

2011

Minimalism and Experimentalism in the Administrative State

Charles F. Sabel
Columbia Law School, csabel@law.columbia.edu

William H. Simon
Columbia Law School, wsimon@law.columbia.edu

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship



Part of the [Administrative Law Commons](#)

Recommended Citation

Charles F. Sabel & William H. Simon, *Minimalism and Experimentalism in the Administrative State*, 100 *GEO. L. J.* 53 (2011).

Available at: https://scholarship.law.columbia.edu/faculty_scholarship/735

This Article is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact scholarshiparchive@law.columbia.edu.

Minimalism and Experimentalism in the Administrative State

CHARLES F. SABEL & WILLIAM H. SIMON*

This Article identifies and appraises the two most promising alternatives to the “command-and-control” style of public administration that was dominant from the New Deal to the 1980s but is now in disfavor. The first—minimalism—emphasizes public interventions that incorporate market concepts and practices while also centralizing and minimizing administrative discretion. The second—experimentalism—emphasizes interventions in which the central government affords broad discretion to local administrative units but measures and assesses their performance in ways designed to induce continuous learning and revision of standards. Minimalism has been prominent in legal scholarship and in the policy discourse of recent presidential administrations, but its practical impact has been surprisingly limited. By contrast, experimentalism, which has had a lower profile in academic and public discussion, has visibly influenced a broad range of critical policy initiatives in the United States and abroad. Indeed, key initiatives of the Obama Administration, including the Food Safety Modernization Act and the Race to the Top education program, are virtually unintelligible from any other perspective. We argue that, in practice, minimalism suffers from an excessive preoccupation with static efficiency norms and price signals, and from insufficient attention to learning and “weak signals” of risk and opportunity. Experimentalist intervention is a more promising approach in the growing realm of policy challenges characterized by uncertainty about both the definition of the relevant problems and the solutions.

TABLE OF CONTENTS

INTRODUCTION	54
I. MINIMALISM AND ITS LIMITS	56
A. GENERAL THEMES	57
B. LIMITATIONS: REGULATION	60
1. The Efficiency Perspective	60
2. Minimalist Intervention	64
a. Cost–Benefit Analysis	64
b. Cap-and-Trade	65

* Maurice T. Moore Professor of Law and Arthur Levitt Professor of Law, respectively, Columbia University. © 2011, Charles F. Sabel & William H. Simon. We are grateful for advice and encouragement to Bruce Ackerman, Lewis Grossman, James Liebman, Jerry Mashaw, and Daniel Richman.

C.	LIMITATIONS: SOCIAL WELFARE	69
1.	The Changing Policy Context	69
2.	Minimalist Intervention	71
	<i>a. Norm Simplification</i>	71
	<i>b. Vouchers and Choice Architecture</i>	74
II.	EXPERIMENTALISM	78
A.	BASIC ARCHITECTURE	79
1.	Decentralization	79
2.	Signals and Norms	80
3.	Incentive Design	81
4.	Stakeholder Participation	82
B.	REGULATION	83
C.	SOCIAL WELFARE	89
	CONCLUSION	93

INTRODUCTION

The administrative style that characterized American public law from the New Deal to the 1980s has been out of favor in recent years. This style, pejoratively called command and control, is identified with rule-bound bureaucracy and deference to ineffable expertise. The efforts of recent years to escape the limits of command-and-control administration have taken diverse forms, and most remain controversial. In this Article, we suggest that these efforts can be understood in terms of two broad models of public intervention, and we assess their advantages relative to each other and to command-and-control administration.

The first model can be called minimalism. The term is associated with Cass Sunstein,¹ but it usefully connotes the understanding of administration in a broad range of legal scholarship, including especially the influential work of Bruce Ackerman and Jerry Mashaw.² This model seeks to ground policy design

1. See Robert Kuttner, *The Radical Minimalist*, AM. PROSPECT, Apr. 2009, at 28.

2. See, for example, BRUCE ACKERMAN & ANNE ALSTOTT, *THE STAKEHOLDER SOCIETY* (1999), presenting a proposal for comprehensive reform of the welfare system, and JERRY L. MASHAW, *BUREAUCRATIC JUSTICE: MANAGING SOCIAL SECURITY DISABILITY CLAIMS* (1983), for an analysis of the relation of bureaucracy and adjudication in a key welfare program. We focus on the parallels among Sunstein, Ackerman, and Mashaw in their treatment of administrative policy implementation. In the sphere of constitutional adjudication, Ackerman rejects Sunstein's minimalism. See Bruce Ackerman, *The Living Constitution*, 120 HARV. L. REV. 1737, 1801–02 (2007).

in economic concepts and market practices, and to minimize frontline administrative discretion and popular participation in administration. Its key normative reference points are efficiency and consistency. It favors cost–benefit analysis for elite bureaucrats and restrictive rules for lower tier ones. Instead of directive regulations, it prescribes regimes of marketable duties, such as emissions permits, or marketable rights, such as welfare vouchers. Where people are prone to make bad choices, it urges “nudges”—providing information or making provisional choices from which the beneficiary can opt out—rather than mandatory rules.

We call the second model experimentalism, though it bears a strong resemblance to what others call “new governance” or “responsive regulation.” Experimentalism has been influenced by both management theory and democratic ideals, especially as the two converge in John Dewey’s idea of democratic experimentalism.³ Its governing norm in institutional design is reliability—the capacity for learning and adaptation. In experimentalist regimes, central institutions give autonomy to local ones to pursue generally declared goals. The center then monitors local performance, pools information in disciplined comparisons, and creates pressures and opportunities for continuous improvement at all levels. The regimes’ distinctive mechanisms for achieving both learning and coordination emphasize deliberative engagement among officials and stakeholders.

Our analysis highlights and explains a fundamental policy reorientation along experimentalist lines in the United States, the European Union, and elsewhere since the 1990s. Minimalism has been influential in legal scholarship and popular policy discourse, but its impact on actual policy design has been surprisingly limited. President Obama frequently invoked minimalist ideas in his campaign, and he appointed Cass Sunstein, a leading minimalist, as Administrator of the White House Office of Information and Regulatory Affairs.⁴ Yet the Administration’s most important initiatives—in fiscal policy, bank solvency, healthcare, food safety, education, and offshore oil drilling—have reflected these ideas only occasionally or dimly. By contrast, experimentalism has had a lower profile in legal scholarship and popular policy discourse, but its premises are pervasively and saliently (if often imperfectly) manifested in recent regulatory and social-welfare initiatives here and abroad. Some of the Obama Administration’s most important initiatives, including the Food Safety Modernization

3. See JOHN DEWEY, *THE PUBLIC AND ITS PROBLEMS* 203 (1927); Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 267 (1998); see also IAN AYRES & JOHN BRAITHWAITE, *RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE* 4–7 (1992) (discussing responsive regulation); Gráinne de Búrca & Joanne Scott, *Introduction: New Governance, Law and Constitutionalism*, in *LAW AND NEW GOVERNANCE IN THE EU AND THE US* 1, 2–3 (Gráinne de Búrca & Joanne Scott eds., 2006) (defining the concept of new governance).

4. Sunstein and several other Obama appointees are associated with behavioral economics, the scholarly underpinning of nudge strategies. *Time* magazine referred to these advisors as a “behavioral dream team” that was trying to “transform the country.” See Michael Grunwald, *How Obama Is Using the Science of Change*, *TIME*, Apr. 13, 2009, at 28, 29.

Act⁵ and the Race to the Top education program,⁶ can only be understood in experimentalist terms.

We argue that these developments reflect the fact that experimentalist regimes are especially well suited for circumstances in which effective public intervention requires local variation and adaptation to changing circumstances. The central characteristic of these circumstances is “uncertainty” in Frank Knight’s sense—contingency that cannot be known or calculated actuarially or with formal rigor but can only be estimated impressionistically.⁷ In the realm of uncertainty, policy aims cannot be extensively defined in advance of implementation; they have to be discovered in the course of problem solving. Yet the characteristic minimalist interventions presuppose a strong distinction between the enactment or elaboration of public goals and their administrative implementation.

On the other hand, experimentalist regimes depend on the controversial premise that public administration can integrate frontline discretion and stakeholder participation in a disciplined, accountable manner. Distrust of decentralized discretion and participation leads some minimalists to reject experimentalist reforms even in situations in which experimentalists regard them as most promising.⁸ Although this concern can only be resolved in practice, we identify some of the ways in which experimentalism responds to it.

In Part I, we defend our notion of minimalism as a heuristic that usefully identifies influential premises of recent public-law scholarship, and we elaborate its limitations as a guide to policy implementation in regulation and social welfare. In Part II, we elaborate the experimentalist alternative and argue for its comparative advantages in key policy spheres. We do not deny the value of minimalist interventions in some contexts. However, we argue that key changes in the configuration of the problems addressed by both regulatory and social-welfare policy favor experimentalist intervention in many areas.

I. MINIMALISM AND ITS LIMITS

We begin by describing the convergent themes of minimalism and then proceed to analyze its limitations. Minimalism tends to be preoccupied with static efficiency and committed to simple rules. Its favored interventions are thus not well adapted to problems that require continuous and highly local adaptation. Yet such requirements characterize many salient emerging policy areas. In regulatory fields such as industrial or product safety or non-point-source pollution, the goal is to mitigate risks that cannot be effectively priced;

5. See FDA Food Safety Modernization Act, Pub. L. No. 111-353, 124 Stat. 3885 (2011) (to be codified in scattered sections of 21 U.S.C.).

6. See U.S. DEP’T OF EDUC., RACE TO THE TOP PROGRAM: EXECUTIVE SUMMARY (2009), available at <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>.

7. See FRANK H. KNIGHT, RISK, UNCERTAINTY AND PROFIT 19–20, 197–232 (1921).

8. See, e.g., David A. Super, *Laboratories of Destitution: Democratic Experimentalism and the Failure of Antipoverty Law*, 157 U. PA. L. REV. 541, 541 (2008).

appropriate intervention depends on changing technology and local contingencies that cannot be distilled into simple and stable rules. Similarly, social-welfare policy is less focused than it was in the past on identifying and compensating people unable to work and more focused on providing people with the skills and accommodations they need to cope with volatile economic circumstances. The efficacy of these efforts often depends on complex customization of services, which is jeopardized by the homogenizing tendencies of both markets and rule-driven administration.

A. GENERAL THEMES

Minimalism expresses disillusionment with the New Deal practice of vast and vague delegation of authority to expert bureaucratic agencies. At the top, the problem is that, because legislative specification generates excessively rigid instructions and invites rent seeking by special interests, statutes must afford discretion to administrators. At the bottom, the problem is the difficulty of controlling the inevitable discretion of the “street level bureaucrats”—welfare caseworkers, police officers on the beat, classroom teachers—who make on-the-spot, face-to-face decisions that often determine the life chances of citizens.⁹

Minimalism also reflects disillusionment with the more ambitious claims among liberals in the 1960s for participatory democracy. Ackerman emphasizes that people have limited time and energy for politics except in times of crisis.¹⁰ Moreover, Sunstein adds, popular deliberation is not all it’s cracked up to be by Utopian democrats. Serious disadvantages arise from tendencies, on the one hand, toward unreflective conformity (“groupthink”) and, on the other, toward unreflective tendentiousness (“group polarization”). Both phenomena are mitigated in decision-making processes that, instead of striving for deliberative agreement, impersonally aggregate individual views.¹¹ The minimalists believe in the possibility of meaningful public-regarding deliberation by officials and citizens, but democracy for them is mostly indirect. It takes place through electoral representation rather than popular participation.

The minimalists look to economics for norms and practices that obviate official discretion. The core economic norm is efficiency, which prescribes that resources be invested so that at the margin their return is equal and that duties

9. See BRUCE A. ACKERMAN & WILLIAM T. HASSLER, *CLEAN COAL/DIRTY AIR* 38–57 (1981); MASHAW, *supra* note 2, at 49–78.

10. Ackerman makes an exception for “constitutional moments” of heightened popular mobilization and deliberation. See Ackerman, *supra* note 2, at 1763, 1810–11. He is also open to the idea of occasional referenda on specific issues, and he likes the idea of quadrennial “deliberation days” during which citizens can discuss and make recommendations on key issues to elected leaders. See BRUCE ACKERMAN & JAMES S. FISHKIN, *DELIBERATION DAY* 3–5 (2004). However, he uses the constitutional-moment idea more to explain the authority of past events than to inspire mobilization over current issues. And he sees direct deliberation as rare and as engaging issues only selectively or at a high level of abstraction.

11. See Cass R. Sunstein, *Group Judgments: Statistical Means, Deliberation, and Information Markets*, 80 N.Y.U. L. REV. 962, 965, 1004–06 (2005).

be assigned to those who can perform them most cheaply. Basic economics suggests that standard markets will achieve both tasks in the presence of certain preconditions. The minimalists insist that there is nothing inherently conservative about such ideas so long as we take account of the ways in which the preconditions to market efficiency are absent (and so long as efficiency is not permitted to displace distributive and dignitary values).¹²

The minimalists would constrain policymakers at the top through rigorous cost-benefit analysis. They support the presidential orders that have required agencies to engage in cost-benefit analysis before promulgating rules (though they deplore the tendentious misuse of such analysis in the service of political goals). They have drawn on and contributed to a large body of literature that tries to give structure to such analysis, in particular by showing how regulatory benefits such as greater health, longer life, and lower mortality can be taken into account systematically.¹³

In welfare programs, they propose a “social-cost accounting” that balances the benefits of individualized treatment and adjudicatory process against its costs. The availability of trial-type hearings should be weighed against the magnitude of the cost of wrongful deprivation of the benefit in question. In setting substantive standards, administrators should choose between rigid rules and contextual standards so as to maximize the social benefits of “transparency,” “accessibility” (ease of application), and “congruence” (with underlying purposes).¹⁴

At the level of implementation, the minimalists have a preference that, if feasible, regulatory programs take a form that complements or induces markets. The core examples are tradeable emissions permits and school or housing vouchers. Reforms of this kind promise two benefits. First, they incorporate the optimizing tendencies of markets. Second, they reduce the number and complexity of the programs’ norms, thus cabining discretion.

Cap-and-trade regimes respond to defects in the more command-and-control approach of the Clean Air Act. Under the Act, polluters have been obligated to reduce emissions to the level of the “best available” mitigation technology. Such abatement is inefficient, however, because it imposes uniform duties that disregard large variations in the compliance costs of emitters. Moreover, determining the best available mitigation technology for myriad highly technical functions is just the sort of task to which bureaucracies are unsuited, especially because the information needed for sound decisions is held by private firms

12. See Bruce A. Ackerman, *Foreword: Law in an Activist State*, 92 YALE L.J. 1083, 1119 (1983).

13. See, e.g., MATTHEW D. ADLER & ERIC A. POSNER, *NEW FOUNDATIONS OF COST-BENEFIT ANALYSIS* 2-8 (2006); RICHARD L. REVESZ & MICHAEL A. LIVERMORE, *RETAKING RATIONALITY: HOW COST-BENEFIT ANALYSIS CAN BETTER PROTECT THE ENVIRONMENT AND OUR HEALTH* 131-43 (2008); CASS R. SUNSTEIN, *RISK AND REASON: SAFETY, LAW, AND THE ENVIRONMENT*, at x-xii (2002).

14. See Colin S. Diver, *The Optimal Precision of Administrative Rules*, 93 YALE L.J. 65, 66 (1983); Jerry L. Mashaw, *Administrative Due Process as Social-Cost Accounting*, 9 HOFSTRA L. REV. 1423, 1435-36 (1981).

with many incentives to withhold it from regulators.

Instead of attempting to define feasible technical solutions, the minimalists propose that the regulator issue allowances authorizing emissions of specified pollutants in socially acceptable aggregate amounts and then compel emitters to buy such permits in a market in which they can be freely traded. Firms that can abate emissions cheaply do so, and those with high costs buy the allowances. The market thus allocates the burden of a prescribed degree of abatement to the cheapest cost avoiders with a precision that no bureaucracy could achieve directly.¹⁵

In social welfare, the minimalists support school and housing vouchers with which beneficiaries choose for themselves from among a range of providers certified as meeting basic general standards. The beneficiary receives a voucher that she can exchange for housing or schooling, and the chosen provider, upon surrendering the voucher, receives a specified payment from the government. As emissions permits contribute to optimization on the cost side by allocating duties to the cheapest cost avoider, vouchers do so on the benefit side by providing the beneficiary with the housing or schooling she deems most valuable. Moreover, in principle, vouchers put pressure on providers that make them more accountable to beneficiaries. If the beneficiaries are not satisfied, they go elsewhere, and the providers lose the payments.¹⁶

At their most ambitious, minimalists would replace much of the existing welfare state with what Ackerman and Anne Alstott call a “stakeholder” regime in which each citizen, upon turning eighteen, receives a substantial, one-time endowment that can be used in whatever ways she decides—“to start a business or pay for more education, to buy a house or raise a family or save for the future.” The goal of this universal voucher is to provide “a fair starting point for all” while protecting individuals from bureaucratic interference and eliminating the work-disincentive effects of means testing. “Stakeholders are free . . . [to make their own decisions, but their] triumphs and blunders are their own.”¹⁷

Where markets already exist and the problem is people’s limited capacity for effective choice, Sunstein has urged that interventions should often be limited to nudges designed narrowly to compensate for the relevant cognitive impairments. These microinterventions involve altering a default rule or attending to “choice architecture” by providing information to consumers. So, for example, if people appear to be saving too little (as indicated by postretirement poverty or expressions of regret), the solution is not to mandate savings, but to change the default rule for tax-subsidized, employer-run savings plans. The original default

15. See Bruce A. Ackerman & Richard B. Stewart, Comment, *Reforming Environmental Law*, 37 *STAN. L. REV.* 1333, 1342–45 (1985).

16. See Goodwin Liu, *Education, Equality, and National Citizenship*, 116 *YALE L.J.* 330, 403 (2006).

17. ACKERMAN & ALSTOTT, *supra* note 2, at 5; see also MICHAEL J. GRAETZ & JERRY L. MASHAW, *TRUE SECURITY: RETHINKING AMERICAN SOCIAL INSURANCE* 150–51 (1999) (proposing national programs of “cash income replacements or supplements, mandated individual savings, and certain forms of stop-loss insurance” that avoid “highly individualized discretion and decisionmaking”).

rule provided that the employer direct a portion of compensation to savings if the employee requested; the new rule would require the employer to contribute to savings unless the employee requests otherwise. By influencing without preempting choice, nudges protect impaired choosers from their mistakes, while maximizing autonomy and avoiding error costs by leaving a significant range for choice.¹⁸

Where administrative judgments are unavoidable, the minimalists prefer to limit discretion with clear and easily monitored rules. A key example here is the “vocational” determination in the Social Security disability programs. Eligibility requires a finding that the applicant is incapable of performing any jobs that exist in significant numbers. For decades, the decisions were made on the basis of all-things-considered judgments. But “[o]ne disability examiner’s ‘broken down manual laborer’ is another’s ‘malingerer.’”¹⁹ The system of appeals to administrative law judges simply reproduced the problem at a second level. The minimalist solution is a grid that dictates a conclusion on the basis of a series of straightforward findings about age, education, physical capacity, and work experience. The problem of street-level bureaucracy is avoided by keeping the bureaucrats off the streets.²⁰

B. LIMITATIONS: REGULATION

We turn now to our reservations about minimalism, beginning with the minimalist treatment of regulation.

1. The Efficiency Perspective

The main preoccupation of minimalist regulatory scholarship is efficiency or optimization. Yet efficiency is not the only, and often not the most important, regulatory value. Management theorists often espouse a different value, which they suggest is sometimes in tension with efficiency. Reliability is one name for this norm. It connotes a capacity for learning and innovation, or more specifically, for prompt recognition of and adaptation to changing threats and opportunities.²¹

We do not doubt that the minimalists would acknowledge the importance of

18. See RICHARD H. THALER & CASS R. SUNSTEIN, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* 9–11 (2008). Influenced by Sunstein, President Obama proposed to mandate that large employers automatically enroll employees in retirement savings programs, and such a requirement was incorporated in the Patient Protection and Affordable Care Act, Pub. L. No. 111-148, § 1511, 124 Stat. 119, 252 (2010) (to be codified at 29 U.S.C. § 218).

19. Mashaw, *supra* note 14, at 1436.

20. See A. I. Ogus, *Bureaucrats as Institutional Heroes*, 7 OXFORD J. LEGAL STUD. 309, 313 (1987). Summarizing Mashaw’s view, A. I. Ogus writes, “Bureaucratic rationality involves accurate and cost-effective implementation of centrally determined goals; it avoids direct confrontations with the claimant, concentrating instead on the application of detailed rules and guidelines and with internal systems of management and control to ensure consistency.” *Id.* at 311.

21. See Martin Landau & Donald Chisholm, *The Arrogance of Optimism: Notes on Failure-Avoidance Management*, 3 J. CONTINGENCIES & CRISIS MGMT. 67, 67 (1995).

reliability or that reliability could be subsumed under a thick definition of efficiency. However, as the efficiency idea is thickened to embrace reliability concerns, it loses some of the precision and rigor that minimalists see as its core virtues. In practice, the minimalists seem to deploy a static conception of efficiency that risks slighting reliability concerns.

Reliability entails responsiveness, not just to strong signals like prices but also to weak signals such as small anomalies or deviances. Weak signals may not be visible in market prices. They differ from prices and other strong signals in being more numerous and diverse, and in requiring deliberative consideration and complex judgment. Consider an incident at the Davis–Besse nuclear power plant near Toledo, Ohio in the early 2000s. Maintenance personnel found rust particles “mysteriously clogging” the air-conditioning and water filters. They had to replace the filters every two days. The industry standard for replacing such filters was two months, although the workers involved did not know this. After every-other-day replacements went on for two years, different personnel discovered that a metal liner built to contain radioactive material had corroded from its original six-and-one-half-inch thickness to about one-half inch. The corrosion had been generating the rust particles. Karl Weick and Kathleen Sutcliffe assert:

The rust accumulation was a weak signal of plantwide problems that might have been detected sooner had information about industrywide experience been disseminated more thoroughly, had local personnel compared this experience with filter replacement in other parts of the facility, had the purchasing department questioned the large orders for filter replacements, or simply if people had started asking around about whether replacement every forty-eight hours seemed out of line.²²

The best known examples of organizations focused on detecting and assessing weak signals are “high reliability” organizations in areas such as aviation or energy production, in which operational breakdown threatens catastrophe, and “lean production” manufacturing firms, in which competitive pressures arising from short product cycles and customer demand for specialized features call for the kind of discipline that the threat of catastrophe imposes on high-reliability organizations. Perhaps the most influential examples have been the U.S. Navy’s nuclear submarine program and the product-development and manufacturing model associated with Toyota.²³

22. KARL E. WEICK & KATHLEEN M. SUTCLIFFE, *MANAGING THE UNEXPECTED: RESILIENT PERFORMANCE IN AN AGE OF UNCERTAINTY* 46 (2d ed. 2007).

23. See Charles F. Sabel, *A Real-Time Revolution in Routines*, in *THE FIRM AS A COLLABORATIVE COMMUNITY: RECONSTRUCTING TRUST IN THE KNOWLEDGE ECONOMY* 106, 122 (Charles Heckscher & Paul S. Adler eds., 2006). See generally STEVEN J. SPEAR, *CHASING THE RABBIT: HOW MARKET LEADERS OUTDISTANCE THE COMPETITION AND HOW GREAT COMPANIES CAN CATCH UP AND WIN* (2008) (examining the successful organizational cultures of Toyota and the U.S. Navy).

The Navy's nuclear propulsion program developed an ethos and a set of practices that treated even minor, unexpected observations as symptoms of potential systemic malfunction calling for inquiry and possibly reform. The Toyota Production System deliberately stresses its processes (for example, by "just-in-time" parts delivery that eliminates buffer inventory stocks) so that errors will surface quickly. It eliminates end-of-the-line rework departments in favor of practices that stop the line when defects are discovered and require that the "root cause" of the problem be diagnosed and remedied before production resumes.²⁴

Although there is no tension in principle between efficiency and reliability, many have observed that efficiency rhetoric can, in practice, inhibit the pursuit of reliability.

First, efficiency is associated with a preference for simple and few norms and hence crude categories. Such preferences encourage a tendency to "normalize the unexpected," to reframe deviant observations in ways that assimilate them to previous understanding, rather than to treat them as opportunities to revise such understanding. Burn marks on the *Challenger* space shuttle and the rust particles at the Davis-Besse nuclear plant were initially perceived as anomalies but, in the absence of pressure to focus on them, were reconceived as normal. Arguing against "simplification," Weick and Sutcliffe assert, "[T]he diagnostic value of weak signals is lost when . . . details are lumped into crude, general categories." They continue, "Categories may improve coordination [that is, static efficiency], but they harm detection of events not seen before."²⁵

Second, the efficiency view is associated with a preoccupation with short-term cost minimization. The pursuit of reliability requires incurring immediate and measurable costs in the hope of speculative future benefits. Those who have been taught to focus on efficiency sometimes find this hard to do. Reporting on the transformation of nuclear-power safety in the 1990s, Joseph Rees writes that under the old regime, "maximizing the output of electricity meant fixing the reactor *after* it breaks." Conventionally trained managers resisted supporting "operating experience review" engineers who sought to prevent all breakdowns because the benefits of their work could not be measured. The managers were uncomfortable with committing resources to "solve problems that haven't even occurred yet."²⁶ Books on Toyota-style organization assert incessantly that, although optimization and reliability ultimately go together, managers need to defer or sublimate optimization concerns and focus on reliability to achieve

24. See generally Sabel, *supra* note 23, at 119–26 (discussing high-reliability organizations).

25. WEICK & SUTCLIFFE, *supra* note 22, at 64; see also DIANE VAUGHAN, *THE CHALLENGER LAUNCH DECISION: RISKY TECHNOLOGY, CULTURE, AND DEVIANCE AT NASA 77–118* (1996) (attributing the *Challenger* explosion in part to a cultural tendency toward "normalization of deviance" that minimizes weak signals of problems).

26. JOSEPH V. REES, *HOSTAGES OF EACH OTHER: THE TRANSFORMATION OF NUCLEAR SAFETY SINCE THREE MILE ISLAND 130–31* (1994).

success.²⁷

When markets fail, it is sometimes more important for regulation to strive to induce reliability than efficiency (in the static, cost-minimization sense that the term often connotes). The regulators need to develop capacities for learning and adaptation in their own systems, and they need to encourage the firms they oversee to cultivate such capacities in a manner that gives appropriate weight to relevant public values.²⁸

When goals are framed in terms of catastrophe avoidance (as with high-reliability organizations), the need for reliability is obvious, but reliability concerns are also potentially relevant when the relevant harms are more diffuse and frequent. Criminal justice is an important borderline case. Minimalism resonates with the law-and-economics perspective that directs attention to “getting prices right” by correctly calculating penalties. This view emphasizes cost-benefit analysis to approximate penalties to the social value of the relevant harm (adjusted to reflect probability of apprehension and conviction). A competing perspective argues for more attention to and resources for prevention. In this latter view, we should have fewer prisons and more police, and we should allocate policing strategically. We can do this by generating visible police presence in areas where crime is likely, by strictly enforcing against small offenses (for example, “quality-of-life” offenses) that lead to larger ones, or by targeting especially dangerous individuals or groups. The focus here is not on prices or penalties but on weak signals from conditions that are not grave in themselves but are symptomatic of danger. The approach calls for collecting information and coordinating the efforts of public agencies and private organizations to interpret that information and to rapidly adjust responses to it.

The cost-benefit approach dominated discussions of criminal justice in the

27. See MASAOKI IMAI, *KAIZEN: THE KEY TO JAPAN'S COMPETITIVE SUCCESS* 49–50 (1986) (“Japanese managers have found that seeking improvement for improvement’s sake[] is the surest way to strengthen their companies’ overall competitiveness. If you take care of the quality, the profits will take care of themselves.”); RAJAN SURI, *QUICK RESPONSE MANUFACTURING: A COMPANYWIDE APPROACH TO REDUCING LEAD TIMES* 56 (1998) (“[T]he manufacturing world learned [from Japanese firms] that . . . [i]f one focused on improving quality, then cost competitiveness would follow.”).

Neoclassical economics encourages the hope that, if the administrator gets the prices right, the regulated firms will spontaneously develop whatever skills are necessary to succeed in the reoptimized market. But other economic learning shows that we cannot assume that this will happen. Key actors may be able to externalize the costs imposed by the regulator. Notably, corporate agents can shift costs to shareholders, and shareholders, with limited liability, can shift them back to the public. Moreover, as we will illustrate shortly, the learning that occurs through the analysis of weak signals has a public-good quality. It is valuable to others, but it is usually not feasible for the discoverer to charge for it. (Information comes in small pieces of initially uncertain value, and it becomes obsolete quickly.) Under the circumstances, there is a substantial probability that, without further public intervention, private investment in the capacity to detect and interpret weak signals will be suboptimal.

28. Some analyses of the financial crisis that began in 2007 associate regulatory failure with a preoccupation with static efficiency at the expense of reliability. See, e.g., FIN. SERVS. AUTH., *THE TURNER REVIEW: A REGULATORY RESPONSE TO THE GLOBAL BANKING CRISIS* 40–42 (2009). See generally William H. Simon, *Optimization and Its Discontents in Regulatory Design: Bank Regulation as an Example*, 4 REG. & GOVERNANCE 3 (2010) (contrasting “the optimization and reliability perspectives”).

1980s and 1990s. It was, however, challenged in the 1990s by proposals for “community policing” regimes, which were adopted with occasional success in many cities.²⁹ September 11, 2001 made this weak signal reliability perspective more salient. Because it involved catastrophic harm, it shifted attention to prevention, and because the critical performance failures involved sharing and analysis of ambiguous information, initiatives emphasized the importance of the capacity for adaptive response rather than for efficient penalty calibration.³⁰

2. Minimalist Intervention

Both of the minimalists’ favorite regulatory tools—cost–benefit analysis and tradable emissions permits—reflect a preoccupation with static efficiency. Although both tools are valuable, they are not as valuable as the minimalists suggest, and their efforts may have obscured or distracted attention from some important issues.

a. Cost–Benefit Analysis. Any sensible regulatory proposal should consider costs. And it is often a valuable heuristic exercise to compare the costs with the benefits of a particular proposal or the relative return on public expenditures from different proposals, even if the exercise requires controversial and speculative estimations. It is useful to know that the cost per life saved from the EPA’s trihalomethane drinking water standard is \$200,000, while the cost per life saved from its 1,2-Dichloropropane standard is \$653 million.³¹ The comparison, even if plausibly calculated, does not establish that either standard is mistaken, but it gives a reason to revisit both.

Nevertheless, the potential value of cost–benefit analysis is limited by its focus, at least in practice, on static factors as opposed to the capacity for learning and adaptation. First, the proponents have been preoccupied with methodological issues concerning the estimation of benefits (for example, the value of a life or of different kinds of lives, or the right discount rate for future lives), whereas, in practice, conclusions will often be dominated by estimates of future costs, which will necessarily be highly speculative. Richard Revesz and Michael Livermore note a persistent tendency to underestimate cost-saving innovation and overestimate costs. They give examples of estimates that proved to be wrong by “several orders of magnitude.”³² Methodological rigor makes only a slight contribution when the inputs are so soft.

Second, the proponents tend to defend cost–benefit analysis in the abstract without much attention to how it might be institutionalized.³³ The value of

29. See ARCHON FUNG, *EMPOWERED PARTICIPATION: REINVENTING URBAN DEMOCRACY* 63–68, 173–210 (2004).

30. See Daniel Richman, *The Right Fight*, *BOS. REV.*, Dec. 2004–Jan. 2005, at 6, 6.

31. See SUNSTEIN, *supra* note 13, at 30–31 tbl.2.1.

32. See REVEZ & LIVERMORE, *supra* note 13, at 135.

33. An important exception on which we rely here is outlined in Nicholas Bagley & Richard L. Revesz, *Centralized Oversight of the Regulatory State*, 106 *COLUM. L. REV.* 1260, 1312–29 (2006).

cost-benefit analysis as an intellectual discipline does not imply any particular conclusions about institutional structure. To the extent that analysis involves comparison of costs and benefits across different regulations and agencies, it suggests a central role for standardizing methodology and collecting and disseminating data, but this central role could be a subordinate one performed without significant loss of authority to the agencies. Or it could involve a substantive oversight role played with varying degrees of strictness. The latter approach raises difficult questions as to how central administrators can absorb and coherently assess the large volume of information pertinent to the federal regulatory output. The former approach involves a more modest role for cost-benefit analysis than its centrality in the minimalist literature implies.³⁴

In practice, the principal effect of cost-benefit analysis in the federal government in recent years seems to have been to rationalize a shift in authority from the agencies to the White House. This shift has at least occasionally involved heavy-handed, ad hoc, central intervention, sometimes rationalized by tendentious and ideologically driven estimations of costs and benefits. It also appears to have involved significant diminution in the transparency of the regulatory process because the White House regulatory offices are less transparent than the agencies. The Office of Information and Regulatory Affairs, which oversees cost-benefit review, is not subject to judicial review under the Administrative Procedure Act, and it “has a long and well-documented history of secrecy.”³⁵ No doubt the minimalists regret these developments, which are not entailed by cost-benefit analysis in principle. But the developments show the limitations of a literature preoccupied with analytical methodology rather than institutionalization.

Third, as practiced, cost-benefit analysis is associated with the enactment or rule-making stage of policy formulation. For many problems, relevant information will be too scarce and ambiguous to support useful formal analysis at this stage. In these situations, interventions should be designed in part to produce information that will guide adaptation, and, sometimes, they may eventually produce the kind of information that makes formal cost-benefit analysis plausible. If so, the important occasions for cost-benefit analysis will occur after the program is underway. Yet the minimalists seem to accept the current practice that confines it to the point of initiation.³⁶

b. Cap-and-Trade. The most developed regulatory proposal in the minimalist literature is the cap-and-trade emissions permit system, which epitomizes the minimalists’ virtues of static efficiency (allocating burdens to the cheapest cost

34. Bagley and Revesz take something like the modest position. *Id.* at 1313–14.

35. *Id.* at 1282–83, 1309.

36. See *id.* at 1264–68; Cass R. Sunstein, Adm’r, Office of Info. & Regulatory Affairs, Address at American University’s Washington College of Law *Administrative Law Review* Conference: Humanizing Cost-Benefit Analysis (Feb. 17, 2010), available at http://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/cost_benefit_analysis_02172010.pdf.

avoider) and minimizing discretion.³⁷ Emissions trading has proven successful in a few programs involving modest numbers of large-scale industrial producers and well-studied pollutants, especially lead from oil refiners and sulfur dioxide from utilities and factories.³⁸ The idea seems likely to prove valuable in other contexts, perhaps at larger scales, but its range is more limited than the minimalists suggest. In particular, four limitations should be noted.

First, the cap-and-trade approach only works with harms that can be effectively measured.³⁹ Although the environmental harms of air pollution cannot be measured directly, they correlate sufficiently with the amount of emissions of specific pollutants, and, for some sources and pollutants, we can measure that. However, for most regulatory goals, we have no comparably powerful proxies. Regulatory regimes are usually prophylactic; they are designed to minimize risk. With food safety, occupational health and safety, or financial institution soundness, for example, there is no single powerful indicator of risk. In such situations, we are forced to rely on weak signals—a multifarious range of indicators that require interpretation and complex judgment and do not lend themselves to pricing.⁴⁰

Second, even when adequate proxies are available, the information demands of cap-and-trade remain large and sometimes disqualifying. The regulator still has to set an aggregate quantity limit (or in the tax alternative, a price) for each pollutant. Because the efficacy of the regime depends on this single number, the stakes are extremely high. Yet both methodology and data for this determination are likely to be incomplete and disputed. The regulator needs to estimate not only the immediate effects of abatement of the target pollutant but also collateral costs and benefits of behavior induced by the abatement incentives—costs, for example, of unregulated but still harmful pollutants that might be substituted for the regulated ones; benefits, for example, of unregulated but harmful pollutants that might be abated by the same efforts required to abate the

37. See, e.g., Ackerman & Stewart, *supra* note 15, at 1341–42.

38. See NATHANIEL O. KEOHANE & SHEILA M. OLMSTEAD, *MARKETS AND THE ENVIRONMENT* 182–90 (2007); Richard G. Newell & Kristian Rogers, *The Market-Based Lead Phasedown 1* (Res. for the Future, Discussion Paper No. 03-37, 2003), available at <http://www.rff.org/RFF/Documents/RFF-DP-03-37.pdf>. Another promising trading initiative is the Regional Greenhouse Gas Initiative undertaken by ten mid-Atlantic and northeastern states. See REGIONAL GREENHOUSE GAS INITIATIVE, <http://www.rggi.org> (last visited July 19, 2011).

39. See Cary Coglianese & David Lazer, *Management-Based Regulation: Prescribing Private Management To Achieve Public Goals*, 37 *LAW & SOC'Y REV.* 691, 725–26 (2003).

40. Of course, measurement needs to be economically as well as technically feasible. Moreover, trading requires an infrastructure (for example, to identify owners and record trades) and entails transaction costs. These prerequisites will limit the application of the model to actors of more than minimal size or abatements of more than minimal social value. The thresholds are met with respect to the kind of large power-generating and manufacturing facilities covered by current regimes. However, to the extent that the relevant problems involve smaller, more diffuse activities, they may not be covered. For example, the model seems unpromising for the increasingly salient problems of emissions from “non-point sources” such as farm-water-pesticide runoff. See Bradley C. Karkkainen, *Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?*, 89 *Geo. L.J.* 257, 264–67 & n.17 (2001).

regulated ones. Moreover, while the market regime obviates determinations of relative abatement costs among emitters, it requires knowledge of aggregate costs. To determine the efficient price, we need to know the point at which marginal aggregate costs equal marginal aggregate benefits. Yet such determinations are notoriously difficult.

It is no surprise then that the few examples to date of operating cap-and-trade systems include some spectacular examples of getting quantities, and hence prices, wrong. In 2000, quantities proved too low to accommodate an unanticipated surge in demand in the South Coast Air Quality Management District trading regime in California. Prices for nitrous oxide emissions soared from less than \$2,500 per ton in 1999 to \$45,000 per ton in 2000.⁴¹ Another crisis emerged with the inauguration of the European Union's greenhouse-gas-emissions trading regime. Allowance prices collapsed in 2006, when it became apparent that the regulators had overestimated current business-as-usual emissions, and thus set an aggregate target that was too high, resulting in the distribution of too many allowances.⁴²

It is true that market-simulating reforms like cap-and-trade will often require less information and complex judgment than comparable command-and-control regimes. However, this advantage comes at a cost.⁴³ There is a risk that market simulation will involve a sacrifice in flexibility, particularly in the capacity to revise in response to new information. Part of the market simulation idea is to induce private investment by creating marketable entitlements. Investors will claim that they need and deserve some measure of stability with respect to the key parameter of price or quantity. Thus, allowances may be good for several years—often five in the EU—and can sometimes be “banked” for use beyond that period.⁴⁴ The regimes typically place limits on the ability of regulators to dilute the value of the permits by new issuances before the end of a period. Regulatory tightening is apt to trigger protest under any regime but, by encouraging emitters to view the current limits as an entitlement for a fixed period, cap-and-trade may exacerbate the problem.

The third limitation of market-simulating regimes is that, although they are sensitive to variations in abatement costs among emitters, they are insensitive to variations in abatement benefits across the localities where emissions occur. At least in the paradigmatic model, there is a single price for emissions throughout the system. Yet the harm that emissions cause usually varies locally with such

41. See Raul P. Lejano & Rei Hirose, *Testing the Assumptions Behind Emissions Trading in Non-Market Goods: The RECLAIM Program in Southern California*, 8 ENVTL. SCI. & POL'Y 367, 373–74 & fig.4 (2005).

42. See Heather Timmons, *Data Leaks Shake Up Carbon Trade*, N.Y. TIMES, May 16, 2006, at C1, C8.

43. See KEOHANE & OLMSTEAD, *supra* note 38, at 164–68.

44. See Anthony Hobley & Peter Hawkes, *GHG Emissions Trading Registries*, in LEGAL ASPECTS OF IMPLEMENTING THE KYOTO PROTOCOL MECHANISMS 127, 147 (David Freestone & Charlotte Streck eds., 2005).

contingencies as population density, aggregate emissions levels, and land-use patterns. The marginal harm in a heavily polluted urban area is often much higher than in a sparsely populated area with low aggregate emission levels. Moreover, market regimes may encourage the creation of “hot spots.” These occur when polluters minimize compliance costs by clustering in low cost areas (areas with low land costs and limited local taxes and regulation). When this happens, pollution burdens concentrate in particular localities. It is possible that these localities will disproportionately include low-income and otherwise disadvantaged people. The hot-spot problem does not apply to greenhouse gases because their effect operates at the global level, and emissions from any point contribute equally. However, for most pollutants, location matters.⁴⁵

Trading regimes can address the localization problem by subdividing markets. Thus, in the South Coast Air Quality Management District, permits sell at different prices in coastal and inland districts. Yet such subdivision multiplies the number of judgments and the amount of information required.⁴⁶

Fourth, it is not clear how powerful the incentives for the production and dissemination of innovative technology created by trading systems will be. To the extent that innovations take the form of major capital goods, especially with

45. Although the global nature of greenhouse-gas effects on temperatures exempts it from the hot-spot problem, it creates a tragedy-of-the-commons problem. Only global efforts, or at least efforts including all major emitting nations, are likely to be strongly effective. Yet it has proven difficult to induce major emitters, especially India and China, to participate substantially. Two responses show the limits of the market approach:

1. The Kyoto Protocol and other regimes allow emitters within the regime to earn credits by subsidizing projects in nonparticipating countries that reduce emissions (or increase gas absorption) beyond what would have happened in the absence of the project. A major fraction of the reductions credited to Kyoto have involved such efforts.
2. Some observers propose a focus on negotiated deals for large infrastructure projects in which rich countries trade economic benefits for developing-country commitments favoring low-emission energy production (natural gas and nuclear power rather than coal and oil).

A key point about both approaches is that they cannot operate primarily through self-adjusting price signals and impersonal trading. They require complex administrative capacity. Failure to appreciate this point accounts for the early disappointments of the Kyoto Clean Development Mechanism. It appears that a large fraction of the credits certified under the regime involved abatements that would have occurred without the regime or could have been induced at a much lower cost than the regime paid. A core problem is that the regime has encouraged private institutional development that has outpaced the development of public institutional capacity to restrain abuses. See Michael Wara, *Is the Global Carbon Market Working?*, NATURE, Feb. 8, 2007, at 595, 595. See generally James Salzman & J.B. Ruhl, *Currencies and the Commodification of Environmental Law*, 53 STAN. L. REV. 607 (2000) (discussing the optimal architecture of environmental trading markets).

46. In their early proposal, Ackerman and Stewart proposed extensive subdivision. See Ackerman & Stewart, *supra* note 15, at 1350. Jonathan Nash and Richard Revesz speculated on the possibility of a massive computer program aggregating information about local climate effects that would continuously generate local prices for emissions of controlled pollutants. See Jonathan Remy Nash & Richard L. Revesz, *Markets and Geography: Designing Marketable Permit Schemes To Control Local and Regional Pollutants*, 28 ECOLOGY L.Q. 569, 609–10 (2001). It seems ironic that a discussion that started out with the goal of minimizing the information demands on regulators should lead to a proposal that resembles nothing so much as Soviet central planning.

patentable features, they will be marketed and adopted when likely to be profitable. But many innovations take the form of incremental improvements or involve adaptation of technologies to particular local conditions. Without public efforts, small incremental improvements may not be disseminated because of transaction costs. Moreover, technology may not be usable without significant customization, and small and medium producers may lack access to necessary technical assistance.⁴⁷ It is possible that trade associations or business consultants can fill this gap, but it is not inevitable. The example of the U.S. Agricultural Extension Service shows that publicly subsidized technical assistance may be needed to induce small- and medium-sized producers to adopt socially desirable practices and to protect them from the competitive disadvantages that they might suffer vis-à-vis larger producers as a result of regulations requiring technologically complex responses.⁴⁸

C. LIMITATIONS: SOCIAL WELFARE

The minimalist welfare program of nudges (manipulation of default rules and choice architecture), norm simplification, and vouchers is likely to prove helpful with respect to some problems. But the approach seems implausible or question begging with respect to many of the most important problems. It is unresponsive to some salient trends in welfare policy, and the most promising recent innovations have taken a much different approach.

1. The Changing Policy Context

The New Deal welfare system was focused on cash transfers under categorical rules. Its central programs were modeled on private insurance, and they reflected an actuarial conception of risk. The basic reference point was a typical worker: a male breadwinner working for forty years, after an initial period of searching, in the same job or a succession of similar jobs in the same industry. His career would be interrupted by spells of unemployment, but these spells would correlate with systematically calculable phases of the business cycle. The Social Security Act created programs of wage-related benefits for this worker and his dependents during bouts of unemployment, after retirement, in case of his disability, and for his surviving dependents after his death. A parallel program of means-tested public-assistance benefits for people who could not work and lacked sufficient employment history to qualify for social insurance was expected to wither away as the economy recovered from the Depression

47. See ERIC VON HIPPEL, *DEMOCRATIZING INNOVATION* 67 (2005) (explaining that the costs of transferring even “fully developed process technology” to new settings can be high).

48. Atul Gawande has argued that in healthcare, “New laboratory science is not the key to saving lives. The infant science of improving performance—of implementing our existing know-how—is.” A new drug with a fraction of the effectiveness of his simple but counterintuitive surgery protocol would have attracted a vast amount of private capital, but developing and demonstrating the “checklist” required public support. See ATUL GAWANDE, *BETTER: A SURGEON’S NOTES ON PERFORMANCE* 242–43 (2007).

and breadwinners built up their social insurance entitlements.

Of course, the public assistance programs did not wither away. The problems that they addressed persisted and grew even in periods that approached full employment. This basic development has required extensive rethinking of these programs and their relation to the social insurance programs. An important body of minimalist scholarship has struggled with this task.⁴⁹ As we've noted, this scholarship has sought to preserve the New Deal emphasis on standardized, rule-defined cash benefits while broadening the scope of both the social insurance and public assistance programs.

These efforts, however, have not taken full account of changes in the social context that underlay the New Deal commitment to standardized, rule-defined benefits, and they have largely ignored the most distinctive policy responses to them. The key social changes concern increased diversity and volatility. The distinction between labor-market participants and nonparticipants has eroded. People whom the New Deal model considered normally outside the labor market—the elderly, disabled, housewives, single mothers—are now thought to have a right, a duty, or both to participate effectively. Long-term careers in the same job or industry are less common; the skills that jobs demand vary more widely and change more rapidly than before. Long-term marriage and two-parent child rearing, which the New Deal took as normal, have diminished, and family structures have become more diverse. The social backgrounds and linguistic capacities of citizens vary increasingly. Some social problems, including many involving mental health and substance abuse, that were not prominent in the New Deal discussions are highly salient today.⁵⁰ And the welfare system has gradually assumed responsibility for healthcare under circumstances in which both relevant technology and needs are changing rapidly.

The most distinctive policy innovations induced by these developments focus on capacitation rather than maintenance or compensation. Where actuarial risk pooling breaks down in an uncertain world, the welfare state strives to equip people with skills to deal with the disruptions they face, whether in the market or the family. The services of the new programs are typically customized and bundled across disciplines and problems. For example, employment-related services may be combined with family support services, or family support services may combine education, mental health, and medical services.

At the same time, persistent and labile problems call for continuing and adaptive intervention. Responding to unemployment will often require not just temporary income support but also reskilling to respond to changing configurations of job opportunities. Chronic educational and psychological problems—

49. See, e.g., ACKERMAN & ALSTOTT, *supra* note 2, at 15–16; GRAETZ & MASHAW, *supra* note 17; Super, *supra* note 8, at 541–42.

50. See INST. OF MED. & NAT'L RESEARCH COUNCIL, THE DYNAMICS OF DISABILITY: MEASURING AND MONITORING DISABILITY FOR SOCIAL SECURITY PROGRAMS 45–46 (Gooloo S. Wunderlich et al. eds., 2002); Mark Peyrot, *Cycles of Social Problem Development: The Case of Drug Abuse*, 25 Soc. Q. 83, 88–91 (1984).

for example, substance abuse, family dysfunction, and learning disabilities—require recurring interventions that take account of the likelihood of relapse.

The design of services reflects changes that have occurred in professional disciplines. One is the greater capacity to diagnose and respond to idiosyncrasy. For example, in medicine, it appears that genetic indicators will eventually permit the customization of drugs to particular patients. More mundanely but at the moment more importantly, effective treatment for many behavior-dependent diseases such as diabetes is now understood to include the construction and monitoring of an individualized plan that covers medication, diet, and activity.⁵¹ Or to take another example: educators have recently come to understand that the debate between “phonics” and “whole language” approaches to reading was misconceived. Each approach has techniques that are effective with some students. The teacher should not choose wholesale between them but should assess each student to determine the combination of interventions most helpful to him or her.⁵²

Reforms in most key areas of social welfare, including education, child welfare, disability rights, and mental health, have reflected these trends in professional practice. They typically require individual assessment and customization of services; complex, heavily diagnostic judgment; and ongoing reassessment and adaptation.⁵³ The minimalists have not taken account of these developments.

2. Minimalist Intervention

We can assess the limitations of welfare minimalism in terms of the two most characteristic minimalist projects in this area: norm simplification and vouchers or stakeholder funds.

a. Norm Simplification. There is a plausible role for programs based on simple norms that obviate pervasive, complex judgment. In his minimalist critique of experimentalism, David Super mentions models of the kinds of

51. See CLAYTON M. CHRISTENSEN ET AL., *THE INNOVATOR'S PRESCRIPTION: A DISRUPTIVE SOLUTION FOR HEALTH CARE* 56–61 (2009).

52. See Diana Jean Schemo, *Federal-Local Clash in War over Teaching Reading*, N.Y. TIMES, Mar. 9, 2007, at A16.

53. Statutes requiring either individual service plans or “reasonable accommodation” include the Education of the Handicapped Act Amendments of 1990, Pub. L. No. 101-476, § 101(e), 104 Stat. 1103, 1104 (codified as amended at 20 U.S.C. § 602(a)(20)) (“Individualized Education Program”); Personal Responsibility and Work Opportunity Reconciliation Act of 1996, Pub. L. No. 104-193, § 408(b), 110 Stat. 2105, 2140 (codified as amended at 42 U.S.C. § 608(b)) (“Individual Responsibility Plans” for public assistance recipients); Adoption Assistance and Child Welfare Act of 1980, Pub. L. No. 96-272, § 101(a)(1), 94 Stat. 500, 503 (codified as amended at 42 U.S.C. § 671(a)(16)) (“case plan” for children in state custody); Fair Housing Amendments Act of 1988, Pub. L. No. 100-430, § 6(a), 102 Stat. 1619, 1621 (codified as amended at 42 U.S.C. § 3604(f)(3)(B)) (“reasonable accommodations” of handicapped residential housing purchasers); Americans with Disabilities Act of 1990, Pub. L. No. 101-336, § 102(b)(5)(A), 104 Stat. 327, 332 (codified as amended at 42 U.S.C. § 12112(b)(5)(A)) (“reasonable accommodations” for disabled workers).

programs he favors: food stamps and Supplemental Security Income (SSI).⁵⁴ These programs provide easily defined benefits (food vouchers or cash) to broadly defined categories of indigent people—"households" (virtually any person or group that lives outside an institution and prepares its own meals) for food stamps, and aged and disabled people for SSI. If we put aside the complicated issues associated with disability determination in SSI, these programs are based on relatively straightforward rules, and as far as they go, they are successful.

The problem is that these programs do not go very far. They can get by with uniform benefits and simple rules because each addresses only a single and simple need—food or money—and they avoid making judgments at the margin by undersatisfying the need. The programs provide benefits only to the most disadvantaged, and the benefits they provide are meager. Relative to the broader purposes of poverty relief, the programs are spectacularly underinclusive. If food stamps and SSI constituted the whole of America's response to poverty and social need, they would be reviled as abominations. They can only be considered successful as a component of a broader array of programs that provide benefits related to health, education, job training, housing, and other matters. These benefits typically take the form of services, and their providers increasingly aspire to the complex customization that we noted above.

The simplification project in the welfare area to which scholars have devoted the most attention is the vocational-grid regulation in the Social Security Disability programs. These programs condition eligibility on permanent and total disability, which entails a finding that there is no substantial number of jobs that a person in the applicant's medical condition could perform. For decades this "vocational assessment" was made through an all-things-considered judgment, first, by a vocational expert, and second, if there was an appeal, by an administrative law judge. The process was costly, and the resulting judgments were wildly inconsistent. The response was to replace expert judgment in a broad range of cases with a grid consisting of a set of rules that dictated eligibility conclusions on the basis of a small number of easily determinable facts about age, physical capacity, education, and work experience.⁵⁵

Minimalists have supported the grid as an efficient trade-off of "congruence" (fit with substantive purposes) for "accessibility" (less expensive process and more formal consistency). The main objection to the grid is that its simplicity makes for underinclusion. There are some cases that would be deemed eligible under an all-things-considered judgment for which the grid denies eligibility. The minimalists respond that perfect justice is impossible and that the quest for it must yield to resource constraints. Moreover, they suggest that the costs of under- and over-inclusion will be relatively small if, as appears, the relevant

54. See Super, *supra* note 8, at 585, 609.

55. See 20 C.F.R. pt. 404, subpt. P., app. 2 (2010). The grid is discussed in *Heckler v. Chaney*, 470 U.S. 821, 831–32 (1985), and Diver, *supra* note 14, at 88–92.

cases are borderline or marginal. To the extent that incorrectly denied cases resemble correctly denied ones and incorrectly accepted claims resemble correctly accepted ones, the social costs of error are likely to be relatively low.⁵⁶

Two limitations of this line of analysis are becoming salient. First, the grid approach is out of touch with changes in the social understanding of disability. When disability programs were added to the Social Security regime in 1956, it was generally assumed that severely incapacitated people were necessarily and appropriately excluded from the workforce. Moreover, the idea of permanent and total disability was considered an observable, self-defining category. But these notions have been discredited as both normative and empirical matters. Normatively, the inclusion of disabled people in the broadest feasible range of social activities has become a central goal, one that is sometimes given precedence over compensation. Empirically, we have discovered that a combination of rehabilitation services and reasonable accommodation at the workplace can enable a large fraction of those once considered incapable of significant work to perform a broad range of jobs. Thus, a variety of programs for disabled people provide or mandate customized services or workplace accommodation.⁵⁷

It is no surprise then that, as Jerry Mashaw reports, many urge that the Social Security program move away from its compensation focus and its categorical norms toward “a community-based and multidisciplinary approach that would deploy financial assistance, medical care, and rehabilitation and transportation services, among other things, to promote the overall well-being and highest possible functioning of disability beneficiaries.” This revised approach, Mashaw emphasizes, “would demand highly discretionary judgments.”⁵⁸

Second, the most commonly asserted arguments for the grid are static cost-benefit arguments. By contrast, modern management theory emphasizes that detection and correction of individual errors involves opportunities for systemic improvement.⁵⁹ Errors can be corrected in a specialized department that addresses only the particular defective product. Alternatively, they can be addressed diagnostically: the causes of a particular defect can be mapped, and the systemic flaws can be remedied along with the individual problem. As a purely logical matter, either approach might be efficient, but a full assessment has to consider not only accuracy and decision costs but also the opportunity costs and benefits of learning and systemic improvement. Companies that feel pressures and opportunities to innovate often decide that they do better to sacrifice static

56. See Diver, *supra* note 14, at 90–92; Jerry L. Mashaw, *How Much of What Quality? A Comment on Conscientious Procedural Design*, 65 CORNELL L. REV. 823, 824–28 (1980).

57. See generally Alan M. Jette & Elizabeth Badley, *Conceptual Issues in the Measurement of Work Disability*, in THE DYNAMICS OF DISABILITY: MEASURING AND MONITORING DISABILITY FOR SOCIAL SECURITY PROGRAMS, *supra* note 50, at 183, 183–210 (providing a conceptual definition of disability).

58. Jerry L. Mashaw, *Accountability and Institutional Design: Some Thoughts on the Grammar of Governance*, in PUBLIC ACCOUNTABILITY: DESIGNS, DILEMMAS AND EXPERIENCES 115, 154–55 (Michael W. Dowdle ed., 2006).

59. See JAMES P. WOMACK & DANIEL T. JONES, *LEAN THINKING* 133–34 (2003).

optimization in order to strive for continuous improvement. When they do so, they tend to adopt explicitly superoptimal norms such as “total quality” or “zero defects,” refusing to consider (at least rhetorically) that any level of error could be optimal.

In Social Security disability, that cases under- and over-included by the grid regulation tend to be borderline has different implications from static and dynamic perspectives. In a static sense, the proponents are right to suggest that the social cost of incongruence is relatively low. But from a dynamic perspective, the opportunity costs of formal rules in these cases may be relatively high because these may be the cases in which the potential for learning and systemic gain from contextual examination may be highest. Even in the current system, which is largely preoccupied with sorting people into categories of disabled or nondisabled, contextual examination in marginal cases might have exceptional diagnostic value if the characteristics ignored or overemphasized by the rules are shared by relatively large numbers of applicants. For example, close examination might lead to the discovery of powerful indicators of the incapacitating effects of a particular medical condition that would be useful in other cases involving the same condition. Moreover, if we imagine a program reconfigured along the lines of recent criticisms to combine income support with rehabilitation, the learning potential from contextual inquiry may be highest in borderline cases because these may turn out to be the cases in which rehabilitation has the most promise.

b. Vouchers and Choice Architecture. Housing and school vouchers are a natural extension of minimalist logic. David Super expresses particular enthusiasm for housing vouchers. Aside from SSI and food stamps, he points to the Section 8 housing voucher program as a model.⁶⁰ Under this program, eligible recipients get certificates they can use to induce landlords to rent to them. The tenant pays a percentage of her income, and the government pays the difference between that amount and the officially determined “fair rental value” of the apartment. The idea is to give the tenant a broader range of choices and more leverage than traditional public housing, in which there are few choices and, once a tenant moves in, she usually loses her subsidy if she moves.

The evidence on housing and school vouchers remains limited, and we are open-minded on the subject. However, we note some basic limitations to the model. Some of these limitations resemble those observed with emissions permits. Foremost among these are information demands. Perhaps it is easier to design a voucher than to run a school or housing project directly, but voucher programs require administrators to set a price. If the housing voucher price is too low, landlords refuse to participate. If the price is too high, they get a windfall. The Section 8 program Super admires has suffered from both problems. During some periods, a large fraction of certificates awarded to applicants

60. See Super, *supra* note 8, at 609.

were returned unused because the holders could not find a suitable apartment for which the landlord would accept the certificate. At the same time, it has been asserted that in some localities, the certificates gave landlords more than the market value of their apartments.⁶¹

There are further problems with vouchers in the welfare context. Beneficiaries may lack either adequate information or adequate analytical capacity to make good choices with their vouchers (as some minimalists themselves emphasize). In addition, providers have strong incentives to “cream,” that is, to seek the voucher holders who can be served most easily. Instead of competing by trying to provide better services, they can compete by better identifying and attracting low-cost clients. Landlords will look for the most stable families; schools for students with the most promising preexisting abilities and behavioral traits; and medical facilities for the healthiest people.⁶²

Moreover, the effectiveness of schools and housing is widely thought to be influenced by the social mix of the local population. Average achievement may be higher when classrooms have a mix of ability levels than when they are segregated by ability. Disadvantaged people may do better in racially and economically integrated neighborhoods. Integration may require a complex balance of groups, achieving a “critical mass” needed for minority group members to feel comfortable, on the one hand, and avoiding the “tipping point” that triggers the flight of the privileged, on the other. If integration or social balance is a goal, then it seems unlikely that it can be achieved through unaided individual choice. Aside from discrimination by providers and local governments, there is a coordination problem.⁶³

Some combination of information and coordination problems seems to account for the finding that “if given vouchers, poor people choose . . . segregated areas similar to the ones they left.”⁶⁴ The most admired effort to achieve housing integration with vouchers—the Gautreaux program in Chicago—depends heavily on professional services to match applicants with apartments and to counsel them about adjusting to life in their new neighborhoods.

61. See Jenifer J. Curhan, *The HUD Reinvention: A Critical Analysis*, 5 B.U. PUB. INT. L.J. 239, 243–45, 250–51 (1995). The danger of setting the price too low could be avoided by giving the voucher a flat base value and allowing the tenant to add any amount she wants to it. However, some worry that this approach, if widely used, might contribute to rent inflation.

62. Participants in voucher and other school-choice programs tend to have disproportionately high socioeconomic status even though the programs usually require schools to choose among applicants by lottery. Advantaged students are more likely to apply and more likely to remain in these schools. This probably is due, in part, to parental motivation and information but also is likely a function of the schools' outreach and recruiting and the degree of support they provide to students who struggle. See DIANE RAVITCH, *THE DEATH AND LIFE OF THE GREAT AMERICAN SCHOOL SYSTEM: HOW TESTING AND CHOICE ARE UNDERMINING EDUCATION* 129–47 (2010).

63. See James S. Liebman, *Voice, Not Choice*, 101 YALE L.J. 259, 292 (1991) (reviewing JOHN E. CHUBB & TERRY M. MOE, *POLITICS, MARKETS, AND AMERICA'S SCHOOLS* (1990)).

64. James E. Rosenbaum & Stefanie DeLuca, *Does Changing Neighborhoods Change Lives? The Chicago Gautreaux Housing Program and Recent Mobility Programs* 5 (Inst. for Policy Research, Nw. Univ., Working Paper No. WP-09-01, 2009).

Healthcare cost control is the major social-welfare challenge of the day. We doubt that the minimalists' approach can play more than a secondary role in effective reform, and we worry that it will discourage critical reforms that depend on the kind of administrative capacity they are skeptical of.

The central American public healthcare programs, and even private insurance, already have voucher-like characteristics. The main difference between Medicare or Medicaid entitlements and a housing or school voucher is that there is no cap on the amount of benefits that the beneficiary of the health programs can demand. The system thus gives providers an incentive to overtreat and to use relatively expensive treatment that may not even serve the medical interests of the patient, much less social efficiency.

The economists' favored approach to the problem involves making consumers more selective by tightening their budget constraints (for example, by cutting back tax subsidies and encouraging healthcare savings accounts and employer self-insurance).⁶⁵ A competing response involves administrative review of practice decisions under continuously updated standards.⁶⁶

The minimalists are unlikely to be comfortable with either perspective. The economists' vision relies on individual choice to an extent the minimalists know is implausible. The alternative perspective calls for exactly the kind of complex administrative capacity that the minimalists despair of creating.

The difficulties for minimalism in this area can be seen in the chapter on health insurance in Sunstein and Richard Thaler's *Nudge*.⁶⁷ It focuses on the Medicare Part D prescription drug plan inaugurated in 2003. The program requires most beneficiaries to choose among a variety of private plans. Sunstein and Thaler summarize evidence that shows two crippling problems. First, left to themselves, many people will not choose at all, and if the default is no coverage, they will lose coverage. Second, most people do not seem to choose the plan that is most beneficial to them. We know this because researchers can get an idea of what people's drug needs are from the records of their plans and then compare the coverage (deductible, copayments, ceilings) they chose with the coverage they could have gotten under alternative plans.

Sunstein and Thaler propose a classic minimalist response: manipulation of default rules and choice architecture. Instead of no coverage, the default for nonchoosers should be a plan chosen by the government. And the government should make these default choices, not randomly (as is now done for some beneficiaries), but by "intelligent assignment." The government can use available data to calculate which plan is most likely to be the best for nonchoosers and assign them to it provisionally. For choosers, the key is to present the data at the time of decision in a way that helps the beneficiary choose effectively. For

65. See generally JOHN F. COGAN, R. GLENN HUBBARD & DANIEL P. KESSLER, *HEALTHY, WEALTHY, AND WISE: FIVE STEPS TO A BETTER HEALTH CARE SYSTEM* (2005) (proposing market-based healthcare tax, regulatory, and insurance reforms).

66. See CHRISTENSEN ET AL., *supra* note 51, at 56–65.

67. See THALER & SUNSTEIN, *supra* note 18, at 161–74.

example, the choice document can contain a list of the beneficiary's drug uses in the past year and the net cost to the patient of such drugs under alternative plans.

Note that, for their only discussion of health insurance, the authors chose a relatively small and manageable portion of the picture. Even if we saw the kind of assisted choice they propose as effective here, it would be harder to see it working with respect to the more complicated choices involved in, say, choosing a doctor or a hospital or a comprehensive insurance plan.⁶⁸

And, most strikingly, even in this relatively simple case, choice ends up doing very little work. Sunstein and Thaler argue convincingly that, by using sophisticated data aggregation and analysis, the government can make plausible decisions about the interests of the large fraction of people disinclined to choose. They suggest further that the government can use the same tools to present data to choosers in a way that will make their choices effective. It seems quite likely that, for many choosers, the presentation of the data will more or less determine the choice. To the extent that the government can accurately determine the best choice and can present the data that underlies that determination clearly, choosers will simply ratify the government's decision. To the extent that choosers depart from the government prediction, it could mean that they understand their situation better than the government, or it could mean that they are confused. In the former case, choice would serve as a valuable check on the government's analytical apparatus (which presumably could be continuously improved as the government investigates the reasons for the discrepancies). In the latter case, choice is just noise and nuisance. But in either case, the key intervention involves much more than a nudge; it is a kind of complex administration that minimalism tends to acknowledge only as an afterthought.⁶⁹

68. In a 2010 speech, Sunstein, as Administrator of the Office of Information and Regulatory Affairs, described the Obama Administration's "approach to regulation" as focused on, in addition to cost-benefit analysis, "[d]efault [r]ules and [s]implification." See Sunstein, *supra* note 36. Aside from encouraging automatic employer enrollment in retirement savings plans, the only examples he cited involved better disclosure to consumers or workers of the risks of products such as chemicals, tobacco, or bank loans. All the interventions seem sound but, with the exception of financial-products disclosure, they seem minor in comparison to the reforms adopted or proposed by the Administration in such areas as food safety, bank solvency, education, healthcare, and offshore oil drilling. *Id.* The Consumer Financial Protection Board created by the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 appears to be the only major initiative strongly influenced by minimalist ideas. See Elizabeth Warren, *Unsafe at Any Rate*, DEMOCRACY, Summer 2007, at 8, 17-18 (arguing for the importance of consumer-protection interventions in finance based on nudges and choice architecture).

69. You would never know from Sunstein and Thaler's breezy discussion of the potential contributions of minimalism that Medicare Part D involves a policy fiasco on a scale of the SO₂ scrubber legislation that Ackerman and Hassler famously denounced in *Clean Coal/Dirty Air*. See ACKERMAN & HASSLER, *supra* note 9, at 1-5. In larding unnecessary subsidies on the drug industry and precluding government price negotiation or foreign imports, Medicare Part D, like the scrubber legislation, represents the victory of special-interest rent seeking and regulatory dogmatism over sensible public policy. See, e.g., Louise M. Slaughter, *Medicare Part D—The Product of a Broken Process*, 354 NEW ENG. J. MED. 2314, 2315 (2006). There is one difference: the relevant dogmatism in the case of the scrubber legislation was New Deal command and control. With Part D, it is market fetishism. To be

II. EXPERIMENTALISM

Experimentalism takes its name from John Dewey's political philosophy, which aims to precisely accommodate the continuous change and variation that we see as the most pervasive challenge of current public problems. Policies should be "experimental in the sense that they will be entertained subject to constant and well-equipped observation of the consequences they entail when acted upon, and subject to ready and flexible revision in the light of observed consequences," he wrote.⁷⁰ At the same time, he rejected standardized bureaucratic solutions and urged responses that combined respect for local context with centralized structure and discipline. Aside from education, in which he was an active reformer, Dewey offered little by way of concrete illustrations of his proposals, but postwar developments in business management and recent reforms in many public policy areas connect fruitfully with Dewey's theories.

In the United States, some experimentalist institutions have emerged in sectors such as nuclear power and food safety, in which technological and economic change has outstripped the capacities of established market and bureaucratic safeguards to protect key public interests. Others have emerged in settings such as public education or child protective services, in which decades of apparently fruitless and interminable debate have created strategic uncertainty among the contending parties and opened them to interventions aimed at learning from promising local experience.

In the European Union and in many international regimes, experimentalist institutions have arisen from the realization that mutual evaluation and learning from diverse national practices is sometimes the only feasible way of coordination in the absence of a conventional national sovereign with presumptive authority to fix common goals.

Not all of these regimes have complete or well-configured architectures, but we consider them experimentalist to the extent that they are designed to achieve local adaptation and aggregate learning by combining discretion with duties to report and explain, and by pooling information. We do not suggest that any of these regimes has proven its efficacy over alternative approaches to the problems they address. In most cases, experience is limited or has not been thoroughly studied. But each of the regimes has functional and jurisprudential properties that make them a distinctively promising response to the dimension of uncertainty in these problems. Our optimism about experimentalism is based most generally on this fit between structures and problems.

sure, the minimalists are not to blame for the bad faith and unreflective uses of their ideas, but the fact that Sunstein and Thaler can discuss their market-mimicking reform without any indication of how dependent it might be on a deeply dysfunctional infrastructure suggests that even well-considered discussions can have distorting effects.

70. DEWEY, *supra* note 3.

A. BASIC ARCHITECTURE

We start with a relatively abstract model of experimentalism in which the basic constituents are a “center” and a set of “local units.” In practice, the center is sometimes the national government, and the local units its federated states or municipalities. Or the center could be a government agency, and the local units the private actors it regulates or the public or private service providers with which it contracts. Or the center might be a single public or private organization, with the local units its (territorial) subdivisions: a state department of child welfare services and its regional districts, as one example, or a school district and its individual schools as another.

These relations are often nested, with an entity such as a school district at once the local unit of a broader (state) jurisdiction and the center of a territorial unit of its own; but the relation between contiguous “higher” and “lower” units is the same, regardless of where they are located in the overall system. Together, the center and local units set and revise goals, and the means of pursuing them, in an iterative process with four basic elements.

First, framework goals (such as “adequate education” or “good water status”) and provisional measures for gauging their achievement are established, whether by legislation, administrative action, or court order, through consultation among the center and local units and relevant outside stakeholders. Second, local units are explicitly given broad discretion to pursue these ends as they see fit.

But third, as a condition of this autonomy, the local units must report regularly on their performance and participate in a peer review in which their results are compared with those units employing other means to the same general ends. These reviews require the local units to describe and explain their efforts to peers and superiors; to show that they have considered alternatives; and to demonstrate that they are making progress by some jointly acknowledged measure of success, or are making plausible adjustments if not. The center provides services and inducements that facilitate this disciplined comparison of local performances and mutual learning among local units.

Finally, the framework goals, performance measures, and decision-making procedures themselves are periodically revised on the basis of alternatives reported and evaluated in peer reviews, and the cycle repeats.

This model involves the following characteristics.

1. Decentralization

Like minimalism but in contrast to command and control, experimentalism combines decentralization of operative control with central coordination of the evaluation of results. John Braithwaite’s observation on the most successful private mine safety efforts exemplifies this feature of the experimentalist architecture:

[W]hile all these companies have decentralized control over safety, they also all have centralized assessment of the safety performance of line managers.

All carefully monitor each mine and each district to ascertain whether their accident and fatality rates are improving or worsening in relation to the performance of previous years and to the performance of other mines and districts. . . . [For all the local managers,] the sense that the head office is watching their safety performance is pervasive.⁷¹

2. Signals and Norms

Experimentalism has been influenced by twentieth-century innovations in economic organization. In particular, it reflects the practices of “learning by monitoring” and “continuous improvement” exemplified by the Toyota Production System.⁷² This system breaks strongly with organizational premises shared by both command-and-control and minimalist administration. Toyota-style organization rejects any strong distinction between making rules, on the one hand, and applying them, on the other. It contemplates that rules will be continuously revised in the course of application. It treats rule departures diagnostically as symptoms of systemic problems and opportunities for systemic improvement. Sometimes, instead of minimizing or isolating them, it tends to induce and expand them (for example, by eliminating “buffer inventory stocks” and imposing superoptimal standards like “zero defects”) in order to make the system’s weaknesses transparent.

The normative structure of regimes of this kind differs from minimalism in regulating in detail and across many dimensions. Experimentalist norms govern both performances or outcomes (such as aggregate CO₂ emissions or scores on educational tests) and processes (such as the calibration of a furnace or the use of a particular math exercise). However, the experimentalist regimes differ from command and control in that a large fraction of their norms are indicative or presumptive rather than mandatory. These regimes have mandatory norms requiring planning, reporting, monitoring, and, often, minimally satisfactory performance. But many norms take the form of indicators that measure performance above mandatory minima, or presumptive rules that need not be followed when doing so would frustrate their underlying purposes. In such cases, the agent can depart from the rule but only if she signals her departure and explains her reasons to peers or superiors. In a phrase used in the EU, her duty is to “comply or explain.”⁷³

The function of nonmandatory rules is not to control discretion but to make practice transparent. They facilitate diagnosis and improvement. The narrow-dimension performance norms emphasized in minimalist regimes are measures of how well an actor is doing overall but often say little about the causes of his

71. JOHN BRAITHWAITE, *TO PUNISH OR PERSUADE: ENFORCEMENT OF COAL MINE SAFETY* 64 (1985).

72. See generally William H. Simon, *Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes*, in *LAW AND NEW GOVERNANCE IN THE EU AND THE US*, *supra* note 3, at 37, 44–55.

73. See Christopher Hogg, *The “Comply or Explain” Approach to Improving Standards of Corporate Governance*, QFINANCE, <http://www.qfinance.com/contentFiles/QF02/g9l6zjv5/18/0/the-comply-or-explain-approach-to-improving-standards-of-corporate-governance.pdf> (last visited July 19, 2011).

success or failure. The more broadly dimensional experimentalist norms aspire to provide a rich array of weak signals that guide adaptation. They assist in self-assessment and comparison with peers, but the only further obligation associated with them is typically an attempt to improve. Moreover, because rules are kept in tight conformity with practice, a newcomer or an outsider can reliably read practice from the rules. This makes learning easier for insiders, and it facilitates accountability to outsiders.

3. Incentive Design

Key norms in experimentalist regimes are designed to induce local actors to participate in the regime. They can involve either coercive sticks or enticing carrots. Coercive norms characteristically take the form of “penalty defaults” that specify fairly harsh consequences if the parties fail to come up with an acceptable alternative.⁷⁴ For example, the Endangered Species Act precludes a developer from improving land where there are endangered species unless the developer concludes a Habitat Conservation Plan with local stakeholders. Enticements might take the form of conditional grants. For example, the Obama Administration’s Race to the Top program awards large grants for educational reform in a competitive process in which state applications are judged on the extent to which they demonstrate capacities to plan and self-assess, to share and make use of information about their own and peer performances, and to coordinate with “critical stakeholders” in both the public and private sectors.⁷⁵ Experimentalism is sometimes associated with “soft law” and reliance on informal social pressures rather than material sanctions, but this view is misleading.⁷⁶ Some experimentalist sanctions are harsh, such as the no-development rule under the Endangered Species Act. Schools, nuclear power plants, and food processors can all suffer the harsh sanction of shutdown under experimentalist regimes if they deliberately or persistently fail to meet mandatory performance norms. The distinctive characteristic of experimentalist incentive design concerns the behavior it is calculated to induce. In command-and-control regimes, incentives are calculated to induce compliance with specific conduct the regulator has determined to be optimal. In minimalist regimes, sanctions are designed to induce performance along one or a very few easily measured dimensions, such as CO₂ emission reduction. The distinctive goal of experimentalist incentive design is to induce actors to engage in investigation, information sharing, and deliberation about problems with multiple dimensions that are only dimly understood. This aim sometimes leads regulators away from sanctions likely to

74. See Karkkainen, *supra* note 40, at 302–04.

75. See U.S. DEP’T OF EDUC., RACE TO THE TOP PROGRAM: EXECUTIVE SUMMARY 6–7 (2009), available at <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>.

76. In the EU and international regimes, resort to soft law is encouraged by the limited capacities for conventional coercive enforcement. See David M. Trubek, Patrick Cottrell & Mark Nance, ‘Soft Law’, ‘Hard Law’ and EU Integration, in LAW AND NEW GOVERNANCE IN THE EU AND THE US, *supra* note 3, at 65, 77–80.

cause actors to hide information or to take defensive adversarial postures in dealings with regulators. Thus, an experimentalist regime is less likely to impose harsh monetary or criminal penalties for unintentionally harmful conduct under vaguely defined standards and more likely to impose them for failing to make or execute plans or to report performance.⁷⁷

4. Stakeholder Participation

Experimentalism has a more sanguine attitude than minimalism toward participation, but the type of participation it relies on is not the universalist and voluntarist kind that minimalists disparage. Experimentalism emphasizes stakeholder participation to elicit and reconcile the diverse views and interests of people distinctively affected by and knowledgeable about the matters in issue. The problems of apathy and ignorance that constrain participation in generalist processes are less severe here. Moreover, some cognitive biases may be mitigated when people have strong incentives to win the cooperation of people with whom they disagree. We expect people stranded in lifeboats to figure out a mode of collaboration even if they have severe differences, and there may be a more general “lifeboat effect” from local interdependence in some contexts addressed by experimentalism.⁷⁸ Moreover, we’ve seen that designers of experimentalist regimes do not need to rely on common interests or cooperative dispositions; they can alter default rules or allocate rewards in ways that motivate responsible participation. Finally, the nature of the issues and circumstances that prompt experimentalist regimes may improve the prospects of effective participation. These regimes are responses to uncertainty. They arise in situations in which public officials are uncertain about both the definition of the problem and its solution. In these situations, stakeholders may find it difficult to map proposed solutions onto their selfish interests. To this extent, they may be more open to good-faith deliberations focused on public interests.

77. Harsh sanctions deter disclosure, first by raising its cost when disclosure increases the likelihood of discovery and sanction, and second by “crowding out” voluntary cooperation when it induces actors to take a more calculated attitude toward the regime. See Jennifer Arlen & Reinier Kraakman, *Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes*, 72 N.Y.U. L. REV. 687, 692–93 (1997); Uri Gneezy & Aldo Rustichini, *A Fine Is a Price*, 29 J. LEGAL STUD. 1, 1 (2000).

Since we are trying to portray both minimalism and experimentalism in their most promising forms, our model of experimentalism excludes the superficially similar approach known as New Public Management. New Public Management emphasizes performance measures and competition among local units. However, its performance measures tend to be strong signals intended to allocate rewards and punishments more than weak signals intended to diagnose problems and indicate directions for improvement. And its idea of competition tends to be harshly Darwinian and to exclude the emphasis on collaborative inquiry and peer review that is central to experimentalism. See Carol Harlow, *Accountability, New Public Management, and the Problems of the Child Support Agency*, 26 J.L. Soc’y 150, 163 (1999); Simon Head, *The Grim Threat to British Universities*, N.Y. REV. BOOKS, Jan. 13, 2011, at 58 (book review).

78. See ELINOR OSTROM, *GOVERNING THE COMMONS: THE EVOLUTION OF INSTITUTIONS FOR COLLECTIVE ACTION* 211 (1990) (concluding that a shared sense among stakeholders that all “will be harmed if they do not adopt an alternative rule” is favorable to agreement on solutions to common-pool resource problems).

B. REGULATION

A variety of regulatory structures in the United States and abroad are best understood in experimentalist terms. One is “management-based regulation” by which the regulator requires each regulated actor to develop a plan to mitigate specified harms; assesses the adequacy of the plans; monitors their implementation; and, through a combination of tangible penalties, technical assistance, and public shaming, induces the laggards to comply with minimum standards and the frontrunners to improve continuously. Examples discussed in recent scholarship include the Hazards Analysis and Critical Control Points (HACCP) program for meat and poultry at the Department of Agriculture; the Massachusetts Toxics Use Reduction Act; the power-plant safety program of the Nuclear Regulatory Agency; and the mining safety programs of Queensland and New South Wales, Australia.⁷⁹

Another structure involves public-private contracting in which an agency enters into a series of bilateral contracts with large regulated actors. The contracts commit the actor to observable performance on one or more regulatory dimensions, and the regulator may in return agree to waive otherwise applicable requirements or use less frequent or costly monitoring. The EPA’s Project Excel and OSHA’s Voluntary Protection and Strategic Partnership Programs are examples.⁸⁰ Or in another variation, the government oversees, ratifies, or oversees and ratifies negotiations among regulated actors, NGOs, and local stakeholders, as occurs in the Habitat Conservation Plan program administered by the Department of the Interior under the Endangered Species Act.

A third structure involves “multi-level governance” in which regulatory authority is decentralized to lower tier governments in ways designed to permit adaptation to local conditions, regulatory competition, or both. The Clean Air and Clean Water Acts and the Occupational Safety and Health Act all give states the option of assuming important regulatory responsibilities, subject to meeting minimum national standards.⁸¹ Some international trade regimes are pushing participating nations in experimentalist directions. Principles of mutual recognition or “equivalence” found in many trade treaties require importing countries to accept products produced in compliance with the regulatory regimes of the exporting countries if these regimes provide protection substantially equivalent

79. See Coglianese & Lazer, *supra* note 39, at 697–99; Karkkainen, *supra* note 40, at 354–56; REES, *supra* note 26, at 123–50; NEIL GUNNINGHAM, MINE SAFETY: LAW REGULATION POLICY 6–7 (2007); Jody Freeman, *Collaborative Governance in the Administrative State*, 45 UCLA L. REV. 1, 6–7 (1997).

80. See Orly Lobel, *Governing Occupational Safety in the United States*, in LAW AND NEW GOVERNANCE IN THE EU AND THE US, *supra* note 3, at 269, 276–77. See generally ENVIRONMENTAL CONTRACTS: COMPARATIVE APPROACHES TO REGULATORY INNOVATION IN THE UNITED STATES AND EUROPE (Eric W. Orts & Kurt Deketelaere eds., 2001) (examining contracts as a means of environmental regulation). Environmental regimes have generated the largest body of self-consciously experimentalist legal scholarship. For discussion and references, see Cameron Holley, *Facilitating Monitoring, Subverting Self-Interest and Limiting Discretion: Learning from “New” Forms of Accountability in Practice*, 35 COLUM. J. ENVTL. L. 127, 139–49 (2010).

81. See 42 U.S.C. § 7413 (2006) (Clean Air Act); 33 U.S.C. § 1313 (2006) (Clean Water Act); 29 U.S.C. § 667 (2006) (Occupational Safety and Health Act).

to the importing countries' regimes.⁸² Such provisions require trading partners to engage in a kind of peer review of each other's regimes. International organizations within and outside the World Trade Organization promote information exchanges and discussions of regulatory standards among member states.⁸³

Multilevel governance is especially prominent in the European Union. As the EU's authority has expanded from its initial focus on trade barriers to embrace almost the full range of its member states' public policies, it has assumed a distinctively experimentalist cast. Early EU intervention commonly took the form of harmonizing directives, mandating that states adopt substantially uniform rules. Recent interventions have more often taken the form of framework directives, such as the 2000 water quality mandate, or variations on the "open method of coordination," such as those prescribed for macroeconomic and employment policies.⁸⁴ These interventions articulate general goals and metrics for measuring progress toward them; mandate that states formulate their own "action plans" for attaining the goals; provide for monitoring of member efforts, exchange of information among states, and peer criticism orchestrated by the European Commission; and contemplate periodic reassessment and rearticulation of goals and metrics.

The ground-level configuration of experimentalist regulation depends in part on whether the risks it addresses take the form more of concentrated, large-scale catastrophes, such as a nuclear reactor meltdown, or of relatively diffuse dangers, such as tainted food. Regimes focused on catastrophe avoidance emphasize event notification—the duty and capacity of regulated actors to identify and report weak signals of systemic problems and to respond to such reports.

In U.S. nuclear power safety, for instance, plants must initially satisfy a demanding set of structural requirements and then must report all potentially dangerous significant operating events, ranging from interruptions in power generation to unexpected deterioration of equipment, to the Nuclear Regulatory Commission. The NRC evaluates the reports and notifies all operators of possible hazards. Thus, the failure of a noncritical valve in one facility can lead to a notice alerting all the others to inspect similar valves, some of which may indeed be performing critical functions. Peer reviews are conducted frequently to evaluate the facilities' responses to such notices. In addition, plants are ranked annually on their overall performance with regard to operating safety. These rankings, along with other safety information, are provided directly to the

82. See Kalypto Nicolaidis & Gregory Shaffer, *Transnational Mutual Recognition Regimes: Governance Without Global Government*, 68 *LAW & CONTEMP. PROBS.* 263, 264 & n.3 (2005).

83. See generally JOANNE SCOTT, *THE WTO AGREEMENT ON SANITARY AND PHYTOSANITARY MEASURES* (2007) (examining the role of the SPS Agreement in promoting information exchanges).

84. See Council Directive 2000/60/EC, 2000 O.J. (L 327) 1 (EU); Council Decision 2003/573/EC, 2003 O.J. (L 197) 13, 17 (EU). See generally Charles F. Sabel & Jonathan Zeitlin, *Learning from Difference: The New Architecture of Experimentalist Governance in the EU*, 14 *EUR. L.J.* 271, 271 (2008) (describing the importance of framework goals in EU governance).

highest level managers and the operator's board of directors to ensure that criticism is not dulled by passage through a managerial hierarchy.⁸⁵

By contrast, in food processing, in which the range of acceptably safe arrangements is much wider than in nuclear power generation and the risk of catastrophe less, responsibility for identifying and mitigating hazards rests more with the operators. In Hazard Analysis and Critical Control Points programs, processors specify the points in their own production processes in which precise control is necessary to prevent contamination by pathogens, develop hazard-mitigation plans for each plant describing how control should be exercised, and design a testing regime to confirm the efficacy of the measures they propose. The regulator reviews and approves the plans and monitors the firms' compliance with their plans. Finally, plans increasingly include provisions for tracing products as they are processed, so that when failures occur it is possible to identify the entity and the procedures that failed.⁸⁶

The Food Safety Modernization Act of 2010 generalizes the HACCP approach, which was previously confined in the United States to processors of meat, poultry, and a few other products, to virtually all food processors above a minimum size. The core of the bill is a mandate that the Food and Drug Administration (FDA) require processors to formulate, validate, and continuously update HACCP plans (now called Hazard Analysis and Preventive Controls). The Act requires the FDA to allocate its inspection resources in accordance with an assessment of the relative riskiness of regulated facilities. Among the specified risk factors are the adequacy of a facility's plan and its implementation, and whether the facility has been certified by a private auditor that is in good standing under an accreditation regime.⁸⁷

In practice, the distinctions between the event-reporting and risk-planning approaches blur as systems based on risk-mitigation planning increasingly focus on event notification and vice versa. Thus, in food safety, the Centers for Disease Control and Prevention coordinates a network that continuously analyzes data from healthcare providers about food-borne illnesses, which federal and state regulators use to target their efforts.⁸⁸ At the same time, the ability to respond effectively to the information disseminated by event-notification systems of the type familiar from the nuclear-power-generating industry is seen as depending on the kind of careful and continuing review of facility operations associated with the implementation of HACCP plans. The convergence of the

85. See REES, *supra* note 26, at 123–50.

86. See COMM. ON THE REVIEW OF THE USE OF SCIENTIFIC CRITERIA & PERFORMANCE STANDARDS FOR SAFE FOOD, NAT'L RESEARCH COUNCIL, SCIENTIFIC CRITERIA TO ENSURE SAFE FOOD 133–75 (2003).

87. See FDA Food Safety Modernization Act, Pub. L. No. 111-353, §§ 103, 201, 307, 124 Stat. 3885, 3889–90, 3923, 3962 (2011) (to be codified in scattered sections of 21 U.S.C.).

88. See Richard L. Arkin, *PulseNet, FoodNet, NARMS; Tools To Fight Disease, Protect Public Health*, 22 FDA VETERINARIAN, no. 6, 2007 at 6, 6, available at <http://www.fda.gov/downloads/AnimalVeterinary/NewsEvents/FDAVeterinarianNewsletter/UCM055723.pdf>. The FDA Food Safety Modernization Act requires FDA efforts to strengthen “foodborne illness surveillance systems.” § 205.

reporting and planning approaches can be seen in the 2010 report of the National Commission on the Deepwater Horizon Oil Spill and Offshore Drilling. The Commission's recommendations, modeled in part on the nuclear safety regime, call for both government-orchestrated and -enforced planning at the facility level and a system of peer information exchange and review.⁸⁹

Another interesting area in which experimentalist regulatory practices are converging is drug regulation. The system of drug regulation established during the New Deal requires "approval" by the Food and Drug Administration prior to "marketing." Approval typically involves an expensive and rigorous testing process focused around clinical trials. The FDA then makes a decision whether to permit marketing on the basis of a general assessment as to whether likely benefits to people within the prescription criteria outweigh likely costs such as inefficacy and side effects. This framework initially had little to say about what happened postapproval and, until 2007, postapproval duties of both the manufacturers and the FDA were much more limited and ambiguous than preapproval ones.⁹⁰

This system has been subject to two salient criticisms. First, critics asserted that the approval process was too lengthy and expensive and, in particular, that it unnecessarily delayed access to potentially beneficial drugs. Second, they decried the scantiness of postapproval regulation, pointing to a series of scandals in which approved drugs continued to be marketed despite evidence that they were doing unanticipated harm.⁹¹

Pressure on the New Deal structure was also increased by trends in medicine toward greater individuation of treatment. These include "personalized medicine," which uses genetics to customize treatments and drugs to the characteristics of the patient, and "behavioral medicine," which focuses on patient education, communication, and monitoring.

Another important set of developments involves the aggregate analysis of patient treatment data. A significant current health policy priority is to develop information systems for standardized recording and transmission of individual treatment data (while protecting the patient's privacy).⁹² Aside from facilitating coherent patient care by making past history transparent to current caregivers, such systems could contribute to the continuous assessment of treatments, including drug efficacy. Drug efficacy has been studied in a controlled environment because that has been the only way to get sufficiently precise data on the

89. See NAT'L COMM'N ON THE BP DEEPWATER HORIZON OIL SPILL & OFFSHORE DRILLING, *DEEP WATER: THE GULF OIL DISASTER AND THE FUTURE OF OFFSHORE DRILLING* 217-91 (2011). See especially the discussion of the nuclear-safety regime.

90. See U.S. GOV'T ACCOUNTABILITY OFFICE, *GAO-06-402, DRUG SAFETY: IMPROVEMENT NEEDED IN FDA'S POSTMARKET DECISION-MAKING AND OVERSIGHT PROCESS* 18 (2006); CHRISTENSEN ET AL., *supra* note 51, at 62.

91. See, e.g., Laura B. Faden & Christopher-Paul Milne, *Pharmacovigilance Activities in the United States, European Union and Japan: Harmonic Convergence or Convergent Evolution?*, 63 *FOOD & DRUG L.J.* 683, 685-89 (2008).

92. See CHRISTENSEN ET AL., *supra* note 51, at 133-47, 177-78.

baseline condition of the patient and the nature and effect of the intervention. But with sufficiently precise and standardized routine treatment reporting, the line between clinical trials and routine treatment would blur.

If current trends continue, the key judgment for drug regulators will cease to be a binary call about whether to permit general marketing of a drug. Instead, regulation will focus on ensuring that the drugs are properly customized to the patient and properly integrated with other medical and behavioral interventions. As the line between research and treatment erodes, a one-time, global, cost-benefit judgment will be far less important than an ongoing capacity to monitor data for signals of problems and opportunities.

These trends are reflected in the Food and Drug Administration Amendments Act of 2007. Among other things, the Amendments increase the FDA's authority to require specified, ongoing research as a condition of approval; require that the Agency enhance its Adverse Event Reporting System; and mandate that it establish a system of proactive monitoring of drug efficacy through targeted inquiries of healthcare records.⁹³

The advent of personalized and behavioral medicine suggests the likelihood of a trend toward regulatory regimes configured to particular drugs. Steps in this direction are evident in the Risk Evaluation and Mitigation Strategies (REMS) encouraged by the 2007 Amendments. As a condition of approval, the FDA can require the promoter to formulate and implement a REMS. An example is the highly regarded System for Thalidomide Education and Prescribing Safety (STEPS). Thalidomide was once prescribed widely for a variety of common conditions like insomnia and anxiety, but it was banned after the discovery that it severely increased the risk to pregnant women of birth defects. Since then, research has indicated that the drug is uniquely effective with certain conditions associated with leprosy. It can now be prescribed to patients with these conditions but only in accordance with STEPS protocols that require registration and training of the prescribing physician, the dispensing pharmacist, and the patient; birth control and pregnancy testing for women of child-bearing age and close monitoring of all patients; and regular reporting of treatment and patient conditions. The program is administered by a multidisciplinary committee within Celgene, the manufacturer, and another one within the Slone Epidemiological Unit of the Boston University School of Medicine. The latter committee has patient representatives, including delegates from the March of Dimes and the Thalidomide Victims Association of Canada. The committees enforce compliance in part through audits and "mak[e] recommendations to the FDA for changes in the program based on real-time data."⁹⁴

Drug regulation has long been experimental in the sense that it has based key

93. See Food and Drug Administration Amendments Act of 2007, Pub. L. No. 110-85, § 921, 121 Stat. 823, 962 (codified as amended at 21 U.S.C. § 355(k)).

94. See Jerome B. Zeldis et al., *S.T.E.P.S.TM: A Comprehensive Program for Controlling and Monitoring Access to Thalidomide*, 21 *CLINICAL THERAPEUTICS* 319, 324 (1999). See generally INST. FOR ALT. FUTURES, *OPTIMAL FUTURES FOR RISK EVALUATION AND MITIGATION STRATEGIES (REMS)* (2010)

judgments on disciplined empirical assessments. But as long as experimentation was confined to a specialized realm and associated with one-time, binary judgments, it lacked the features associated with experimentalism. As it moves to emphasize and integrate event notification with contextual risk planning, it increasingly resembles experimentalist regimes such as nuclear power and HACCP.

It is not easy to account for such regimes in minimalist terms. Static notions of optimization capture very little of the central thrust of the system, which is to facilitate learning and continuous revision. Moreover, by blurring the line between rulemaking and implementation, the system precludes the kind of accountability that binds local actors to centrally promulgated rules.

In comparison to the minimalist approach, the experimentalist approach to regulation promises the following advantages.

First, these regimes create pressures to respond to weak signals with self-diagnosis and corrective action. Moreover, they can provide a range of incentives, including rewards and subsidies, reputational effects, and conventional penalties.

Second, experimentalist regulation reduces the information burden on the regulator in determining initial levels of acceptable performance. To the extent that performance standards are based on questions of technical or economic feasibility, they can be set based upon observed experiences across the full population of regulated actors. The regulator can require that plants that fail to perform at some fraction of average industry performance be subject to remedial or punitive consequences, or both. Under the HACCP regime for meat and poultry safety in the United States, only the poorest performing two percent of plants—those whose products most frequently test positive for conveniently measured proxies of pathogens—are at risk of closure. But as laggards improve or exit the industry, the average performance rises and the minimum standards could “ratchet up” without the need for special investigations or proceedings of the kind necessary, for example, to adjust allowance allocations in a cap-and-trade regime.⁹⁵

Third, experimentalist regimes accommodate diversity by leaving extensive discretion to local actors while making their activities mutually transparent. The key mechanisms involve metrics that measure performance. Experimentalism co-opts the paradox of standardization: by imposing uniformity along a few dimensions, we permit more variations on others.⁹⁶ Microbiological testing for both serious toxins and specified proxy organisms permits us to compare food processing plants and identify the leaders and laggards, while leaving them

(discussing REMS in the context of the Food and Drug Administration Amendments Act of 2007 and healthcare system evolution).

95. Unfortunately, as implemented, the U.S. HACCP regime lacks this ratcheting up feature. See 9 C.F.R. § 310.25 (2010).

96. See DEWEY, *supra* note 3, at 215 (“Uniformity and standardization may provide an underlying basis for differentiation and liberation of individual potentialities.”).

substantial discretion over their operations. Minimalist market-simulation regimes aspire to achieve the same goal, but their logic focuses effort around a single indicator: price. Because experimentalist regimes rely on a range of indicators, they are less vulnerable to the failure of one. And their indicators typically have more diagnostic value than prices.⁹⁷

Fourth, an experimentalist regime readily propagates relevant technical and organizational advances that may not circulate freely under bureaucratic or minimalist regulation. Event notification and risk planning tend to make technical knowledge accessible to both regulators and peers. The agency can provide technical assistance to firms in developing risk-management plans or can disseminate quick fixes to problems registered in event-notification systems. It can mandate publication of risk-management plans and can encourage or mandate information exchanges through peer review. It can subsidize demonstration projects in which progressive firms get assistance in customizing new technology to local conditions so that its potential can be publicized to the firm's peers.⁹⁸

C. SOCIAL WELFARE

In the welfare area, promising programs with key experimentalist traits can be found in at least two broad areas.

The first includes programs that provide individualized services. These include child protective services, healthcare, "special" and general education, job training, mental health services, and disability capacitation. They typically require highly individuated planning, pervasive performance measurement, and efforts to aggregate and disseminate information about effective practices.

The second area involves public participation in the design of local public goods. Notable examples are community policing and community economic development. In community policing, police consult with neighborhood groups in order to set priorities in local law-enforcement objectives, to obtain local knowledge to configure strategies, and to coordinate with private activities in implementing enforcement plans. Success is monitored in terms of various indicators of safety. In community economic development, government agen-

97. Robert Kaplan and David Norton propose the use of a "Balanced Scorecard" consisting of parallel sets of outcome and performance measures: "Outcome measures without performance [measures] do not communicate how the outcomes are to be achieved. . . . [P]erformance [measures may] . . . fail to reveal whether the operational improvements have been translated into . . . enhanced financial performance." ROBERT S. KAPLAN & DAVID P. NORTON, *THE BALANCED SCORECARD: TRANSLATING STRATEGY INTO ACTION* 150 (1996).

98. See DANIEL P. CARPENTER, *THE FORGING OF BUREAUCRATIC AUTONOMY: REPUTATIONS, NETWORKS, AND POLICY INNOVATION IN EXECUTIVE AGENCIES, 1862-1928*, at 238-54 (2001); VON HIPPEL, *supra* note 47, at 93-132; Atul Gawande, *Testing, Testing*, *NEW YORKER*, Dec. 14, 2009, at 34, 38-39; Kathleen G. Noonan, Charles F. Sabel & William H. Simon, *Legal Accountability in the Service-Based Welfare State: Lessons from Child Welfare Reform*, 34 *LAW & SOC. INQUIRY* 523, 545 (2009); Charles Sabel, *Self-Discovery as a Coordination Problem* 10-11 (2010) (unpublished manuscript) (on file with *The Georgetown Law Journal*).

cies and private charities make financial and technical assistance available to community-based organizations to implement locally produced plans for housing, job, and business development. Again, projects are evaluated in terms of an array of criteria, and the more successful organizations receive priority in later rounds of funding.⁹⁹

These programs redefine the conventional relation between the center and the frontline. The center's role is no longer merely to monitor frontline compliance with promulgated standards. It is responsible for providing the infrastructure and services that support frontline efforts. Thus, the role of the principal in the experimentalist school is not just to verify that the teacher's class is studiously at work but also to organize the specialized services and framework conditions—remedial reading, testing to diagnose learning difficulties, coaching in team building—on which the teacher's team must rely in formulating and implementing individual learning plans.

In child welfare, caseworkers rely on a center that trains and otherwise qualifies foster parents, facilitates contracting with outside specialists, and marshals resources to respond to the unexpected needs of particular families or sudden community-wide problems (for example, a tornado or an epidemic of methamphetamine abuse). In community policing, the precinct and its subunits count on a department that provides timely information about the migration of crime and criminals from one neighborhood to another (perhaps in response to local success in policing), helps orchestrate cooperation with community groups that work with youth recently released from detention and at risk of reincarceration, and provides mediation and consultation services in the aftermath of racial incidents.

The key influence in the move toward experimentalism is the view that services need to be tailored to the needs of beneficiaries. Such tailoring requires the active participation of beneficiaries because effective intervention depends on their cooperation or because they have information essential to diagnosis and planning. As Dewey wrote, “[t]he man who wears the shoe knows best . . . where it pinches.”¹⁰⁰

Tailoring also requires an understanding of local context. A child welfare worker putting together a plan for an obese child may be able to include a bicycle in the plan if she knows that the family's church can be persuaded to come up with one if credibly assured that it will fill an important need. Police dealing with a high-crime neighborhood can be more effective if they learn from local residents that a poorly maintained house from which drugs are sold is a magnet for nonresident deviants.

The solitary street-level bureaucrat, who was discovered in the organizational

99. See FUNG, *supra* note 29; WILLIAM H. SIMON, *THE COMMUNITY ECONOMIC DEVELOPMENT MOVEMENT: LAW, BUSINESS, & THE NEW SOCIAL POLICY* (2001) (examining the institutional configuration of Community Economic Development practice).

100. DEWEY, *supra* note 3, at 207.

literature of the 1960s and has since haunted the minimalists, does not appear in the emerging experimentalist regimes. The street-level bureaucrat exercised tacit discretion under the radar of her superiors in the broad interstices of poorly enforced rules.¹⁰¹ Experimentalist design departs from the organizational features that gave rise to the street-level bureaucrat in four important ways.

First, difficult frontline issues are more likely to be decided by a team than by a single worker. The social professions increasingly see individual problems as functions of multiple and diverse causes that call for interdisciplinary diagnosis and intervention. In the most highly regarded child-protective-service programs, the case worker's chief responsibility is to form and periodically convene a team that typically includes key family members, a health professional, lawyers for the child and the state, a therapist, and perhaps a teacher.¹⁰²

In schools, analogous interdisciplinary teams—the classroom teacher, the reading specialist, and the behavioral therapist—formulate plans for students with learning difficulties. In community policing, the precinct officers meet regularly with community members and their representatives, school officials, and landlord and tenant associations, among others, to determine which crimes are currently most disruptive and threatening, and to explore ways of reconfiguring public services to reduce them—perhaps rerouting a bus line to preempt encounters between rival gangs or demolishing an abandoned house that has become a magnet for drug dealers.

Group decision making promotes accountability in two ways. Team members act under the gazes of a shifting array of peers. Thus, the informal pressures of pride and shame are triggered. Furthermore, collaborative decisions require articulation, and the fact that the team members are diverse in backgrounds means that matters that might be taken for granted in a more homogeneous setting have to be explained and subjected to examination.

The second feature of experimentalist service provision that distinguishes it from street-level and other bureaucracies is a distinctive form of monitoring. Like event-notification practices in experimentalist risk regulation, social-services monitoring engages in intensive scrutiny of individual cases to reveal systemic problems. But whereas event notification is triggered by unexpected disruptions, core monitoring in experimentalist service provision is part of the organizational routine. A particularly well-developed example is the Quality Service Review (QSR) used in child welfare programs in Utah and several other states. The QSR begins with selection of a stratified random sample of cases. A two-person team, including an agency official and an outside reviewer, examines each case over two days, beginning with a file review and proceeding to interviews with the child, family members, nonfamily caregivers, professional team members, and others with pertinent information.

101. See MICHAEL LIPSKY, *STREET-LEVEL BUREAUCRACY: DILEMMAS OF THE INDIVIDUAL IN PUBLIC SERVICES* 3–12 (1980); JAMES Q. WILSON, *VARIETIES OF POLICE BEHAVIOR* 7–9 (1968).

102. See Noonan, Sabel & Simon, *supra* note 98, at 541–42.

The reviewers then score the cases numerically in terms of two sets of indicators, the first concerning “child and family status”—the wellbeing of the child and his or her family—and the second concerning “system performance”—the capacity to build teams, make assessments, formulate and update plans, and execute the plans. The initial scoring is then refined in meetings among the reviewers as a group, and then between reviewing teams and the caseworkers and supervisors whose decisions they have reviewed. Following these clarifications, there is a second meeting to draw general conclusions and a final meeting between reviewers and personnel from the region to discuss the systemic significance of their findings. The final report sets out the aggregate scoring, identifies recurring problems, and illustrates these from specific cases.

The QSR is as much a process of norm elaboration as of compliance enforcement. Agency goals like child safety and family stability (permanence) are, in the abstract, indeterminate. The QSR helps establish paradigmatic instances of what the goals mean and how the processes should be carried out. Participation by officials from the child welfare department’s central administration promotes consistency across regions. The integration of outsiders from other states or consultants with national practices promotes consistency across states. Similarly, QSR data functions as a measure of performance and as a diagnostic tool of systemic reform. The scores can be compared over time and (in principle though not yet in practice) across states, giving rough but serviceable indications of where attention and remedial effort should be focused.

Third, rules have a different relation to accountability in experimentalist administration than in conventional legal thought, including in minimalism. Workers often have discretion to depart from rules when they believe it would be counterproductive to follow them. This discretion, however, is limited by the requirement that a worker do so transparently in a manner that triggers review and, if her judgment is sustained, prompt rewriting of the rule to reflect the new understanding. These regimes challenge the premise of the street-level bureaucracy literature that the only escape from the rigidity of mechanical rule following is low-visibility, ad hoc, frontline discretion.

Fourth, another important accountability mechanism in many experimentalist regimes is coinvestment among different government and private actors. Child welfare case planning in Utah-type regimes includes an assessment of available “private supports.” Private supports may come from relatives, friends, churches, and employers connected to the child’s family. They may come from institutions in the community with charitable missions. Community-based housing development grants are structured so that projects typically require the support of multiple governmental and private institutions. Job training programs aspire to enlist private employers in supporting and designing their curricula. From one perspective, the coinvestment practice is an extension of the team concept, further marshaling diverse perspectives and expertise. At the same time, it serves as a check on agency discretion. The willingness of independent institutions to invest in the plans and projects provides some reassurance of their soundness.

Coinvestment can also occur between the programs and their beneficiaries, and this kind of coinvestment potentially responds to concerns about beneficiary responsibility. Given his skepticism about complex, ongoing administrative judgments, Ackerman favors a scheme—stakeholder grants—that is generous upfront in a standardized way—but necessarily stingy at the back end for individuals whose choices do not work out.¹⁰³ But experimentalism makes possible an approach that responds to needs throughout a person's life. The moral-hazard problems are muted because the key benefits are services that officials and beneficiaries have jointly chosen and because the individual coinvests her efforts and perhaps also resources. The Utah child-welfare model, in which parents invest in service plans, is an example. A more large-scale example is the Danish “flexicurity” model that provides intensive job training and placement services, as well as generous income support, to workers at whatever stage of their careers that they need it. The model is controversial, but is widely credited with neutralizing the employment disincentive effects of high benefits.¹⁰⁴

CONCLUSION

Much current public law scholarship seems to reflect the minimalist preoccupations with static optimization and administrative simplification at the expense of experimentalist concerns with reliability and learning. At the very least, this bias amounts to a failure of description. The number of significant, recent policy initiatives that reflect distinctive minimalist concerns is small compared to the number that employ unmistakably experimentalist rhetoric and architecture. And the experimentalist initiatives have been central in a broad array of the most pressing policy areas, including pollution, occupational safety, food safety, policing, and nearly every major sphere of social service.

A common feature of these problems is that they arise in circumstances of fluidity and diversity, and they call for interventions that are capable of adaptation and contextualization. Interventions of the sort that minimalism tends to favor—nudges, market simulation, rule simplification—have scant capacity for adaptation and contextualization. By contrast, experimentalist regimes have developed forms of nonbureaucratic administration that try to combine accountability with local initiative in ways that facilitate learning and individuation. They strive for accountability less through simple rules than through peer review of local discretion. The aspiration is that pooled learning will discipline local autonomy while generalizing its successes. Perhaps minimalism could accept such regimes as appropriate to categories of problems other than those on which it has focused, but, in order to do so, it would have to soften its resistance to frontline discretion.

103. See ACKERMAN & ALSTOTT, *supra* note 2, at 39–40.

104. See Torben M. Anderson & Michael Svarer, *Flexicurity—Labour Market Performance in Denmark*, 53 CESIFO ECON. STUD. 389, 389–93, 421–24 (2007).