997; Mirowski, 1989) and mainstream sociology (Snijders & Bosker, 1999), is characterized by a stable mean and finite variance (Greene, 2002). Here, outliers are either ignored or somehow linearized so that their effects are minimized (Greene, 2002). In this world, whether or not they foreshadow extreme outcomes, outliers neither shape managerial expectations nor do they drive organizational adaptation. Often, however, assumptions of independence and additivity must be replaced by assumptions of connectivity, multiplicativity, interaction, and positive feedback. These dynamics give rise to PL distributions that have "long tails," potentially infinite variance, unstable means, and unstable confidence intervals. Three Mile Island, Microsoft, Enron, 9/11, the Internet, the liquidity crisis of 07/08, and so on, all

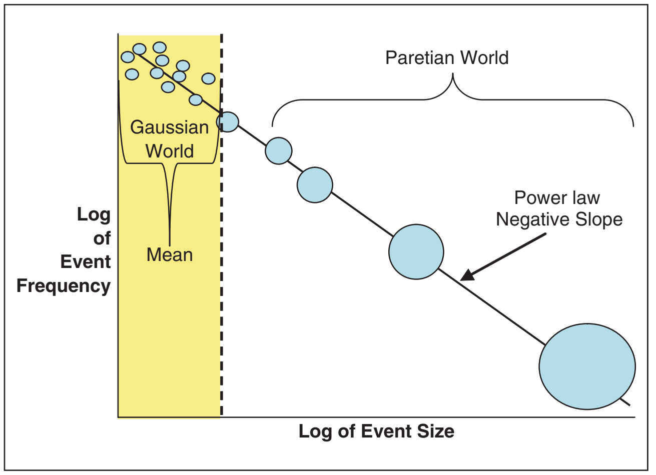

Figure 1. Stylized Pareto distribution on log-log scale

point to the possibility that the "normal" world that we believe ourselves to be inhabiting is but a small and stable corner of a much broader and more complex one in which we are actually immersed. Indeed, the most "interesting" possibilities for action, those potentially incurring the highest costs and generating the highest returns, often reside in the tails of a PL distribution (Kirchgaessner & Kelleher, 2005).

As stylized in Figure 1, a well-formed rank/frequency (Pareto) distribution plotted in terms of double-log scales appears as a PL distribution—an inverse sloping straight line. PLs often take the form of rank/size expressions such as $F \sim N^{-\beta}$, where F is frequency, N is rank (the variable), and β, the exponent, is constant. The market capitalization of the 30 largest U.S. publicly traded firms is PL distributed, as are various industries (Glaser, 2009). Moreover, Stanley et al. (1996) find that the size distribution of manufacturing firms in the United States is PL distributed and exhibits a fractal structure, as does Axtell (2001). Fractals, like cauliflowers, are self-similar and exhibit the same causal dynamics at multiple levels of resolution. Barabási (2002) shows how networks in the physical and biological worlds, and social capital networks in the social world, are fractally structured, yielding a rank/frequency of network connections that is PL distributed, with many social "loners" at one extreme and a single, highly connected "star" at the other. Scale-free theories explain the appearance of fractal structures. Though scalability lies at the core of complexity science (Brock, 2000), scale-free theories have only recently gained attention (Andriani & McKelvey, 2009b; Newman, 2005).

Until recently, the default assumption of many scientists was that most natural phenomena are distributed according to the bell curve or *normal* distribution. However, PLs are now being discovered in such a great number and variety of phenomena that some scientists are calling them "more normal than 'normal'." In the words of mathematician Walter Willinger and his colleagues, "The presence of [PL] distributions in data obtained from complex natural or engineered systems should be considered the norm rather than the exception" (Mitchell, 2009, p. 269).

In his now classic book *How Nature Works*, Per Bak (1996) explains PL distributions by looking at how sandpiles build up: Falling grains of sand are allowed to slowly accumulate in a pile. Eventually, the sandpile becomes high enough and its slope steep enough to trigger sand avalanches of varying sizes. These restore stability to the slope. The steepness of the slope depends on two elements: (a) gravity and (b) the sharp, irregular shape of the individual sand grains. Take away gravity and there is no force causing the grains to slide down past each other—call the influence of this force the *tension* effect. Take away the irregular shape of the individual grains, and they become frictionless, unable to resist the downward force exerted by gravity—somewhat like smooth M&M peanuts, they will then scatter, unable to cohere enough to build up a pile. Call the influence of friction the *connectivity* effect. Bak observed that sand grain movements varied from the frequent but barely perceptible movement of a few isolated grains to the rare but gigantic avalanches in which thousands of sand grains move in unison. The size and frequency of sand grain avalanches is PL distributed (Bak, Tang, & Wiesenfeld, 1987).

The nonlinear tensions and connectivities that lead to extreme outcomes (the largest avalanches) are key elements of complexity science. Bak (1996) labeled the results of the nonlinear interplay of tension and connectivity "*self-organized criticality*"—when the force of gravity encounters the friction-induced resistance of irregularly shaped grains of sand, these will move so as to maintain the sandpile's slope in a precarious state of equilibrium. The rate and volume of sand moving at any given instant is (a) nonlinear, (b) unpredictable, and (c) occasionally productive of extreme events. In addition to the normally distributed phenomena characterizing much of physical science and described by Gaussian statistics (data points assumed to be *i.i.d.*), then, researchers have discovered an ever-increasing number of phenomena—from physical to biological to social—that are best described by attributes (a)–(c) above. These attributes are associated with "tiny initiating events" (what Holland [2002] terms "*small 'inexpensive' inputs*" or "*lever point phenomena*," p. 29) and result in rank/frequency distributions of outcomes that are PL distributed and best explained by scale-free theories. Andriani and McKelvey (2007, 2009b) list 40 physical and biological phenomena, and 101 social and organizational phenomena that are PL distributed.

In the real world, Gaussian and PL dynamics coexist. Before researching and offering advice to managers, we academics have to know for sure which dynamics are relevant. At present, however, both academics and managers show a clear bias toward Gaussian thinking. What price do organizations and managers pay for this bias? Drawing on both Gaussian and PL thinking, we explore the issue in the next two sections, framing it in terms of adaptation.

Ashby's Law of Requisite Variety

A Cognitive Interpretation of Ashby's Law

Biological behavior, like that of certain artifacts, is driven by informational as much as by mechanical action (Boisot, 1995). Ross Ashby, one of the founders of cybernetics, was interested in the range or variety of situations that an animal or a machine could respond and adapt to. His law of requisite variety states that "*only variety can destroy variety*" (Ashby, 1956, p. 207; our italics). The range of responses that a living system[3] must be able to marshal in its attempt to adapt to the world must match the range of situations—threats and opportunities—that it confronts. In some cases, the response might be wholly behavioral and often outside a system's cognitive control—as in the case of a hormonal response or a reflex. In other cases—for example, requiring a fight-or-flight decision—the response might be a blend of behavior and cognition that is contingent on the system successfully categorizing a stimulus as foreshadowing, say, the presence of a foe. It will then respond to *representations* of its environment that are constructed out of such categorizing activity rather than to its environment directly (Plotkin, 1993). Gell-Mann (2002, pp. 16-17) sees representations as "schema"—descriptions of real world "*regularities*" that form the basis for predictions and what he signifies as *effectively complex* responses.[4] In science they may appear as equations; in culture as laws, customs, and memes; and in management as strategies and practices. What advantage do better schema confer?

Not everything in a living system's environment is relevant or meaningful for it. If it is not to waste its energy responding to every will-o-the-wisp, a system must distinguish schema based on meaningful information (stimuli conveying "important" real-world regularities)—Gell-Mann's (2002) effective complexity—from noise (meaningless stimuli). Note that what constitutes information or noise for a system is partly a function of an organism's own expectations and judgments about what is important (Gell-Mann 2002)—as well as of its motivations—and hence, of its models of the world. Valid and timely representations (schema) economize on the organism's scarce energy resources (Ball, 2004; Vermeij, 2004; Zipf, 1949).

We can illustrate this interpretation of Ashby's law by means of a simple diagram that we label the *Ashby space* (Figure 2). On the vertical axis, we place the real-world stimuli that impinge on an organism. These range in their variety from low to high. A low-variety stimulus might be an image of

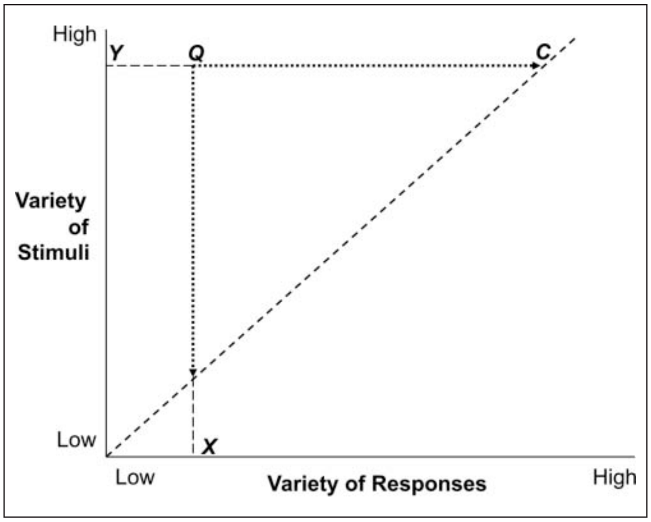

Figure 2. The Ashby space

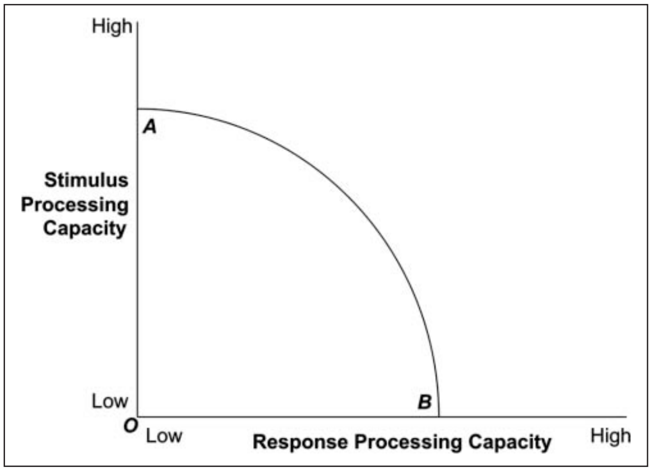

Figure 3. An adaptation budget

the moon; a high-variety stimulus might be the trajectory of individual insects in a swarm. On the horizontal axis, we place the variety of the system's *responses* to the stimuli. These also range from low to high. A low-variety response to the moon as stimulus would simply be to stare at it, meditate, and otherwise do nothing. A high-variety response to the insect swarm, by contrast, might be to chase after each individual insect flying past. The first type of response saves energy, the second wastes it. The diagonal in the diagram indicates the set of points at which variety can be considered "requisite," that is, where the variety of a system's response matches that of incoming stimuli in an efficiently adaptive way.

Ashby stressed the need to reduce the flow of some forms of variety from the external environment to certain essential processes in a living system. This was the role of regulation, and, as Ashby pointed out, the amount of regulation that can be achieved is bounded by the amount of information that can be transmitted and processed by the system (Ashby, 1956). The variety that it then has to respond to depend in part on the system's internal schema development and transmission capacities and in part on the operation of tunable filters—controlled by the system's cognitive apparatus—used by the system to separate out regularities from noise (Clark, 1997). The more intelligent a system, the higher will be the cognitive component in its response relative to the purely behavioral one. There is, thus, a trade-off between the behavioral and the cognitive resources that a living system has to marshal to be adaptive. Birds mostly act according to genetically derived behavioral instincts; monkeys show both behavioral and cognitive responses; humans exhibit higher-level cognitive skills.

Matching stimulus and response variety on the diagonal is only adaptive; however, if it occurs inside the *OAB* region of Figure 3, a region which defines the energy budget available to a living system, either to process incoming stimuli or to generate responses to them. To the right of this region, the variety of responses it generates, whether adaptive or not, causes the system to waste too much of its energy budget, eventually leading to its physical disintegration. Above this region, the data-processing resources required to register incoming stimuli, to interpret them, and to formulate adaptive responses also exceeds the system's budget, eventually leading to its cognitive disintegration. Cognitive and physical disintegration, however, are not mutually exclusive alternatives: the first will sooner or later lead to the second. The challenge for an adaptive system, then, is to locate itself at some point on the diagonal of Figure 2, but remaining within the budget area *OAB* of Figure 3.

An intelligent system, however, can use its data processing and transmission capacities to convert a high-variety stimulus into a low-variety one or vice versa. It does this by interpreting the stimulus and by distinguishing which part of the variety associated with it is information bearing and which part is noise. In doing so, it moves either down or up the vertical dimension of the Ashby space. Consider a system located at point *Q* in Figure 2, corresponding to some previous activity level, *X*, along the horizontal axis, which registers a stimulus at point *Y*, along the vertical scale. It could respond to the high variety associated with point *Y* directly in a purely "mindless" behaviorist fashion by simply moving horizontally to the right until it hits the diagonal at *C*—no cognitive simplification of the stimulus will be involved, just a behavioral response to it. But in doing so, the system might move outside its budget area *OAB* in Figure 3 and would rapidly deplete its energy resources. Call this a *headless chicken strategy*. Alternatively, the system could respond in a purely cognitive fashion by moving vertically down the diagram until it approaches the horizontal axis. In this case, it treats *all* incoming stimuli as already known regularities and noise and, thus, doesn't need any *new* behavioral response. This is the response of the "been-there-done-that" person who over-confidently feels no need to actually *do* anything different. Call this a *routinizing strategy*. However, because any downward movement calls for an interpretation and classification of the incoming stimuli, whether this second response is adaptive depends on how well the resulting schemas match the real-world regularities confronting the system—that is, how *effectively complex* they are.

Intelligent adaptive systems are best off locating on the diagonal in Figure 2 while remaining within the budget area *OAB* depicted in Figure 3. That is, they first need to *interpret* the stimuli impinging on them—a cognitive move either up or down the vertical scale in the diagram that attempts to extract relevant information about regularities from the noisy incoming signals that register as data with them. Then they need to develop relevant schema and respond with some *action* to the regularities so extracted—a behavioral move horizontally across the diagram toward the right that is only adaptive if it stops when it meets the diagonal and does so within the budget area. A cognitive move up the Ashby space expands the range and variety of stimuli that a living system will need to process before responding—in an exploratory learning process, schema become more complex (Gell-Mann, 2002; March, 1991). A cognitive move down the Ashby space draws on prior learning to reduce both the range and variety of stimuli and simplify the schema required—an exploitative learning process (March, 1991).

The trajectory of a living system through the Ashby space reflects its "intelligence," that is, its capacity to discern meaningful regularities, develop adaptive schemas that interpret these, and then generate *effectively complex* responses to them. Given the limited number of stimuli that a bird can "make sense" of, the trade-off between energy and data processing favors drawing predominantly on its energy resources. The variety of stimuli that a human being can respond to adaptively, by contrast, is much greater so that the trade-off favors drawing predominantly on its data-processing resources. For many living systems and especially for human beings, the budget area *OAB* of Figure 3 is constantly being expanded outward from the origin by means of artifacts (Clark, 1997), cultural transmission (Boyd & Richerson, 1985), and through collective organizational action (Corning, 2003). These simultaneously increase a system's effective complexity and adaptive capacity.

Gauss and Pareto in the Ashby Space—Three Ontological Regimes

Is there any limit to the expansion by human beings of their data-processing and schema-building resources—that is, their budget area *OAB*? Computational theory teaches us that

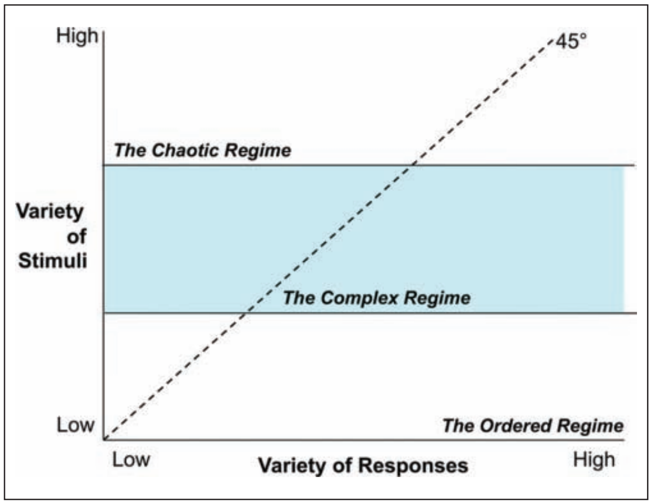

Figure 4. Three regimes in the Ashby space

problems whose size grows much faster than that of their inputs may require what amounts to an infinite amount of data processing for their solution (Chaitin, 1974; Sipser, 1997). If we take a problem's input size as a proxy measure of stimulus variety for an intelligent organism such as a human being (Grünwald, Myung, & Pitt, 2005), we can locate the different input sizes of various threats and opportunities along the vertical axis of Figure 2 and partition the axis to give us three distinct *ontological regimes* confronting the organism: the chaotic, complex, and ordered. We show these in Figure 4.

In discussing the phenomena that fall under the three regimes, Gell-Mann (2002) distinguishes between two fundamentally different "underlying generative processes" (Bhaskar, 1975):

Type 1 Regularities: The reductionist law-like regularities that drive the causal processes of normal science. These are predictable and easily represented by equations (Gell-Mann, (2002)—the focus of classical physics and neoclassical economics. They are also the point attractors of chaos theory—equilibria made possible by the conservation of energy.

Type 2 Regularities: The scale-free regularities resulting from an accumulation of random "tiny initiating events" that are amplified and propagated over time by positive feedback effects and that may become "*frozen accidents*" having lasting effects (Gell-Mann, 2002, p. 20). These are the strange attractors of deterministic chaos theory—they offer a different kind of regularity that never repeats and is nonlinear but does allow some predictability, hence the label "deterministic chaos."

Type 1 regularities generate the processes characterizing the *ordered regime* at the bottom of Figure 4. These may be confidently schematized and allow predictions that then become the basis of prescriptive solutions. Stimuli appearing in this regime appear to be, or are presumed to be, linear in nature and are experienced as relatively unproblematic by an intelligent organism—they are the stuff of everyday experience and normal science. The ordered regime fits a Gaussian world that can sometimes seduce one into seeing and responding to regularities that may not exist. Type 2 regularities are the source of Gell-Mann's (2002) tiny initiating "butterfly events." They are unpredictable and can produce significant nonlinear—and possibly extreme—outcomes. Type 2 regularities are mostly beyond the reach of normal science. It is hard to extract useful information from stimuli that appear in the *chaotic regime* at the top of Figure 4. They may therefore be judged computationally intractable, partly because of their variety and partly because they are *experienced* as chaotic. Unless nature "shows its hand" in good time or luck intervenes, an intelligent organism drawing on conventional representations and unaware of chaotic dynamics can typically make no sense of such stimuli within an adaptive time frame.

Stimuli appearing in the *complex regime* of Figure 3 are experienced as a blend of Gell-Mann's (2002) two regularities—a mix of partly law-like and partly unpredictable butterfly events—plus noise. Schema development in this regime is challenging to be sure, but once methods for separating out the two kinds of regularities from noise are available, it becomes computationally tractable. The butterfly effects that may arise in this regime call for PL thinking.

The larger the number of phenomena people can unproblematically classify as ordered, the more easily they can economize on scarce data-processing or energetic resources, holding these in reserve for more challenging phenomena—that is, parsimonious representations of phenomena help them to minimize the distance that they have to travel to the right in Figure 2 in quest of an adaptive response. In their quest for compact explanations and efficient, routinizable responses, humans thus have an obvious interest in steering phenomena down into the ordered regime wherever possible. Yet they can overdo it. If their interpretations are oversimplified and hence mistakenly allocated to the ordered regime, their response will be ill adapted. It often happens that either complex or anomalous situations are construed as being routine and responded to accordingly—sometimes with disastrous results.

Clearly, the first step in schema development with respect to some impinging real-world phenomenon is to identify the ontology appropriate for dealing with it. We outline three possibilities in Figure 5. If, for example, a system interprets a phenomenon as being ordered, it will find itself on a Gaussian trajectory, that of the cognitively routinizing strategy. This is

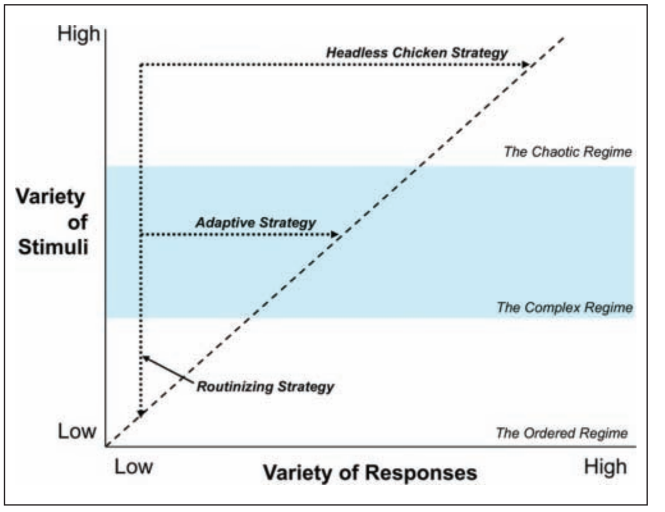

Figure 5. Three schema judgments in the Ashby space

the least-cost trajectory within the budget area *OAB* of Figure 3 because regularities allow a speedy extraction of schemas from the data. If, by contrast, the system views the phenomenon as chaotic, it will find itself on a purely behavioral trajectory that we have labeled the headless chicken strategy. Here, any latent law-like regularities are obliterated by the surrounding noise so that scale-free theories and PL thinking have little purchase. One can only wait for nature to eventually show its hand. On this trajectory, therefore, the system cannot make sense of anything and so responds mindlessly, sometimes expending so much undirected energy that if nature fails to show its hand and clarify things in good time, it can end up outside its budget area where it eventually disintegrates.

If the system takes the phenomenon to be complex—that is, neither so ordered that it can mobilize a least-cost response nor so chaotic that it can mobilize no meaningful schema at all—it is on a PL trajectory, one defined both by butterfly events, frozen accidents, and nonlinearities as well as by many of the attributes characterizing the ordered regime. Here, because schema development combines both law-like and butterfly events, an adaptive response will be feasible, but it will be more expensive than what is offered by a Gaussian trajectory. The system, however, will now be able to move more successfully to the right toward the diagonal while still remaining within its budget frontier.

Which ontology will turn out to be adaptive for a system depends on the level of adaptive tension imposed on it. As with Bak's (1996) sandpiles, increasing tension often increases the level of connectivity between hitherto unconnected phenomena, thus stealthily transforming a seemingly Gaussian ontology into a PL one. With the tension connecting up its constituent elements, tiny initiating butterfly events will rapidly propagate throughout the system, amplifying into

PL-distributed outcomes. To illustrate, imagine a fishnet lying loosely crumpled up in a pile. Cut the net between any two nodes and the rest of the net will remain undisturbed. Now place the net under tension by stretching it taut. If stretched taut enough, a single cut could then initiate a tear that would spread instantaneously from one side of the net to the other. A similar dynamic underlies the power blackouts that occasionally afflict power grids when the utilities, by temporarily shutting down one overloaded station—that is, a station under tension—trigger a cascade of further shutdowns throughout the grid system. In sum, under tension, a tiny butterfly event can rapidly propagate throughout a network of connections to produce an extreme outcome. An adaptive strategy in the complexity regime of the Ashby space, thus, requires an epistemology appropriate to the PL ontology that it is called on to deal with and an approach to the processing of data (stimuli) that supports it. How might we frame such an approach?

Data, Information, and Pattern Processing in the Ashby Space

In developing schemas to represent the regularities judged important in real-world stimuli, we extract information from data—that is, we "link up dots" (data points) to build intelligible patterns that represent our efforts to interpret the incoming stimuli (Boisot & Canals, 2004).

The different patterns that we can construct and stabilize out of linked dots—that is, information—then make up the knowledge—that is, schemas—that we build up to make sense of the phenomena generating the dots (Boisot & McKelvey, 2006). To illustrate the meaning of variety, Ashby (1956, pp. 124-125) made use of the following series: "*c*, *b*, *c*, *a*, *c*, *c*, *a*, *b*, *c*, *b*, *b*, *a*." In this series, *a*, *b*, and *c* repeat, so that there are effectively only three distinct *elements* present, that is, three kinds of variety or three degrees of freedom. In the language of patterns, however, this is the variety achieved at the level of the "dots" alone. Suppose that instead, we wish to define variety in terms of the number of *patterns* that can be derived from a given number of dots, then, using the formula in Table 1, we see that although four dots would allow us to generate up to six links and 64 patterns, the number of possible patterns achievable increases exponentially as we slowly increase the number of dots. Approximately 73.8 quintillion patterns could be extracted from combinations of 12 dots. Variety indeed! Although the vast majority of patterns will be quite meaningless and should therefore be ignored, trillions of suggestive ones will still be left, and one will still not necessarily know, up front, which ones will represent important regularities and which will represent noise.

If the processing of *dots* underpins most routinizing strategies, the processing *patterns*—that is, developing schemas—better fit PL-driven adaptive strategies. In effect,

Table 1. Relation of Dots to Links and to Patterns

Number of dots (N)	Number of possible links: $L = N(N - 1)/2$	Number of possible patterns: $P = 2^L$
4	$L = 6$	$P = 64$
10	$L = 45$	$P = 35$ trillion
12	$L = 66$	$P = 73.8$ quintillion

processing dots amounts to little more than processing data, a low-level cognitive activity. Processing patterns, in contrast—pattern recognition—is acknowledged to be a high-level cognitive activity, one that involves a judicious selection of relevant patterns from among myriad possibilities (Kelso, 1995; Thelen & Smith, 1994). If we view human organizations through a schema-building lens, then we see that those following Gaussian principles will tend to operate hierarchically, generating data (dots) at the base, selectively linking data items to each other in the middle reaches of an organizational hierarchy to produce information, and leaving the task of constructing and processing meaningful patterns out of the myriad possibilities available—that is, interpretation—to senior managers located at the top of the organizational pyramid (Galbraith, 1973). This sequential process of upward filtering works well in a stable world in which the relevant patterns are few in number, familiar, or readily discovered.

Figure 6 depicts this data-processing strategy as a pyramid, *A*, that acts in a top-down fashion as an informed selection device. Its approach to learning is exploitative (March, 1991). Linear programming and optimization models work well here, as do reductive mathematical equations in general (Gell-Mann, 2002). However, when the pool of potentially relevant dots is large, novel, and nonlinear—as, for example, when they stem from butterfly events and frozen accidents—then the task does not lend itself to a hierarchical data-processing approach. The organizational task now appears as an inverted pyramid, *B*, in which even a limited amount of data at the base generate an unimaginable number of candidate patterns of potentially overwhelming complexity for processing. In effect, framed in conventional data-processing terms, the task becomes for all practical purposes computationally intractable.

Figure 6 highlights the nature of the adaptive challenge by presenting data-processing strategies and pattern-processing tasks as two intersecting pyramids. Although processing the huge number of patterns in the inverted pyramid, *B*, requires the bulk of an organization's data- and pattern-processing resources to be located toward the top, under the hierarchical data-processing regime of pyramid *A*, the available data-processing resources are mostly located at the bottom and a very modest allocation of pattern-processing resources is located at the top. The resulting imbalance will

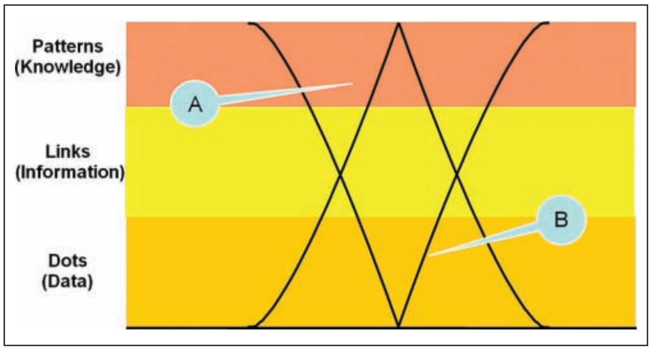

Figure 6. The pattern processing challenge

then often become a source of cognitive inertia, "groupthink" (Janis, 1972), and "threat-rigidity" (Staw, Sandelands, & Dutton, 1981).

Research Implications for PL-Based Management Inquiry

In the ordered regime, Gaussian science offers the highly refined quantitative methods presented in textbooks such as Greene's (2011). In the chaotic regime, we encounter the rich, ethnographic, or historical case studies (Yin, 1989) that are believed by some to lack the scientific legitimacy of the quantitative methods that the ordered regime is able to draw upon. How might management inquiry meet the challenge of reconciling them?

Start with the observation that elephants are not just larger versions of mosquitoes. Although both are made up of the basic biological building blocks, cells, that aggregate and connect up in different ways and quantities, they exhibit quite different properties. The complex cells that they are composed of, the eukaryotes, have themselves been built up over time from simpler cells, the prokaryotes (Margulis, 1993). If mosquitoes occupy the upper left-hand corner of the PL distribution of Figure 1, elephants occupy the lower right-hand corner of the same distribution. "More is different" says Anderson (1972, p. 393), and the myriad alternative ways in which cells can aggregate and connect up under adaptive tension generate a size/frequency distribution of species, some of which (elephants) will be more complex and longer lived than others (mosquitoes). In the case of mosquitoes, the variety required for adaptation and survival is achieved through fast breeding and genetic transmission at the species level. The ill-adapted die off, but there are plenty more where they come from. In the case of elephants, in contrast, a good part of the variety needed to adapt is achieved through the acquisition of a flexible behavioral repertoire at the level of the individual organism. In some species, not necessarily the largest, the behavioral repertoire is augmented and directed by the exercise of intelligence, thus extending the species' adaptive potential as indicated in the Ashby space.

Applying the above more concretely to human social and economic conditions, in the upper left region of Figure 1, we have millions of tiny Ma&Pa stores—the mosquitoes of the retailing world—with little in the way of assets, annual sales of less than US$1million, and no paid employees. They are fiercely independent—often this, rather than money is the reason for setting up shop—and show little growth, serving only very local customers. Such "atomized" stores are presumed *i.i.d.* and treated in a Gaussian fashion. Like mosquitoes, most live and die with little capacity for adaptation to changes in their environment. Over time, however, through the strategic application of intelligence, some of these will acquire a more flexible behavioral repertoire and will therefore grow larger. A few like Walmart, the world's largest retail firm—with its billions of dollars in assets and profits, hundreds of thousands of employees, and thousands of products, it is less an elephant than the *T. Rex* of the industry—will eventually end up in the lower right-hand region of Figure 1. Between Ma&Pa stores and Walmart, we can locate thousands of firms on the PL distribution that will vary in size. The "average" of these firms may look nothing like a Walmart, but it will look nothing like a Ma&Pa store either. As Axtell (2008) puts it, in a rank/frequency distribution "the typical firm does not exist." Within the retailing sector, therefore, these firms in effect constitute different species.

What follows from this? In a PL distribution, differences rather than similarities dominate so that heterogeneity, not homogeneity, is in the driving seat. Deleting the five largest retail firms located on the lower right of Figure 1, as if they were outliers, could significantly change what the "average" retailer looks like without having any effect on what the modal firm looks like. Deleting 1,000 stores located on the upper left of the figure would also change little. From the foregoing, we see that the characteristics of the two tails and those of the middle are vastly different. Simply put, applying Gaussian statistical methods to PL distributions and getting rid of outliers—a process called "Winsorizing"—can produce quite misleading results and is thus inappropriate and blatantly distorting.

Given that "more is different," what characteristics of higher level species might legitimately be inferred from the study of lower level ones in biology? And in the case of human organizations, what could we learn about the few giant Walmarts from a study of the cloud of tiny Ma&Pa stores? In other words, what kinds of inferences allow us to move up or down the PL curve of Figure 1? A well-tested and validated theory allows us to use *deductive* strategies to predict the next instance of an event. Our theory of gravity, for example, allows me to deduce that if I let go off the ball I am holding in my hand, it will fall to the ground. Even in the absence of a well-tested and validated theory, providing that events can be construed as being independent of each other, we can focus on the similarity between events and use *inductive* methods to predict the next instance of one. My dog always runs up and barks at the postman when he delivers the mail; I hear the dog barking and I can reasonably infer that the postman has arrived. Of course, such an inductive inference is less epistemically robust than what deduction can deliver because it could turn out that it is the arrival of some person other than the postman that is causing the dog to bark. In a PL distribution, no less than in a Gaussian one, some variables will remain similar in their effect from one tail to the other—for example, how subordinates deal with superiors, feelings of relative deprivation, tendencies toward groupthink, and so on do not necessarily vary with organizational size. Researching these variables offers lessons that retain their validity across the entire distribution. In such cases, Gaussian methods will work just fine.

The problem is, of course, that many Gaussian processes are nested inside PL ones so that the challenge becomes one of deciding which type of distribution will be appropriate for a given problem. Some variables, after all, will be unique to each tail in a PL distribution, and those that are operational at one end of the distribution will contribute little or nothing to what goes on at the other end. Here, winsorizing will not help. The millions of Ma&Pa stores have many things in common, as do groupings of larger kinds of firms—giant firms, for example, have a financial and political clout unavailable to small store owners. Within both "groupings," therefore, Gaussian statistics may apply, but they cannot be applied to the sum of these groupings. Some variables, however, will turn out to be scalable and with some modifications will apply to stores anywhere in the Pareto distribution. Such scalable variables are critically important to an understanding of whether a given firm will grow. The founding technologies and organizing strategies of firms such as Intel, Microsoft, Walmart, and GE under Jack Welch, for example, have proved to be scalable, helping the firms to grow from small beginnings to their current giant size. Scalability effects are especially sensitive to the presence of tensions and connectivity within a system, and in the case of the firms just mentioned, the tension may be imposed on the organization by the motivation and goals of the owners or senior managers. Complexity theory holds that such tension leads to increased connectivity and that connectivity, in turn, triggers interactivity, positive feedback, amplification, and other scale-free dynamics.

This said, if events are so heterogeneous that no similarity between them can be discerned, are not repeated frequently enough, are not independent of each other, or if one lacks a theory to interpret them, then we may effectively be dealing with samples of one (March, Sproull, & Tamuz, 1991). Such will often be the case in the complexity regime in the Ashby space. Here, although information may be neither chaotic nor beyond understanding, it is neither ordered nor easily understood. As noted earlier, one then faces a large pool of relevant potentially interpretable patterns that may be novel (stemming from butterfly events and frozen accidents), nonlinear (the causal processes generating extreme outcomes are difficult to recognize), and nonrepeatable. Here, neither deduction nor induction will deliver valid inferences. The appropriate inferential strategy will be *abductive*. According to Peirce (1933), "abduction . . . consists of examining a mass of facts and in allowing these facts to suggest a theory" (p. 205). Abduction is an "inference to the best explanation" that underpins Popper's (1935) "logic of discovery" (Ben-Menahem, 1990; Harman, 1965) and that, in contrast to induction, uses *all* available data, no matter how heterogeneous, to generate coherent patterns (Hanson, 1958). Abduction rests on context-dependent intuitions; it complements the inductive and deductive processes (Lipton, 1991) that actually connect the dots of Table 1 and is thus well suited to the inferential challenges posed by the processing of the myriad patterns that populate a PL distribution.

Of course, if it were simply a question of the "size" of the pattern-processing problem, Gaussian approaches to inference would still apply. The scalable regularities discussed by Gell-Mann, however, require a new inferential approach—call it *scalable abduction*. How does scalable abduction differ from Gaussian abduction? Applied across PL distributions, it exploits the scale-free character of extreme outcomes, offering an inferential approach that builds on whatever regularities can be discerned in the multilevel dynamics that underpins scale-free causal processes. One identifies such regularities by focusing on the many "tiny initiating events" that occur among the smaller entities in the upper left region of Figure 1 and looking for meaningful patterns. The great majority of patterns will turn out to be meaningless and hence irrelevant. Some, however, will constitute butterfly events that will propagate outward and explode into the larger outcomes in the lower right of the figure. Once scalable regularities are uncovered, they will point to the causal relationships and theories that, under tension, are capable of generating extreme outcomes.[5]

In the absence of an adequate data-processing capacity, the earlier prescription will strike many as a counsel of perfection. After all, how far up the PL curve of Figure 1 can one travel before the differences between countless events begins to get blurred? If one lacks the data-processing resources to properly distinguish one event from another across the population as a whole, they will begin to look *i.i.d.*, leading one to resort to Gaussian methods that will legitimate traveling further down the Ashby space of Figure 2 in search of parsimonious representations and responses. Yet, the big story at the beginning of the new millennium is that the necessary data-processing resources are now rapidly becoming available, allowing us to deal with samples of one on a scale hitherto unimagined (Hofer, McKee, Birnholtz, & Avery, 2008). As noted earlier, many firms located on the upper left region of Figure 1 have customers who live in small microniches, have unique tastes, and may now be catered to in a customized fashion using Internet-based strategies and technologies—as done by Amazon, eBay, and Google (Anderson, 2006). Medical science is also moving toward customization, defining genetic profiles that are unique to individuals and "microdesigning" individualized drugs in response ("Briefing: Genetics," 2007). This new ability to extract useful information from individual idiosyncrasies located in the population variance—and then to connect these up to each other as Amazon does when it tries to build up sales momentum by telling one customer what other customers are reading—allows one to travel ever further up the PL distribution of Figure 1 without having to draw on Gaussian methods. Here, as elsewhere, an idiographic "science of the unique" is emerging that contrasts with the nomothetic "science of the Gaussian average" that has hitherto dominated management inquiry. Unlike reductionist approaches, scale-free abductive ones rest on Gell-Mann's "middle level" causal mechanisms—those that spread across the multiple levels of a rank/frequency distribution—scalable abduction can actually travel in both directions up and down the PL slope of Figure 1 and come to replace much of the Gaussian thinking that characterizes the upper left-hand region of the figure with qualitative context-dependent thought.

Deductive and inductive strategies characterize what Reichenbach has called *the context of justification*—they provide an epistemic justification on which actionable beliefs, once developed, can rest. Abductive strategies, in contrast, characterize what he calls *the context of discovery* (Reichenbach, 1938)—they provide a basis for the development of actionable beliefs. Although robust scientific theories, therefore, might be the product of abductive strategies located in the context of discovery, they only become established through deductive and inductive strategies located in the context of justification. The former are epistemically weaker than the latter. They only warrant anticipation, not prediction. When might anticipation be adaptive? When, under circumstances that often hold in the complexity regime of Figure 2, uncertainty is irreducible. Beware, then, of the siren calls of normal distributions! The attraction of the top left-hand corner of Figure 1 is that it allows you to move with some confidence down the vertical axis of the Ashby space of Figure 2, thus economizing on scarce resources of attention and energy. Yet, the top left-hand corner is plagued by ambiguities. If it creates false positives when we treat what is, in effect, a Gaussian distribution as if it was PL distributed, it also creates false negatives when we treat what is a PL distribution as if it were Gaussian. False positives—an example of which would be the Y2K panic at the end of the last millennium—wastes resources by leaving you further up the Ashby space than you need to be and somewhat more impoverished. False negatives, however, fail to mobilize the resources necessary for timely adaptation—the lack of preparations made for hurricane Katrina or the overrelaxed regulatory response to the subprime mortgage crisis illustrate the point.

The quest for predictability when confronted with irreducible uncertainty is not just futile, it is often dangerous. We must therefore learn to treat outliers as weak signals that all may not be well with our assumptions. Rich accounts of butterfly events—outliers—in books such as those by Perrow (1984), Vaughan (1996), and Weick and Sutcliffe (2001) highlight the importance of researching their emergence and consequences. The first two books are excellent examples of what happens when butterfly events are ignored; Weick and Sutcliffe's focus on high-reliability organizations, in contrast, shows how they forestall disasters by identifying and managing "tiny initiating events" before they get amplified and run out of control. The ability of collectivities of heterogeneous human agents to discover things and make sound decisions—that is, to display collective intelligence—is now well established. Surowiecki (2004) and Page (2007) show that when it comes to puzzle solving, agent diversity beats expertise much of the time. In exploiting this diversity, the challenge is to achieve a trade-off between the *ex ante* costs of anticipating and the *ex post* costs of responding in adaptive ways to high-impact, low-frequency outcomes. Adaptive responses, in turn, require a trade-off between an ability to make sense, *ex post*, of the extreme outcome and an ability to respond adaptively to what has been made sense of *ex post*.

The challenge is both social and cognitive. The opportunity or threat posed by extreme outcomes first appears as small butterfly events to which heterogeneous agents, initially endowed with zero-order connectivity, respond by searching for and connecting to other agents (Barabási, 2002; Holland, 2002). Through such interactions, and providing that they can overcome any tendencies to passive dependence (Argyris, 1957) or to "groupthink" (Janis, 1972), the agents' sensing processes will often be capable of reaching beyond the atomized outcomes suggested by Gaussian interpretations to anticipate the PL dynamics at work. As the pool of agents enlarges to straddle multilevel hierarchies, the chances in finding butterfly levers and emergent dissipative structures go up. If much of what is relevant to practitioners results from negative butterfly events, positive ones are also possible. Entrepreneurs, for example, will exploit Schumpeter's (1934) "gales of creative destruction" and Christensen's (1997) disruptive technologies to create as well as to destroy. The appearance of microprocessors at Intel (Burgelman, 2002) and of genetic engineering at Monsanto (Day & Colwell, 2006) both illustrate the positive butterfly events that get amplified and turned them into strategic opportunities by well-thought out yet emergent sense-and-respond schemas. Unfortunately, butterfly events and their outcomes—whether positive or negative—remain mostly outside mainstream management inquiry. The implications are clear. Managers need to learn enough about PLs to avoid both the potential incoherences of the chaotic regime as well as the tempting oversimplifications that so often plague the ordered regime in the Ashby space.

What tools are available for pursuing abductive strategies? There are many, but the binding agent between the heterogeneous events that such strategies draw on is likely to be narratives, stories, case studies, and histories and not just numbers. If Gaussian thinking focuses on *probability*, PL thinking focuses on *plausibility*. Both are sources of beliefs on which organizations might be willing to act, but although probability requires well-defined repeatable and identical events, plausibility is achieved by pattern-matching skills that draw on heterogeneous prior experience—situated, uniquely human, and historical. Given its analytical tractability, management inquiry mostly stays in the upper left-hand region of Figure 1, where probabilities have traction. Here, it focuses on events that have high-enough frequencies to meet statistical sampling requirements and that are mostly similar enough in size not to violate *i.i.d.* requirements. In contrast, given their need for relevance and for coping with extremes, practicing managers have been drawn to "vivid stories," case studies, and scenarios (Perrow, 1984; Sheffi, 2005; Weick, 1993), which build on dynamics originating in the lower right-hand region of the figure. Such narratives offer opportunities for retrospective sense making, which act as inputs for either current or future situations. The lessons of history may not be fully repeatable, but useful patterns can nevertheless often be extracted from them. John F. Kennedy, for example, relied heavily on his reading of *The Guns of August*, Barbara Tuchman's account of the butterfly events that sparked off World War I, to guide him through the thickets of the Cuban missile crisis of October 1962 (Allison, 1973).

How can scale-free butterfly levers facilitate responses to extremes? The social sciences in general and management inquiry in particular enjoy options that may not always be available to the natural sciences. For example, each year California experiences some 16,000 small quakes (4 or less on the Richter scale) which add up to more than 2 million "average" quakes since the last great quake of 1857. Although some are still trying (Sornette, 2002), most geologists have just about given up attempting to predict large quakes (Main, 1997). Their problem is that the small quakes that initiate them—located up to 400 miles underground—are inaccessible and therefore hard to study. In organizations, however, the small butterfly events that foreshadow extreme outcomes are often at a scale that makes them readily accessible to direct human observation. Organizational actors thus stand a better chance of detecting their scale-free causes and then acting either to forestall or to exploit them.

Given that one's experience often limits what one looks for and sees, the key idea here is to switch from Gaussian methods that look for regularities and convert them into "averages" to PL methods and scale-free regularities.[6] The switch involves shifting one's focus from the upper left of Figure 1 to the leverageable, scalable regularities that straddle multiple levels of a hierarchy. If the upper left-hand side of Figure 1 supports Gaussian ontologies and *managerial* behaviors and the lower right-hand side supports PL ontologies and *entrepreneurial* behaviors. Gaussian ontologies are atomistic. They favor a managerial approach that uses deduction and induction to reduce uncertainty by converting it into measurable risk. PL ontologies, in contrast, are connectionist. They favor an entrepreneurial approach, one that in the lower regions of Figure 1 will produce one of Schumpeter's far-from-equilibrium creative destroyers (Schumpeter, 1934). Schumpeter's creative destroyers absorb uncertainty so as to be able to better exploit it. Managers want to travel up the PL curve of Figure 1 and down the vertical axis of Figure 2 to achieve *justification*. Entrepreneurs want to travel down the PL curve of Figure 1 and up the vertical axis of Figure 2 to secure either *discovery* or *creation*. Their epistemic requirements are thus different. A more connected—and hence more complex—world will favor entrepreneurial strategies over managerial ones. How so?

Seeing potentially relevant connections, whether actual or merely possible ones, is a *creative achievement*, a process of pattern matching not unlike what an artist does (Gombrich, 1960). However, it does not come naturally. Connecting up ostensibly independent phenomena is cognitively expensive. We are object-oriented animals with a bias toward what is immediately visible, given, and hence can be taken for granted (Botero, 1999). This perceptual inertia releases attention for redeployment toward things that interest us. However, because not everyone is interested in the same things, a manager faces a different problem to that of the entrepreneur, namely, that of getting others aligned behind the pattern that he or she sees to generate an organized response. An entrepreneur in quest of a nonlinear competitive advantage may not actually *want* others to see what he or she sees—at least not until later.

Conclusion

We began by defining key elements underlying PLs. Then, drawing on Ashby's (1956) law of requisite variety, we partitioned the Ashby response space into chaotic, complex, and ordered regimes. Next, we argue that practitioner-relevant research in the latter space calls for a "PL science" and suggested how management inquiry could become more PL oriented. Managers and entrepreneurs, each in their own way, organize the elements of connected network processes, some of which yield adaptive structures and some of which do not. Our theoretical approach has the effect of naturalizing management inquiry and bringing it closer to recent trends in physics and biology—shades of econophysics (McCauley, 2004). It makes it less dependent on the "object-oriented" engineering from which it sprang and more in tune with interactive networks (see, for example, Newman, Barabási, & Watts, 2006; West & Deering, 1995).

Physics operates at the highest level of generality achievable in the natural sciences. *Econophysics*, as the name implies, involves the application of some methods of physics to PL-distributed economic phenomena. By focusing on PL representations of Pareto distributions, PL science increasingly uncovers phenomena—from atoms to galaxies, from DNA base-pairs to species extinctions, from small stores to giant corporations—that are best characterized by PL "signatures"—a negatively sloping straight line emerges when Pareto distributions are plotted in double-log graphs. Scale-free processes result from causal dynamics that cross multiple levels of analysis. Andriani and McKelvey (2007, 2009a, 2009b, 2011) now extend research into these processes to organizations and management.

In a world that is becoming ever more connected and complex, however, management inquiry is yet to take up the PL challenge. It lacks an adequate conceptualization of extreme outcomes and so, unsurprisingly, is unable to incorporate them into its current theorizing. We have therefore drawn from the complexity sciences to introduce a more penetrating treatment of PLs, extreme outcomes, and scale-free theory into management inquiry. We believe that PL thinking will ultimately connect the concept of organization used in the social sciences with those used in the kindred fields of biology and ecology (Boisot & Cohen, 2000).

Most of orthodox economics is attuned to normally distributed, independent-additive data points. So is much of management inquiry, especially when conducted by researchers under the spell of econometric statistics (Greene, 2011). Faced with the phenomenon of increasing connectivity, both are beginning to look at networks. However, they do so while retaining an object-oriented focus on stable structures (e.g., Wasserman & Faust, 1994) that fails to take into account the nonlinear dynamics of network evolution (an exception is Newman et al., 2006). The result is an excessive focus on those aspects of organizations that are "*thing-like*" as opposed to "*connected*" and hence, like living systems, resilient under changing conditions. The advent of an information-rich postindustrial economy, however, is changing both the nature of organizations as well as the way we think of them. Organizations are becoming more flexible and changeable, and their boundaries are becoming more blurred. They are becoming less object-like and more like extensible networks unconstrained by boundaries. This requires a shift in theorizing first initiated by Weick (1979) three decades ago; from *organizations* as things to *organizing* as an unfolding process.

With the world becoming ever more densely connected through new information and communication technologies, adaptive tensions can now spread globally in minutes. Moreover, though strong ties are also needed for cohesion, trust, and efficiency (Granovetter, 1973; Janis, 1972), weak-tie connections lead to faster paced innovation and change (Granovetter, 1973), In such circumstances, we can no longer afford to confine ourselves to an independent-additive Gaussian epistemology that screens out the emergent complexity and the associated PL rank/frequency distributions now challenging us. Drawing on the Ashby space, our analysis suggest that, to be adaptive, responses increasingly need to be both more varied and more rapid (McKelvey & Boisot, 2009) than what is made possible by the "object-oriented" assumptions of traditional Gaussian thinking. To be sure, these will remain necessary in many circumstances. But they are no longer sufficient.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Our term *butterfly effect* simply recognizes Ed Lorenz's famous phrase which is the title of one of his presentations, "Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?" (Lorenz, 1972)
2. Power law (PL) distributions are Pareto distributions that are graphed using log X and Y axes to simplify their representation of a complex phenomenon.
3. Although Ashby talked of "systems" *tout court*, we shall be dealing with "living systems." We use the term as Miller (1978) does to cover systems made up of bio-organisms and ecologies, people, groups, and organizations.
4. For Gell-Mann (2002), science is concerned with separating what are judged meaningful regularities from random or irrelevant noise. Although the complexity that we see or sense is a mixture of regularities and noise, Gell-Mann's "effective complexity" (p. 13) is the simplest or most concise description of the regularities judged important. Viable living entities—labeled *complex adaptive systems*—are able to develop parsimonious "schema" useful in understanding contextual regularities and in adapting to them.
5. Given the negatively sloping PL straight line, and extreme statistics, this can be done with considerable accuracy (Baum & McKelvey, 2006).
6. They are defined in Andriani and McKelvey (2009b).

References

Allison, G. (1973). *The essence of decision*. Boston, MA: Little Brown.

Anderson, C. (2006). *The long tail*. London, UK: Random House Business Books.

Anderson, P. W. (1972). More is different. *Science, 177*, 393-396.

Andriani, P., & McKelvey, B. (2007). Beyond Gaussian averages: Redirecting organization science toward extreme events and power laws. *Journal of International Business Studies, 38*, 1212-1230.

Andriani, P., & McKelvey, B. (2009a). Avoiding extreme risk before it occurs: A complexity science approach toward incubation. *Risk Management: An International Journal, 12*, 54-82.

Andriani, P., & McKelvey, B. (2009b). Extremes and scale-free dynamics in organization science: Some theory, research, statistics, and power law implications. *Organization Science, 20*, 1053-1071.

Andriani, P., & McKelvey, B. (2011). From skew distributions to power-law science. In P. Allen, S. Maguire, & B. McKelvey (Eds.), *Handbook of complexity and management.* London, UK: SAGE.

Argyris, C. (1957). *Personality and organization.* New York, NY: Harper.

Ashby, R. W. (1956). *An introduction to cybernetics*. London, UK: Methuen.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science, 293*, 1818-1820.

Axtell, R. L. (2008, February 8). *Nonexistence of a typical firm in the U.S. economy: Extremely heavy tails in firm size and growth.* Presented at Organization Science Winter Conference XIIIV, Squaw Valley, CA.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York, NY: Copernicus.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of 1/*f* noise. *Physical Review Letters, 59*, 381-384.

Ball, P. (2004). *Critical mass: How one thing leads to another*. London, UK: Arrow Books.

Barabási, A. L. (2002). *Linked: The new science of networks*. Cambridge, MA: Perseus.

Baum, J. A. C., & McKelvey, B. (2006). Analysis of extremes in management studies. In D. J. Ketchen & D. D. Bergh (Eds.), *Research methodology in strategy and management* (Vol. 3, pp. 123-197). Oxford, UK: Elsevier.

Ben-Menahem, Y. (1990). The inference to the best explanation. *Erkenntnis, 33*, 319-344.

Bennis, W. G., & O'Toole, J. (2005). How business schools lost their way. *Harvard Business Review, 83*, 96-104.

Bhaskar, R. (1975). *A realist theory of science*. London, UK: Leeds Books.

Boisot, M. (1995). *Information space*. London, UK: Routledge.

Boisot, M., & Canals, A. (2004). Data, information, and knowledge: Have we got it right? *Journal of Evolutionary Economics, 14*, 43-67.

Boisot, M., & Cohen, J. (2000). Shall I compare thee to . . . an organization? *Emergence, 2*, 113-135.

Boisot, M., & McKelvey, B. (2006). Speeding up strategic foresight in a dangerous, complex world: A complexity approach. In G. G. S. Suder (Ed.), *Corporate strategies under international terrorism and adversity* (Chapter 3, pp. 20-37). Cheltenham, UK: Edward Elgar.

Botero, J., (1999). The immediately given as ground and background. In J. Petitot, F. Varela, B. Pachoud, & J. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 440-463). Cambridge, MA: MIT Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Briefing: Genetics, medicine and insurance. (2007, August 25). *The Economist*, *384*, 69-71.

Brock, W. A. (2000). Some Santa Fe scenery. In D. Colander (Ed.), *The complexity vision and the teaching of economics* (pp. 29-49). Cheltenham, UK: Edward Elgar.

Burgleman, R. A., & Grove, A. S. (2002). *Strategy is destiny: How strategy-making shapes a company's future*. New York: Free Press.

Chaitin, G. J. (1974). Information-theoretic computational complexity. *IEEE Transactions on Information Theory, 20*, 10-15.

Christensen, C. (1997). *The innovator's dilemma*. Boston, MA: Harvard Business School Press.

Clark, A. (1997). *Being there*. Cambridge, MA: MIT Press.

Corning, P. (2003). *Nature's magic*. Cambridge, MA: Cambridge University Press.

Day, D., & Colwell, K. (2006, August 15). *The Monsanto paradox.* Atlanta, GA: Academy of Management Conference.

Friedman, M. (1953). *Essays in positive economics*. Chicago, IL: University of Chicago Press.

Galbraith, J. R. (1973). *Designing complex organizations.* Reading, MA: Addison-Wesley.

Gell-Mann, M. (2002). "What is complexity?" In A. Q. Curzio & M. Fortis (Eds.), *Complexity and industrial clusters* (pp. 13-24). Heidelberg, Germany: Physica-Verlag.

Glaser, P. (2009). *Fitness and inequality in an increasing returns world: Applying the tools of complexity economics to study the changing distribution of US stock market capitalizations from 1930 to 2008.* Term paper, UCLA Anderson School of Management, Los Angeles, CA.

Gombrich, E., (1960). *Art and illusion*. New York, NY: Phaidon.

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

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Grünwald, P., Myung, I., & Pitt, M. (2005). *Advances in minimum description length*. Cambridge, MA: MIT Press.

Hanson, N. R. (1958). *Patterns of discovery*. Cambridge, UK: Cambridge University Press.

Harman, G. (1965). The inference to the best explanation. *Philosophical Review, 64*, 88-95.

Hofer, E., McKee, S., Birnholtz, J., & Avery, P. (2008). High-energy physics: The large hadron collider collaboration. In G. Olson, A. Zimmerman, & N. Bos (Eds.), *Scientific collaboration on the internet* (pp. 143-151). Cambridge, MA: MIT Press.

Holland, J. H. (1995). *Hidden order.* Cambridge, MA: Perseus Books.

Holland, J. H. (2002). Complex adaptive systems and spontaneous emergence. In A. Q. Curzio & M. Fortis (Eds.), *Complexity and industrial clusters* (pp. 24-34). Heidelberg, Germany: Physica-Verlag.

Janis, I. L. (1972). *Victims of groupthink*. Boston, MA: Houghton Mifflin.

Kaplan, A. (1964). *The conduct of inquiry*. San Francisco, CA: Chandler.

Kauffman, S. A. (1993). *The origins of order*. New York, NY: Oxford University Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kirchgaessner, S., & Kelleher, E. (2005, July 11). UK bombs drive insurance debate. *Financial Times Europe,* p. 5.

Lawson, T. (1997). *Economics and reality.* London, UK: Routledge.

Lewin, K. (1951). *Field theory in social science*. New York, NY: Harper.

Lipton, P. (1991). *Inference to the best explanation*. New York, NY: Routledge.

Lorenz, E. N. (1972). *Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?* Paper presented at the 1972 meeting of the American Association for the Advancement of Science, Washington, DC.

Main, I. (1997). Earthquakes—long odds on prediction. *Nature, 385*, 19-20.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*, 71-87.

March, J. G., Sproull, L. S., & Tamuz, M. (1991). Learning from samples of one or fewer. *Organization Science, 2*, 1-13.

Margulis, L. (1993). *Symbiosis in cell evolution*. New York, NY: W. H. Freeman.

McCauley, J. L. (2004). *Dynamics of markets: Econophysics and finance.* Cambridge, UK: Cambridge University Press.

McKelvey, B., & Boisot, M. (2009). Redefining strategic foresight: "Fast" and "far" sight via complexity science. In L A. Costanzo & R. B. MacKay (Eds.), *Handbook of research on strategy and foresight* (pp. 15-47). Cheltenham, UK: Edward Elgar.

Miller, J. G. (1978). *Living systems*. New York, NY: McGraw-Hill.

Mirowski, P. (1989). *More heat than light*. Cambridge, UK: Cambridge University Press.

Mitchell, M. (2009). *Complexity: A guided tour.* New York, NY: Oxford University Press.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*, 323-351.

Newman, M. E. J, Barabási, A. L., & Watts, D. J. (Eds.). (2006). *The structure and dynamics of networks*. Princeton, NJ: Princeton University Press.

Page, S. E. (2007). *The difference: How the power of diversity creates better groups, firms, schools, and societies.* Princeton, NJ: Princeton University Press.

Peirce, C. S. (1933). *Collected papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

Perrow, C. (1984). *Normal accidents*. Princeton, NJ: Princeton University Press.

Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management: Learning & Education, 1*, 78-95.

Plotkin, H. (1993). *Darwin machines and the nature of knowledge.* Cambridge, MA: Harvard University Press.

Popper, K. R. (1935). *Logik der Forschung*. Vienna, Austria: Springer-Verlag. (Reprinted as *The logic of scientific discovery*, New York, NY: Harper & Row, 1959.)

Reichenbach, H. (1938). *Experience and prediction*. Chicago, IL: University of Chicago Press.

Schumpeter, J. (1934). *The theory of economic development*. Boston, MA: Harvard University Press.

Sheffi, Y. (2005). *The resilient enterprise*. Cambridge, MA: MIT Press.

Sipser, M. (1997). *Introduction to the theory of computation*. Boston, MA: PWS Publishing.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis*. Thousand Oaks, CA: SAGE.

Sornette, D. (2002). Predictability of catastrophic events. *Proceedings of the National Academy of Sciences, 99*(Suppl. 1), 2522-2529.

Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., . . . Stanley, H. E. (1996). Scaling behavior in the growth of companies. *Nature,* 379, 804-806.

Staw, B. M., Sandelands, L. E., & Dutton, J. E. (1981). Threat-rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly, 26*, 501-524.

Surowiecki, J. (2004). *The wisdom of crowds*. New York, NY: Doubleday.

Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press

Vaughan, D. (1996). *The challenger launch decision*. Chicago, IL: University of Chicago Press.

Vermeij, G. (2004). *Nature: An economic history*. Princeton, NJ: Princeton University Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis.* New York, NY: Cambridge University Press.

Weick, K. E. (1979). *The social psychology of organizing*. Reading, MA: Addison-Wesley.

Weick, K. E. (1993). The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly, 38*, 628-652.

Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected*. San Francisco, CA: Jossey-Bass.

West, B. J., & Deering, B. (1995). *The lure of modern science: Fractal thinking*. Singapore: World Scientific.

White, H. C. (1963). *An anatomy of kinship.* Englewood Cliffs, NJ: Prentice Hall.

Yin, R. K. (1989). *Case study research: Design and methods*. London, UK: SAGE.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. New York, NY: Hafner.

Bios

Max Boisot (PhD, Imperial College) is a professor at Escuela Superior de Administración y Dirección de Empresas in Barcelona (ESADE); an associate fellow at the Said Business School, Oxford University; and senior research fellow at the Snider Center for Entrepreneurial Research, the Wharton School of the University of Pennsylvania. Between 1984 and 1989, he directed the China–Europe Management Program in Beijing, which has since evolved into the China–Europe International Business School in Shanghai. He has published in *Academy of Management Review*, *Administrative Science Quarterly*, *Organization Science, Research Policy*, and *Journal of Evolutionary Economics*. He is currently researching the knowledge base of the ATLAS experiment on the large hadron collider at European Organization for Nuclear Research (CERN).

Bill McKelvey (PhD, MIT, 1967) is a professor of strategic organizing and complexity science at the University of California, Los Angeles' (UCLA) Anderson School of Management. He has directed more than 170 field study teams on 6-month projects concerned with strategic and organizational improvements to client firms. His book, *Organizational Systematics* (1982), remains the definitive treatment of organizational taxonomy and evolution. He chaired the building committee that produced the US$110,000,000 Anderson complex. In 1997, he initiated activities leading to the founding of UCLA's Center for Human Complex Systems and Computational Social Science. He coedited *Variations in Organization Science* (with J. Baum, 1999), a special issue of *Emergence* (with S. Maguire, 1999), and a special issue of *Journal of Information Technology* (with J. Merali, 2006). He is the coeditor of *Sage Handbook of Complexity and Management* (2011), Editor of *Routledge Major Work: Complexity* (2012; five-volumes, 2000 pages.). He has 65 publications on complexity-related topics applied to organizations and management since 1997.