2
A Real-Time Revolution in Routines
Charles F. Sabel

Introduction

We are living in an age of institutional revolution and democratic doubt. Our institutional capacity to solve complex problems—including especially the problem of providing ever more crucial services, such as education, to an ever more differentiated public—is increasing. But this capacity seems disconnected from, perhaps even at odds with our capacity for self-rule by familiar, parliamentary means.

Pessimism about a mismatch between our problem-solving capacities and our capacity for self-rule is congenital to organizational theory and a recurrent theme in modern reflection on democracy. Weber famously saw our passion for practical mastery of the world leading down the path lit by purposive reason to an inescapable, iron cage of bureaucracy (Weber 1958).¹ John Dewey, writing in The Public and its Problems (1927) in the late 1920s, worried that the American republic was threatened by the rise of the large organization controlled by manipulative elites. Democracy might save republican freedom, he thought, but only if the public could learn to organize itself to challenge again and again its own habits of thought. Public primary schools in which pupils, alone and together, learned to make independent sense of immediate experience became in this program a key bulwark against propagandistic manipulation (Dewey 1915). Writing twenty years later and deeply sympathetic to Dewey, Herbert Simon seemed to close

* This essay benefited greatly from incisive comments by Michael Cohen, Gary Herrigel, Peer Hull Kristensen, Karim Lakhani, William Simon, Yane Svetiev, Duncan Watts, Jonathan Zeitlin. The root cause of failures in error detection and correction is me.
the door on such reflective independence. In *Administrative Behavior* he showed that modern organizations were efficient precisely because they built on, indeed reinforced habits and routines (Simon 1947). At the end of his career Simon did raise, wistfully, the possibility of creating a universal language of design by which technical elites could coordinate organizational choices with each other while learning from and reporting to the public at large (Simon 1969). But the very need for such a political lingua franca underscored the background assumption of a disjuncture between the constraints of organization and the constraints of public choice—an assumption that has remained largely unquestioned down to the present.

And yet, and yet. The organizational revolution of today is different. It seems to tie organizational success to the very ability to challenge habit and routine that Weber could not imagine at all and Dewey and Simon could intuit as theoretical perspectives with, at best, limited practical prospects. In so doing today’s revolution in organizations opens the way to a reconsideration of our pessimism about democracy, if not a democratic renewal. Such, at least will be my conclusion. To arrive there I will have to make good on the claim that this organizational revolution is different from the one that produced the huge public and private bureaucracies that came to define ‘modern’ organizations a century ago, and different because it makes a kind of permanent uprising against habit the key to survival in an otherwise unmanageably turbulent world. To show this I will have to present a theory of the new organizations that supplies new answers to the key questions that classical organizational theory taught us to ask. That will be my main task.

To begin, consider this schematic contrast between the canonical form of organization in the period from the late nineteenth century to roughly 1980, and their canonical form thereafter. In the first period organizations are hierarchical and closed. Headquarters sets goals. Successively lower levels of managers decompose the broad goals into narrower and narrower ones. Eventually organizational routines specify in great detail how to parse and execute tasks, and verify their execution. The organization is thus a vast machine for generating one set of rules for decomposing a broad goal into countless small, easily mastered steps, and another set of rules for checking compliance with the first. Subordinates are rewarded for following the rules (Chandler 1977).

After roughly 1980, the canonical organization is federated and open. Decisions of higher units are shaped by lower ones and the lower units can be formally outside the organization. Or, to capture the idea that information in the new organization flows up and down as well as sideways, organizations
are said to be networked. General designs are set provisionally by the highest level and revised in light of proposals by internal and external ‘lower’-level units responsible for executing key modules or subsystems. The organizational routines define methods for choosing provisional, initial designs and production set-ups, and revising them in the light of further review and operating experience. Collaborators are rewarded for achieving broad goals according to standards defined as part of the process by which the goals themselves are set. Rule following entails—paradoxically, given the older understanding of compliance—the obligation to propose a new rule when the current one arguably defeats its purpose (Nohria and Eccles 1992).

Network organizations manifestly outperform hierarchies in volatile environments, where goals change so quickly that reducing them to a seamless set of task specifications is highly risky, if it is possible at all. Specifically, in such environments the open, federated organization can produce a more useful and resilient design for a product or service by canvassing more alternatives in less time than a hierarchy with a like purpose (Eisenhardt and Tabrizi 1995; Ward 1995). Put another way, the network organization can achieve simultaneously three goals—cutting development time while increasing the utility and reliability of designs—which hierarchies, we will see, are thought to have to trade against each other, and that only in stable environments.

The new organizations, moreover, can be built in widely different settings. They were pioneered in the rich countries, beginning, of course, in Japan, with its highly explicit culture of reciprocity and trust, understood as a moral prohibition on exploiting the vulnerabilities created by intimate collaboration (Dore 1983; Nishiguchi 1994). But network organizations have been mastered and in some ways improved by Americans exalting individualism (MacDuffie 1997), Danes with social democratic solidarities rooted in Protestantism (Kristensen and Zeitlin 2004), and Irish with traditions of social inclusion derived from Catholicism (OECD 1996). The new organizations have diffused rapidly from the rich countries to developing and transition economies: some of Volkswagen’s most advanced factories, for example, are in Brazil (Ramalho and Santana 2002) and the Czech Republic (Dörr and Kessel 1997). The network organizations are found in virtually every industry in the private sector, from capital-intensive automobile plants (Womack et al. 1990) to labor-intensive garment factories (Abernathy et al. 1999), and in many firms providing design, engineering, or computing services to business or consumers (Spreenberg et al. 1995). They are also, as we will see in the example of the reform of US public schools, playing an increasingly important role in the reorganiza-
tion of the public sector. There is nothing new in such institutional iso-
morphism across the public–private divide: The large bureaucracy emerged
in Europe first in public administration and the military, then became a
model for private firms (Kocka 1981), while the sequence was, by and large,
reversed in the USA (Chandler 1977).

To say that the network organization is potentially as generally applic-
able as the hierarchy is not, however, to say that there is ‘one best way’ to
build the requisite class of network, any more than there was ‘one best
way’ to build the classic hierarchy. The large German organization could
make do with fewer layers of hierarchy and less precise rules than its
French counterpart, because the German school system, to a greater extent
than the French, prepared the workforce to fill gaps in the formal instruc-
tions (Maurice et al. 1982). Given many analogous and persistent differ-
ences in institutional context we should thus expect that there will be
local, regional, and national variants of the network firm as well.

Some limits of the new organizations are, finally, as conspicuous as their
competitive advantages over hierarchies in many settings. For all their
counterintuitive accomplishments, and notwithstanding their propensity
to empower subordinates and outsiders in ways that hierarchies systemat-
ically limited, these organizations are not utopia realized. Suppliers and
customers cooperate more intimately in the new organizational world
than in the old; but despite the intimacy, and the heightened mutual
vulnerabilities it creates, suppliers rightly fear their customers will some-
times betray them (Herrigel 2004; Whitford forthcoming). Subsidiaries of
multinational firms cooperate with, but also scheme against, each other,
and strife frequently interrupts their cooperation with headquarters (Kris-
tensen and Zeitlin 2004). Newspapers periodically report on reforming
schools that cheat on test scores to outperform peers from whom they
should be learning. Although, as we will see, the form of the power
struggles in these organizations is in some ways as novel as the institu-
tional routines whose control is their object, they are selfish and mean
struggles none the less.

These stylized facts about the new, networked organizations prompt two
immediate questions. First, how do networked organizations work? At a
minimum, how do they manage to cut development times and errors
while increasing the quality of designs through extensive canvassing of
possibilities? Second, how are networked organizations governed? How, in
other words, can they survive without the task specifications and compli-
ance checking that controlled opportunism in the hierarchical corpor-
ation? Some form of governance must control betrayal and the fear of it,
Charles F. Sabel

or the institutions would not function at all; but whatever governance there is must be imperfect, or the risk of opportunism would not be widely registered as worrisome.

The task of this essay is to furnish elements of a response to these questions. The starting point is the classic theory of the organization developed by Simon in *Administrative Behavior* as supplemented by the work of Oliver Williamson on transaction costs (Williamson 1975; Coase 1937). This work provides a model for linking analysis of individual and institutional behavior to form a theory of organizations. It has also provided, almost as an afterthought, an explanation of the current organization revolution. Looking at what is wrong with that explanation will help clear the way to an alternative. As my aim is to put this body of work to explanatory use, not contribute to its intellectual history, I present the classic theory as augmented by the research it provoked, not as originally formulated.

‘Dual’ organization theory: human, all too human hierarchy and its flexibly informal twin

Classic organization theory was ahead of its time in being resolutely behavioral or human oriented. Where much economics of the 1950s (and of course still today) assumes that human agents are, for their purposes, practically omniscient and have limitless foresight, organization theory more plausibly assumes that humans have limited abilities to acquire and process the information needed to choose ends given means. This is the condition that Simon calls bounded rationality (Simon 1982). Where important currents in law, philosophy, and political theory assume human agents are powerfully motivated by principles of fairness and justice, organizational theory takes human agents to be appetitive and self-interested (with the calculation of individual self-interest itself limited by bounded rationality).

Important counter-currents within mainstream organization theory of course went even further, asserting that so bounded is the rationality of organizations that they are incapable of learning in the sense of improving decisions by deliberation on experience. Thus the assumption that decision makers ‘survey’ only the first feasible choice immediately accessible to them at the moment of decision, and ‘prefer’ that choice to any other or inaction, yields ‘garbage-can’ models of organizations, in which decisions result from collisions between decision-makers and solutions (Cohen et al. 1972). The assumption that decision makers can
compare only a few current solutions to their problem, and prefer the one
that best meets their needs, but cannot draw from this decision any
analytic conclusions regarding subsequent choices, turns organized deci-
sion making into muddling through (Lindblom 1959). The assumption
that circumstances impose one and only one choice on decision makers at
crucial moments, and that once made these choices are long unalterable
because of self-reinforcing effects (network externalities or other increas-
ing returns) or because of ‘lock-ins’ due to technical interdependencies,
yields path dependency: where you come from determines where you go
(Perrow 1984; Arthur 1994; David 1985). At the limit mainstream theory
was challenged by the assumption that organizations are not engaged in
problem solving at all. Rather, on the institutional isomorphism view,
they are themselves locked in the iron cage of national or global capital-
ism, whose structural logic and moral imperatives they are powerless to
resist, let alone influence (Zucker 1987).

Models of organizations as irrational had some currency in the study of
the public sector, where, until recently, many important decisions,
whether capriciously changeable or practically immutable, seemed
equally impervious to learning. 2 But this qualification notwithstanding,
classical organization theory was able convincingly to portray large in-
sstitutions as capable of purposive behavior: less the plaything of fate
than some sociologists and political scientists feared, though hardly the
sovereign masters of all their possibilities that many economists took
them to be. Indeed the grand achievement of organizational theory
and transaction cost economics is to show how, in relatively stable
environments, agents with human features can construct institutions—
principally hierarchies—capable of ‘satisficing’: achieving by joint action
an acceptable minimum of some goal beyond their individual reach,
without being undone by the opportunism of individuals pursuing selfish
interests at the cost of the organizational good (Simon 1955, 1956,
1957). In drastic synthesis the argument goes in these three steps.

The response to the problem of bounded rationality, under the import-
ant condition of relative stability, is the successive decomposition of
complex tasks into ones simple enough for human agents to execute. At
the limit the task is so simplified and routine that execution is habitual,
requiring almost no conscious attention at all. This decomposition naturally
yields an organizational hierarchy, where each subordinate takes
instructions from, and passes the output of her efforts to, one superior.
The assembly line is the emblem of this form of collective production,
linear programming its mathematical counterpart.
Charles F. Sabel

The hierarchical decomposition of tasks is efficient because of economies of scale: The more a product is made or a service delivered, the lower the cost of producing it. ‘Economies of scale’ is the name of a complex bundle of efficiency-increasing elements whose nature singly and in interactive relation to each other is—still—very imperfectly understood. Some of the sources of these efficiency gains are direct: repetition or habituation increases dexterity, reducing the time to complete a task. Staying fixed on one or a few tasks also and obviously reduces the expenses associated with switching tasks—what Smith called the problem of ‘sauntering,’ and we would today count as a transaction cost—the cost of organization of a transaction. Others are indirect and complexly cognitive: simplifying tasks makes it easier to invent machines that replace human producers (Smith 1976; Williamson 1975). But whatever their exact source and interaction, the elements bundled as economies of scale demonstrate the enormous problem-solving capacities of individuals despite (in part even because of) their bounded rationality.

Hierarchical decomposition creates, but also provides the means for managing, a governance problem, where governance can be thought of as the system or method, with associated transaction costs, of giving all collaborators incentives to act in the joint interest. The governance problem arises because of the mutual specialization intrinsic to hierarchy. What I make only has value when combined with the part you make; my asset is dedicated or specific to yours. Hence if I can damage the organization without greatly harming myself by withholding my part or asset-specific investment, I can extort greater returns to my collaboration. This is the famous hold-up problem of non-vertically integrated but highly specialized mass producers (such as GM in the 1920s) with regard to specialized but external suppliers (such as Fisher Body) (Klein et al. 1978). The opportunism or shirking of internal units or employees are attenuated variants of the same problem.

The workable response is for the principal—those in control of the hierarchy—to set the penalties for opportunism and the rewards for compliance so high that the subordinate agents will not forgo self-interested strategies. If the agents are outsiders, as in the classic hold-up case, the customer will have to acquire ownership of the independent supplier in order to exercise this discipline. Hence the drive for vertical integration or institutional closure. For insiders the solution is a contractual arrangement rewarding compliance with or sanctioning disregard of the principal’s goals.
But notice that this solution is just workable or manageable; it is demonstrably imperfect. This is the inadvertent finding of economists initially bent on showing the reverse, namely that it is possible to set compatible incentives for the organization (or its owners) and its employees. In one version of these inadvertent impossibility theorems, each employee is promised the full return of her marginal unit of effort, and so has the incentive to act to the benefit of the whole organization. Of course all team members can’t plausibly be promised the exclusive benefit of something they produce jointly—the marginal output of the organization as a whole. So all but one will have to be disappointed. You might think that the owner/CEO could sidestep the problem of joint or team production by signing a forcing contract with the team of employees in the hierarchy: The contract makes all (slightly) better off than otherwise if they successfully collaborate to meet the organization’s goal, otherwise each suffers the same penalty. But this shifts the problem from ensuring the fidelity of the agents to the joint task to ensuring the fidelity of the principal. For under the forcing contract the owner/CEO has an incentive to bribe—one of the team members to under-perform just enough to block the group from achieving its goal. This increases the owner’s proceeds—the profit or residual—from the venture, but at the cost of limiting organizational performance (Holmstrom 1982). And so on.

The principal-agent governance conception of organization theory resonated with contemporaneous conceptions of society-wide, democratic governance. The public (conceived in the USA as the voters, in Europe as corporatist stakeholders) sets through its political or organizational representatives the goals of public policy and the attendant public administration, and periodically compares the promised platform to the results. Because efficiency requires hierarchy, and hierarchy can only (and, given the stability requirement, only occasionally) be controlled from the top, democratic participation in governance is necessarily limited to such periodic reviews. Hence the fears of Dewey and Simon, among many others, that a polity that has chosen to reap the rewards of efficient, large organizations can, at best, escape tyranny, not achieve true self-rule. Network governance, we will see below, offers a way beyond this minimalist or Schumpeterian view of democracy as the people’s power to throw the bum rulers out, because it is rooted in forms of accountability that do not depend on the principal-agent relation (Schumpeter 1976).
Informality and flexibility—the elasticity of classical theory

Despite its preoccupation with hierarchy, classical organization theory, we saw, grew up in opposition to the Weberian idea of perfectly functioning bureaucracy. In addition to assuming lifelike agents with bounded rationality and a propensity for opportunism, the theory assumed from the start that few if any environments are stable enough to allow for perfect parsing of complex tasks; that the limits on parsability require improvisation; and that in understanding how humans act in organizations—administrative behavior—requires attention to this endemic improvising.

Three settings regularly called for such improvisations. The first is at its inception or start up, when the organization does not have a goal (or, rather, it must choose which one to pursue among the many that are possible) (Simon 1958: 170–88). In this phase of self-programming it or its leaders are searching for and evaluating objectives, not parsing them. This search behavior is perceived as prior to and constitutive of the organization (the job of equity owners, CEO, and legislators is to have goals). The search is thus distinct from the organization’s routine operation, except insofar as it is established as a separate activity—an industrial research laboratory, for example—outside the normal operating hierarchy.

But maturity—the second setting—does not eliminate the need for improvisation. Even in well-established hierarchies task definitions do not accurately specify what actually needs to be done. So individuals, moved by camaraderie, reciprocity, or other forms of sociability not reducible either to the functional imperatives of the organization or calculations of self-interest, must cooperate fluidly to correct the plan. This is in classical theory the realm of the informal; and social value—solidarity in one form or another—is its coin (Dalton 1948; Crozier 1964; Burawoy 1979). For Chester Barnard, one of the pioneers of the classic theory, the paramount task of the executive was precisely to coordinate the formal and informal organizations so as to achieve the organizational goal (Barnard 1938). Given what we just saw about the impossibility of writing incentive-compatible contracts or otherwise solving the governance problems of principal–agent organizations in principle, you might say we have worked our way back to a methodologically sophisticated, game-theoretic version of his position (Gibbons 2003).

The third setting is the volatile environment, where changes in background conditions call for so much adjustment that the organization is in effect always starting up. Already in the 1960s some niches in the ecology
of large organizations were becoming so changeable that it was patently self-defeating to parse tasks hierarchically. The effective response, captured in Burns and Stalker’s idea of the organic organization, was a kind of institutionally acknowledged informalism. In contrast to mechanic hierarchies, organic organizations are characterized by the adjustment and continual redefinition of individual tasks through interaction with others: a network structure of control, authority, and communication. The sanctions which apply to the individual’s conduct in his working role derive more from presumed community of interest with the rest of the working organization in the survival and growth of the firm, and less from a contractual relationship between himself and a non-personal corporation, represented for him by an immediate superior; omniscience is no longer imputed to the head of the concern; knowledge about the technical or commercial nature of the here and now task may be located anywhere in the network, this location becoming the ad hoc centre of control authority and communication; a lateral rather than a vertical direction of communication through the organization, communication between people of different rank, also, resembling consultation rather than command (Burns and Stalker 1961: 121).

As the role of hierarchy decreases, there is a corresponding change in the values that bind the organization together. The loyalty of individuals to a particular institution, and the solidarity of those who share the same loyalties, declines. Instead ‘commitment,’ like ‘that of the professional scientist to his work, becomes the glue of the organization.’ So central is such ‘commitment’ to the direction of the institution that ‘it becomes far less feasible to distinguish “informal” from the “formal” organization’ (Burns and Stalker 1961:122). Lawrence and Lorsch substantiate these claims with more detailed evidence some years later (Lawrence and Lorsch 1967).4

The upshot is that by the late 1960s the classic theory of organization became a close analog to the contemporaneous dual theory of labor markets and industrial organization (Doeringer and Piore 1971). Both identify a dominant or primary sector operating in stable environments. In the economic models this is the market segment where demand is growing predictably. This sector is complemented by a secondary one serving the remaining, volatile niches, where the nature and level of demand changes rapidly. In both cases the primary-sector organizations are hierarchically structured to take advantage of economies of scale. The secondary sector, more informal and fluid in both, is held together by professional commitments in large organizations and the craft ethos of small firms. In both cases too the informal, secondary sector is at once
vanguard and rearguard of the primary. As vanguard it produces innovative products: think of the organic organization or the small firms that supply machine tools to the mass-production economy. But, as rearguard it also produces the shoddy goods and services that can be sold when transient peaks in demand allow small firms to sell inferior goods at high prices without inducing large firms/organizations to make long-term investments in expanding efficient production capacity.

When, as happened from the late 1970s, the self-reinforcing effects of what we now very loosely call globalization and technological change made the general environment abruptly more volatile, classic organization in its dual variant had a ready explanation for the institutional changes that ensued: The subordinate form of organization becomes more salient, if not dominant, and its informal fluidity—its routine rejection of routine—becomes, improbably, at least as much norm as exception.

Schematizing greatly, and leaving aside a detailed, descriptive literature aimed mostly at practitioners, research on this organizational ‘insubordination’ has gone in two directions. The sociological investigation has been directed towards a renewed appreciation of the suppleness of ‘informal’ coordination and the inventive capacity of human sociability. Consider the idea of ‘communities of practice,’ exemplified by the Xerox repairpersons who learn their craft while exchanging stories of difficult cases around the water cooler, not from the company’s perpetually out-of-date and maddeningly incomplete service manuals (Lave and Wenger 1991; Brown and Duguid 1991). Or recall Weick’s work on the mutual sense making at the heart of face-to-face interactions as an indispensable resource for ‘organizing’ (without need of a formal organization) a resilient response to especially unruly problems such as firefighting or launching planes from the pitching deck of an aircraft carrier (Weick 1995). 5 Weick likens organizing to jazz improvisation: the disciplined, aesthetically pleasing, and collaborative recombination of the familiar and the novel in changing circumstances, including of course circumstances created by improvisation itself. ‘To watch jazz improvisation unfold is to have palpable contact with the human condition’ (Weick 1998: 551). Contemporary organizations may ‘redistribute improvisation rights’ to speed response to an ever more changeable world (Sutcliffe and Sitkin 1996, quoted in Weick 1998: 549). But this only accentuates the deep sociological truth—long obscured by classic organization theory—that innovation can only become routine when the innovators are ‘loosely coupled’ (Weick 1976)—intimate enough to learn from nuance, but detached enough to break with convention and the habits of the group.
By taking humans to be inherently sociable, and sociability to include the capacity to make mutual sense of individual ambiguity, this line of sociological argument provides an important corrective to the relentlessly strategic characterization of human motivation in the behavioral assumptions of classic theory. But, as we will see in a moment, by limiting the effective expression of inventive sociability to informal groups or communities, writers in this school discounted the possibility of ‘sense making’ in formal organizations, and so directed their key insight away from a domain where, I will argue, it is especially illuminating: the networked organization.

The second, business-school line of research is not concerned with the bedrock character of sociability, but rather the problem of managing the relations between formal and informal—mechanical and organic—organizations in a world where the need to balance the two, or switch rapidly from one to the other, is becoming ever more acute. As developed by the production-engineering wing of the Harvard Business School, the early versions of this literature focused on the emergence of cross-functional product development teams within the traditional hierarchical organization. Through the 1970s product development meant the execution, by narrowly specialized experts, of components specified in a master plan. At the limit, in some US and European automobile companies, for example, the latches for car doors were designed by engineers who did nothing but design door latches, model generation after model generation. The product manager in this setting chaired a committee whose members represented and liaised with—but had no authority over—the functional departments. This ‘lightweight’ product manager was thus in reality a ‘glorified clerk,’ who used formal meetings and reports to track the progress of events he could scarcely influence.

As market changes, often made visible by the success of Japanese firms, obstructed traditional forms of planning and made reliance on engineering tradition more dangerous, firms worldwide (often following the Japanese lead) shifted to ‘heavyweight’ product managers: high-ranking company officials with contacts to customers and marketing as well as production and testing, with a staff of rising stars from the functional departments who could intervene at all stages of development with bench-level engineers. Characteristically, however, the relation of both the ‘heavyweight’ team to the traditional organization and the ‘heavyweight’ product manager to his collaborators were informal: intervention in the functional departments was tolerated, even encouraged, but not formally authorized, and the ‘heavyweight’ managed by walking around...
selling ideas, discussing solutions—almost never holding meetings or writing reports (Clark and Fujimoto 1991).7

The current version of this approach turns the problem of balancing or hybridizing organizational forms into a near impossibility—a ‘dilemma,’ though not quite yet a tragedy. The argument, as developed by Clayton Christensen (Christensen 1997; Christensen and Raynor 2003),8 begins with the now-familiar assumption that organizations adapted to routine, predictable, or ‘sustaining’ development require disciplines antithetic to those that must innovate or ‘disrupt.’ Managers risk tragedy because markets and firms built to address sustaining competition can always and (almost) unpredictably be disrupted by innovators (and vice versa, though we can ignore this limb of the argument here). The danger arises through the combination of a fact about technology and a fact about the cognition of incumbent managers: The most masterful producers and users of the dominant method know in the abstract that there is almost always a superior alternative to the currently dominant know-how (the technological fact). But given their routine disposition to improve on what they already know and find flaws in upstart challengers, the incumbents ignore the threat (the cognitive trap). Disruptive technologies therefore take hold in peripheral markets of no interest to the dominant players, as when electric-arc mini-mill steel producers started making low-grade rebar to reinforce concrete construction. Or they can arise from the ‘non-consumption’ of groups cut off from the existing markets for certain products, as when Sony sold the first, tinny transistor radios to teenagers who could not afford higher-fidelity, tube-powered sets, and in any case wanted to listen to music away from home. Once proven by ‘outsider’ firms, they are generalized to core domains of application, dislodging the established producers. Radical and generalizable production processes, such as lean production, can arise in the same way.

The managerial response to all this, in Christensen’s view, is informalism at the highest level: He suggests that top managers, guided by elements of ‘disruption theory,’ can remain vigilant about the need to shift from one organization form to another, and even devise routines for facilitating the shift. But given the pervasiveness of the threat, the speed with which it can eventuate, and the hold of the cognitive trap, there is as much artful, on-the-spot fusing of old and new—improvisation in Weick’s sense—as science in this response. Above all, and again characteristically, the repertoire of organizational forms remains unaltered: ‘Institutions can’t disrupt themselves,’ Christensen writes (Christensen and Raynor
So, subtleties aside, the new world was generally taken to be the old one gone permanently out of tilt, if not topsy-turvy. In retrospect this seems like an easily understandable mistake. It is easily understandable because the increasing weight of the ‘subordinate’ sector was just what contingency theory predicted in an increasingly volatile environment, and researchers don’t often look a gift confirmation in the mouth. I say this from experience: Michael Piore and I reacted the same way to the developments of the 1980s, reworking the categories of dual labor market theory into the notion of flexible specialization: the use of general-purpose or craft skills and general-purpose machines (now programmable and standardized, and therefore mass-producible) to respond to fragmented, volatile markets (Piore and Sabel 1984). The emphasis on craft community as the source and ultimate regulative of flexibility resonates with the new sociological informalism; the emphasis on the coexistence of and changing relation between mass and craft production at all levels of economic activity resonates with the attention to organizational hybrids and flip-flops in the managerial informalism.

But understandable or not, the move is a mistake. It saves the core of the old theory while illuminating some aspects of the new developments. But it does this at the price of obscuring those novel features in the current situation that most directly and profoundly challenge settled assumptions. The tip-off is the formalism of the new organization. On anything but cursory inspection there are simply too many formal procedures—routines—to square with the notion of the networked organization as organized informality. But the routines of the new organization look counterintuitive or simply self-defeating from the vantage point of classic theory because they extend the organization’s search for answers beyond familiar domains, rather than limiting the searches through detailed task parsing. To show the generality of these routines, and thereby to underscore the need for a re-examination of the theoretical underpinnings of organizational sociology, I present three examples of these new disciplines, drawn from separate strands of writing on the new organizations which quite obviously should at least talk to each other, but just as plainly do not.

Three types of pragmatist organization

Take first the deliberately innovative organization. This is the canonical case of the new, networked organization: the firm—it almost always is a
firm—whose very purpose is to produce innovative products or services. A well-documented example is of design and production in the automobile industry (Helper et al. 2000).

The process starts when, say, the new-van design team in an automobile firm sets the general performance characteristics of the vehicle by benchmarking the best features of current vans and exploring which innovations under development can be incorporated in its design. To benchmark the potential of developmental work the team may ask for engineering simulations of possible outcomes, testmarket a product embodying a potentially valuable feature, or otherwise try to test the actual reaction of buyers to some approximation of the design they are exploring. Assessing the results of these probes, and again guided by reference to leading examples and comparison of possibilities, the team next provisionally subdivides or, to take a term from cognitive science, ‘chunks’ its general goals into subtasks—the design of an engine, or heating, ventilation, and air-conditioning system—and chooses a specialist team from inside or outside the parent company to realize the initial specifications. It may seem unduly fussy to refer in this connection to ‘chunking’ rather than the more familiar ‘modularizing,’ especially since the latter is often used loosely in the automobile industry and elsewhere to mean the former. But modularization, more strictly speaking, aims at the creation of fixed, black boxes whose performance is durably defined in an interface listing the output returned for any input. Where the corrigible provisionality of design choices being described here is useful, modularization strictu sensu is, as we will see below, risky, even impossible (Sabel and Zeitlin 2004; MacDuffie and Helper, this volume).

After this initial chunking, separate project teams elaborate all the provisional subsystems concurrently, applying to that task the same kind of evaluation of competitors’ successful efforts and developmental possibilities used in the van team’s first round of benchmarking. In addition, they benchmark the production processes central to their eventual products to ensure that the methods employed will meet or surpass the efficiency of their most capable competitors. Engine plants, for instance, will have to produce engines that are at least as cheap and warrantable as those of competitors making similar engines in comparable volumes.

Then the initial overall goals are modified by the methods of simultaneous or concurrent engineering, e.g. the engine-design group may find a way to better its target specifications or to cut its manufacturing costs if it can persuade other component groups that design characteristics should be modified accordingly.
A Real-Time Revolution in Routines

Refinement of this iterated co-design continues once production begins by means of just-in-time and the error-detection and correction methods associated with it. In just-in-time production, parts are supplied to each workstation only as needed: ideally, one at a time. Hence disruptions are immediately visible. A breakdown at one station halts production by stopping the flow of parts to downstream operations.

To assure the flow of production, therefore, the source of disruption must be identified. This typically requires tracing long causal chains back to improbable origins by insistent questions sometimes called the ‘five whys.’ For example:

- Why is machine A broken? No preventive maintenance was performed.
- Why was the maintenance crew derelict? It is always repairing machine B.
- Why is machine B always broken? The part it machines always jams.
- Why does the jam recur? The part warps from heat stress.
- Why does the part overheat? A design flaw. (MacDuffie 1997: 494)

Thus error detection and correction, like benchmarking and simultaneous engineering, reveal possibilities for improvement in unexpected (mis-) connections among the parts of complex endeavors; and the cumulative effect of these results is captured in improvements in the benchmark standards for various production processes.11

Benchmarking, simultaneous engineering, and error-detection methods like the ‘five whys’ are counterintuitive from the classical vantage point. In classical theory, we saw, routines are questioned only in exceptional circumstances. It is simply oxymoronic from this perspective to try to do what the new institutions do as a matter of course: routinely question the suitability of current routines for defining and solving problems. We can think of these new institutions as pragmatist in the sense of the philosophy of Peirce, James, and Dewey: They systematically provoke doubt, in the characteristically pragmatist sense of the urgent suspicion that our routines—our habits gone hard, into dogma12—are poor guides to current problems. Or we can think of benchmarking, simultaneous engineering, error detection, and the other disciplines grouped under the anodyne heading of ‘continuous improvement’ as institutionalizing, and so making more practically accessible, the deep pragmatist intuition that we only get at the truth of a thing by trying to change it (Unger forthcoming).
As a second instance of such pragmatist institutions consider the High Reliability Organizations (HROs) (Rochlin et al. 1987; La Porte and Consolini 1991). HROs are well but incompletely designed to perform without fail such extraordinarily demanding tasks as generating electric power through nuclear fission, launching and recovering jet aircraft rapidly from and back onto pitching, greasy flight decks, launching and recovering space shuttles, or fighting forest fires as they race through rough terrain. The designs are necessarily incomplete because the conditions under which the specified tasks are to be accomplished are changing continuously in more or less subtle ways. If the organization is not adjusted accordingly, it fails, catastrophically. In the pragmatist terms just introduced, HROs become disastrously unreliable if they assume that routine, (nearly) invariant success is the result of following invariant routines; and the organizational challenge is to avoid accidents day after day without imperceptibly making this assumption (Weick et al. 1999).

Key to this are error-detection and root-cause analysis disciplines of the sort broadly familiar from the production/operations level of the innovative organization. In the setting of HROs, the most important and characteristic of these is near-miss reporting and analysis. Near misses of course are accidents that only accidentally didn’t happen. So the near misses—and beyond that ‘out-of-control’ sequences that nearly produced near misses—are the urgent analog in the HRO to the line stoppages in a just-in-time system. Both trigger root-cause analysis meant not only to uncover the proximate cause of the incident, but to eliminate, through redesign of the organization if necessary, the background conditions which generated the immediate source of danger.

Note here a circumstance to which we return later when considering governance—the root-cause analysis regimes are often regarded as a key element in the governance of HROs, and when that is the case they are subject to regulatory scrutiny. For example nuclear power plants in the USA are rated in part on how effectively they respond to reports of anomalies at other plants that might be of relevance given their own set-up, and so on (Rees 1994; Perrin 2005).

New Public Services are a third case of pragmatist organizations. This type of institution is little discussed in the organizational sociology literature, but is receiving increasing attention in various domains of policy studies (Elmore and Burney 1997, 1998; Liebman and Sabel 2003, with references to the extensive literature). Well-documented examples are schools in Texas, Kentucky, North Carolina, and elsewhere in the USA that actually teach poor children of color to read and do mathematics...
with proficiency comparable to that attained by rich, white pupils. The new organizations are like HROs in being designed to achieve a single, complex task—teach children to read and use mathematics—and in relying on error detection to compensate for design deficiencies. But there is an important difference. Where the design gaps in HROs are, roughly speaking, small in relation to the overall structure (that is why we say it is well designed), in the new public service the situation is reversed: there is lots of ‘gap’ and little by way of consolidated design. No one, in other words, knows how to install a turn-key school system that produces the desired result. Instead the institutional solution is to build an organization that uses error detection and correction at the lowest (classroom) levels to find out what works, and then adjust the higher (school and district) level structures to generalize that behavior and encourage more refined error detection, and so on.

With regard to reading, for example, all students learn by some idiosyncratic combination of decoding strings of letters/phonemes (phonics) and derivation of the meaning of words and sentences from context (whole language method). Teachers identify the strengths and weakness of each student’s mixture of strategies by sampling their skills in brief, daily sessions, and suggest improvements. (This might be called first-order error detection and correction.) The performance of students in the same grade is measured periodically state-wide by a standard test, allowing for the comparison of the performance of teachers schools, and districts (second-order error detection).

The job of principals in this system is to create conditions in the school for generalizing the successes of the most successful teachers, and the job of the principals’ superior—the district supervisor—is to create conditions for diffusing the successes of the most successful principals. By these means the reformed school is invented through the piecemeal, but eventually comprehensive improvement of an arbitrary, provisional structure, supposing only agreement on a very broad (but still non-vapid) goal: educational achievement (by mainstream measures) should not vary across groups in culturally salient hierarchies, and gaps should be closed by leveling up, not down.

Accepting now as a stylized fact that the new organizations operate successfully by application of these methods, and leaving for another time discussion of the morphology of and techniques for constructing the networks presumed in what follows, we return to the classic questions of bounded rationality, efficiency, and governance and try to provide a consistent and conceptually plausible account of this success.
Bounded rationality, efficiency, and governance in the pragmatist organization

Rethinking bounded rationality

The centrality of search routines to the new organizations suggests an alternative to the decomposition of tasks as a solution to the problem of bounded rationality: Under volatile conditions, when no one can know the whole answer to the design question we face, the response to bounded rationality is to find actors who are already solving (part of) the problem we will turn out to be trying to solve. On a very general level of course something of this sort must be possible: just as language is used to correct the defects of language, and thought the defects of thought, so the answers to current problems, even imprecisely defined ones, will eventually prove to be derivable from recombination of other, previous or current answers. But just as obviously this alternative search response can’t be simply a matter of looking around, any more than parsing tasks hierarchically is simply a matter of breaking big problems into little pieces, or having new ideas is just a matter of thinking about old ones. Indeed, as a practical matter, the notion that the solution to our own limitations is to find collaborators who are less limited seems perilously close to wishful thinking, akin to the ridiculous strategy of the stranded economist who intends to open a can of tuna fish on a desert island by assuming there is a can opener at hand.

To be workable at all any such problem-solving search (regardless of where in the overall design or production sequence it is conducted) will have to (1) disentrench enquiry, directing attention away from habitual answers and towards unfamiliar solutions, while (2) it produces sufficient information about advantages and disadvantages of rival possibilities—the solution space—to suggest a (provisionally) acceptable solution, and so becomes, for the moment, self-limiting. If it is not disentrenching the search will be uselessly redundant: When you don’t know what you are looking for, it is especially unproductive to look once more where you always look. If it is not fruitfully self-limiting the search will be outright destructive, sending the actors off on an endless fool’s errand.

Benchmarking and error detection meet these conditions on solution-generating searches by obliging the actors to look for answers in a novel, initially open, but ultimately bounded space of possibilities. For benchmarking this is the set of best current or potential designs; for error detection it is the activity chains that might have caused a particular
breakdown. The initial canvas of design solutions is necessarily novel and open because it must consider responses that are ‘like,’ but potentially better than, current practice on at least some of many dimensions, and ‘like’ has no determinate meaning in advance of a particular search. Indeed, to search for likes is to invite surprises. The point of the search is precisely to uncover an unsuspected but highly informative resemblance.

Consider by way of example this telling episode in the evolution of the modern paintbrush. A company is trying with little success to produce the traditional house painter’s brush with artificial bristles. Even an accomplished artisan can’t apply a smooth coat of paint with the synthetic bristles. The question arises: what is a paintbrush like? If a trowel, then the brush’s performance is determined by the shape of the bristles’ tip, and the company’s problem is to find an artificial equivalent for the split ends of natural hairs. If, however, paintbrushes are like pumps, their performance is determined by the contours of the channels formed by the bristles, and the next step is to reposition the artificial bristles to channel paint correctly (Schoen 1979: 257–9). In this case the brush was more like a pump than a trowel—but this result, of course, is provisional. In more artfully inclined hands a brush may need to be less a pump than a trowel, sling-shot, or eyedropper.

As in this example, each possibility considered in metaphoric benchmarking provides a vantage point from which to evaluate the strengths and weakness of the others. Hence the search produces something ‘like’ a provisional taxonomy or map of accessible solution strategies in relation to each other. In this way it is self-limiting: Once you have a serviceable map of the solution space, you stop doing cartography and decide where you want to go.

Before turning the implications of this search solution to bounded rationality for the efficiency and governance of the new organization we need to consider a possible objection. The classical theory is, we said, insistently behavioral, and the notion of bounded rationality is its core. Is there any warrant to think that the account just given of learning through comparison is a characteristically human form of learning? Or is it only one of those ‘just-so’ stories that evolutionary biologists and functionalist social scientists invent to ‘explain’ outcomes that nothing else in their theoretical repertoire would have predicted? In responding it is useful to break the search process into two phases or stages.

The first, ‘chunking,’ or the initial subdivision of complex tasks into a manageable number of components or subsystems, is wholly uncontroversial. There is very good evidence that we are hard-wired to divide all
complex tasks into no more than seven plus or minus two ‘chunks’ because that is the number of slots we have in short-term or working memory (Miller 1956). Complex tasks like designing a new van are thus chunked in both the classic approach and the alternative search method. The crucial and controversial issue is what happens next.

In the classic system the initial distinctions, once fixed, become the frame for subsequent ones. The model is traditional taxonomy in biology, where, say, each phylum contains families, each family genera, and each genus species; and the proper definition of a species never perturbs the definition of higher-order sets. Hence in the classic view it is considered a grave design flaw when lower-level, finer-grained subdivisions force reconsideration of higher-level, coarser-grained ones. The analog to this view of classification in product design is of course the idea of perfect modularity: interfaces are so stable that components can be substituted within any module without disturbing the operation of the more encompassing system of which the altered module is itself a component (Ulrich 1995). Chomsky’s generative grammar I (Chomsky 1965) and Simon’s own closely related work on list processing in AI is a mathematical restatement and generalization of the traditional view of taxonomy. Because categories have, in this view, a determinate meaning and unambiguous empirical referents, they can be represented by formal symbols. Any argument can in turn be represented as a string of symbols, and all reasoning by manipulation of such strings by appropriate rules.

In the alternative method, in contrast, chunking is heuristic—the goal of an initial partition is to see how, from the vantage point of each piece, it and the others might be reconfigured. The revision of categories is a desirable and expected outcome, not a failure of intelligence. On the contrasting, semantic holism or cognitive view, the categories of language and thought are inherently ambiguous. Members of each category share common features. But subsets of them have features common to members of other categories, and boundaries blur the same way among subsets. The only way to provide a usefully rich account of any one category, therefore, is to see it in relation to the web of likenesses and differences that connect it to the others: holistically (see generally Fodor and Lepore 1992). On this view metaphoric comparison is the essence of language, not a poetically licensed abuse of linguistic precision: Unless we engage in mutually clarifying ‘sense making,’ the ambiguities of our own thoughts render them too imprecise to be reliably useful. This view implies that interpretation of ambiguity within a language is in principle no different from reconciliation of differences across languages—translation (Davidson 1974). So if
collaborative sense making and metaphoric comparison are indeed at the heart of language, we should be able to identify others already solving problems like our own, and having found them, construct jointly a language for collaboration—notwithstanding deep differences of original approach or paradigm, to use Kuhn’s word.  

Though the social construction of meaning is less a finding in cognitive science than an assertion of philosophic anthropology (Markell forthcoming), an active research topic in cognitive anthropology (Hutchins 1995), and a fruitful hypothesis in the history of science (Galison 1997) and many other disciplines, recent work in semantics and cognitive psychology and classification in evolutionary biology (see Lakoff 1987 for a masterful summary) puts the metaphoric understanding of categories and language on substantial empirical foundations, suggesting that this revision of categories is as ‘human’ as chunking. The prototype effect is a seminal example of this research (Rosch 1978, 1983). Unless formally trained to do otherwise, we individuate categories, and categorize individuals, in relation to concrete prototypes or exemplars that embody in fullest form all the features for which a category stands. We experience an eagle as ‘more’ birdlike than a penguin, because eagles soar and glide while penguins flap their wings as they hop and waddle. But this means that to identify a bird as a bird we have to be thinking of winged creatures that are par excellence, just barely, and not quite birds—benchmarking the bird, if you like.  

Thus the search process by which networked organizations address the problem of bounded rationality seems if anything more native to our thought—more behaviorally accurate—than the symbolic processing of classical theory.

Human problem-solving techniques are not, of course inherently efficient. On the contrary, behavioralists sometimes revel in the imperfections of human problem solvers revealed by comparison with the performance of truly rational beings. So having grounded the behavior plausibility of the new problem-solving methods, we have to ask after their efficiency as well.

Robustness as highly dynamic efficiency

Before we can ask how organizations can be efficient in high-volatility environments, however, we have to clarify the question. Efficiency is a static concept, or rather a concept for stable environments. The efficient solution is just the least-cost means of reaching a known end. When the
ends are various, and by some measure closely related, it can still be meaningful to speak of ‘an’ efficient solution or a set of these. But when the ends become more various and less well specified it no longer makes sense to speak of optimum or efficient solutions. Instead we can ask whether our problem-solving technique is robust, meaning that it can be expected to produce workable answers in turbulent task environments. Think of robustness as highly dynamic efficiency.

To underscore this point let me note here the misleading incompleteness of Simon’s well-known story about two watchmakers, Tempus and Hora, whose competition is intended to show the robustness (resilience in the face of disruption) of hierarchical modularization. Tempus and Hora made similarly complex watches of about 1,000 parts each. But Tempus, the craftsman, assembles each watch from 1,000 pieces, while Hora, the mass producer, makes subassemblies of ten pieces each and then combines these into the final product. Simon shows easily that unless the probability of disruption per unit time is extremely low, Tempus almost never completes a watch, and the typical disruption destroys an extremely valuable, nearly finished piece. Hora completes much more work and loses much less in each disruption, so modularization seems robust (Simon 1969: 90–2). And so it is in comparison with craft production of large numbers of extremely complex machines.

But modularization is plainly not robust in our current world, where key module makers—meaning here sole-source suppliers of indispensable components—are routinely destroyed by natural catastrophes or the complete devaluation of a core capacity through a competitor’s innovation. The robust strategy in this world is one where there is no difference between setting up the production system and repairing or replacing it in case of disruption; or, put another way, where the same principles that generate the initial set-up also generate the capacity to respond to disruptions of the first design as circumstances require. These are the principles that our robust producer, Quaesitor, follows when she searches for collaborators who are good at searching for collaborators all the way down. In this system, if any of the producers ‘breaks,’ the broken one and its collaborators find at least one replacement or substitute.

Two pieces of anecdotal evidence from the automobile industry illustrate the fragility of module makers and, conversely, the robustness of Quaesitor-type re-chunkers. Thus it was until recently widely believed in the industry that the mega-module makers—producers of large systems such as complete power trains or complex seating—would be the great beneficiaries of the disintegration of the closed, hierarchical firm: Their
capacity to integrate disparate components combined with the capacity to anticipate future developments would allow them to impose their design choices on their customers, the final assemblers, while their power as monopsonist purchasers would allow them to impose favorable terms on their own suppliers. But in fact the mega-module makers, having invested in the capacity to produce relatively fixed systems, are today constantly wrong-footed by the fluidity of actual designs (see MacDuffie and Helper, this volume). Federal Mogul is in bankruptcy, while Dana and Tenneco teeter on the brink. The real winners turn out to be firms that federate diverse and flexible makers of specialized components. These federated component producers are more robust in iterated-co-design settings than mega-module makers simply because components are ingredients in many systems while large systems—modules as they are, not as they are imagined to be in a world of interchangeable black boxes—must be tailored to only a few products. And flexible component makers can switch from one production process to another as innovation demands, hedging themselves and so their parent firms against the risk of technological obsolescence. Realizing the limits of modularization, the head of Dana’s Automotive Systems Division recently announced the transformation of his unit ‘from a component producer capable of supplying systems to a systems producer capable of supplying components’ (Carol 2002).

The destruction, by fire, and the almost instantaneous, collaborative reconstruction of the capacity of a key Toyota supplier illustrates the robustness of networks of searchers seeking searchers (Nishiguchi and Beaudet 1998). When it burned down in 1997, Aisin’s Kariya plant 1 in Kobe was the extraordinarily efficient, sole-source supplier to Toyota of P-valves, a relatively simple, but high-precision component of an anti-skid break-control mounted on all of the assemblers’ makes and models. True to its just-in-time discipline, Toyota had only two-days’ worth of P-valves in inventory on the day the fire broke out; when that stock was exhausted, the assembly lines stopped. But by three days after the fire a congeries of a few Aisin suppliers and many other firms in the Toyota group—very few with prior experience in manufacturing this particular part, all presumably motivated by solidarity and the self-interested desire to return to normalcy and distinguish themselves on the way—were producing more than a hundred types of P-valves in their own facilities. By five days after the fire two of Toyota’s Kobe plants were reopened; three days later car production was at more than 90 per cent of the pre-fire level; a week later there was no shortage of P-valves at all. Ultimately 62 firms became emergency producers of P-valves; these producers themselves relied on
more than 150 emergency suppliers. So devastating had been the fire that none of the participants had access to the precision tooling that Kariya plant 1 used to make the original part. Aisin did contribute to the recovery by installing thousands of additional phone lines to respond to requests for information. But neither it nor Toyota attempted to direct the reconstruction effort.

The key to the extraordinary success was the participants’ vast, common experience of the Toyota variant of iterated co-design: Aisin could characterize the part and the production process in general terms—chunk them; the emergency producers could devise, starting from their own experience, many different ways (some quite innovative) of achieving these ends. Then they could chunk the processes they were contemplating so that their suppliers could do the same with respect to subtasks. So, as Quaesitor would have expected, the same disciplines that generated a network of ‘module makers’ continuously searching for ways to improve also generated the capacity to search out alternative solutions if one of the network’s module-making nodes fails.

In a malign world this distinction between efficiency and robustness would translate into an impossible choice: we would have to either increase efficiency or increase robustness, without knowing from one day to the next which is called for. In a more benign world we would expect that increased robustness would increase efficiency as well, so that measures that help us enhance performance in a volatile environment also improve operations under more stable conditions. This is in fact what we observe with the new organization. I suspect that the overlap between robustness-enhancing and efficiency-enhancing mechanisms explains a good deal of the explosive diffusion of this new institutional type. Here is a list of some of these mechanisms; and like the list of mechanisms that contribute to economies of scale, it is illustrative, not exhaustive.

The first is benchmarking and the metaphoric learning associated with it. In a volatile environment we expect tomorrow’s conditions to differ sharply from today’s. So minimal prudence demands that we continuously benchmark new developments to avoid disastrous surprises. But we just saw that benchmarking searches inform us about all the possibilities in the—expanded—solution space. That means that so long as the environment has not cooled down to the organizational equivalent of absolute zero, where benchmarking itself is unnecessary because conditions never change, these comparisons tell us new things about solutions we already know even as they reveal solution strategies we had not foreseen at all. This is surely an important part of the reason why comparison
of additional variants often improves the reliability of the eventual designs rather than degrading it.

Something similar is probably at work with the second mechanism: concurrency. Concurrent development—in which ‘upstream’ and ‘downstream’ steps proceed simultaneously, each taking account of the (changes in the) requirements of the other—is almost unavoidable when rapid change puts a prohibitive penalty on missing or being late for a market. But like benchmarking, concurrent development sheds new light on familiar designs and practices even while illuminating new environments. In particular concurrency calls into question taken-for-granted assumptions about the relation among components in any sub system by prompting chains of what-if questions about how perturbations in the environment would reverberate through the inner structure of the complex. So in all but the most stable settings, some concurrency probably leads to efficiency-improving discoveries about shortcomings in familiar arrangements while also increasing robustness.

A third class of mechanisms with advantages in stable as well as volatile worlds are the techniques of what might be called flexible formalization: the tools or methods for making tacit, lived knowledge explicit, but in a way that permits rather than discourages further exploration and revision. Metaphoric benchmarking and iterated co-design generally depend on, and generate, devices for characterizing imperfectly theorized alternatives well enough to allow comparison and provoke some combination of action and reflection that may eventually improve the characterizations. But this is of course also true of methods, such as the five whys, which are the production, or stable-world, analogs to these design disciplines. In the five whys each response clarifies an obscure circumstance and demands scrutiny of its own obscurities. Even the ‘final’ results of such iterated questions—the root causes of problems subjected to root-cause analysis—are frequently formulated as rebuttable constraints on practice. An example is the book of lessons learned often found in one or another variant in Toyota or Honda plants (MacDuffie 1997). These compilations log design or processing steps—stamping curves of more than a certain radius in sheet metal—that bump up against existing technical constraints. They mark the limits of current practice, not the limits of the possible. Aisin’s astonishingly quick recovery from the destruction of its P-valve facility suggests, we saw, a close connection between the flexible formalisms used to characterize ongoing processes rigorously for purposes of continuous improvement in efficiency and the flexible formalisms that allow the on-the-fly assessment of design alternatives required for
robustness. As we will see next, this kind of rule making as ruling in or out for now has important analogs in rule and law making by the state as well.

**Governance**

Governance in the classic organization, we saw, is on the principal–agent model, or its stakeholder variant. In this model the choice of projects is taken to be unproblematic. Principals are simply assumed to be endowed with the ability to identify the projects best for them. The task for governance, rather, is to motivate the agent’s faithful execution of the project, punishing her for opportunistic, self-regarding use of gaps or imprecision in the principal’s plan, and rewarding her for using her discretion to fill gaps and correct mistakes in the plan to achieve its goal.

In the pragmatist search organization the choice of goals is at least as great a concern of governance as the control of opportunism. In these organizations there can be no clear distinction between ‘principals’ who make initial plans and ‘agents’ who are expected to revise or remake those plans in the course of ‘executing’ them under volatile conditions. Choice of goals and the broad projects embodying them are as much the product as the starting point of organizational activity. In a world of fallible, appetitive decision makers, the stakeholders in the organization—those who stand to benefit from its success or suffer from its failures—will therefore surely want to establish some mechanism for checking that this goal-determination process has not been subverted by ignorance or greed. Likewise, taking broad goals or projects as at least temporarily fixed, the stakeholders will want institutionalized assurance that collaborators are not making opportunistic use of the wide discretion routinely granted them in pursuing current projects. Here I want to argue that, just as hierarchies call forth (in the threat of hold-ups) a distinctive governance concern, and furnish (in vertical integration) a correspondingly distinct response, so the pragmatist organization’s error-detection and correction regimes afford responses to its characteristic governance risks.  

Take first the problem of controlling the opportunistic exploitation of discretion. By making each party’s facility with shared and highly revealing problem-solving techniques transparent to the others, pragmatist organizations make current collaboration (or the exploration of possible joint work) richly informative about the potential and risks of partnership. Put another way, at least some of the information needed for the substance of collaborative problem solving in particular cases can be used for benchmarking...
the abilities and probity of current and potential partners. Thus a firm
that easily ferrets out the source of errors with five-why methods will
presumably do better at solving new problems, and hence be more a reliable
 colaborator, than one that gets lost in the maze of its own confusions.
Similarly, the mutual transparency that results from co-design disciplines
makes it possible to detect, and attempt to correct, potential problems
before they become disasters. Call the fusion of substantive and evalu-
active knowledge that allows an organization to learn as it monitors
accountability (and vice versa) learning by monitoring (Sabel 1996).

The explosive diffusion of pragmatist disciplines and the success of the
new organization generally strongly suggest that such learning by mon-
itoring is in fact workable at the level of individual projects or the opera-
tion of relatively small units, where ‘workable’ here means only that
show-stopping opportunism and ineptitude are detected, not that all
interests are perfectly aligned, eliminating every trace of power. On the
contrary: In the real world, the immediate collaborators in any given
project are seldom so buffered from the concerns of related activities,
with perhaps conflicting goals, that their relations to partners reflect
only their joint experience with them. Thus design engineers or purchas-
ing managers may fully expect to build long-term, co-design relations with
their current suppliers, yet come under irresistible pressure from market-
ing or other departments to cut costs in the short term, even at the price of
subverting the collaborative relation. Fearing this kind of betrayal, sup-
pliers may hedge their bets with customers, raising questions about their
own dedication and loyalty, and so on.

But this turbulence, and the institutional fragilities it surely creates, no
more excludes iterated co-design among pragmatist organizations than
the impossibility of writing incentive-compatible contracts excludes the
possibility of successful hierarchical organization in suitable environ-
ments. Just as managers used the camaraderie of the hierarchy to close
the gap between formal instruction and actual need, so, we may suppose,
the supervisors in the network organization are using the social values
underpinning or generated by pragmatist institutions to ‘manage’ the
dangers of a new kind of opportunism. What those values are is, to be
sure, a very open question. Solidarity based on common, long-term mem-
bership in a stable organization which values the members’ identification
with one another and the institution is unlikely to flourish in a setting
where entrance and exit from the organization is common at all levels,
and where the organization’s own identity or purposes, let alone the
identity of its collaborators, are routinely in question. Professionalism in

A Real-Time Revolution in Routines
Charles F. Sabel

its traditional form, though not linked to boundaries of particular operating units, is also an unlikely source of mutual dedication in the pragmatist organization. Professionalism in disciplines such as medicine, law, and teaching is based on the mastery of particular tools and techniques through apprenticeship with master practitioners; professional mastery is precisely the ability to apply those tools and techniques independently to challenging problems. The mutual monitoring within and across professional groups inherent in iterated co-design and other pragmatist routines thus often affronts the dignity of professionals trained that only beginners and incompetents must continually account for themselves. I have used the term studied trust to characterize the mutual reliance—neither calculating *modus vivendi* nor blind allegiance—that emerges among actors who come, through repeated, closely monitored exchanges, to count regularly on another’s probity and capacity (Sabel 1993). MacDuffie finds just such a process of repeated and repeatedly assessed exchange generating ‘collaborative community,’ as defined in this volume, whose ‘core value or standard for trust is not performance to preset targets, but the willingness and capability to engage in discussion about how to work towards solutions of problems’ (MacDuffie, this volume). But such nuance of characterization aside, if participation in the new disciplines, and the values to which it seems to give rise, did not facilitate detection of dangerously incapable or opportunistic partners, the spread of pragmatist organizations would have been quickly stopped by some variant of the hold-up—the collaboration-destroying use of collaborative discoveries—that classical organization theory long took to exclude just the sort of vertically disintegrated co-design commonplace today.

Yet the very connection between problem-solving substance and partner evaluation that makes learning by monitoring a—workable—instrument of governance on the project or small-unit level disqualifies it as a mechanism for governing large and complex institutions, such as multinational firms, public service providers, or regulatory systems, that must periodically reallocate resources among various, perhaps competitive projects, or add entirely new ones (Fujimoto 1997; Cusumano and Nobeoka 1998). The demonstration by each and every project group in such an organization that it and its collaborators can develop promising new goals from current ones plainly bears on, but cannot alone decide, which goals the encompassing organization should pursue.

By bringing disentrenching searches to bear on its choice of strategy or goals, however, the pragmatist organization can use the information generated by the lower-level exploration of possibilities to inform and
discipline higher-level decisions. Recall that the aim of metaphoric benchmarking, and the self-limiting searches it generates, is precisely to encourage exploration of solutions beyond the boundaries marked by routine. The same device can be used to elaborate and reconsider the understandings and commitments underpinning the organization’s goals. Alternative goals embodied in bundles of projects advance rival interpretations of the institution’s purpose and values; choosing amongst these alternatives the organization simultaneously considers, and fuses, questions of principle (ends) and questions of technical feasibility (means).

The process of re-evaluating goals parallels the process of searching out and assessing narrower design choices. Thus the ‘center’ of the encompassing pragmatist institution—acting on behalf and with the help of representatives of the individual subunits—metaphorically or openly benchmarks its overall objectives, looking for goals ‘like’ the current ones, but arguably better on some dimension. Each of the federated units then does the same with respect to the broad subgoal for which it is responsible; and the general institutional goals are, if necessary, revised in the light of the interim results. Through iterated pooling of the benchmark goals the organization and its units set the general priorities with which to rank projects.

Accountable behavior in this setting no longer entails compliance in the sense of rule following, but rather provision of a compelling explanation for choosing, in the light of fresh knowledge, one way of achieving the common (sub)goal over others. At the limit principal–agent accountability gives way to peer review, in which decision makers learn from and correct each other even as they set goals and performance standards for the organization.

Notice that this theoretical response to the encompassing governance problem is not proof against manipulation, any more than learning by monitoring is an absolute bar to opportunism. The very existence of a goal-selection process of the kind described can tempt units and project groups to manipulate the center into adopting the rules that show their own efforts in the best possible light. If one succeeds, metaphoric benchmarking of goals will be a sham, and the organization’s goals will depend on the outcome of struggles to gain control of the rule-making power. But, at least for now, it seems that this vulnerability too can be ‘managed,’ though, as noted a moment ago, the management necessary differs because the nature of the vulnerabilities does.

Brief accounts of a range of pragmatist governance mechanisms from the private and public sectors indicate how the general principles of
pragmatist organizations (though not their managerial complements) apply in particular contexts. Take first the case of the Illinois Tool Works, a mid-size conglomerate with $9 billion in annual sales and nearly 50,000 employees worldwide. One of its divisions makes plastic and metal components, fasteners and assemblies, industrial fluids and adhesives, fastening tools, and welding products for the construction, automotive, and consumer durables markets. Another makes consumer and industrial packaging, as well as product identification and quality assurance equipment for the food retail and service, food and beverage, and capital goods industries. The firm is organized as a federation of some 600 units. These are financed by headquarters, which also provides corporate research and development facilities. But within these limits the units are largely independent. Each is responsible for setting its own goals, and is held accountable for outcomes. Because of this focus on components and small subsystems, and the flexibility of its individual units, ITW has flourished where mega-module makers, in their rigidity, have floundered.

The key governance rule for goal setting in ITW is derived from Pareto's rule of thumb that only a small fraction—about 20 per cent—of all the causes of an outcome account for some 80 per cent of the total effect. The 80/20 governance rule obligates the units to regularly redetermine which 20 per cent of their activities account for (roughly) 80 per cent of their profits. These disproportionately beneficial activities are to be generalized and developed, and the others spun off or abandoned. Generalizing the successful activities is not, however, straightforward. Each unit's sales to its customers bundle many different kinds of goods and activities: particular co-design services; unusually short product-development cycles: innovative use of particular materials or the processes by which they are worked. So, following the 80/20 rule, the unit's first task is in effect to do a root-cause analysis of the grounds of its own current successes, and redirect its strategy according to the results. That done, the unit must metaphorically benchmark the new strategy by exploring potential uses of its newly recharacterized capabilities that are 'like' the currently successful ones. Because judgements about the viability of strategies discovered in this way always contain a speculative element, a subsidiary governance rule provides that managers are not penalized for being wrong (once or twice), but face immediate sanctions if they are caught pursuing strategies that have not been disclosed and justified to headquarters. This second rule gives managers an incentive to (temporarily) immunize themselves against the risks of incorrect decisions by increasing the transparency of...
their decision making to headquarters, and so to other units in the firm and outside stakeholders.21

A second illustration is the governance system of Cisco Systems, the leading maker of network routers, switches, and interface devices. To remain competitive in one of the world’s technologically and commercially most volatile markets, Cisco invests heavily in research (roughly 17 per cent of sales in recent years). But the distinctive aspect of its strategy is a policy of acquiring technologies or products pioneered by other, usually much smaller firms, and then working with the managers of acquired firms to develop them (Mayer and Kenney 2002: esp. 24). As of mid-2001, Cisco had incorporated some seventy-five units through A & D. This policy depends in turn on two governance rules. The first, disentrenching rule obligates business units to conduct a ‘make or buy’ review when preparing their annual business plan. This requires each unit to compare the strengths and weaknesses of its current produce or service, and closely related variants of these, to those alternatives under development or already produced by competitors. Because Cisco’s headquarters has rich knowledge of the changing needs of end users through its sales force, and many ties to the research community, business unit managers have every reason to identify and evaluate potential acquisition targets before they come to general attention. In case of an acquisition, a second governance rule provides that the inside managers are rewarded for retaining the managers of the target firm, and integrating them into Cisco. Together the two rules not only encourage (as at ITW) regular reassessment and occasional brusque changes of strategy at the business-unit level, but also, by making outsiders into insiders, increase the cognitive diversity of management generally, and so facilitate the next rounds of assessment and change. And of course business-unit changes can in combination lead to large changes in overall strategy.

The public-sector equivalents of these disentrenching governance rules take the general form of an obligation that each unit measure itself against some general goal or performance standard—reliable and safe operations, continuous improvement in service provision—and correct shortfalls revealed by comparison with the performance of others facing similar situations. A straightforward regulatory application is the requirement to undertake near-miss analysis in the US nuclear power-generating industry. Utilities in the industry must report disruptions in their operations to the Institute for Nuclear Power Operation (INPO), an industry-funded entity ultimately responsible to the Nuclear Regulatory Commission. INPO officials sift these reports to distinguish harmless disruptions from dangerous
ones. Thorough analyses of the causes of the dangerous disruptions, and ways of preventing them, are then circulated as Significant Operating Experience Reports, or SOERs. Industry Operating Experience Reviews then periodically assess the ability of particular plants to effectively use the SOERs and other means to improve their own affairs. For purposes of this review, a team of specialists in a variety of areas evaluates the plant’s troubles since the last INPO inspection, paying particular attention to the plant’s own reports on how it has responded to SOERs (Rees 1994, but see also Perrin 2005).

Where minute variations in daily operations are less likely to signal the possibility of substantial hidden risks, the government can require the regulated entity to scan periodically for possible hazards and present a plan for mitigating those that it identifies. A regulatory oversight body then evaluates the adequacy of the plans, and the steps to realize them, against the benchmarks set by the best performers. The shift in the USA in the 1990s from poke-and-sniff (organo-leptic) methods of ensuring food safety to the hazard analysis of critical control points (HACCPs) shows the drift of developments (US Department of Agriculture, Food Safety and Inspection Service 1996). In the organo-leptic method an inspector from the Federal Safety and Inspection Service examines every head of cattle or chicken being disassembled in a slaughterhouse for quality defects and especially signs of pathogens. The limitation of the method is that some pathogens may not be detectable by the usual examinations, so that meat products leaving the processing plant are not assuredly safe; and even if they are, pathogens introduced at later stages of the food supply chain would not be noticed. Under the HACCP regulations introduced by the US Department of Agriculture, meat and poultry processors have to identify all the points in their production processes where pathogens are likely to be introduced, detail how they will reduce these risks, and verify, by testing, the success of the adequacy of their measures. The HACCP plan must be complemented by a Standard Sanitation Operations Plan detailing the plant’s regular housekeeping measures. The role of federal inspectors shifts from direct examination of animal carcasses to verification of the processors’ hazard reduction systems. This verification starts with assessment of the adequacy of the HACCP plan (and the companion Standard Sanitation Operations Plan) and includes review of the plant’s test results as well as independent testing by the inspectors. Eventually the HACCP system is to cover every link in the food supply chain from farm to plate.  

The US Securities and Exchange Commission is shifting the regime governing financial disclosures by publicly traded companies in an
analogous direction. Having rediscovered in the recent stock-market
darket bubble that complex rule systems are easily gamed by managers with
powerful financial incentives to do so, the SEC is requiring firms to
practice 'critical accounting': In reporting their financial results, firms
will have to identify the critical accounting issues where their choice
among arguably legitimate but rival methods of valuing results made a
material difference to their statement of overall performance. For each of
these issues the firm will have to document the alternatives considered;
the valuation that would have resulted from the application of each; the
reasons for the choice of the method actually used, and the key partici-
pants in the decision-making process that produced the final result; and
even an assessment of the risks to the regulatory system as a whole that
would eventuate if the firm's preferred method were broadly adopted.

With regard, finally, to complex public services, the disentrenching
governance rules typically require operating units to formulate and peri-
odically revise strategies for increasing rates of improvement towards a
general end. In the case of pragmatist public school reform discussed
above, for example, the general goal is to reduce and eventually eliminate
the difference in performance in key subjects such as reading and math-
ematics between affluent, white students and poor students, who are also
often of color. In Texas the state governance regime accordingly requires
periodic testing in these subjects by means of sophisticated standard tests
that (now) reward the ability to conceptualize rather than rote learning.
Each school must report the results of these tests disaggregated by eco-
nomic and ethnic groups, and the state pools the data so that parents with
children in a particular school can compare the rate of improvement of the
relevant subgroup in that school to the thirty-nine other schools in Texas
demographically most similar to their children's. Further disaggregation at
the district or school level then yields information about the performance
of particular teachers and administrators that can guide further reorgan-
ization.

In elaborating these general features of pragmatist governance and illus-
trating their practicality the aim, again, is to make plausible the claim that
the new search networks are governable in practice. this is not to suggest
that such networks are, in virtue of the pragmatist institutions on which
they are based, somehow proof in principle against all disruptions of
governance, or that they are already well governed. In fact, leaving aside
cases of egregious financial wrongdoing, it is striking that there is not a
single, generally recognized example of sustained, organized, and well-
characterized good governance in a multinational firm, including of course
the many indisputably successful ones (Kristensen and Zeitlin 2004). One searches in vain for a canonical example of pragmatist governance in the public sector as well. The lesson, perhaps, is that behind the mask of prudence and sobriety, decision makers are irresponsibly reckless, all too eager to take a flyer on any promising fad without due consideration of the risks they are running. Or, as I suspect, the lesson may be that, in a world so volatile that failing to learn rapidly is almost surely fatal, it is more prudent to build the requisite kind of learning organization, and at the price of managing governance issues along the way, than to stay put until the governance of pragmatist institutions has been reduced to a textbook.

**Democracy**

A final contrast between classical organizations and pragmatist alternatives returns us to the opening theme of this essay: their relation to democracy. A fuller treatment would have to sketch in detail a pragmatist democracy, tracing the role of the legislature, administration, and judiciary in a polity that deliberately governed itself by framework laws intended to be revised in the light of diverse efforts to implement them (Sabel and Cohen 1997, 2003; Sabel and Dorf 1998; Sabel and Simon 2004). Here I only want to indicate how pragmatist institutions invert the very features of classic hierarchies that made them an encumbrance on, if not an outright obstacle to democracy: If the class organization reasonably occasioned pessimism about the prospects of democracy, then in its networked mirror image should, all else equal, occasion optimism about a democratic revival.

In classic theory, we saw, the routines of the large organization were the bane of democracy. Whether rooted in actual technical necessity, or imposed as technical necessities through the manipulations of self-interested, technically versed elites, these routines so limited individual and group autonomy as to reduce self-rule to the periodic power to change one set of rulers for another. Hence public school became (for Dewey) a kind of incubator of citizen autonomy, a last-ditch defense against the encroachment of the elites, while the idea of a universal language of design was (for Simon) a fanciful means of connecting fundamentally disparate technical elites, and allowing them at least to communicate with the masses.

But in the pragmatist organization, we saw, the questioning of routine at the level of individual projects and more generally has itself become institutionalized. In this sense the lesson of the Deweyan school and the world of work surely overlap, even if they are not identical: In both rule
following elides with rule making, and individual autonomy is explicitly linked to group decision making. Reform of the current, bureaucratic public school system on pragmatist lines further blurs the distinction between education and other forms of problem solving. Meanwhile Simon’s language of design has been transformed in pragmatist institutions from a forlorn, academic hope into an everyday necessity: the many, interconnected protocols of iterated co-design are in effect so many (partial, but intercommunicating) design languages, allowing actors with diverse expertise and different background assumptions not only to exchange ideas jointly but also to develop new tools for mutual understanding. More yet: in assuming all current expertise to be importantly limited, and hence the corresponding need to develop corrigible institutions through peer review and local experimentation informed by lay knowledge, pragmatist institutions directly challenge the traditional equation of efficiency with rule by unquestionable professionals and technical experts. By their nature, therefore, these institutions invite the individuals and groups that together form civil society to participate in new ways in the decisions that shape their lives. Long aware of the limits of principal–agent governance in volatile circumstances, and increasingly aware of emergent alternatives that allow for institutional learning in the absence of master plans, mayors and local administrators—in Denmark, for instance (Sørensen 2002)—and high civil servants and cabinet-level politicians—in, for example the Netherlands (Wetenschappelijke Raad voor het Regeringsbeleid 2004)—are beginning, but only just, to think openly about the implications of a shift to pragmatist public problem solving for parliamentary democracy (Engelen and Ho 2004).

Sidney Hook, one of the great philosophic wits of the last century, famously quipped that pragmatism was good in theory, not so good in practice. But his is almost surely not the last laugh. The deep surprise of the current organizational revolution is that pragmatism institutionalized—put rigorously into practice—for once, in the reality of our own time, seems to be confounding our inveterate theoretical pessimism, expanding our capacities for problem solving while inviting us to exercise our capacities for self-rule.

References

Charles F. Sabel


A Real-Time Revolution in Routines

Charles F. Sabel


Ernst, Dieter (2004). ‘Complexity and internationalisation of innovation: why is chip design moving to Asia?’ Honolulu: East West Center, Working paper #64, 8 Nov.


A Real-Time Revolution in Routines


Charles F. Sabel


Lakhani, Kosim, and Jeppeson, Lass Bo (work in progress). ‘Broadcast search and solution attraction in scientific problem solving.’


A Real-Time Revolution in Routines


Charles F. Sabel


A Real-Time Revolution in Routines


Charles F. Sabel


A Real-Time Revolution in Routines


151
Notes

1. Weber (1958: 181): ‘In Baxter’s view, the care for external goods should only lie on the shoulders of the “saint like a light cloak, which can be thrown aside at any moment”. But fate decreed that the cloak should become an iron cage.’

2. For recent criticism of path dependency models for understating the capacity for large-scale institutional innovation, see Thelen (1999) and Crouch and Farrell (2004). For criticism of garbage-can models see Bendor et al. (2001).

3. For a reinterpretation of the Fisher Body acquisition as a forerunner of rather than alternative to pragmatist organizations, see Helper et al. (2000).

4. Here is how Lawrence and Lorsch summarize Burns and Stalker:

In mechanistic systems the problems and tasks facing the concern as a whole are broken down into specialisms. Each individual pursues his task as something distinct from the real tasks of the concern as a whole, as if it were the subject of a subcontract . . . Organic systems are adapted to unstable conditions, when problems and requirements for action arise which cannot be broken down and distributed among specialist roles within a clearly defined hierarchy of formal rules . . . We have found . . . that effective organizational units operating in stable parts of the environment are more highly structured, while those in more dynamic parts of the environment are less formal. (Lawrence and Lorsch 1967: 187–9)

5. See also Hedlund’s idea of the Nearly Recomposable—not Decomposable—System. A key example is the university, with its flexible, project-based research groups and informal coordination mechanisms (deanship is a burden to be rotated; ‘leading’ positions confer honor, not the authority to boss others around), in Hedlund (1999).

6. This approach rehabilitates and applies to the explanation of highly organized activities the processual psychologies and ontologies of James and Bergson, with their emphasis on the dominance of intuition (now called tacit knowledge) over formal method, and of becoming (organizing) over being (organization). See Tsoukas and Chia (2002: 567).

7. Studies of the personal computer and industrial design industries from the mid-1990s come to similar conclusions, often characterizing the observed organizational form as a hybrid of the Burns and Stalker types. See the study of IDEO, a leading US industrial design firm, by Brown and Eisenhardt (1997); see also Brown and Eisenhardt (1995: 364); and Tushman and O’Reilly (1996: 8). In characterizing the new system not by its structure but rather by the respect the manager commands, this literature unwittingly anticipated the personalism that would prove a crucial flaw in the ‘hybrid’ regimes: Heavyweight managers
used their authority to advance their own projects, but also to obstruct inter-project coordination. On the relations between such and the later travails of the Japanese heavyweight project management see Fujimoto (1997).

8. For antecedent argument breaking down the distinction between incremental and radical innovation, and thus between stable and volatile markets, see Henderson and Clark (1990). They focus on the reconfiguration of familiar components to achieve new systemic effects, as in the repositioning and modifications of the motor, fan blades, and housing that change a ceiling fan into a portable floor fan.

9. Historians of one of the key precursors to iterated co-design—the ‘systems management’ used by the US military and civilian contractors to develop ballistic missiles and manned space-flight vehicles—have also been struck by the (new) formalism of the process of concurrent innovation. ‘We have found that concurrency [in design] is as unforgiving to inept management principles as a high performance aircraft is to pilot error. In fact, it requires more formality, not less.’ Lieutenant Colonel Benjamin Bellis (1962), cited in Johnson (2002: 47).


11. For examples of how NUMMI (the GM-Toyota joint venture) handles the tension between standard operating procedures and continuous improvement, see Adler et al. (1999).

12. Dewey (1922: 65–7) is careful to distinguish habit, as the creative disposition to respond, almost unselfconsciously, to new situations on the basis of long experience, from routine, as habit frozen into a compulsion to repeat prior responses despite novelty in the situations that typically trigger them. The difference between habit and routine is manifest in the difference between the artist and the mere technician: ‘The artist is a masterful technician. The technique or mechanism is fused with thought and feeling. The “mechanical” performer permits the mechanism to dictate the performance. It is absurd to say that the latter exhibits habit and the former not. We are confronted with two kinds of habit, intelligent and routine. All life has its élan, but only the prevalence of dead habit deflects life into mere élan’ (Dewey 1922: 66). What Dewey calls habits, modern cognitive neurobiology calls zombie agents: non-conscious but trainable processes directing purposive action such as tennis forehands or the darting, Saccadian motions by which the eye acquires its next target. Just as intelligence, or intelligent habit, is for Dewey the ability to reflect on and reshape habit, so cognitive neurobiology sees consciousness or awareness as an ‘executive summary’ of brain processes presented so as to enable choice among different plans of action. Learning achieved through conscious choice can then be ‘automated’ by training zombie agents to perform tasks that once required awareness. See Koch (2004: 233–7).

13. While early formulations of HRO theory just referenced made it seem that, given high enough stakes, organizations could almost always learn to learn,
normal accident theory (Perrow 1984), with its emphasis on tight, not loose, coupling and technical lock-ins, seemed to exclude learning regardless of the stakes. Later exchanges produced a judicious, if theoretically inconclusive outcome: Normal accident theorists conceded that complex technical systems could learn (Perrow 1994), while HRO theorists agreed that they did invariably do so (La Porte 1994 and La Porte and Rochlin 1994). This essay aims to advance understanding of how organizations learn without claiming that effective learning mechanisms cannot be thwarted, most especially by contests for power. See below, pp. 132–5.


15. See Galison (1997) for the view that scientists do indeed create pidgins to communicate across conceptual divides and that novel theories can emerge from these pidgins. For the contrary view of distinct languages and conceptual schemes as mutually unintelligible or incommensurate see Kuhn (1962). For Kuhn’s extremely nuanced late views, which emphasize the role of metaphor in bridging ‘paradigms,’ see Kuhn (1979).

16. Debates over classification in evolutionary biology are particularly revealing. The traditional approach yields undecidable conflicts of classification: Cladists create categories on the basis of shared, historically derived features that reflect the sequence of evolutionary branching points. Pheneticists look to overall similarities in function, form, and biological role. Evolutionary biology aims for forms of categorization that anticipate the transformation of some subgroups of one species into another, even if it is impossible to specify the transformations in advance (Lakoff 1987: 118–21).

17. More precisely we can say that our disposition for cognitive self-entrapment is just as natural as our capacity to extricate ourselves from the traps we create. Thus recent writing in behavioral law and economics focuses on the many ways human decision-making rules or heuristics are systematically biased by the context of decision making. Recent events, for instance, have an undue influence on our expectations of future occurrences simply because they are more accessible to our memory. We are inclined to risk more to recover losses than to earn equivalent gains because we are, unreasonably from the point of view of decision theory, attached to our endowments. Deliberating groups may sacrifice the potentially complementary knowledge of individual members to ‘groupthink,’ or ignore conclusions reached by groups deliberating about related problems and so on. (For overviews see Jolls et al. 1998; Sunstein et al. 2002; and Mendelberg 2002.) But of course the extent of this context dependence is itself context dependent: Some contexts trigger biasing mechanisms, while others allow us to ‘de-bias’ judgements. In ‘near real’ experiments, for instance, teams of medical diagnosticians in which the team leader periodically reminded the group of information dispersed among the members outperformed teams whose leaders periodically underscored common knowledge.
A Real-Time Revolution in Routines

(Larson et al. 1998). The shift to pragmatist organizations can be thought of as a deliberate effort to create an institutional context that allows for rapid detection and correction of context-induced cognitive biases.

18. For an incisive discussion of the challenges that the new, collaborative pursuit of innovative possibilities within and across firm boundaries poses for the traditional theory of the firm and finance economics, see Zingales (2000). If the principal—shareholders or CEO—can write contracts completely specifying the tasks of all agents, Zingales argues, then the firm is a nexus of contracts guaranteeing all the firm’s input providers earnings equal to their opportunity costs. Assuming that initial investments are sunk, and can be disregarded, the value of the firm to its equity owners is then just the discounted value of future payouts minus contractual obligations. When innovation matters, so tasks cannot be specified precisely in advance, and exploration of profitable possibilities must be coordinated by implicit contracts, traditional theory can no longer be used to adjudicate conflicting claims on the firm’s earnings. The firm’s value to equity holders, moreover, depends less on its assets in place—current activities—than on the value of the ‘real options’; or investment opportunities created by collaborative exploration. But, Zingales argues, traditional finance theory cannot evaluate these real options. The governance mechanisms described above can be thought of as the means institutions actually use to ‘write’ the implicit contracts postulated in Zingales’s account to value the ‘real options’ they help generate.

19. Compare Simon’s claim, founded not on Dewey but on logical positivism, that ‘value elements’ pertain to questions of policy, while ‘factual elements’ pertain to matters of administration (Simon 1947: 45–60).

20. From this point of view almost all broad organizational decisions are, like the idea of law or beauty, what Gallie calls ‘essentially contested concepts’ (ECC). In his 1956 formulation, essentially contested concepts are jointly defined by seven elements:

1. concept must be appraisive; accredits valued achievement;
2. achievement must be of an internally complex character, so its worth is attributed to the whole;
3. the accredited achievement is initially variously describable;
4. the accredited achievement must be of a kind that admits considerable modification in light of changing circumstances;
5. each party recognizes the fact that its use of the contested concept is contested by those of other parties;
6. ECC involve derivation from an original exemplar that conflicted parties agree on;
7. the claim of continuous contestation enables the original exemplar’s achievement to be sustained or developed.
Charles F. Sabel

21. See Herrigel (2004). Notice that in casting about for an organizational solution to the innovator’s dilemma of managing ‘sustaining’ or routine lines of business without being blindsided by ‘disruptive’ threats, Christensen seizes on the example of a company—Nypro—whose structure is similar to ITW’s, but whose governance procedures are in crucial ways less formal and effective. What ITW is to metal components and small systems, Nypro is to plastics: a federation of many small units providing specialized products in various industries. Nypro, like ITW, provides central research and development services to its largely independent business units. The key difference is that Nypro has no equivalent to ITW’s 80/20 disentrenching rule. Instead the firm uses periodic meetings among unit managers to spread ideas and develop a cooperative culture despite their rivalry for the headquarters’ approval. In addition the CEO keeps close touch on innovations that diffuse among several units, and requires that the most successful of these be incorporated into the firm’s standard operating procedures. There are, however, no systematic incentives, as there are at ITW, to scrutinize present practices for clues about new markets. This, for Christensen, remains the job of the CEO (Christensen and Raynor 2003: 271–5).

22. For the limits to the current implementation of the HACCP, see Petersen and Drew (2003). But for an illustration of the ways this system produces the information needed to pinpoint deficiencies in its own implementation, see US Department of Agriculture, Office of Inspector General (2003).