Chapter 14
Dynamic Robustness and Design in Nature and Artifact

Thomas Nickles

14.1 Introduction

My chapter applies a bit of qualitative risk analysis to systems of inquiry and their products. It extends Charles Perrow's theses about "normal accidents" in technological systems to epistemic systems, that is, to humanly constructed (explicitly or implicitly designed or engineered), evolved and evolving complex technological systems of *inquiry* and their products. My focus is on enterprises of ongoing scientific research at innovative frontiers. My central claims are: (1) Although the robustness (in Wimsatt's sense) of a scientific research program and/or its products is obviously highly desirable, no improvements in robustness can render these processes or their products invulnerable to failure. (2) On the contrary, such improvements can often, as far as we know, make inquiry systems vulnerable to new kinds of failure, sometimes worse failures than before. Robustness in one dimension can render a system more vulnerable to catastrophic change in another dimension. (3) Thus robustness is a relative rather than an absolute concept. Rather than vanquishing fragility, complex robustness can shift its location. More than that, increasing robustness (e.g., by adding new experimental or conceptual linkages) can actually generate fragility where none existed before. Accordingly, we cannot expect to make uniformly cumulative progress toward risk reduction. My argument can be construed as a (relatively new?) attack on foundationism and also on strong forms of convergent epistemic realism in the sciences. (4) In evolving epistemic systems with lookahead, *prospective* robustness is crucial to decision-making, contrary to traditional, retrospective empiricist theories of confirmation and robustness. I shall employ a broadly Kuhnian conception as my representation of mature science, that is, science with well-established problem-identifying and problem-solving routines. Any particular choice is, of course, controversial as well as arbitrary, but Kuhn's model is the best known.

T. Nickles (✉)
Department of Philosophy, University of Nevada, Reno, NV, USA
e-mail: nickles@unr.edu

L. Soler et al. (eds.), *Characterizing the Robustness of Science*, Boston Studies in the Philosophy of Science 292, DOI 10.1007/978-94-007-2759-5_14,

14.2 The Robustness-Fragility Tradeoff

In recent years the term ‘robustness’, its cognates and neighbors (solidity, persistence, hardiness, reliability, resilience, viability, flexibility, healthiness, etc.) have been applied to just about everything. In fact, ‘robust’ has become a buzzword in popular culture that can be applied to anything that exhibits strength of some sort. We want our aircraft to be as resilient as possible to turbulence and to various types of mechanical failure. We want our electrical grids to be robust to local failure, to unusually high power demands, and to insults such as severe storms. We want our nuclear plants to have monitoring and backup systems reliable enough to prevent the expected occasional malfunctions from leading to catastrophic accidents. We want our scientific research programs and experimental systems to be sensitive to empirical signals yet tolerant of modeling flexibility (abstraction, idealization, simplification, and approximation) and even errors of certain kinds. And so on.

We have made a lot of progress on all of these fronts. Thus the following question arises. While perfection is unattainable, can we not expect to approach it as an ideal, so that the only failures will be relatively small ones? Will not our progress toward greater robustness be cumulative as we successively minimize existing sources of error? Will not our successful error-reduction efforts converge on perfection in the limit? Let us term this *the cumulative fragility-reduction thesis* or *convergent risk-reduction thesis*, labels that are clumsy but descriptive. Surely the thesis is true and thus identifies a realizable methodological goal?

No, said organizational sociologist Charles Perrow in *Normal Accidents* (1984), not for technological systems (cf. Gertstein 2008). Perrow argued that the components of tightly coupled, complex technological systems will normally experience unexpected, untested, and practically untestable interaction effects. Such “interactive complexity,” as he called it, becomes apparent in accidents involving multiple failures, accidents such as the Apollo 13 space-module problems of 1970 (problems that, fortunately, were handled so as to avoid disaster); the DC-10 air crashes, including the Turkish Airlines plane over Paris in 1974; the partial meltdown of the nuclear reactor at Three Mile Island, Pennsylvania, in 1979; and the disastrous explosion at Chernobyl in 1986. A more recent example is Hurricane Katrina’s devastation of New Orleans in August 2005. In this case the very water control systems previously constructed by the U.S. Army Corps of Engineers and other agencies significantly worsened the failure by channeling the water in a destructive way. Still more recently, it is likely that the spring 2011 earthquake, tsunami, and nuclear disaster in Japan will be another case, once the details come to light. The deep recession of 2008–2011 was a failure of internationally linked economic systems apparently triggered by Wall Street’s opaque bundling of risky financial derivatives as well as governmental regulative laxity. Mutual defense treaties have similar features. The very attempt to improve a nation’s security can make it vulnerable via an attack on one of its friends.

Perrow’s second major contention is that these “accidents” should be considered a *normal* part of the operation of such systems rather than as highly contingent insults from outside the system. They are endogenous, not exogenous. They are intrinsic to the system and are to be expected, in a generic sense, although not, of course, specifically. (A hybrid case is accidents triggered by an external event, in which the response makes the situation worse.) Third, the operators of such systems should not be saddled automatically with the blame for failure, even if a precise sequence of procedures could have saved the day; for such failures are extremely confusing. Human beings are not omniscient and, typically, the event cascades are rapid and the incoming data and advice that the operators do get (as given by meter readings, indicator lights, warning horns, from experts employed by the manufacturers, etc.) are conflicting or otherwise unreliable. Failures of this sort, Perrow says, should be regarded as system failures, not operator errors.[1]

Perrow's fourth main point is that adding further safeguards against such errors only adds to the complexity and hence introduces new sorts of fragility into the system, even as it increases robustness elsewhere. For such additions typically increase exponentially the number of possible interaction effects, which, by nature are nonlinear.

Perrow's ironic conclusion is that there is a direct coupling of robustness to fragility. An increase in robustness does not mean an absolute decrease in fragility. Striving for greater robustness increases complexity that, in turn, commonly creates new paths for potential failure, including major malfunctions. Thus even successfully reducing the malfunction rate from expected sources does not result in a cumulative gain, as far as we can tell. There is a tradeoff. The very effort to eliminate fragility and catastrophic failure is, to a degree, self-undermining.

My purpose in this chapter is to extend Perrow's insight to epistemic systems, to systems of inquiry and their products. For they, too, can be regarded as complex technological systems. A theory or model can be regarded as a design, but so can a research program. A field can be represented by linked networks of several kinds involving personnel, equipment, "natural" materials, research designs, social support and demand systems, published papers, and the like. Or so I shall assume without argument.[2] I shall also assume that Wimsatt's insightful analysis captures a broad sense of robustness in the sciences, at least at a relatively high level of description (Wimsatt 1981, 2007, Chapter 7; see Chapter 10). The sense of robustness with which I shall be most concerned is one in which a system is sufficiently responsive to a variety of internal and external shocks that may befall it as to maintain its viability. This conception overlaps Wimsatt's in the sense that both involve a kind of invariance under changing circumstances.

[1] See also Perrow (1972, 2011). It is also relevant to mention here Herbert Simon's seminal work on bounded rationality and the behavior of complex organizations (Simon 1947). Bill Wimsatt and I share an admiration for Simon's work. As I once told Bill, I regard Simon as one of the great American pragmatists.

[2] Although I do not have space to defend this claim, most of the people represented in this volume surely accept some version of it. Incidentally, while large portions of the design are deliberately engineered in response to inputs from "nature," as is usual in cases of large human constructions, we should expect that modern sciences and their research programs contain important elements that were not explicitly designed, some of which we are surely unaware.

My central claim (combining claims 1 and 2 of my Introduction) is that the convergent risk-reduction thesis is false when applied to epistemic systems. The very attempt to control certain kinds of risk or error can generate new sources of risk that are rarely realized but that are generally more difficult to predict and to handle when they do. Robustness is engaged in an intricate dance with fragility. The implication for us is a heightened sense of the fallibility of our research systems and their products, an implication that makes my position on issues such as epistemological realism broadly compatible with those of Pickering and Soler (Pickering 1980, 1995; Chapter 10). I certainly do not deny that there has been rapid scientific progress on many fronts. However, I am not confident that even our most mature sciences tell us which entities and processes *really* populate our universe at bottom. A robust scientific realism that tells us what the universe is really like (thereby replacing the metaphysics of yore) is not at hand, especially when it comes to the very large and the very small. And the complexity-theory worries that I sketch here raise difficulties even for theories and models "of the middle range."

How is it possible that the risk-reduction thesis is false? How can increasing robustness at the same time decrease robustness? The obvious answer is that we must relativize robustness to specific dimensions or types of failure. And this is my third thesis. Robustness is not simply a matter of degree, as it is sometimes depicted). It is *relative* to specific kinds of external insults and perturbations and to internal breakdowns or constructive changes as well. Robustness is a relational property, represented by a two-place logical relation at a minimum and better as a three- or four-place relation. We can say that system s is robust to perturbation p to degree d except when c. In symbols, $R(s,p,d,c)$. The last clause may be optional, considered to fall under the usual ceteris paribus clause. However, operating manuals often provide explicit exceptions that seem stronger than ceteris paribus clauses.

I cannot of course prove that these theses hold in every case. Indeed, I don't think that they do. But I claim that they do hold for important kinds of epistemic systems, especially those such as we find in the sciences that place a premium on innovation and which, as a result, experience dynamic change over time. That is enough, I believe, to challenge the cumulativity thesis.

Whereas Perrow considered accidents in relatively static systems, once they are designed and built, we must consider also the crucially important case of systems that are designed to be dynamic in a strong sense—that contain intrinsic, endogenous sources structural change. How is it possible that such systems can be viable, that they can robustly maintain functionality throughout the change? For living systems, this question has both an ontogenetic and a phylogenetic form. How is it possible that biological organisms are robust to sometimes extreme developmental changes, as in the life-cycle of a butterfly? And how is it possible that genetic and other systems continue to operate reliably through sexual-recombinant and mutational change? Over time our human ancestors changed from small mammals to large, upright creatures, not to mention the earlier lives of plants and, long before that, single prokaryotic cells. In this sense we still face a version of the fixity-of-species problem: How is it possible that species can evolve? As Wagner (2005) points out, changing the genome (including patterns of gene regulation) changes the basic operating instructions for organisms, and these changes are heritable by offspring; so the fate not only of a particular individual but also that of an entire species hangs in the balance.[3]

The problem, then, is to understand how biological nature or intelligent beings can design a system robust to *internal* design changes, to technological upgrades, so to speak, occasionally even in deeply *entrenched* subsystems, without loss of functionality (Wimsatt 1981), and to appreciate the risk of catastrophic failure in such systems. Once we reject the foundationist impulse, we realize that epistemic research systems are strongly dynamic in this sense. Presumably, this holds for any enterprise that places a premium on suitably adaptive, creative innovation. As inquiry proceeds, even the deepest principles can be overturned and the enterprise restructured. Historically, many such enterprises have failed as a result, but others have survived major transformations. The Internet is a recent technological example of the latter sort (Willinger and Doyle 2005; Doyle et al. 2005).

Since the 1960s, the debate about scientific revolutions has turned partly on these issues. Those sympathetic to Thomas Kuhn's view in *The Structure of Scientific Revolutions* (1962, 1970) regard scientific revolutions as radical events akin to those political revolutions, such as the French Revolution, that throw out the old social order and send the community off in an unexpectedly different direction. These analysts are in turn challenged by those who stress continuity rather than failure. In any case, as Kuhn himself emphasized, scientific revolutions are extremely creative episodes in which the old paradigm is regarded not as a complete failure that ends the enterprise but, rather, as a failure only relative to a promising new approach that somewhat reinvents the enterprise, returning it to robustness as a progressive site of ongoing research. In the words of the economist Schumpeter (1942), we can identify "waves of creative destruction" in the history of science as well as in economic and technological history. Here we can distinguish enterprises that are truly left behind from those that reinvent themselves, maintaining a continuity of some sort. In the sciences, perhaps more than in technology and the general economy, the transformation is often enterprise-preserving creative destruction. I return to Kuhn's work in Section 14.9.

The topic of strongly dynamic epistemic systems leads me to my fourth thesis. Unlike the rest of nature, the human designers of epistemic systems possess a degree of lookahead. We humans can think about future prospects and consequences, make plans, and, to some degree, shape our own opportunities. Scientists in particular are capable of assessing the future promise of various alternatives and of making corresponding decisions today and then proceeding to reprogram themselves, so to speak. For systems involving creative inquiry, then, I want to suggest that there is a

[3] Wagner and co-researchers are engaged in an extensive research program to explore these questions involving robustness, embryological development, evolution, specialization, modularity, and the like. For example, Martin and Wagner (2008) discuss the tradeoffs in genetic networks that serve more than one function. They ask to what extent the need to serve several functions constrains the network architecture—and what effect such compromises may have on robustness.

prospective dimension to robustness, one sometimes so powerful as to trump *retrospective* judgments of robustness based on empirical track records such as reliability tests and multiple derivations of a result (Wimsatt 1981).

As a designer, biological evolution has major advantages over us in some respects. For example, almost every single organism (of the zillions on earth, past and present) "field tests" a distinct design variation; and nature is not constrained by our human horizons of imagination. But we have some advantages as well. For example, biological nature is more limited than human designers are in the respect that humans can often start over from scratch, undertake fundamental revisions, and combine previously distinct technologies. Furthermore, many of our artifacts do not require continuous functionality: we can tinker with them in the laboratory without releasing them into the field. The epistemic systems of the sort that I ultimately want to consider as part of a larger project are those that have managed to preserve functionality under transformative change, sometimes on the basis of lookahead. That is why I speak of *inquiring* systems: to bring out the fact that these are dynamic enterprises. The most important part of epistemology, in my opinion, is "frontier epistemology," the study of how knowledge practices grow at the frontiers of research and how the various modes of inquiry often manage to survive the surprises encountered there.

Perrow's claim about normal accidents is interesting for other reasons, for a normal accident is, somehow, less contingent, less accidental, than an accident as commonly understood. This point about relative contingency can be generalized and has important implications for several important problems, especially those concerning dynamical change in a system. For example, I believe that Perrow's insight concerning normal accidents as intrinsic to complex systems has a happier bearing on the problem of endogenous innovation. Accidents are usually considered exogenous—system insults from outside, mere contingencies rather than *systemic* features. Whereas for Perrow most accidents are normal, part of the behavior arising from within such systems, part of the normal background noise of complex systems. Given the extreme nonlinearity of such systems, an ordinary event such as a valve getting stuck or accidentally being left closed after testing can cascade into the meltdown of a nuclear reactor.

I extend this insight toward an endogenous account of *innovation*, which is also highly nonlinear and unpredictable, but dependent on a sort of ordinary background noise—that is, variation—intrinsic to the system. In some cases, a relatively ordinary or "normal" sort of innovation can have revolutionary consequences, producing, now, an exciting cascade of problem *solutions* or technical *innovations* rather than disaster. Here we have returned to the problem of Kuhnian revolutions. Members of the old guard may indeed consider a revolution a disaster, as destruction, while those in the vanguard see it as a creative continuation of the overall enterprise. There can, of course, be breakthroughs that are not destructive of the received platform.

14.3 Tradeoffs Between Robustness and Fragility: A Variety of Examples

First a political example, the reign of Louis XIV (d. 1715), who supposedly asserted *L'état c'est moi* and insisted that a united, coherent nation-state must have *un roi, une loi, une foi*. At the time the nation-state was a relatively new entity and violent power struggles were common. (In his social contract theory of the state published a few decades before, Hobbes had noted that a single powerful ruler was a stabler arrangement than a ruling body of two or more persons.) Under these circumstances, the rigid hierarchy that Louis imposed created a robust system that minimized the danger of both external and internal threats via a triple structure of military, legal, and religious power. And yet that very hierarchy made the system fragile or brittle in another respect. Strongly hierarchical, "command and control," hereditary regimes such as Louis XIV's are robust to confusions about chain-of-command; but they are fragile to questions of succession and to failure at the top, for there is nothing in the system to guarantee a strong leader. This turned out to be the case in French history. After Louis' death, the succeeding Louises were weak and vacillating. The story is of course very complicated, but the cumulative result of many factors was the French Revolution. Complexity theorists Levin et al. (1998) observe that, still today, rigidity is often taken to be the mark of robustness in social systems. Such systems withstand the forces of change, but this ultimately makes them seem antiquated. When they fall, they tend to fall quickly. The fall of the Soviet empire is a recent example.

As noted, one of Perrow's own primary examples was nuclear power plants. Designing complex safety mechanisms and backup systems certainly improves robustness against anticipated failures, but it creates new routes to unexpected failures, e.g., when a failure of one system masks failures in others, including human error.

The "Six Sigma" (6σ) quality control program in industry aims to reduce variation in manufactured products so that there are fewer than 3.4 failures per million.[4] Here 'robust' production clearly means elimination of unwanted variation, noise, and waste. Later we shall have reason to question the assumption of a normal distribution in some such processes. However, the following illustration has a different slant. Consider the recent change in the 3M Corporation (formerly Minnesota Mining and Manufacturing), makers of many kinds of tape and thousands of other products. 3M has a history of encouraging innovation at all levels, e.g., by building innovation time into many employees' contracts and setting as a company goal that 30% of annual sales will be of products introduced within the past 4 years. The result has been a robust corporation that has flourished by repeatedly reinventing itself for over one hundred years, far outlasting many competitors. In an effort to make the corporation even more robust, its products more reliable, the new CEO of

[4] Actually, this failure rate calculates to 4.5 "sigmas" or standard deviations from the mean value, assuming a normal distribution. The extra 1.5 sigmas are imposed to allow for long-term variation, given that the approach is evidence-based and that most evidence is short-term.

3M recently implemented Six Sigma quality control and a rigid system of employee evaluation. Unfortunately for 3M, there has been a tradeoff: less time and incentive for innovative *intra*preneurship (Hindo 2007). Innovation is typically a process that looks sloppy and inefficient from the outside, since research and development depends heavily on variation-selection processes with many false starts. It requires slack. It requires time. The primary cost of new products is in the design stage, not in final production cost per unit. Rigid quality control in the generative process can have serious long-term consequences for innovation.

For a classic epistemological example, consider Descartes. His foundationism had the goal of building knowledge structures that are impervious to future criticism and thus guaranteed to last. The model of a finished system was Euclid's geometry, an axiomatic deductive system.[5] Although such a system is robust to failures of logical entailment and can deductively systematize a remarkable amount of content by means of a few postulates and inference rules, we can identify several sources of fragility or brittleness in such systems—and in Descartes' in particular. (1) His starting principles were not as perfectly robust ("clear and distinct") as he believed. (2) Several of his inferential steps are deductively questionable, and there is the well-known problem of circularity in the early steps of his system. (3) The extreme verticality of the Cartesian foundational hierarchy is itself cause for concern, as Fig. 14.1 immediately suggests. The system teeters in unstable equilibrium.

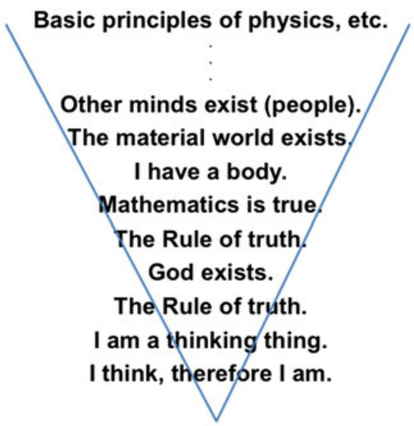

Fig. 14.1 The supposed Archimedean point of Descartes' system is the *cogito ergo sum* at the *bottom*, which allegedly defeats the most severe forms of skepticism. Then come, in order, the other steps of *Discourse on Method* and *Meditations*: I am a thinking thing, the Rule of Truth (Everything that I perceive clearly and distinctly must be true), etc. The entire enterprise is supposed to provide an escape from history (thus from historical path-dependence other than its own, intrinsic, cumulative history of logical development) in the sense of giving us epistemic foundations so robust that they are impervious to the ravages of time and circumstance. (Figure by the author)

[5] Descartes admitted algebraic derivations as well as geometrical ones, and his use of 'derive' was more liberal than that allowed by the later, formal concept of deduction.

(4) Shockingly, failure of types (1) or (2) propagates uncertainty *instantly* through the remainder of the system, even if—or especially if—those couplings are tight, that is, deductively valid. Ironically, the very thing that makes the system epistemologically well founded, highly integrated or systematized, and economical, with instantaneous transmission of derivability from starting points to manifold end points—and thus so robust in these respects—makes it fragile in another respect. “Instantaneous logical action-at-a-distance,” as we might call it, can be a very good thing—or a very bad thing. Tradeoffs! The couplings of the propositions are so tight that an error anywhere produces a disastrous cascade of failure through everything else that depends on that step. There is no way to stop it. The system is vulnerable to epidemic contagion of error, or at least uncertainty (which was tantamount to error for Descartes-the-foundationist). However, things are not so bad for methodologists who believe that empirical support comes from consequential testing rather than from antecedent premises. For them strong deductive coupling has the advantage that a predictive failure propagates backward in such a way as to help them to root out fundamental error, modulo Pierre Duhem’s fault-assignment problem.[6]

(5) Our confidence in a system that pretends to be failsafe plummets when even a small failure is detected. This is surely one reason why some critics find it (too) easy to discredit modern science. In their view, scientific knowledge claims are supposed to be nearly infallible; so, for them, every reversal of scientific judgment is damning.

The chapters of this volume contain several examples from recent experimental research of control systems for robustness. Soler and other contributors, adding to the growing science studies literature, are advancing our understanding by showing how many, how esoteric, and how intricately coupled are the investigative techniques, regulative procedures, standards, and professional judgments involved in establishing or negotiating an experimental claim. This in contrast to those early philosophers of science who considered “experimental observation” relatively unproblematic, epistemologically uninteresting, and requiring no special attention. Concerning one stage of the Gargamelle search for weak neutral currents in which experts checked the tracks of nearly 300,000 photographs, Soler writes:

> [T]he aim is to extract a set of tracks whose interpretation is globally more reliable, in such a way that it becomes less likely to make mistakes in the counting of the potential tracks of NC [processes identifiable as neutral currents].
>
> But if this is the aim, experimentalists are never sure that it is indeed achieved. The filtering operations aim at eliminating some confusions, but they can themselves be sources of mistakes.

The elaborate control systems for detecting and avoiding error that we build into our experimental technologies open up new routes of potential error.

In epistemology and methodology, philosophical thinkers have tried for centuries, without success, to develop fail-safe systems, systems and methods so robust

[6] I refer to Duhem (1954) on the need for auxiliary assumptions in any predictive inference and the resulting difficulty of pinning the blame for failure on any one premise.

that all threats, all sources of failure, are eliminated.[7] Operationism as a method of concept formation is a recent example.[8] By now, however, most thinkers have declared themselves to be fallibilists of one kind or another, ranging from still-confident, strong epistemological realists to skeptical antirealists. But the strong realists would appear to be weak fallibilists in the sense that they apparently believe that, although we cannot totally eliminate error, we can identify its sources and cumulatively reduce its occurrence without thereby creating new sources of error. It is this cumulativity thesis of error reduction that I am combating, following Perrow and the complexity theorists discussed below. If I am right, then, despite the tremendous advances of recent, mature science, we must remain thoroughly fallibilistic—in part *because of* those very successes! Karl Popper was a thoroughgoing fallibilist (e.g., Popper 1963), but even he based his claims for verisimilitude (approach to the truth) and realism, in part, upon this idea of gradual error elimination.

Wimsatt (1981), citing Feynman (1965), distinguishes linear, "Euclidean" intellectual structures from "Babylonian" structures (see Chapter 8). The latter are multiply connected. In a Babylonian structure, such as a truss bridge, if one link fails, the structure does not collapse. In the intellectual counterpart, if you forget one way of deriving a result, you can resort to other ways, since they are interconnected. Your "failure" is contained. Moreover, unlike the Euclidean intellectual model, the support is mutual to varying degrees. Justification does not flow in one direction only, from foundational axioms to theorems (or inductively, from good test results to hypotheses). As Putnam (1962) observed long ago, it can go in many ways at once.

Thus we come to the idea that there can be types of virtuous circularity as opposed to the vicious circularity inherent in Descartes' system. In logical-semantic structures as in causal structures we can have mutual support. Pickering (1995) develops this idea in terms of symbiosis. Today's network analysts frequently deal with what they sometimes call "circular causality."

[7] Here we can include not only philosophical and scientific systems based on reason and evidence but also ethical and religious systems based on revelation, mind-control, and various other self-protecting contagions and social viruses discussed by meme theorists. The most secure, deeply entrenched of these is, according to the line being developed here, vulnerable to catastrophic collapse. Gaye McCollum-Nickles reminds me here of Oliver Wendell Holmes, Jr.'s poem, "The Deacon's Masterpiece or, the Wonderful 'One-hoss Shay': A Logical Story," about the sudden collapse of Calvinism in late 19th-century America. Pickering (Chapter 13) regards standalone, material machines as the hallmark of modernity. Also falling within modernity is the extension of this idea to Bacon's and Descartes' (unsuccessful) attempts to mechanize scientific procedures (the alleged discovery of "the scientific method"). As a response to an unpredictable world, the idea of "standalone" ethical systems invulnerable to the arrows of fortune is very old. Such were the systems of the ancient Stoics and Epicureans.

[8] Bridgman (1927), one of the founders of operationism, blamed the need for the relativity revolution on physicists' failure to operationally define their concepts (especially simultaneity) prior to theorizing—as if we could work out a robust system of concepts prior to theory in work at the frontier of research!

We must be careful, however, for, insofar as a Babylonian *intellectual* structure is tightly coupled by deductive relations (or something similar), error or uncertainty can spread rapidly there as well—in fact "in all directions at once" in which there are tight structural connections. To be sure, scientists possess various devices—buffers, spacing, and security walls—for containing failure in order to avoid devastating cascades. This point has been noted by several authors, e.g., Quine (1951) on the web of belief with its centralities and priorities, Lakatos (1970) on research programs with their "protective belts," and Wimsatt (1981, 2007) on walling off difficulties. But if I am right, nothing that we can humanly do can prevent occasional, surprising avalanches of failure. At the very least, the topic deserves further study.

14.4 Highly Optimized Tolerance (HOT)

Jean Carlson and John Doyle (physical scientists at the University of California, Santa Barbara, and Caltech, respectively) utilize the latest tools of network theory and percolation theory to develop a Perrow-like thesis (Carlson and Doyle 1999, 2002). They speak of "spirals" of complexity. Attempts to improve robustness lead to greater complexity which in turn generates new kinds of fragility, which lead to new security mechanisms, and so on. This amounts to a different sort of "arm's race" than the traditional one between predator and prey, or at least provides a different perspective on the latter as a special case.[9] Carlson and Doyle see the quest for robust viability as the driving cause of complexity. For them pressure for greater robustness *explains* complexity in a sense that generalizes evolutionary biology to include technological design: greater robustness is greater "fitness."

So why not keep it simple? Why start the complexity spiral in the first place? Why not just stick to simple, reliable systems?

Carlson and Doyle make an obvious but useful distinction between two kinds of robustness: (1) simple systems made of a few highly reliable components and (2) complex systems of "sloppy," cheap components, where the robustness is an emergent, systemic feature deriving from backup systems, extensive monitoring, automated, computerized control, and so on. The problem with simple robustness is that it is just too simple. Such systems operate in too narrow a range, with systemic failure looming beyond. Relative to a desired wider range of operation, complex systems can be far more robust. As Carlson and Doyle point out, a Boeing 777 is more robust to variable weather than a two-person airplane with simple instruments. Besides, nothing is failsafe anyway. Even simple systems sometimes fail. When a

[9] Since robustness includes resilience to environmental shocks in addition to predator-prey issues, even in biology, the phenomenon is more general than the usual sort of arms race in which, e.g., long legs for greater speed increase the fragility of the legs. Our primary concern here is with human technological systems, including epistemic systems. We can construe the race as an attempt to identify and avoid possible new sorts of accidents before they happen, or before they happen again.

component of a simple system fails, there can be no cascading failure, since there is no complexity to support a cascade; but the failure is typically disastrous in any case since simple systems are not likely to degrade gracefully. Complex systems with redundancy, networking, and/or high-tech monitoring and feedback control can typically minimize the damage from isolated, random component failure, because they do not demand consistently high tolerance in individual components.

Highly designed systems are more expensive, but often the additional efficiency or yield more than compensates. This is a second important feature of their HOT systems, those manifesting highly optimized tolerance, namely, the pressure for optimization. The physiology of an elephant is far more efficient than that of an ant, and a Boeing 777 is more efficient (in load carried per amount of fuel) than a Piper Cub.[10] And, again, complex systems can employ sloppier, hence cheaper components.

The upshot of all this is that we cannot keep it simple: we must deal with complex systems. The demand for robustness requires complexification of the control structures. The question then becomes which architectures are more robust than others. The general strategy that Carlson and Doyle take is not to avoid all failure, which is impossible, but to prevent *anticipated* types of failure from becoming so highly contagious that an epidemic or failure cascade ensues. And the way to do that is to maintain adequate spacing, metaphorically speaking, around danger areas, to install buffer zones in order to confine the damage. Their favorite illustrative model (a standard one) is forest fires, where the spacing is quite literal. Zones subject to high lightning strike rates or to heavy human (mis)use should have buffer zones around them to keep any fire from spreading very far.

Installing large enough buffer zones will, of course, nearly always prevent devastating fires. However, that recourse also cuts down on timber production (or the amount of forested land preserved for any reason, e.g., as an ecological system). So the problem becomes how to optimize "throughput" or productivity or yield (the harvested wood in their toy model) while maintaining adequate robustness. Maximizing productivity or overall "fitness" is where the 'highly optimized' of their Highly Optimized Tolerance (HOT) research program comes from. The tolerance refers to the robustness.

Carlson and Doyle maintain that maximal productivity does not occur in merely physical systems, for it requires *design*, either biological evolutionary design or deliberate engineering design as in human technology.[11]

---

[10] See Geoffrey West's lecture, with slides, available at http://online.itp.ucsb.edu/online/pattern_i03/west/. Power laws apparently characterize the metabolism rate, energy use, extinction rates, and many other aspects of the animal and plant worlds. For criticism, see Downs et al. (2008). For a survey of leading models of species extinction, see Newman and Palmer (2003).

[11] In my opinion the distinction between novel "design" by natural selection and intelligent design by human engineering is usually exaggerated. While there are important differences, at bottom both are selectionist processes, that is, variation-plus-selection processes. See Nickles (2003 and forthcoming).

> The central claim from our model is that the essence of this robustness, and hence of biological complexity, is the elaboration of highly structured mechanisms that create barriers to cascading failure events. (Zhou et al. 2002, p. 2053)

They argue that highly specialized design can provide remarkable robustness for anticipated failures but is extremely fragile to *unanticipated* kinds of changes, external and internal, as well as to design errors. For example, biological species that track their environment very closely become "specialists" vulnerable to environmental change, at which point the "generalists" will survive or move in and take over (Calvin 2002).

Carlson and Doyle defend their claims with a technical analysis that I cannot repeat here. A central point is that the "size" of failures in complex systems in HOT states, that is, systems that are highly optimized in the indicated manner, have distributions with "heavy tails" (or "fat tails" or "long tails," as they are also called). Heavy tails of just the right shape are the signature of power laws. Heavy-tailed probability distributions decay at slower rates than the exponential drop-off characteristic of Gaussian normal distributions. The former are sub-exponential over at least part of the tail.

With a Gaussian distribution the three main kinds of averages (mean, median, and mode) coincide, and the rapid drop-off means that very few cases exist that are more than three standard deviations from the average (the more so as the peak is narrower). Thus it makes good sense to speak of a "typical" item or event, namely, one close to average. This determines the "scale" of the phenomenon. The mean together with a well-defined variance or spread of the distribution provide a neat, two-parameter summary of the phenomenon in question.

Not so with heavy-tailed distributions. These are characteristic of scale-free (scale-invariant) phenomena. Here one can get large events as far out on the tails as you please, with non-negligible probability. There is no such thing as a typical size. The variance is no longer well defined; in effect, it becomes infinite, modulo the limits of system size. The bad news is that, when the events in question are failures, failures of any magnitude are realistically possible. There is no preventing occasional disaster. To illustrate the difference: human height is normally distributed. It is extremely improbable to encounter a human who is eight feet tall and virtually impossible to meet one ten feet tall; but huge earthquakes, although rare, are not "virtually impossible." And—to revert to a previous example—insofar as the reliability of industrial processes can be characterized by distributions with heavy tails, it makes no sense to speak of "six sigma" reliability. For heavy-tailed processes, traditional risk analysis underestimates the true danger—the probability that an event with a large negative utility will occur.

In sum, Carlson and Doyle offer a technical understanding of Perrow's compromise. The cost of controlling robustness in such a way as to maximize throughput is susceptibility to unanticipated sorts of disasters. But Carlson and Doyle perhaps go further than Perrow to consider also constructive changes internal to the complex systems in question. The problem of how a system can withstand variability of its components is part of the problem of managing failure of traditional, static systems; but how account for the viability, the flexibility, the adaptive capacity or evolvability, of complex systems that undergo basic design changes that could improve performance? After all, the process of optimization is itself an ongoing process. As noted above, Andreas Wagner confronts this problem for biological evolution. Doyle and Carlson, being physical scientists, take the Internet as their chief example of such a system (Doyle 2005; Willinger and Doyle 2005), showing how the system can be robust to rapid technological changes at the applications level. A similar problem arises for an ongoing scientific research program, e.g., in what Kuhn calls normal science. Philosophers have worried most about continuity through scientific revolution (incommensurability and all that), but for Kuhn it was the strong continuity of normal scientific research that most required explanation. Kuhn's own solution to this problem was an overly static treatment of normal science, one that minimized internal change (Nickles, forthcoming).

14.5 Power Laws and Their Implications

A simple form of a power law distribution is $N(x) = ce^{-\delta x}$, which means that the number of events N of size x or greater equals a constant c times the exponential function. δ is also a constant, a scaling parameter. Power laws[12] and their distributions are scale invariant or "scale free." As indicated, earthquakes provide a familiar example. According to the Gutenberg-Richter law in its simplest form, $N(m) = ce^{-m}$, where m is the magnitude of the earthquake. If we take the logarithm of both sides of power law equations, we get the formula of a straight line as shown in Fig. 14.2. Thus a straight line on a log-log graph is the signature of a power law.

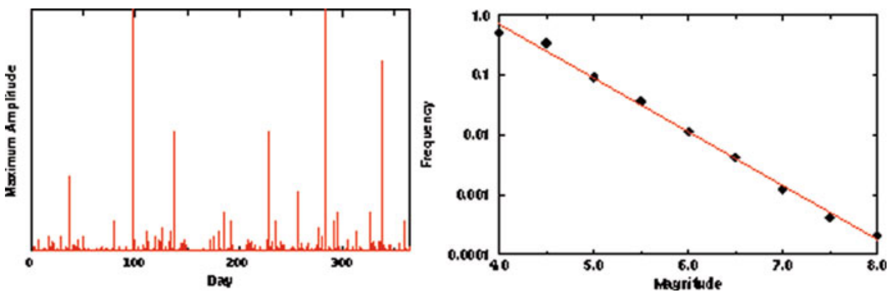

Fig. 14.2 On the *left* is the picture of a time series of earthquakes, plotted according to size. On the *right* is the corresponding log-log graph. Note that the slope parameter δ in this case is approximately –1. Apparently, the slope of the graph varies a bit, according to geographical region, but it is always around –1. From http://simscience.org/crackling/Advanced/Earthquakes/GutenbergRichter.html

[12] A generic form of common power laws is $f(x) = cx^{\delta} + o(x^{\delta})$, where c is the constant, δ is the scaling factor, as before, and o is an asymptotically small function that captures small deviations or uncertainties.

Scientists and other observers long thought of many natural phenomena as the product of large numbers of small, independent events, uncorrelated as to "direction." More recently, this view has begun to give way to a conception of nature as a more highly integrated, complex system than even the "mechanistic universe" of modern physics implies. Especially since the 1960s, power laws have been turning up seemingly everywhere—at least everywhere that complex systems are found—with biological organisms, human artifacts, and social systems being the primary examples. For example, the metabolic rate of biological organisms, from smallest to largest, possibly fits a power law (but see Downs et al. 2008). Ditto for heart rates and so on and on. Hungarian physicist Albert-László Barabási and his group have discovered many (alleged) power laws governing the Internet and various human social networks such as scientists working in a particular specialty area. Power laws turn out to characterize physico-chemical processes at phase transition points (critical points), as when water is about to freeze or turn from liquid to steam or when a piece of iron is about to become magnetized.

The frequent appearance of power laws in complex systems and the importance of phase transitions in many fields raises the interesting question of whether there is a common mechanism behind them, a deep structure underlying the natural phenomena about which we can formulate a general theory. Barabási is convinced that there is such a science of order and connectivity, a topic to which I return in Section 14.8.

14.6 HOT Versus SOC

Several scientists and mathematicians have attempted to formulate a general theoretical approach to complex systems. Given that complexity of one sort or another appears across many scientific disciplines, especially in the biological and social sciences and engineering, a trans-disciplinary theory of complex systems would be an enormous achievement. However, there is wide disagreement about the prospects for such a theory and even about what a complex system is (Gershenson 2008). In this section I briefly compare the approaches of Per Bak's self-organized criticality (SOC) with Carlson and Doyle's highly optimized tolerance (HOT). In the following sections I look more closely at the work of Barabási.

From the 1980s the late Danish physicist Per Bak argued for the importance of "self-organized criticality" (SOC) as the successor to early work on autocatalytic systems and other forms of self-organization at the edge of chaos (EOC) proposed by Stuart Kauffman (1993) and others. Although Bak worked with Kauffman for a time at the Santa Fe Institute, he denied that Kauffman's well-known NK model attains true criticality.[13] Bak's central claim, defended at length in *How Nature*

[13] See Kauffman (1993), Chapter 2 *et passim*. The NK model is Kauffman's start toward "a statistical mechanics of fitness landscapes" (1993, p. 40). "N refers to the number of parts of a system—genes in a genotype, amino acids in a protein, or otherwise. Each part makes a fitness contribution which depends upon that part and upon K other parts among the N."

Works (1996), is that all truly complex systems exist in a critical transition state between order and chaos. Ordinary chaos theory cannot explain complexity, he said. All and only complex systems exhibit his self-organized criticality. His favorite model was the sand pile. Grains of sand are dropped one by one, forming a pile. Eventually, the slopes become steep enough that one additional grain will start an avalanche, usually a small one but occasionally a very large one. When a histogram of avalanches from smallest (displacement of a single grain of sand) to the largest (when perhaps a third of the pile slips away) is plotted, the result is a straight-line, log-log graph, signifying the working of a power law.

Power laws are pervasive in biological nature because evolutionary adaptation drives systems to a critical point, Bak argued. “Biological evolution is a self-organized, critical phenomenon” (1996, p. 150). (But, as noted above, some nonbiological, non-intentional systems, including growing sand piles, also exhibit this sort of behavior, namely at critical points where phase transitions occur.) Moreover, for Bak the power laws indicate a scale-free phenomenon that signals self-similarity and that possesses a fractal signature. Bak explicitly rejected models that are specially “tuned” by an outside designer to achieve critical states. That’s where the ‘self-organized’ comes in: the systems must themselves move toward the critical state, where a certain sort of equilibrium obtains.

Carlson and Doyle (among others such as physicist Geoffrey West, former president of the Santa Fe Institute) contend that Bak’s account of complexity is now dated. Carlson and Doyle naturally claim that their HOT supersedes Bak’s SOC. They insist that massive empirical information from the physical, biological, and engineering-control worlds shows that SOC theory does *not* capture “how nature works.” They disagree with Bak even about what counts as a complex system. Bak clearly wanted to include not only biological but also purely physical systems, systems that manifestly involve no design of any kind. But in throwing out design as a necessary condition, contend Carlson and Doyle, Bak threw out the baby with the bathwater. Their HOT systems, after all, are highly designed, either by biological evolution or by human engineering.

> Highly Optimized Tolerance (HOT), which links complexity to robustness in designed systems, arises naturally through Darwinian mechanisms. . . . [R]obustness tradeoffs [are] a mechanism that drives complexity in biology. (Zhou et al. 2002, p. 2049)

> [The model is] sufficiently general that it could be equally well motivated by competition and evolution in other settings, such as between technologies and companies in an economic setting. (Ibid., p. 2050)

While both SOC and HOT differ from statistical mechanical accounts of power laws by linking the power laws to internal structure, Carlson and Doyle (1999) go on to locate HOT at the opposite extreme from SOC, in several ways. (a) A sand pile is boringly self-similar, a mere aggregate, but a HOT system is not. The specialized, highly-engineered systems of a Boeing 777 resemble neither each other nor the airplane as a whole. Ditto the specialized subsystems of biological organisms.

(b) Robust, optimized natural and artificial systems designed by a selection process or by engineers do not necessarily exist at a point of criticality between order and chaos. In general, HOT states are not critical states. (c) HOT systems typically optimize several parameters at once, not just one as in Bak's models. (d) Thus robustness is an *emergent* property of the growth of complex systems, whether these systems are designed by human engineers or by natural selection. (e) The power laws of HOT are "steeper" than those of SOC.

> [T]he HOT power laws are steeper and extend to larger event sizes than the critical power laws, which are very flat. Large events at criticality are fractal, resulting in no macroscopic losses in the limit of large lattices. This is in contrast to both our model and the fossil record, which show losses that are a large fraction of the total organisms or species. (Zhou et al. 2002, p. 2054)

Carlson and Doyle note the policy implications of our choice of complexity model:

> There is much at stake in this debate. If ecosystems are in a SOC/EOC state [i.e., state of self-organized criticality/edge-of-chaos—TN], then observations of massive species extinctions and global warming could be attributed to the natural behavior of the system. In this scenario, large fluctuations emerge and recede as a natural consequence of the internal dynamics, and would not be attributed to man made causes. This would support a policy in which humans could be relatively cavalier about their interactions with the environment, because the system would be fluctuating as observed regardless of our behavior. Alternatively, if ecosystems are in a HOT state then we expect the system to be robust, yet fragile. Heavy tailed distributions are expected, but the system is also hypersensitive to new perturbations that were not part of the evolutionary history. (Carlson and Doyle 1999, p. 1426)

Barabási likewise rejects Bak, saying:

> Networks are not en route from a random to an ordered state. Neither are they at the edge of randomness and chaos. Rather, the scale-free topology is evidence of organizing principles acting at each stage of the network formation process. There is little mystery here, since growth and preferential attachment can explain the basic features of the networks seen in nature. No matter how large and complex a network becomes, as long as preferential attachment and growth are present it will maintain its hub-dominated scale-free topology. (2002, p. 91)

We'll return to preferential attachment in a moment. My present observation is that while Bak's sand pile model was a dynamic approach to self-organized complexity, apparently it was not dynamic enough. According to its critics, it misses crucial details of growth. It pays insufficient attention to both the ongoing processes of generation and corruption and the specialization of the resulting structure. To be fair, Bak claims (p. 150) that biological evolution moves organisms toward what he terms a critical state, but, as we have seen, his critics deny that sandpile type models can really capture this.

14.7 A Bit of Network Theory

As mentioned above, Barabási claims that his research group has founded a new, rigorous scientific inter-discipline of complex systems, one that unveils the deep structure common to all such systems, from biology to business.[14] The brilliant Hungarian mathematicians Paul Erdős and Alfréd Rényi gave the field a start in the 1950s with a series of papers on the connectivity characteristics of random graphs, that is, graphs whose nodes or vertices are connected randomly. Adding links to such graphs, one by one, between randomly chosen nodes, brings the net to a critical state in which adding a couple more links suddenly transforms it into one that is fully connected, i.e., each node can now reach every other node through a path consisting of a series of links or "edges." (One thinks here of the "tipping point" phenomenon popularized by Malcolm Gladwell, 2000.) Conversely, subtracting just a few links randomly can turn a well connected network into a highly fragmented one. This rapid increase or decay is exponential. Thus the normal curve (more precisely, the Poisson distribution in this case) with its exponential drop-off is the signature for random networks just as for random, independent events in nature.

Later Steven Strogatz at Cornell and his former student, Duncan Watts of Columbia University, became interested in applying graph theory to real-world systems. They were especially interested in "small world phenomena" in which almost any two people or points are at most a few links apart but in which there are local clusters of highly interconnected nodes. Efficiency in many systems requires a small number of "degrees of separation," but, for that very reason, such systems are also vulnerable to epidemic contagion. To achieve small worlds with clustering, Strogatz and Watts started from a ring lattice of nodes connected only to their neighbors as in Fig. 14.3, then randomly added a few long-distance links across the circle (Strogatz 2003; Watts 1999). The remarkable result was that just a few additional links tremendously reduced the average degree of separation of the network. By combining the order of the initial lattice with a dash of randomness, they achieved small world networks that were also clustered. However, the element of randomness was enough for them to retain an exponential signature.

The third stage in this development was accomplished by Barabási and his students (especially Réka Albert and Hawoong Jeong), who concluded that real networks are usually even more clustered, containing a few "hubs" with a very large number of connections. This network topology proved to be far more robust than the others. Since the vast majority of nodes are not hubs, now randomly disconnecting a majority of links would produce nothing worse than a graceful degradation of the network. And the signature of these network topologies is . . . power laws and their fat-tailed distributions! In other words, these networks are scale free (up to the limit

[14] There is already a large literature, both technical and popular, on these developments. Newman et al. (2006) is a collection of many of the most influential papers to that date. See also the Carlson and Doyle articles, Watts (1999), and Jen (2005). Among the popular or semi-popular works, see Buchanan (2002, 2007), Miller and Page (2007), and Barabási (2002). The websites of many of these people often contain additional resources.

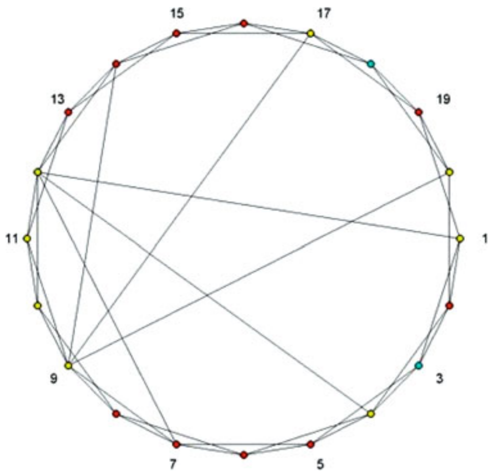

Fig. 14.3 A graph of the Strogatz–Watts type that combines the features of local clustering (by starting from a periodic ring lattice in which only nearest neighbors are connected) and randomness (by adding a few random connections)[15]

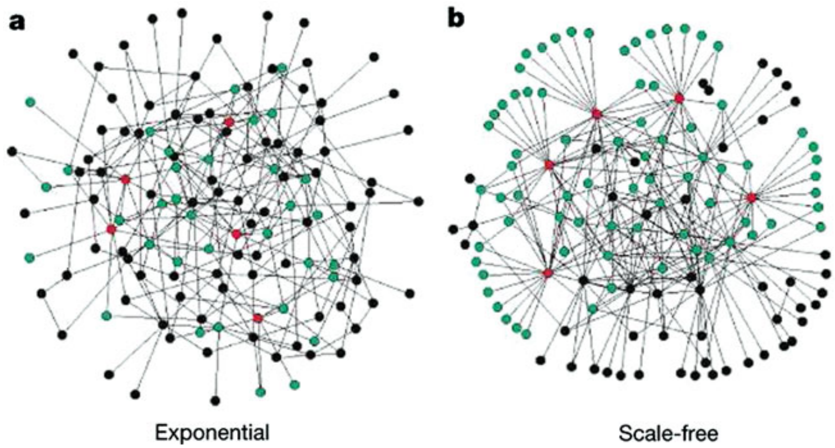

Fig. 14.4 (**a**) The exponential [random] network is homogeneous: most nodes have approximately the same number of links. (**b**) The scale-free network [with hubs] is inhomogeneous: the majority of the nodes have one or two links but a few nodes have a large number of links, guaranteeing that the system is fully connected. *Red*, the five nodes with the highest number of links; *green*, their first neighbours. Although in the exponential network only 27% of the nodes are reached by the five most connected nodes, in the scale-free network more than 60% are reached, demonstrating the importance of the connected nodes in the scale-free network. Both networks contain 130 nodes and 215 links. The lighter nodes in greyscale are the green ones. The red and black ones are not distinguishable. [The network diagrams are taken from Albert et al. 2000, p. 379.][16]

[15] The graph can be found at www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=4206&objectType=file.

[16] For the color version, the reader may consult Albert et al. (2000) or http://www.computerworld.com/s/article/75539/Scale_Free_Networks.

of their size). They look the same at all scales. Figure 14.4 is a comparison diagram from Barabási's group (Albert et al. 2000), with their own explanation.

The Barabási group and other experts around the world soon found numerous examples of this sort of network in the biological, technological, and social worlds. According to them, biological food webs, gene networks, and protein networks have this structure. So do commercial airline routes, national electrical grids, and the Internet. And in social worlds, AIDS spreads via such a network, which is extremely efficient in connectivity yet highly robust to random failure.

Is there still a fragility compromise? Yes there is. Where is such a network vulnerable to attack? Answer: in the hubs. While scale-free networks are robust to random failure and also to various kinds of upgrading, as, for example, with the Internet (Willinger and Doyle 2005), they can fail disastrously under targeted attack. Attacks on the hubs can quickly produce a fragmented residue of a once well-connected network. That is how a few "crackers" (the label given to evil hackers as opposed to good hackers) have been able to bring the Internet to a halt with some simple viruses and worms. The net structure allows bad stuff to spread just as efficiently as good stuff.

The happy side of this worrisome feature is that sometimes we do want to destroy networks. For example, Barabási and his colleagues have suggested that the most effective way to treat AIDS and similar diseases, given the expense of the available drugs, is to go after the hubs. The hubs are Malcolm Gladwell's "connectors" (Gladwell 2000, Chapter 2), the people who have done far more than any others to spread the disease. This strategy might also apply to attacking international terrorism.

The fat tails phenomenon suggests another bit of good news for our efforts to understand innovation more endogenously (see Section 14.10 below). Just as systemic failure can always take us by surprise, so can systemic success, including unimagined successes of large magnitude. The trouble is, there are a lot more ways to fail than to succeed. And, again, we should not confuse the two types of failure—getting efficient results that we do not want (such as the spread of a virus) versus disintegration of the network itself.

A second word of caution is in order here. Although these investigators have made great progress in exploring the mathematical and causal structures that underlie networks, the nets themselves are much too simple to capture a lot of rich interaction. In real-world networks, as opposed to their human representation in idealized models, not only are the nodes typically directed (corresponding to a directed graph) but also there are nodes and links of many different kinds. For example, not all human-to-human links are equal. This becomes obvious immediately when we play the "degrees of separation" game. "How many degrees of separation do you have from Einstein?" Well, that obviously depends on what counts as a direct *connection* with (i.e., one degree of separation from) another person. I once shared an elevator ride with Kurt Gödel, who was one degree from Einstein; but as a young graduate student I was too shy to say a single word. Does that count as a connection? Am I only one degree of separation from Gödel and two from Einstein? Obviously not if that means any sort of personal acquaintance, let alone being on a "first name basis," which is the criterion sometimes used. Newman (2001) defines two scientists as directly connected only if they have co-authored a paper.

Nonetheless, the simple networks explored so far have shed considerable light on the topic of robustness and fragility and tend to confirm Perrow's original insight that we must always consider tradeoffs. This approach has given real direction to research programs across the natural and human sciences, engineering, and business.

However, a third important caution is that critics such as John Horgan and Evelyn Fox Keller argue that the enthusiasm for a general complexity theory is premature and that claims for the ubiquity of power laws and scale-free networks are unjustified, amounting to something of an intellectual fad (Horgan 1995; Keller 2005; Downs et al. 2008; Mitchell 2009). This is the primary fear of complexity theorists themselves (Gershenson 2008). Sornette (2003, p. 208) points out that power law distributions are difficult to extract from data sets, given their similarity to other distributions, that many different mechanisms can produce power laws, and that log-periodic features are sometimes more reliable indicators of the underlying mechanisms. There are, of course, many distributions with exponential terms, including one with the technical name "the exponential distribution." For simplicity, I focus on power law versus normal distributions.

14.8 Dynamic Networks

Normal distributions signal randomness, a lack of structure, in some cases the product of entropic processes that break down structure. Conversely, then, we are invited to look at processes that *produce* structure for the origin of power law distributions. But not all structures give rise to power laws. Simply being non-Gaussian is not enough. Why the seemingly pervasive existence of scale-free power-law distributions? And does the explanation have something to do with robustness and/or complexity? Carlson, Doyle, West, and Barabási all think so.

For Bak, as we have seen, the explanation is nothing more special than what his sand pile model suggests. Simple aggregation of ordinary events can produce critical states with nonlinear consequences. For West, who is closer to the Carlson-Doyle camp, power laws themselves imply the existence of a robust design or mechanism that produces them, subject to physical constraints.

In effect, Carlson and Doyle develop Perrow's old theme that the drive for robustness generates spiraling complexity vulnerable to unexpected, cascading failure as signaled by power law distributions with their fat tails. Carlson and Doyle add a second drive—for optimal throughput. And the mechanism in both cases is natural selection (or a human design analogue), which always involves a compromise among many factors.

Meanwhile, Barabási proposes a more detailed *growth* model for networks exhibiting power laws. The two principles in play here are "the early bird principle" and "the rich-get-richer principle," resulting in a network with distinct hubs. In the growth of such a network, being an early node increases the probability of becoming a hub simply because later links must be made to existing nodes. But preferential attachment also plays a role in the development of many systems, meaning that linking with an already well-linked node is more probable than linking with a relatively isolated node. The rich get richer.[17]

A recent move in this particular debate involves a disagreement between the HOT model and the preferential attachment model. The HOT team claim that preferential attachment does not fit the empirical data as well as their model does. To mention one example: D'Souza et al. (2007) state that preferential attachment stands in the Pareto, Polya, Zipf, and Simon tradition, whereas the optimization approach goes back to Mandelbrot's statistical study of language.[18] D'Souza et al. propose to reconcile the two approaches with their "tempered preferential attachment model," a model that attempts to explain preferential attachment in terms of a cost function rather than simply assuming it. In their model, nodes can become saturated with links, leading to an exponential cutoff, and not all attempts to create new nodes are successful. The overall result is that a tempered preferential attachment emerges from the net formation process subject to the cost constraints.

That is only one example. Various investigators have recently pointed out that there are many mechanisms for producing power law distributions, that preferential attachment is only one of them (and is not really original with the Barabási group). Accordingly, anyone who claims the evidence for a power law distribution proves an underlying preferential attachment mechanism is guilty of the fallacy of affirming the consequent. (I do not say that the Barabási group make such a strong claim.) So, again, we must proceed cautiously here and realize that gestures toward real-world applications are highly conjectural.

14.9 Application to a Kuhnian Model of Science

To apply these ideas in detail to models of scientific research would require a book examining citation networks, actor networks of Bruno's Latour's sort (Latour 1987), semantic networks, and so on. This is obviously not the place to tackle the major task of working the growing literature into an explicit model. Instead I shall point out possible connections to a broadly Kuhnian conception of mature science. Kuhn's model is problematic, of course, but so are all such models. For present purposes I shall stick close to Kuhn's, as the one best known. My analysis suggests that future study of various types of scientific networks and their transformation will offer support for some aspects of Kuhn's model while undermining others.

In *The Structure of Scientific Revolutions* (1962), Kuhn already offered a proto-complexity model that fits my extension of Perrow's thesis to scientific inquiry. Research scientists seek to design structures (research programs, theories, models,

[17] Both principles would seem to play a role in Wimsatt's conception of generative entrenchment as applied to developing systems, either biological or humanly devised.

[18] Among other attempts to explain some power laws, see Fabrikant et al. (2002). The study of phase transitions at critical points is another locus of such efforts.

experimental systems, etc.) that are robust to anticipated kinds of failure. The conviction that a more robust account of their scientific domain is not only possible but also now accessible to inquiry is what drives Kuhnian normal scientists to design increasingly intricate and esoteric structures and practices, as the other chapters in this volume attest. What Popper considered the very core of scientific research—formulating bold conjectures followed by vigorous attempts to falsify them—Kuhn recast as threats to be controlled. A well-designed Kuhnian paradigm is robust to both threats. Normal scientists who are too bold will be disciplined by the community, Kuhn said; and Popper's "falsifications" of major principles are really only routine anomalies that provide new research puzzles. The basic principles are not up for test in the first place. Kuhnian normal science is far more tolerant of what Popper considered mistakes than is even Popper's account of science.[19] The resulting solidarity of the research community over what counts as legitimate problems and solutions permits not only routine problem-solving success but also productive exploration of new, increasingly esoteric research puzzles.

However, Kuhn then surprised the philosophical community by emphasizing the fragility of paradigms.[20] The very robustness and productivity of a successful Kuhnian paradigm in the indicated respects makes the enterprise increasingly vulnerable to major failure, he said.

> Anomaly appears only against the background provided by the paradigm. The more precise and far-reaching that paradigm is, the more sensitive an indicator it provides of anomaly and hence of an occasion for paradigm change. . . . By ensuring that the paradigm will not be too easily surrendered, resistance guarantees that scientists will not be lightly distracted and that the anomalies that lead to paradigm change will penetrate existing knowledge to the core. The very fact that a significant scientific novelty so often emerges simultaneously from several laboratories is an index both to the strongly traditional nature of normal science and to the completeness with which that traditional pursuit prepares the way for its own change. (1970, p. 65)

The very robustness of ordinary science (whether or not precisely Kuhnian) makes it increasingly vulnerable to revolutionary overthrow. What is it about a mature science that makes it ripe for revolution? A cognitive psychological point is that strong focus by the community on narrow, esoteric matters provides guidance to the community at the frontier of research but blinds it from other things. A social

---

[19] Popper attempted to remove the fear of making mistakes. His motto was "We learn from our mistakes." In his methodology of scientific research programs, Lakatos (1970) followed Kuhn in the respect mentioned in the main text while retaining some Popperian elements. Kuhn (1970) objected to Popper's talk of falsifications as mistakes.

[20] In the chapters on normal science, Kuhn invites us to look at scientific work from the point of view of the normal-scientific practitioner, who, according to Kuhn, is convinced that s/he is uncovering the truth about the world. In the chapters on scientific revolutions, Kuhn invites us instead to look at the history of science from above and to note the contingency involved in the revolutionary passage to a new paradigm. As a rough generalization, philosophers have tended to take a normal scientific, realist view whereas sociologists have distanced themselves from the normal science perspective. In my opinion, contrary to Kuhn, taking the normal scientists' viewpoint does not require conviction that the paradigm is on the track to a final truth about the world.

psychological point is that routine anomalies become more serious over time as they resist the attempts of the best people to resolve them. Confidence in the resources of the paradigm falls rapidly (unless, I would add, there is compensating, rapid progress on other fronts). The emergence of a possible alternative approach can also produce a crisis by amplifying the importance of extant anomalies (Kuhn 1970, p. 86). But the Kuhn quotation suggests an additional, non-psychological mechanism. As a science matures, the linkage between its components becomes more complete and more entrenched and the research more systematic. Thus a minor measurement discrepancy or a new discovery that does not quite fit can now penetrate deeply. The maturation of the paradigm has given the anomaly more leverage. It is now more threatening. The possible error that it represents can no longer be contained. It increasingly propagates through the system, in ways somewhat reminiscent of Descartes' system, discussed above. Previously, it could be bracketed, perhaps treated by analogy with the "God of the gaps" tactic familiar from theology; but now the gaps have closed.

Per Bak would say that normal science has reached a critical state in which even a "normal" result could trigger a revolution, and that this becomes explicit in a Kuhnian crisis.[21] However, although normal science is cumulative, according to Kuhn, there is far more structure to the accumulation than Bak's sandpile model allows. For many analysts a sandpile epitomizes a mere aggregation rather than a complex system (Wimsatt 2006). For this and other reasons, the HOT model of Carlson and Doyle fits Kuhnian science better. For the goal of science is to maximize problem-solving productivity while eliminating normal (expected) kinds of error—and to do so precisely by improving the design of theory and data structures. Moreover, the development of normal science would seem to be broadly evolutionary (Nickles, forthcoming).

We should credit Bak with two other important insights that Carlson and Doyle also appropriate and incorporate in the HOT model. One is the point already stated above: that in a highly mature, rigorous, science even a seemingly ordinary result can trigger a cascade of developments that lead to revolution. Scientific work is highly nonlinear in this respect. It does not take a big, revolutionary cause to trigger a process that produces a revolutionary effect.[22] Again, most anomalies begin as

[21] Bak does not mention Kuhnian revolutions, though he does relate his work to other models of transformative change such as Gould-Eldredge punctuated equilibrium and mass extinction (Bak, Chapter 1). Sornette (2003, Chapter 3) contends that the causes of stock market crashes are not ordinary events, that crashes are outliers with a special statistics of their own that call for special explanations. However, his model is not totally different from those under discussion. The underlying processes involve increasingly correlated phenomena of complex systems, driven by positive feedback, that send the system to a critical point, where it becomes unstable. At this point a normal change can tip the system one way or the other. As an emergent phenomenon of a complex system, such a disruption, like a Kuhnian revolution, is holistic. It cannot be analyzed into component parts.

[22] This point is nearly explicit, however, in Kuhn's other major work, his history of the quantum theory. (See my discussion of his Planck case in Nickles, 2009 and forthcoming.) Critics complained that Kuhn failed to integrate his quantum history with the model of *Structure* (Klein et al. 1979).

small disruptions that normal scientists have every reason to believe they can handle with the resources available to them. Roughly speaking, the more mature (robust) the science, the greater the nonlinearity, since small discrepancies now have more leverage.

Bak's other insight is that these disruptions occur on all scales. Applied to Kuhn, this suggests that there is no principled distinction between normal and revolutionary science, that normal science is more dynamic than Kuhn's account allows, and that revolutions are simply normal disturbances writ large (cf. McMullin 1993, Wray 2007).[23] But perhaps a better way to express the scaling point is this. Kuhn states that there are small revolutions within specialty and subspecialty fields as well as large, highly visible revolutions. A small revolution may look like cumulative change to those practitioners working in other fields who notice it at all. If Kuhn is correct, this suggests that the structure of the general field is modular with the connections between some of the specialty areas and even between them and the core rather weak. In such a case the large field is, in Herbert Simon's term, nearly-decomposable, at least to some extent (Simon 1981; Wimsatt 2007, Chapter 9).

We can also notice a connection to Barabási's preferential attachment model. In Kuhn's account of normal science (especially as amplified in his "Postscript—1969"), exemplars come to function as hubs. New research puzzles are solved by relating them to one or more genealogies of puzzles and solutions that the community takes as exemplary. Accordingly, the discovery that an entrenched exemplar is defective or that it has exhausted its ability to yield new problem-solving insights sends shocks through much of the corresponding normal science. Further elaboration of this point would bring in Wimsatt's aforementioned work on generative entrenchment. Given the historical contingency of which specific exemplars are, in effect, selected as the "early birds," the ensuing scientific work under that paradigm is likely to remain contingent in important respects owing to historical path-dependence, even though scientific practitioners do often succeed in reworking older material in such a way as to eliminate some contingencies (Nickles 1997). This point directly connects with Soler's concern with the historical contingency versus inevitability thesis in her chapter in this volume.[24]

[23] It is then open to a Kuhnian to reply that genuine revolutions are precisely those disturbances that could *not* be contained within the bounds of normal science and that resulted in overturning the old approach, that making the distinction a matter of degree violates the hierarchical nature of his model. One response would be that even Kuhn, qua historian, agrees that classical mechanics went through several phase changes between Newton and Einstein, changes that transformed it almost beyond recognition as it incorporated the so-called Baconian sciences; later adopted the latest Lagrangian, Hamiltonian, and other mathematical techniques; rejected the ideas of action-at-a-distance and that all forces are central forces; became statistical-probabilistic, etc. The very concept of mechanics was transformed in the process. So why count all of this as normal science?

[24] See Soler (Chapter 10, Section 10.16) as well as Soler (2000, 2004). Kuhn (1962) stated that paradigm change is almost inevitable given the unavoidable contingency of its formation. After all, at the frontiers of a new domain of research it is most obvious that scientists cannot yet know much about the structure of that domain. It is thus exceedingly improbable that a given paradigm will

Are mature branches of science that lack the theoretical integration of mechanics equally subject to Kuhn-style revolutions? The chemical and biological sciences, for example, possess a less theory-centered structure. Are they therefore less subject to catastrophic failure of the relativity or quantum theory variety and thus supportive of a more robust scientific realism in their domains? This question is worth further exploration. On the other hand, inspired by Kuhn's own talk of exemplars as practically making a theory structure unnecessary, writers such as Giere (1988, Chapter 3, 2008), Teller (2001, 2008) and Rouse (2003) adopt a more exemplar-centered than theory-centered conception of Kuhnian scientific practice, even in mechanics. On this view a "theory" is really a collection of models.

Finally, the network approach helps us to make sense of how Kuhn can pass so quickly from talk of incommensurability (e.g., of relativity theory with classical mechanics) to the following claim:

> We may even come to see [the relativity revolution] as a prototype for revolutionary reorientations in the science. Just because it did not involve the introduction of additional objects or concepts, the transition from Newtonian to Einsteinian mechanics illustrates with particular clarity the scientific revolution as a displacement of the conceptual network through which scientists view the world. (*Structure*, p. 102)

Talk of the same concepts is odd, since Kuhn has just made his strong claim that meaning change prevents literal limit relationships between the relativity theory and classical mechanics. But we can understand what he is getting at in terms of networks. In the background is a quasi structuralist account of meaning according to which the individual linguistic units are meaningless in themselves (or have arbitrary meaning), their technical meaning deriving from their place in the structure or network, combined with the logical empiricist idea of implicit definition. With the coming of relativity theory, the linkages among the concepts changes, thereby producing a holistic change in the meaning of the concepts themselves, a change that cannot be analyzed as a piecemeal change in a single "definition" or two. (There will be a corresponding cognitive change for those who understand the language in the new way.) Kuhn is treating the concepts as syntactic nodes the meaning of which depends on their place in the network.[25] And it is this claim and perhaps also the behavioral economics of the Kuhnian scientific community that suggest that self-organizing systems may be in play. (The earlier point about the nonlinear vulnerability of mature normal science is rather different from the

---

be able to anticipate future results so as to get everything right. Thus Soler's treatment of contingency nicely complements my own about maturation as increasing vulnerability to transformative change in the system. Sociologists of science from Latour and Woolgar (1979), Knorr-Cetina (1981), and Pickering (1984) to the present have given far more attention to the contingency of scientific decisions than have philosophers of science.

[25] This last is a familiar point often made about Kuhn and Feyerabend. Papineau (1979) provides an excellent discussion. Kuhn retained a more limited sort of meaning holism later in his career (Kuhn 2000). Soler (2000, 2004) interprets Kuhn's account of meaning and meaning change in terms of the structuralism inspired by Ferdinand de Saussure's work.

point about self-organization, but ultimately related.) A Kuhnian scientific revolution, like a transformative change of a complex system, is an emergent, holistic macro-phenomenon that cannot be analyzed adequately at the level of its component parts. It remains unclear (at least to me) to what extent a Kuhnian scientific community, with its imitation or herding tendencies, can be regarded as a *self-organizing* complex system in which positive feedback can take the system through a critical point and into a new regime.

14.10 Prospective Robustness

Robustness is a concept typically applied to specific research results, based on their track record of empirical and theoretical support, especially in Wimsatt's (1981) sense that robust results can be derived and/or checked in multiple ways. I have extended the concept of robustness to entire epistemic systems, regarding them as designed problem-finding and problem-solving systems that, when successful in surviving shocks, evolve toward states of increasing robustness in some respects but also increasing vulnerability to failure in other respects. And I have relativized robustness to specific dimensions of failure. Typically, these dimensions of potential failure are somehow "anticipated." Admittedly, this is a difficult idea to apply to biological systems without lookahead, but it is crucially important to human inquiring systems. Accordingly, I want to suggest a further extension of the concept of robustness, to more explicitly include designing for the future, as the network theory we have canvassed suggests. My thought, inspired by Kuhn's, Pickering's, and Wimsatt's work on heuristic fertility, is that a research program (for example) is more robust than another insofar as its long-term prospects for fertile development are better.[26]

It is this prospective "heuristic appraisal" (as I call it) that enables Kuhnian paradigms and Lakatosian research programs to be robust to the anomalies that Popper regards as falsifications. In making science safe for failure with his emphasis on learning from our mistakes, Popper took an important step forward. But it was simultaneously a step backward: in excluding all sorts of ad hoc hypotheses, and in saying that scientists should reject a "falsified" theory, Popper amplified the destructive effects of anomaly. He made scientific work more robust and less robust at the same time and in the same respect, to the failures that he called falsifications.

Unlike nonhuman biological evolution, human inquiry can take advantage of some degree of lookahead, and this prospective orientation marks a major difference between Kuhn and the logical empiricists and Popperians.[27] Traditional

[26] Hume's problem of induction implies that we should be fallibilists about the future, but the problem is so general that it undercuts all enterprises more or less equally. More specific considerations involved in what I call heuristic appraisal provide differential appraisals of future prospects and thereby make a difference (Nickles 2006).

[27] Lakatos (1970) and his students made heuristics an important part of search programs.

confirmation theory and Popper's theory of corroboration treat the epistemic merit of a theory or research program as a function of its empirical track record to date and regard empirical failure as epistemic death (falsification). Kuhn fundamentally disagreed. For Kuhn (1962, 1970) it is prospective fertility that determines the viability of a research program, not past success or failure. Since all decisions are about the future, future prospects are obviously critical. The point is both sociological and technical-scientific. Scientists are attracted to a paradigm in the first place because they can see how to use it to investigate interesting problems and thus contribute to that specialty and, in the process, to build their careers and maintain their self-identity as productive scientists in a specific field.

An indication of Kuhn's departure from the received view is his rejection of the claim that "context of discovery" (as a general label for innovative activities) is not epistemologically interesting. To be clear: Kuhn did not believe there is a logic of discovery or a rigid scientific method of any kind, including a method of justification. However, in his view a good research program provides strong indications of where and how future work is likely to bear fruit. In the best cases, the paradigm practically guarantees that any research puzzle that can be formulated in its terms is solvable by means of its resources, which also explains the rapid drop-off of confidence when failure is evidence.

This prospective versus purely retrospective conception of robustness carries over to smaller-scale units such as an investigator's experimental system (Rheinberger 1999) and even to research proposals. A good system or proposal is one that has a clear direction but also one that is flexible enough to be opportunistically adaptable in various ways in response to anticipated possibilities of failure and success. It is not fragile in the sense of falling flat at the first sign of failure. It is one with strong heuristic promise.[28]

14.11 Concluding Summary

I have argued for a broadened conception of robustness, one coupled to fragility and one that takes into account the future, at least where some degree of lookahead is possible. Robustness is not absolute but relative to types or sources of failure. Increasing robustness in one dimension is typically coupled to increased

[28] Although the two are intricately linked (Rescher 1977), we need to distinguish between the retrospective and the prospective (or heuristic) robustness of a product and the robustness of the process that produced it. A product can be robust even though the process that produced it is now regarded as "played out," sterile, unlikely to produce fruitful new results, thus not robust in the heuristic sense. And a process may be evaluated as robust in the heuristic sense even if has not yet produced much in the way of warranted products. Thus we need to interpret my basic expression $R(s,p,d,c)$ generously to allow that robustness of system s in the dimension(s) of future promise might be of degree d great enough to withstand both endogenous perturbations p (anomalies, adjustments to the research system) and less endogenous ones such as the complaint of underdevelopment, compared to the competition.

vulnerability in others. Thus I reject the cumulative fragility-reduction/risk-reduction thesis, a point that at least somewhat undercuts the appeal of some strong realists to the success of (usually recent) "mature science" as the basis of their claims.[29] Major failures may always result from unexpected exogenous events (a large meteorite hitting the earth, a large cut in project funding, a major breakthrough by a competitor), but they may also result from endogenous developments owing to the nonlinearity inherent in complex systems.

Since the same point holds for major successes, this is a step toward an endogenous account of scientific innovation ("discovery"). After all, we do want our inquiry systems (both processes and products) to change in positive ways. In accordance with the literature on complex systems, I have extended robustness attributions beyond individual empirical claims to entire systems, whether natural or artificial, and I have focused on robustness in the sense of ability to respond adequately to shocks. I have suggested that the philosophy of science literature on robustness needs to go beyond standard confirmation theory to include the prospective evaluation of future promise. Thus I should like to shift the emphasis from the dominant, retrospective, analytic epistemological view approach that is most concerned with the degree of justification of claims already on the table, and of the processes that produced them, to an approach more in tune with scientists' own, future-oriented points of view at the creative frontiers of research. The first approach remains important, of course, but equally important is the second; and, as I have attempted to show, there is an interaction between the two. Our evaluations of future prospects are, of course, fragile in obvious ways; but our evaluations of past results are also fragile—in less obvious ways—precisely *because of* the necessity of anticipating future changes!

Acknowledgements Thanks to Léna Soler for organizing the conference on robustness at Nancy 2, to the members of the Poincaré Archives for their hospitality, to the participants, especially Léna, Bill Wimsatt, an anonymous referee, and Gaye McCollum-Nickles, for helpful comments on either my presentation or a previous draft. I am also generally indebted to Andy Pickering for his attention in his publications to what I call heuristic appraisal and to his pragmatic outlook on the sciences generally. For discussion of Kuhn I am indebted to my students, Jared Ress and Jonathan Kanzelmeyer.

References

Albert, Réka, Hawoong Jeong, and Albert-László Barabási. 2000. "Error and Attack Tolerance of Complex Networks." *Nature* 406(27 July):378–81. Reprinted in Newman et al. (2006), 503–6.

Bak, Per. 1996. *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus, Springer.

Barabási, Albert-László. 2002. *Linked: The New Science of Networks*. Cambridge, MA: Perseus.

Bridgman, P.W. 1927. *The Logic of Modern Physics*. New York: Macmillan.

[29] A detailed discussion of scientific realism is not possible in this already long chapter. For the importance of appeals to the success of mature science in recent defenses of realism, see, e.g., Laudan (1981), Leplin (1984), and Psillos (1999).

Buchanan, Mark. 2002. *Nexus: Small Worlds and the Groundbreaking Theory of Networks*. New York: Norton.

Buchanan, Mark. 2007. *The Social Atom: Why the Rich Get Richer, Cheaters Get Caught, and Your Neighbor Usually Looks Like You*. New York: Bloomsbury.

Calvin, William. 2002. *A Brain for All Seasons: Human Evolution and Abrupt Climate Change*. Chicago: University of Chicago Press.

Carlson, Jean, and John Doyle. 1999. "Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems." *Physical Review* E60:1412–27.

Carlson, Jean M., and John Doyle. 2002. "Complexity and Robustness." *Proceedings of the National Academy of Science* 99:2538–45.

Downs, Cynthia, J.P. Hayes, and C.R. Tracy. 2008. "Scaling Metabolic Rate with Body Mass and Inverse Body Temperature: A Test of the Arrhenius Fractal Supply Model." *Functional Ecology* 22:239–44.

Doyle, John, et al. 2005. "Robustness and the Internet: Theoretical Foundations." In Jen (2005), 273–85.

D'Souza, Raissa, C. Borgs, J.T. Chayes, N. Berger, and R. Kleinberg. 2007. "Emergence of Tempered Preferential Attachment From Optimization." *Proceedings of the National Academy of Sciences* 104:6112–7.

Duhem, Pierre. 1954. *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press. Translated from the French edition of 1914.

Fabrikant, Alex, Elias Koutsoupias, and Christos Papadimitriou. 2002. "Heuristically Optimized Trade-offs: A New Paradigm for Power Laws in the Internet." In *Proceedings of the 29th International Colloquium on Automata, Languages, and Programming (ICALP)*, edited by P. Widmayer, et al., 110–22. Berlin: Springer.

Feynman, Richard. 1965. *The Character of Physical Law*. Cambridge, MA: MIT Press.

Gershenson, Carlos, ed. 2008. *Complexity: 5 Questions*: Automatic Press/VIP/Vince.

Gertstein, Marc. 2008. *Flirting with Disaster: Why Accidents Are Rarely Accidental*. New York: Union Square Press.

Giere, Ronald. 1988. *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Giere, Ronald. 2008. "Comment on Teller: Incommensurability from a Modeling Perspective." In Soler et al. (2008), 265–9.

Gladwell, Malcolm. 2000. *The Tipping Point: How Little Things Can Make a Big Difference*. Boston: Little Brown.

Hindo, Brian. 2007. "At 3M, a Struggle Between Creativity and Efficiency." *Business Week*, June 11.

Horgan, John. 1995. "From Complexity to Perplexity." *Scientific American* 272(June):74–9.

Jen, Erica, ed. 2005. *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies* (Santa Fe Institute Studies in the Sciences of Complexity). Oxford: Oxford University Press.

Kauffman, Stuart. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.

Keller, Evelyn Fox. 2005. "Revisiting 'Scale-Free' Networks." *BioEssays* 27:1060–8.

Klein, Martin, Abner Shimony, and Trevor Pinch. 1979. "Paradigm Lost?" *Isis* 70:429–40.

Knorr-Cetina, Karin. 1981. *The Manufacture of Knowledge*. Oxford: Pergamon Press.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. 2nd ed. with "Postscript—1969", 1970. Chicago: University of Chicago Press.

Kuhn, Thomas. 1970. "Logic of Discovery or Psychology of Research?" In Lakatos and Musgrave (1970), 1–20.

Kuhn, Thomas. 2000. *The Road Since Structure*, edited by James Conant and John Haugeland. Chicago: University of Chicago Press.

Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In I. Lakatos and A. Musgrave, 91–195.

Lakatos, Imre, and Alan Musgrave, eds. 1970. *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts*. Beverly Hills, CA: Sage.

Laudan, Larry. 1981. "A Confutation of Convergent Realism." *Philosophy of Science* 48:19–49.

Leplin, Jarrett, ed. 1984. *Scientific Realism*. Berkeley, CA: University of California Press.

Levin, Simon, et al. 1998. "Resilience in Natural and Socioeconomic Systems." *Environment and Development Economics* 3:225–36.

Martin, O.C., and Andreas Wagner. 2008. "Multifunctionality and Robustness Tradeoffs in Model Genetic Circuits." *Biophysical Journal* 94:2927–2937.

McMullin, Ernan. 1993. "Rationality and Paradigm Change in Science." In *World Changes: Thomas Kuhn and the Nature of Science*, edited by Paul Horwich, 55–78. Cambridge, MA: MIT Press.

Miller, John, and Scott Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.

Mitchell, Melanie. 2009. *Complexity: A Guided Tour*. New York: Oxford University Press.

Newman, Mark. 2001. "Scientific Collaboration Networks." *Physical Review E* 64, paper 016131.

Newman, Mark, and Richard Palmer. 2003. *Modeling Extinction*. Oxford: Oxford University Press.

Newman, Mark, Albert-László Barabási, and Duncan Watts, eds. 2006. *The Structure and Dynamics of Networks*. Princeton, NJ: Princeton University Press.

Nickles, Thomas. 1997. "A Multi-Pass Conception of Scientific Inquiry." In *Danish Yearbook of Philosophy 1997*, vol. 32, edited by Stig Andur Pedersen, 11–43. Copenhagen: Museum Tusculanem Press.

Nickles, Thomas. 2003. "Evolutionary Models of Innovation and the Meno Problem." In *International Handbook on Innovation*, edited by Larisa Shavinina, 54–78. Amsterdam: Elsevier.

Nickles, Thomas. 2006. "Heuristic Appraisal: Context of Discovery or Justification?" In *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*, edited by Jutta Schickore and Friedrich Steinle, 159–82. Dordrecht: Springer (Archimedes Series).

Nickles, Thomas. 2009. "Scientific Revolutions." *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/scientific-revolutions/.

Nickles, Thomas. Forthcoming. "Some Normal Scientific Puzzles." In *Kuhn's "The Structure of Scientific Revolutions" Revisited,* edited by Theodre Arabatzis, and Vasso Kindi. London: Routledge.

Papineau, David. 1979. *Theory and Meaning*. Oxford: Oxford University Press.

Perrow, Charles. 1972. *Complex Organizations: A Critical Essay*. 3rd ed. New York: McGraw-Hill.

Perrow, Charles. 1984. *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

Perrow, Charles. 2011. *The Next Catastrophe*. Princeton, NJ: Princeton University Press.

Pickering, Andrew. 1980. "Exemplars and Analogies: A Comment on Crane's Study of Kuhnian Paradigms in High Energy Physics" and "Reply to Crane." *Social Studies of Science* 10:497–502 and 507–8.

Pickering, Andrew. 1984. *Constructing Quarks*. Chicago: University of Chicago Press.

Pickering, Andrew. 1995. *The Mangle of Practice: Time, Agency, and Science*. Chicago: University of Chicago Press.

Popper, Karl. 1963. *Conjectures and Refutations*. London: Routledge.

Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. London: Routledge.

Putnam, Hilary. 1962. "What Theories Are Not." In *Logic, Methodology and Philosophy of Science*, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski, 215–27. Palo Alto, CA: Stanford University Press. Reprinted in Putnam's *Mathematics Matter and Method: Philosophical Papers*, vol. 1, 2nd ed., Cambridge: Cambridge University Press, 1979.

Quine, W.V. 1951. "Two Dogmas of Empiricism." *Philosophical Review* 60:20–43. Reprinted with changes in Quine's *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953, 20–46.

Rescher, Nicholas. 1977. *Methodological Pragmatism*. Oxford: Blackwell.

Rheinberger, Hans-Jörg. 1999. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford, CA: Stanford University Press.

Rouse, Joseph. 2003. "Kuhn's Philosophy of Scientific Practice." In *Thomas Kuhn*, edited by Thomas Nickles, 101–21. Cambridge: Cambridge University Press.

Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy*. New York: Harper & Brothers.

Simon, Herbert. 1947 (later editions). *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Macmillan.

Simon, Herbert. 1981. *The Sciences of the Artificial.* 2nd ed. Cambridge: MIT.

Soler, Léna. 2000. "Le concept kuhnien d'incommensurabilité, reconsidéré a la lumière d'une théorie structurale de la signification." *Philosophia Scientae* 4:2(October):189–217.

Soler, Léna. 2004. "The Kuhnian Concept of Incommensurability Reconsidered in the Light of a Saussurian Framework." *Malaysian Journal of Science and Technology Studies* May:51–75.

Soler, Léna, Howard Sankey, and Paul Hoyningen-Huene, eds. 2008. *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities?* Dordrecht: Springer.

Sornette, Didier. 2003. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton, NJ: Princeton University Press.

Strogatz, Steven. 2003. *Sync: The Emerging Science of Spontaneous Order*. New York: Hyperion.

Teller, Paul. 2001. "Twilight of the Perfect Model Model." *Erkenntnis* 55:393–415.

Teller, Paul. 2008. "Of Course Idealizations Are Incommensurable." In *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities?* edited by L. Soler et al., 247–64. Netherlands: Springer.

Wagner, Andreas. 2005. *Robustness and Evolvability in Living Systems*. Princeton, NJ: Princeton University Press.

Watts, Duncan. 1999. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ: Princeton University Press.

Willinger, Walter, and John Doyle. 2005. "Robustness and the Internet: Design and Evolution." In *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies,* edited by E. Jen, 231–71. Oxford: Oxford University Press.

Wimsatt, William C. 1981. "Robustness, Reliability, and Overdetermination." In *Scientific Inquiry and the Social Sciences*, edited by M. Brewer and B. Collins, 124–63. San Francisco, CA: Jossey-Bass. Reprinted in Wimsatt (2007), 43–74.

Wimsatt, William C. 2006. "Aggregate, Composed, and Evolved Systems: Reductionistic Heuristics as Means to more Holistic Theories." *Biology and Philosophy* 21:667–702.

Wimsatt, William C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Wray, K. Brad. 2007. "Kuhnian Revolutions Revisited." *Synthese* 158:61–73.

Zhou, Tong, Jean M. Carlson, and John Doyle. 2002. "Mutation, Specialization, and Hypersensitivity in Highly Optimized Tolerance." *Proceedings of the National Academy of Science* 99:2049–2054.