

# 10

---

## FROM OVERSIGHT TO INSIGHT

### *Federal Agencies as Learning Leaders in the Information Age*

SHELLEY H. METZENBAUM

Justice Louis Brandeis's observation that "states are the laboratories of democracy" is frequently quoted.<sup>1</sup> Unfortunately, these ostensible laboratories too often lack scientists. Few study state and local experiments taking place across the country. Little attention is directed to documenting key details and distinguishing studies with positive from those with negative or no results. Few resources are devoted to writing up experimental findings, weeding out unsubstantiated conclusions, and distributing lessons to interested parties.

Federal agencies that depend on other levels of government to accomplish their objectives need to play a much stronger role studying experiments in state and local laboratories of democracy or causing such study to occur and then sharing the findings. Especially in the information age, federal agencies that depend on states and localities to accomplish their objectives need to assume both a learning role and a leadership role. They need to be learning leaders.

Federal agencies can promote learning by building the capacity to learn which state and local actions improve outcomes; stimulate experiments that complement those already occurring; disseminate findings; and encourage uptake of effective and discard of ineffective approaches. This implies a much stronger federal role in the creation, management, and transfer of knowledge.

The federal government can exhibit leadership by providing inspiration, a direction for action, motivation, expert coaching, and insights that enable genuine achievement.<sup>2</sup> To function as learning leaders, many in federal agencies need an attitude adjustment. They need to shift their emphasis from conducting oversight to generating outcome-focused, evidence-based insights. They need to pay less attention to ensuring that specific activities take place or specific amounts of money get spent and more to helping states and localities understand problems, then find and adopt effective solutions. Federal agencies need to identify what works, motivate uptake of effective interventions, and encourage the ongoing search for ever more productive ways to prevent, mitigate, and treat problems. That is not to suggest that federal agencies compromise their accountability to Congress and the public nor relax intergovernmental accountability expectations. States and localities will always need to ensure that taxpayer dollars are spent legally, efficiently, and honestly. Federal domestic policy agencies can no longer concentrate on procedural and fiscal matters, such as requiring the establishment of a single fund-receiving agency and maintenance of prior spending levels at the state and local level, to the neglect of performance accountability that “motivate[s] better performance than would otherwise occur” and democratic accountability that promotes government responsiveness to the choices and preferences of its citizens.<sup>3</sup> Instead, federal agencies need to act as intergovernmental learning leaders. They need to enhance performance accountability by helping states and localities learn how to improve outcomes and by motivating adoption of promising practices, avoiding the temptation to push specific means for achieving goals without compelling evidence of their effectiveness. They need to boost democratic accountability by clearly articulating priorities and supporting debate about the appropriateness of targets given available resources.

Incentives are also an important tool in the intergovernmental accountability toolkit. They should not, however, be used primarily to reward governments that meet their targets and punish those that do not, as is so often assumed. Incentives work best as a companion to goals and measurement to motivate the collection of standardized data, win attention to goals, enlist goal allies, and compel the attention of reluctant goal adopters. Incentives must be employed with great care, though, lest they stimulate dysfunctional responses owing to fear, frustration, or goal opposition.

Interestingly, the intergovernmental literature focuses most of its attention on incentives, especially grants, as the primary mechanism of federal influence on other levels of government. It rarely pays attention to goals and measurement as discrete intergovernmental tools. When it does, it is primarily in the

context of a principal-agent relationship, in which goals reflect the principal's objective (assumed to differ significantly from the agent's) and measurement provides the means for monitoring agent fulfillment of obligations to the principal. The power of goals and measurement as tools for inspiration, communication, and knowledge management has only recently begun to garner attention.<sup>4</sup>

More attention is warranted because federal agencies can use goals and measurement to inspire, motivate, and illuminate performance-improving opportunities and pair them with incentives to enlist goal allies, generate measurement, and stimulate analysis. Mastering use of these three intergovernmental tools will improve both performance and accountability while at the same time reducing the rancor that so often arises when federal agencies place primary emphasis on state and local compliance with grant conditions. Of course, federal-state relations will not always be smooth even when federal agencies act as learning leaders. Political perspectives on which problems need attention inevitably vary across people and parties and therefore, not surprisingly, across levels of government. Intergovernmental battles about goals, especially specific targets, will be common. But battles about outcome priorities rather than administrative matters constitute healthy democratic debate rather than wasteful wrangling.

Federal agencies cannot function solely as learning leaders. When Congress sets national standards to protect basic rights or minimum standards of well-being, federal agencies must also serve as basement border guards, ensuring that states (and their residents) not trespass nationally set boundaries. Furthermore, when state and local delivery capacity is inadequate as was the case with Hurricane Katrina in 2005, the federal government may need to provide direct service or catalyze creation of intergovernmental service delivery networks.

### **Strengths and Stresses of the Intergovernmental Governance Structure**

In the United States, multiple levels of government serve the American people for both practical and political reasons. A multiparty governance arrangement is well established in domestic policy areas such as education, the environment, poverty reduction, and transportation.<sup>5</sup>

Well-functioning multilevel operations, if not too rule bound and controlling, have great practical advantage both in government and the private sector. Thus practicality argues for intergovernmental delivery of government

services. Centralized units of well-functioning multilevel organizations or networks open possibilities for economies of scale, expertise specialization, and risk-spreading, while local units afford agility and adaptability. Intergovernmental arrangements also afford a practical means for handling cross-jurisdictional costs such as environmental pollution, cross-jurisdictional benefits such as highways, and cross-jurisdictional free-rider problems that arise when local governments offer redistribution programs. Poorly executed multilevel arrangements can, however, be cumbersome, inefficient, and infuriating.

In the United States, the political argument for multilevel governance dominates the practical one. The country's founders feared that an overly powerful government, acting on behalf of the majority and influenced by the wealthy and well organized, might trample the rights of the minority. They adopted a federalist system to counter this, with separately sovereign national and state governments. Together with the checks and balances of three branches of government, is designed to protect civil rights and deliver more responsive government.<sup>6</sup> When citizens do not like policies advanced at one level of government, they can organize politically to bring about changes at another. Although effective in protecting the minority, the ongoing quest for political responsiveness inevitably creates intergovernmental policy inconsistencies and recurring contention.<sup>7</sup>

The federal government seldom exercises force to assert hierarchical authority. Instead, it has long used grants to persuade states and localities to pursue federally selected goals.<sup>8</sup> When states and localities accept federal dollars to advance a specific objective, they agree to a set of conditions defined by law, regulation, guidance, or grant agreement. The legitimacy of this intergovernmental control mechanism has been challenged, but the courts have consistently "conceded to Congress the right to attach any regulations to any aid Congress provides."<sup>9</sup> Federal agencies dependent on states and localities to accomplish their objectives are not limited to using conditioned grants to motivate intergovernmental action. They can also use goals and measures. When federal agencies use grants to support the leadership potential of goals and the learning potential of measures rather than to control, multilevel governance thrives.

### **Government Performance and Results Act**

In 1993, convinced of the logic of using outcome-focused goals and measurement to improve both performance and accountability, the U.S Congress passed the Government Performance and Results Act (GPRA). The law

requires every federal agency to set strategic (five-year) and annual goals, measure performance, and report performance annually to the public.<sup>10</sup> Perhaps surprisingly, the law “[pays] scant attention to federalism—a fundamental feature of the American political system that profoundly shapes program implementation.”<sup>11</sup> It leaves it to each federal agency to sort out how to use goals and measurement in its intergovernmental arrangements.

In chapter 11 in this volume, Beryl Radin asserts that the GPRA and the Performance Assessment Rating Tool (PART), an executive branch review process introduced by the U.S. Office of Management and Budget to increase agency use of outcome-focused goals and measures as mandated by the GPRA, exacerbate intergovernmental difficulties.<sup>12</sup> Cantankerous conflicts can indeed erupt when the federal government tries to impose goals on states and localities. States and localities do not like another level of government telling them what their priorities should be, even when they agree with the goal itself, and often see measurement requirements associated with federal grants as a drain on resources needed to get the job done. Yet intergovernmental use of goals and measurement can defuse contention and speed program effectiveness when goal debates pertain to program priorities rather than power and when federal agencies use collected measurements to generate performance-improving insights rather than to control.

The GPRA requires every federal agency to set outcome-focused goals and measure progress. Federal agencies dependent on other levels of government to accomplish their objectives therefore must figure out how to use goals and measurement in an intergovernmental context within the bounds of their existing laws. These bounds vary greatly. Some laws set state and local goals; others instruct federal agencies to set them. Some restrict federal agency authority to set goals; and others charge states with goal-setting. Many also explicitly link goals and measurement to incentives.

### **Core Tools of Intergovernmental Performance Management**

The core tools of intergovernmental management for performance are goals, measurement, and incentives. Goals and measurement can operate as intergovernmental tools independent of incentives. The most common incentives are the promise of grants and the threat of penalties.

#### *Goals*

John F. Kennedy intuitively appreciated the inspirational value of goal setting in government when he announced his plan to land a man on the moon

within a decade. Goals, especially those that are specific and challenging, fulfill a remarkably powerful performance-driving function for both individuals and organizations.<sup>13</sup> Goals affect performance through four mechanisms: a directive function, an energizing function, persistence, and a stimulating function that encourages discovery and the use of task-relevant knowledge and strategies.<sup>14</sup> Goals can also frustrate, though, when targets are unrealistic relative to available knowledge, skills, or resources. To inspire effectively, goals need to be challenging, but they also need to be realistic. Otherwise, they simply enrage or discourage.<sup>15</sup>

Goal setting has received limited attention as a tool of government influence. One exception, a study of school mission statements, tentatively concludes that the framing of a mission statement can affect performance. Statements phrased in an active voice, with a few outcome-oriented goals, improve school performance more than those stated passively with either multiple goals or goals focused on processes or behaviors.<sup>16</sup> Goals have received even less attention in the intergovernmental literature, except when combined with incentives.

Experience in the health area suggests the potential power of well-set federal goals. The federal government issued the first Healthy People report in 1979, providing a health vision for the country by setting five specific national outcome targets, one each for five distinct age groups, to be achieved within ten years. This goal-setting report was followed one year later by a report setting specific targets for fifteen priority program areas, such as immunization and accident prevention, to reach the five national goals. The federal government has updated the health goals and objectives for the American people each successive decade, setting targets for the decade to follow.<sup>17</sup>

Healthy People goals and objectives are set based on a review of the evidence about deaths and illness, their causes, and the effectiveness of prevention efforts. The reports select high-level goals, such as the current Healthy People goals of increasing quality and years of healthy life and reducing health disparities, and specific objectives for specific focus areas. Data on risks associated with illness and death guide the selection of objectives and focus areas. Health care system problems, behavioral factors, environmental hazards, and human biological factors are also considered in the selection of objectives. The reports classify preventable risks and tally them to find the best opportunities for risk reduction.

Healthy People goals do not have the force of law, only the power of persuasion. Nonetheless, they have proved contagious, guiding priorities, informing decisions and influencing government spending. Other federal

agencies, states, localities, and the private sector have embraced these goals. The New York City Department of Health and Hygiene uses Healthy People goals to guide and gauge its own performance. Forty-four states, the District of Columbia, and Guam, have replicated the Healthy People model and adopted their own health promotion and disease prevention objectives to guide local health initiatives.<sup>18</sup>

Congress incorporated Healthy People objectives into several laws and programs, including the Indian Health Care Improvement Act, the Maternal and Child Health Block Grant, and the Preventive Health and Health Services Block Grant. Indeed, the positive congressional response to Healthy People goals suggests that federal agencies can use goal setting as a way to start (or continue) a conversation with Congress about appropriate priorities.

The federal government did not assume other levels of government would embrace Healthy People targets; it worked hard to build support. It offered technical assistance to encourage states and localities to undertake their own tailored Healthy People efforts and created the Healthy People Consortium, made up of 350 national membership organizations in addition to state and territorial health agencies.<sup>19</sup>

The Healthy People initiative also stimulated nongovernmental financial support to advance its goals. The Robert Wood Johnson Foundation, for example, awarded a grant in 1997 to direct the attention of American businesses to Healthy People goals.<sup>20</sup> A Healthy People Business Advisory Council was created to encourage attention to Healthy People goals in the workplace.

The intergovernmental trajectory of one Healthy People objective suggests how federal goal setting can influence others. The 1979 Healthy People report identified cigarette smoking as the single most preventable cause of death and adopted a smoking cessation target.<sup>21</sup> Since that time and without central coordination, other federal agencies, states, and local governments have adopted a wide variety of actions to prevent smoking. These include taxes, lawsuits, warning labels, and bans on smoking in public places. Uncoordinated goal-focused intergovernmental action, led at different times by different levels of government, has driven U.S. smoking levels down dramatically since the mid-1960s, when the surgeon general first issued a report warning of the dangers of smoking.<sup>22</sup>

Well-selected federal goals, these vignettes suggest, have a powerful persuasive effect when they deal with issues that concern people and are backed by evidence showing the relative importance of a problem and the existence of effective prevention practices. They inspire effort and investment by others. Experience with the Healthy People initiative suggests that federal agencies

can use the federal bully pulpit to set goals that persuade, even without incentives.

Is Healthy People an exception or does it offer a model for other policy areas? Arguably, four replicable attributes of the initiative caused states and localities to align voluntarily with federally nominated objectives. First, its goals focused on issues that concern the public. Second, goal selection was informed by accumulation, analysis, and publication of data about the relative seriousness of problems needing attention. Third, Healthy People reports tally not only problems but also their preventable causes, suggesting a path for problem reduction. Fourth, the federal government used the goals to reach out broadly to recruit experts and implementers whose independent actions and decisions could improve health outcomes.<sup>23</sup>

It has been suggested that Healthy People is unique because the health field is rich with data and is an area with high goal congruence across government levels. However, Healthy People lacked data for about one-third of the objectives it set when it started. A decade later, it had reduced that gap to 20 percent because agencies started to generate the data they needed. In addition, data shortages in federal agencies are not always as desperate as claimed. Many agencies collect reams of data they never analyze, forgoing opportunities to understand the relative import of problems, their causes, and prevention possibilities.<sup>24</sup>

With regard to concern about goal congruence, intergovernmental goal dissonance undoubtedly exists. Differing values and the cost of goal pursuit make debates about government's goal selection and the appropriate portfolio of goals common. Those conflicts intensify across levels of government, especially when goals set by one level impose costs on another. Goal dissonance is less an issue with hortatory federal goals. Contention rises when Congress makes state or local goal adoption mandatory and threatens penalties. It can also rise when goal adoption is required as a new condition added to existing intergovernmental grants routinely awarded every year.

Even with agreement about specific goals, states and localities can get testy about federal goal setting. When the U.S. Department of Transportation first adopted GPRA goals and selected safety as its number-one priority, states protested. They urged the department to select a customer service target instead. The department declined. Despite their protest, state officials eventually adopted the same shared priority goal the federal government had selected: safety. In this situation, the problem was not goal dissonance but resistance to federal goal selection without state input and fear that the goal would eventually be used in ways that embarrassed states, such as report cards.

Because the federal agency chose a priority goal that concerned the public and promised not to generate report cards, the goal conflict subsided.<sup>25</sup>

The Healthy People and Department of Transportation examples suggest that many federal agencies may miss opportunities to tap the performance-driving power of specific, challenging goals, especially those chosen based on data about problems and their preventable causes. Moreover, when agencies share data about the relative import, characteristics, and causes of the problems they are trying to reduce, they can influence state and local goal selection. Even within constraints of laws they implement and limited data availability, federal managers can turn to specific, challenging goals to activate and orchestrate other levels of government. These are two of the three key skills Lester Salamon identifies as key to managerial success in third-party arrangements.<sup>26</sup> And, as with Healthy People, federal agencies can use evidence-informed goals to converse with Congress about priorities.

#### *Measurement*

Without measurement, goals are merely words. Measurement brings a goal alive. Measurement serves multiple functions that contribute to performance improvement. It motivates, illuminates, and communicates.<sup>27</sup> Measurement also informs the selection of goals, strategies, and tactics. The rich, outcome-focused measurement of Healthy People supported goal setting, goal reinforcement, and goal attainment.

Measurement motivates because people (and organizations) take pride in their accomplishments and like to do well. Measurement illuminates problems, programs, or places needing attention. It reveals which problems have a higher incidence and greater consequences, causal factors linked to the problems, and their frequency. It also reveals malfunctioning processes that need fixing. Measurement also illuminates promising prevention and treatment practices. Comparable measurement of similar performance units identifies top performers worth studying to determine whether their actions merit replication. Tracking changes before and after the introduction of a new government practice suggests possible effects of the practice. Investigating anomalies to understand why they occur can lead to the discovery of unexpected solutions as well as unknown problems.

In addition, measurement communicates. Measuring progress toward priority goals reinforces their importance. The failure to measure progress toward a priority goal signals that the espoused priority is, in fact, not a priority. Measurement also supports cooperation among multiple parties. It shifts attention from turf battles about assumed solutions to an examination

of the evidence, strategy formulation, and assignment of responsibilities.<sup>28</sup> In addition, well-communicated measurement informs electoral and consumer choice, serving as a sort of market mechanism leading to improved societal outcomes.<sup>29</sup>

Measurement also helps organizations better calibrate the ambitiousness of a goal and their strategies. Without measurement, researchers find, people tend to judge their past practices as more effective than they in fact were, resulting in adoption of overly ambitious goals and subsequent investment in wasteful strategies.<sup>30</sup>

Although goals need measurement to be effective, measurement does not always need a goal to be useful. Comparison with the past, peers, or other problems naturally sets de facto goals, although caution must be exercised when using peer comparison to avoid discouraging those who do not like to compete and encouraging others to game the system.<sup>31</sup>

Experience suggests that the kind of information federal agencies collect from states and localities and the way they use it affects both outcomes and intergovernmental contentiousness. When federal agencies gather information primarily for knowledge management, to help states and localities learn from their own and others' experience, measurement is a powerful intergovernmental tool that improves performance and enhances accountability without dampening state and local flexibility. When they use it primarily to determine grant eligibility, document compliance with grant conditions, or motivate improvement by embarrassing low performers, measurement rankles.

The experience of the federal road (now highway) program illustrates the potential of measurement as a tool of influence.<sup>32</sup> When the first federal road office was established in 1904, it made information the center of its strategy for working with states. It gathered information about road conditions, mileage, and program characteristics in every county and conducted experiments to find effective road- and bridge-building practices and then frequently disseminated its findings to other levels of government. Information collection, analysis, and dissemination was and is a core federal highway function.<sup>33</sup> In 2000 the Federal Highway Administration modernized its knowledge management role by launching online versions of its publications and databases to reach more interested users.<sup>34</sup>

Congress requires the secretary of transportation to report annually on the condition and performance of the nation's roads but does not mandate inclusion of state-specific data in the annual report. The Federal Highway Administration opted to establish state reporting requirements through regulation.

Despite the absence of federal law mandating state reporting, states have willingly delivered data to the federal government for more than fifty years. Why? According to one agency official, “We have been doing highway statistical information for 50 to 60 years, and have turned it back [to the states] from the beginning. We have always done some value-added work when we turned it back to the states. It lets them see what other states were doing, and they see the data as a really valuable resource.”<sup>35</sup> By functioning as a knowledge manager serving states and localities, the federal road agency built a performance-improving partnership with states that has flourished more than 100 years.

As automobile use increased, so did accidents. In 1966 Congress responded by adopting a new federal goal: highway traffic safety. It required every state to establish a highway safety program in accordance with uniform federal standards.<sup>36</sup> The National Highway Traffic Safety Administration (NHTSA) funds state employees in every state to review and code incident reports, such as police crash and coroner reports, collected for local purposes. State staff record data about traffic-related fatalities before, during, and after each accident, noting key characteristics of physical, social, and environmental conditions associated with the accident such as the state of the operator, type of equipment, and accident costs. They submit the information to a national highway fatality database. A dozen states voluntarily supplement the NHTSA’s fatality database with their own data about nonfatal crashes.

The NHTSA not only supports and collects measurement, it also functions as the scientist in Brandeis’s laboratories of democracy. It studies the data it collects to look for patterns of problems, such as accident levels correlated with driver age and alcohol use. It determines the most prevalent problems needing attention. It also looks for anomalies and tries to understand their causes. For example, the NHTSA helped one state understand why it had a higher right-angle crash rate than other states: as it turned out, more driveways in the state feed directly onto major thoroughfares than in other states. When states change their laws, the NHTSA compares changes in fatality rates in changed and unchanged states. Studying the effects of changes in state laws enabled the NHTSA to discover that when states allow police to pull people over to check safety belt use (instead of checking only when police stop drivers for other reasons), it increases belt use and lowers fatality rates.<sup>37</sup>

The NHTSA also uses measurement to help states find and replicate effective practices.<sup>38</sup> North Carolina identified an interesting program in Canada that combined a media campaign with a “stop-and-check” blitz to increase

seat belt use. North Carolina replicated the program, with good results. South Carolina wanted to try it but lacked primary enforcement (“stop-and-check”) authority, so the state sought help from the NHTSA to adapt the North Carolina program. The NHTSA developed audience-focused outreach materials to increase public awareness of the importance of wearing safety belts, including sample materials packaged for key target audiences such as schools (for example, morning announcements), law enforcement officials, parents, and others.<sup>39</sup> The adaptation worked; after adoption, belt usage increased 9 percent in South Carolina and other southeastern states.<sup>40</sup> The NHTSA then rolled the approach out nationally, simultaneously promoting state adoption of primary enforcement laws. It used grants to recruit volunteer states to participate in a controlled, measured experiment. Ten states tested the NHTSA Click It or Ticket campaign, four states served as a control group that did nothing, and four states tested programs of their own design. The NHTSA funded observers to measure belt use before and after the campaign, using a common measurement methodology. Belt use increased 8.2 percent in full implementation states, 2.7 percent in states using programs of their own design, and 0.5 percent in the control group of states.<sup>41</sup>

The NHTSA’s data-rich, audience-focused work with states and localities dramatically improves outcomes. Automobile fatality rates have fallen in all but two of the past twenty-five years, and safety belt use climbed from 58 percent in 1994 to 81 percent in 2006.<sup>42</sup>

Intergovernmental measurement efforts are not always so successful. Data collection for early urban grant programs was more troubled because program offices used the data “as a means of control rather than a means of knowing.”<sup>43</sup> Programs used data to determine grant eligibility or calculate grant size, or they required planning and needs assessments for grant applications. Unfortunately, programs did not analyze the plans or the data they collected to understand the nature of urban problems, nor did they search for state and local programs that might be worthy of replication. They measured simply to monitor compliance or justify spending, forfeiting a valuable opportunity to make measurement useful to data suppliers, key decisionmakers, and the public.

The evolution of intergovernmental measurement in the education field is also revealing. States began voluntarily supplying education data to the federal government in 1869, which the federal government organized and published in data tables. Several problems limited the performance-improving value of state education data. It was not standardized across states or even

within a single state. Few states gathered outcome information, so the federal government could not determine whether any had improved outcomes over time. And although looking at local variations in educational outcome is essential to finding effective approaches because education is delivered locally, the federal government did not collect local data.<sup>44</sup>

Changes in education data measurement have evolved to fix these problems. Data standardization began in the 1950s, and the federal government began gathering data from local governments in the mid-1970s. Collection of educational outcome information did not begin until the 1990s, however. As with federal managers of urban grants, managers of federal education programs resisted efforts to collect educational outcome data along with the data they already collected about students and spending because outcomes are hard to control and might show problems that put program funding at risk.<sup>45</sup>

Congress finally decided it needed outcome data to inform its decision-making and created the National Center for Education Statistics in 1965. State education officials initially supported the creation of the office but changed their position as data-supplying proposals emerged. They feared an increased workload, being embarrassed by poor comparative performance, and meddling by Washington. Without strong state support, even with a congressional mandate, the newly formed educational statistics unit could not fight the resistance of powerful “program managers . . . [who] had no intention of relinquishing control over data they were collecting to establish eligibility for federal funding, to monitor state and local compliance with federal regulations, and to justify the program’s existence to Congress.”<sup>46</sup>

The first successful federal attempt to gather educational outcome data came in 1969 with the creation of the National Assessment of Educational Progress (NAEP). Again fearing embarrassment, state program officials sought to water it down. They made sure the NAEP measured educational outcomes only at the national level and could not detect state- and local-level performance differences. This, of course, prevented the National Center for Education Statistics from analyzing variations to find effective practices that would help states and localities improve.<sup>47</sup>

State attitudes about outcome data began to change in the 1980s when federal and state political leaders got exercised about educational quality. In 1981 the U.S. secretary of education convened the National Commission on Excellence in Education. The commission’s 1983 report, *A Nation at Risk*, showed Americans scoring last on seven out of nineteen international academic tests and never scoring first or second.<sup>48</sup> Like the Healthy People

reports, *A Nation at Risk* used measurement to nominate an issue for national attention.

State governors responded. Concerned about the economic consequences of inadequate education for students in their states, the nation's governors overrode the objections of their own state program officials and asked the federal government to administer state-level NAEP tests. Because not every governor wanted this information, the governors requested it on a voluntary basis. By 1996 forty-four states had voluntarily requested some form of NAEP testing.<sup>49</sup>

The 1984 cohort of governors, one of whom eventually became president and two of whom eventually became secretaries of education, pressed aggressively for a stronger federal role in educational outcome measurement throughout the 1980s and early 1990s.<sup>50</sup> In 1994 Congress responded by passing the first federal law mandating that every state measure educational outcomes. It took the governors, acting collectively, to lead the charge for federal action to produce measurement of educational outcomes. Even with the state-initiated push to require every state to measure outcomes, the 1994 law let each state set its own standards and explicitly prohibited federal agency action on national standards, even voluntary ones.<sup>51</sup>

A confluence of political events, including the presidential election of another governor who strongly supported mandated state measurement of educational outcomes, moved common outcome measurement one step further with the No Child Left Behind Act of 2001 (NCLB).<sup>52</sup> The NCLB delicately balances intergovernmental measurement and federal mandates. It lets states set their own standards of educational performance but mandates that every state participate in NAEP testing. The law also strengthens sanctions on states that fail to measure progress toward state-set standards, since many states had not complied with that requirement in the 1994 law, and requires each state to ensure that every school makes adequate performance progress against its own standards.

The NCLB quickly stirred up the perennial intergovernmental tug-of-war. The Connecticut attorney general, citing the high cost of measurement, sued the federal government.<sup>53</sup> Yet despite state pressure to change some provisions of the NCLB, including great unhappiness with NCLB sanctions, the nation's governors urged federal adoption of a common method for measuring high school dropout rates and remain supportive of the federal government's collecting common outcome data. Much to the surprise of their own program officials, several years after the law's passage, governors continued to

endorse the idea of the federal government's holding states accountable for student learning while calling for greater flexibility.<sup>54</sup>

The NCLB's focus on outcomes has not changed the control orientation of federal Department of Education program offices. The law includes statutory deadlines and punishment-triggering accountability mechanisms that keep program offices consumed with control rather than the search for and promotion of replication-worthy practices.

The NCLB has, however, resulted in a stronger research and statistics office that is beginning to play a learning leadership role. The Department of Education now funds an annual compilation of school-level performance measurement from every grade in every school in every state, normalizing school scores by comparison with state median scores for each grade and using the NAEP to put state scores in perspective. The Department of Education makes this normalized school data available to the public, inviting analysis by those outside government. A nongovernmental organization, Education Trust, has accepted this implied invitation every year, releasing an annual report identifying high-performing schools in high-poverty, high-minority areas. Education Trust encourages other researchers to look for distinguishing practices that are evident in these high-performing schools but not in low-performing ones, practices that others might want to adopt.<sup>55</sup> The Department of Education has also built a "what works" clearinghouse and provides online data-mining tools to facilitate comparisons along specific dimensions (for example, math performance trends in grade 8 in urban schools) to aid the search for models with useful lessons.<sup>56</sup>

Block grants that combine existing categorical grants advancing narrowly specified federal goals into a single more flexible grant can be particularly challenging for federal agencies that want to create an outcome-focused measurement system that helps grantees learn from one another's experience, especially when Congress constrains federal agency flexibility in promulgating regulations implementing the laws.<sup>57</sup> Yet even with block grants, federal agencies can work with states to acquire the needed outcome-focused information if they adopt a learning and leadership attitude. In implementing the maternal and child health block grant, the U.S. Health Resources and Services Administration negotiated with states a common set of outcome indicators every state would report. Each state can choose its own priorities, but all states must report the same outcome indicators. Indicator selection is an ongoing process and improves as everyone learns. Healthy People targets provide the context for, but do not dictate, indicator selection. The health

resources agency also posts national and state priorities and performance on the web.<sup>58</sup>

The experience with measurement in urban grant programs, education, transportation, and maternal and child health suggests federal agencies that depend on states and localities to accomplish their objectives would speed performance improvements, reduce intergovernmental contention, and enhance democratic and performance accountability if they made the knowledge management function more central to their intergovernmental relationships, especially since passage of the GPRA. Laws such as the Paperwork Reduction Act, the Federal Education Rights and Privacy Act, and other federal legal restrictions can make this difficult, yet the experience with Healthy People, the NHTSA, the Health Resources and Services Administration, and the Department of Education suggest it is an effort worth making. Agencies should place greater emphasis on building measurement systems that advance understanding of problems, probable causes, and effective prevention, treatment, and mitigation practices rather than monitoring primarily for control.

#### *Incentives*

What happens when incentives, both negative and positive, are used with goals and measurement? Answering this question is critical for containing unconstructive intergovernmental tensions and improving societal outcomes.

Incentives can be remarkably effective. Even with knowledge that fewer fatalities occur at fifty-five miles an hour than at eighty, many more people would speed down the highway if they did not fear the threat of a traffic ticket and higher insurance rates. But incentives can also discourage people and organizations, triggering high levels of stress and dysfunctional responses.<sup>59</sup>

Structuring incentives appropriately is a tricky business. Managers need to master what Salamon calls “modulation” skills “to decide what combination of incentives and penalties to bring to bear to achieve the outcomes desired.”<sup>60</sup> Experience suggests that federal managers should use grants primarily to identify, strengthen, and sometimes create goal allies and to stimulate common outcome measurement across states and localities. They should use penalty threats to compel adoption of goals dealing with serious but locally ignored problems and spur development and implementation of cogent strategies.

GRANTS. Both grants and penalties require prior congressional authorization. Federal grants account for about a quarter of state and local expenditures.<sup>61</sup> Grants are often seen as “the initial tool used to stimulate interest and adoption.”<sup>62</sup> But grants have limited value as a performance driver unless

paired with outcome-focused goals and measurement. Without them, grants risk generating “ineffectual” data and “useless forms of window dressing.”<sup>63</sup> When paired with outcome-focused goals and used to stimulate standardized outcome-focused measurement, though, grants are remarkably effective.

The federal government uses grants “to influence the conduct of state and local governments in such a way as to promote the realization of its own goals.” The simple act of offering a grant, especially a categorical grant for a specific purpose, nominates an item of federal interest for local political attention. Grants can stimulate healthy local debate about goals and the right balance among competing priorities. Federal grants essentially initiate an extended intergovernmental negotiation regarding both ends and means.<sup>64</sup>

One way grants influence the selection of state and local goals is by lowering the cost of pursuing a specific objective, but federal grants seem to work most powerfully in another way: they identify, strengthen, and sometimes even create local goal allies. By applying for a grant, states and localities indicate their shared interest in a problem that concerns the federal government.<sup>65</sup>

The availability of a grant strengthens a local ally’s bargaining position in the battle for local resources, and grant deadlines give local goal allies a reason to move an issue up on the local action agenda.<sup>66</sup> Federal grants have successfully created a core of professional counterparts in the states, often a single state agency, who function as perennial goal allies. Their ability as allies to improve outcomes is limited, however, if grant conditions tend more to administrative matters than to outcomes.<sup>67</sup>

Even when grants are allocated by formula rather than competitive application, they can boost existing local goal allies. Block grants help local allies by providing them with resources to do what they already want to do without local strings attached. Categorical formula grants that routinely get distributed to the same program every year sustain local goal allies’ ability to devote time to federally nominated problems.<sup>68</sup>

In addition to identifying and strengthening goal allies, federal grants can stimulate the generation of standardized data and state and local measurement capacity. When these data aid the public in understanding outcomes, data-stimulating grants enhance democratic accountability. When agencies require measurement simply to determine funding levels or document the completion of required activities, such grants have little democracy-enhancing value.

Federal agencies can also use grants to enrich the value of collected data. The Federal Highway Administration provided grants to states to develop software that transformed data collected for federal purposes to information useful for state budgeting and planning.<sup>69</sup> The Centers for Disease Control

purchased business intelligence software licenses to enable it to work more closely with states on epidemiology, disease detection, and response activities.<sup>70</sup>

Federal grants can also support experimentation and innovation to find more effective and cost-effective interventions. Federal agencies can use grants to attract participants to measured experiments, as the NHTSA has done. Nonrecurring grants are useful as seed capital to test new approaches or to pay one-time costs for equipment upgrades.<sup>71</sup> For these kinds of grants to be useful, however, their effect needs to be assessed. Otherwise, ineffective practices are likely to be repeated.

Perhaps the biggest challenge with intergovernmental grant management arises in setting and managing grant conditions.<sup>72</sup> Grant conditions are established to ensure that the federal government gets what it thinks it is buying for its money, rather than simply relieving the local tax burden or supporting local projects that do not advance federal objectives. Grant conditions can successfully catalyze constructive change in states and localities, but they can also result in wasteful activities.<sup>73</sup>

Federal agencies often stipulate administrative conditions. These are very effective when a clear link has been established between the administrative action required and the target outcome. Administrative conditions have been effective in creating in-state goal allies and measurement capacity and in motivating adoption of practices demonstrated to be effective. Administrative conditions also minimize unethical and illegal grant uses.

Problems arise from the number, nature, and way grant conditions are interpreted. Just staying in compliance with the sheer number of not-always-consistent grant conditions can consume so much grantee time that it precludes pursuit of more productive endeavors. Another problem arises when grant conditions fail “to cover the questions of major substantive importance” and focus instead on administrative, personnel, and fiscal practices.<sup>74</sup> Yet federal agencies opt for administrative rather than outcome-focused grant conditions for several sensible reasons: the need to recruit local allies who might not apply if they fear grant conditions will expect them to change societal conditions they cannot fully control, the relative ease of tracking administrative matters instead of outcomes, and the political difficulty of selecting specific outcome targets, especially values-based ones.<sup>75</sup>

The problem is that administrative grant conditions often become ends in and of themselves that interfere with outcome gains. Many federal laws, for example, require states and localities to write plans as a grant condition. Agencies confirm plan completion but fail to learn from plan content. They

do not look for what the plans say about changes in local conditions that might be instructive for other communities nor for patterns of problems across communities that might respond to common solutions. Sometimes, they do not even expect grantees to use the plans they prepare.<sup>76</sup> As a consequence, many grantees simply complete minimal plans to meet grant conditions but never use them. Grantees submit copious quantities of data as required that are neither analyzed nor returned to data suppliers with enhanced value. When federal agencies use data for control rather than knowing, state and local grantees tend to treat their data-reporting obligation as a task to be completed rather than a contribution to a shared knowledge base from which they will benefit. A vicious cycle is thereby established wherein data suppliers stop worrying about data quality, making the data useless if federal agencies eventually try to analyze it for lessons worth sharing.

Overly rigid interpretation of grant conditions by either the grantee or grantor can also trigger silly decisions. For example, to deal with the worry that federal funds will simply substitute for state or local spending with no net benefit to the community, federal laws and agencies often establish maintenance-of-effort requirements and then try to follow the money trail.<sup>77</sup> This created an absurd situation in one program studied by researchers: a community pulled their educationally disadvantaged students out of the classroom for tutorial instruction. Although “pull-out” programs had been found educationally inferior to in-classroom training and stigmatized students, hiring tutors made it easier to track spending to satisfy federal reporting requirements.<sup>78</sup>

With so many oversight bodies in government looking for problems, grantees seeking to avoid problems must follow rules strictly rather than exercise commonsense discretion. Failure to do so can get them into trouble. One state environmental agency learned this the hard way. The inspector general of the federal Environmental Protection Agency (EPA) criticized the state agency when it let its water quality program borrow a video camera purchased for the air quality program, even though the air program was unable to use the camera full-time.<sup>79</sup>

Following the money also creates the impression that federal budget watchers can calculate return on federal spending. When federal money gets put in a pot with state and local money to advance a shared objective, determining what those federal dollars bought becomes a practical impossibility. To deal with this problem, federal agencies opt to track activities, people, and equipment bought rather than changes in societal outcomes because they are easier to measure. But measuring activities instead of outcomes makes it

harder to find out what federal spending, or even federal spending combined with state and local funds, did or did not accomplish.

In sum, grants can be an effective tool of federal influence. They can lower the cost for localities of tackling specific objectives the federal government has judged as important. They can identify, strengthen, and sometimes create goal allies. They can also stimulate generation of standardized data, state and local measurement capacity, local use of collected data, innovation, and program modernization. But grants do not always have a positive effect. When grant conditions are too numerous and rigid, a problem common to mature grant programs, and when attention to administrative matters overwhelms attention to outcomes, grant conditions can cause wasteful activity and compromise program productivity.<sup>80</sup> To drive performance improvement, Congress and federal agencies should first and foremost use grants to persuade states and localities to collect and report credible and comparable outcome and program information. The federal government should also use grants to encourage adoption of specific, challenging outcome-focused targets by the federal government when there is a reason for national targets and by states and localities when there is not. With passage of the GPRA, Congress has, in most cases, given federal agencies implicit authority to make this happen.

Federal agencies must do more than use their grant authority to encourage states and localities to supply outcome information, though. They also need to organize, analyze, and disseminate the data they collect in ways that help state and local decisionmakers make wiser choices about priorities and program design, choices that are informed by evidence about problems and effective interventions. In addition, federal agencies need to encourage analysis by others as the Department of Education has begun to do by making its databases easily accessible on the Internet with online tools to support public analysis. Federal agencies should place a priority on returning knowledge to state and local data suppliers in a way that catches their attention and aids their decisionmaking. They should also regularly engage grantees in discussions of the data to enhance discovery and the learning process.<sup>81</sup> If agencies fail to do this on their own, the U.S. Office of Management and Budget should encourage them to do so during OMB program reviews. Congress, too, should adopt laws requiring agencies to assume a value-adding knowledge management function.

One final observation on grants: many experts describe the intergovernmental arrangement as a principal-agent relationship.<sup>82</sup> The evidence reviewed here suggests that this is not the appropriate conceptual framework. A principal-agent relationship assumes divergent interests of the federal

government as the principal and state or local governments as the agent. The grant is seen as a mechanism for aligning the interests and objectives of the principal and agent, and measurement is seen as a monitoring device used to confirm that the agent is serving the needs of the principal. The examples and studies reviewed here, however, suggest that grants should instead be seen as a tool for recruiting, activating, and strengthening those whose interests are already aligned, for orchestrating cooperation and learning among them, and for securing the outcome-focused indicators every state must supply. Measurement, in turn, should be seen not so much as a monitoring mechanism but as a device for strengthening goal allies, supporting their ability to recruit and motivate others, illuminating problems and possibilities, and communicating lessons.

Penalties. Grants are a positive, preperformance incentive, whereas penalties are a negative, postperformance motivator. Perhaps the key difference between grants and penalties in the intergovernmental context is that grants invite state and local attention to a goal, whereas penalties compel it. Penalties can be a useful tool to force attention to externalities, the costs one party imposes on another. They are also useful when federal policymakers decide that all states or localities should protect specific rights or ensure a minimum level of well-being for their residents.

Because they are coercive, penalties are far more contentious than grants. Penalties can effectively drive performance gains and stimulate democratic debate, but they can also polarize political positions. In addition, they can galvanize organized opposition that successfully constrains federal program authority and contributes to performance declines. The federal government must therefore exercise great caution when using penalties to avoid positional stalemates, measurement manipulation, implosion of the measurement system, or elimination of its penalty power. It must master its modulation skills.

Penalties can effectively compel others to adopt a target and to take actions necessary to meet that target even when significant expenditure of local resources is needed to do so. With the Clean Air Act Amendments of 1990, for example, Congress succeeded in getting local attention to air quality problems by threatening the loss of lucrative federal highway funds (an exception was made for highway safety projects) and curbs on development. Under this law, sanctions are automatically triggered if a community does not meet various federally set air quality standards by specified dates. Over nearly two decades, the potential loss of this big pot of dollars has successfully convinced states and localities to adopt costly actions, such as vehicle air emission inspection programs. It has also improved air quality. States and

localities have seldom embraced these measures enthusiastically or immediately, but they have eventually adopted numerous practices resulting in improved air quality.

Penalties are useful not only for forcing attention to a problem but also for motivating states to measure outcomes. Little progress was made reducing tobacco sales to minors until Congress adopted a penalty for states that did not use a common and credible method to measure sales. Since then, all fifty states have reported tobacco sales annually, and tobacco sales have steadily declined.<sup>83</sup> Smoking among minors has also declined significantly in the same period, although not solely owing to reduced sales to minors.<sup>84</sup>

Penalties can also increase the use of data by local communities, contributing to better informed decisions. The 1968 Intermodal Surface Transportation and Efficiency Act required states to adopt six different information management systems or lose up to 10 percent of transportation monies awarded under the act. The mandate infuriated states, who successfully lobbied Congress to eliminate it. Ironically, even after the penalty was eliminated, most states ultimately adopted the mandated management systems anyway because they found them useful.<sup>85</sup> Penalties proved an effective short-term way to promote adoption of learning tools that persisted even after the penalties were removed.

Federal penalties to encourage state adoption of universal motorcycle helmet laws have been similarly effective, albeit highly contentious. The Highway Safety Act of 1966 threatened states that did not adopt universal helmet laws with the loss of 10 percent of federal-aid highway funds. Although no states had universal helmet laws in 1966, all but California and Utah had adopted them by 1975.<sup>86</sup> When the secretary of transportation moved to penalize these two noncompliant states, his action galvanized intense opposition. Already furious about the loss of their freedom to ride without a helmet which they valued more highly than the personal safety benefit of a helmet, motorcycle riders convinced Congress to eliminate federal penalty power and restore their freedom to ride as they choose. Soon thereafter, twenty-eight states dropped their universal helmet laws.

Congress again authorized federal penalties in 1991. This time, the National Highway Traffic Safety Administration moved quickly to penalize more than half the states for failure to adopt universal helmet laws. It moved too fast. In 1995 Congress repealed federal helmet penalty authority at the insistence of newly elected Colorado senator Ben Nighthorse Campbell, a motorcycle rider who liked neither helmets nor federal mandates.<sup>87</sup> As of

March 2008, only twenty states and the District of Columbia had universal helmet laws, twenty-seven required younger riders to wear helmets, and three states had no helmet law at all.<sup>88</sup> As the number of states with universal helmet laws has fallen, motorcycle fatalities have steadily climbed.<sup>89</sup>

When do penalties work, compelling local attention to a goal and local actions to attain it, and when do they backfire? Several factors may explain the success or failure of a penalty: the size of the penalty, its structure, the validity of mandated actions (dependent on the quality of the evidence about changed outcomes), the strength of organized opponents compared with that of organized proponents, and the presence of interstate externalities.

The size of the penalty matters. The 1991 motorcycle helmet penalty law allowed the federal government to shift a portion of federal highway construction funds for noncompliant states to state transportation safety programs, a far weaker threat than the 1975 penalty of a 10 percent holdback of federal aid for highways. All but two states eventually changed their practices in response to the 1975 penalty threat; but only two states adopted new universal helmet laws between the passage of the 1991 law and its 1995 repeal.<sup>90</sup>

Penalty structure also matters. A pyramid of escalating responses motivates compliance more effectively than a one-size-fits-all penalty.<sup>91</sup> Under the Clean Air Act as amended in 1990, a state's failure to meet a nationally set air quality target triggers the commencement of an escalating response process, not immediate curtailment of highway funding and development flexibility. A state that does not meet its target must develop a cogent strategy to do so, the state implementation plan. It must get the plan approved and then implement it. Each state develops its own strategy, using air quality data gathered with federal financial support in accordance with national standards and evidence about effective interventions. The federal EPA uses a model to project whether a state's proposed pollution control measures will meet the standards by the target date. When it deems proposed actions inadequate, it negotiates with the state other actions needed. The agency cuts funding and development flexibility only as a last resort, when a state refuses to develop and implement an acceptable implementation plan. The performance accountability approach used in the federal air quality law, exercising the extreme penalty of funding cuts and development restrictions not if a state fails to meet its target but only if it fails to measure air quality, develop cogent strategies, and implement them, has also been used successfully in other federal programs.<sup>92</sup> The approach also corresponds to the performance accountability principle attributed to former New York City police commissioner

William Bratton for precinct captains: “No one ever got in trouble if the crime rate went up. They got in trouble if they did not know why it had gone up and did not have a plan to deal with it.”<sup>93</sup>

One mistake federal agencies sometimes make is to penalize states and localities for failure to adopt specific practices that are not backed by evidence of their effectiveness. Early versions of the EPA model used to review the adequacy of state plans to achieve air quality targets, for example, used assumptions about the effects of air pollution control actions that even EPA regional staff did not understand and could not explain to states. Not surprisingly, this enraged the states. The agency subsequently relieved this problem by updating the models it uses to assess state plans, making the underlying evidence and assumptions transparent, and engaging states in model development. Intergovernmental tensions on plan reviews have subsided significantly. They have not abated completely, of course, because few like having their discretion restricted and being required to obtain another’s approval.

Despite state fury about the initial EPA black-box review and penalty threat under the Clean Air Act, states were unable to eliminate penalties as motorcycle helmet opponents had. The penalty threats of both laws catalyzed organized, powerful, and sustained opposition, but in the case of the federal air quality law, highly organized proponents including the American Lung Association successfully countered the opponents and preserved federal penalty authority.<sup>94</sup>

Another possible reason for the survival of the penalty threat is that states themselves have different views about federal penalties. Downwind states bear health care and compliance costs from upwind polluters. They want the EPA to have the power to reduce negative interstate externalities.

Penalties need not be financial to be effective; they can also be reputational. External advocates commonly use comparative data to try to embarrass government agencies into better performance. Federal agencies, however, are likely to encounter trouble if they try comparisons to embarrass or otherwise penalize low-performing states and localities.<sup>95</sup> Such actions can trigger resistance to outcome-focused measurement and prompt “cream skimming,” whereby programs opt to serve clients that boost their own performance ratings rather than those with the greatest need or potential for gain. In addition, it can lead to submission of data so poor it is useless.<sup>96</sup> Federal agencies are likely to be more successful when they use comparison to find programs worth replicating, not to penalize.

Penalties, like grants, appear to be an effective federal tool for catalyzing local debate about a federally nominated problem. Unlike grants, which

invite the debate, penalties force it. This makes them particularly useful for dealing with interjurisdictional externalities and national minimum standards that are costly for states and localities to address. More than just stimulating debate, federal penalties compel other levels of government to adopt and pursue federally nominated targets, as they did with air quality improvements around the country and, while they lasted, reduced motorcycle fatalities. They can also be effective in getting jurisdictions to use common measurement methods. They compelled measurement of tobacco sales to minors. Penalties also increased state and local use of infrastructure management software tools. But penalties are provocative, especially when they compel actions that seem unreasonable to those asked to take them or violate values held dear. They are especially provocative when a federal agency quickly exercises the most costly version of a penalty, rather than holding it in reserve as an available threat. A modulated, escalating response strategy that uses the harshest punishments only as a last resort is likely to be more effective. If no proponents are strong enough to counter opponents galvanized into action when penalties are used, federal penalty authority can be eliminated as a tool in ongoing intergovernmental negotiations about the right balance of government goals to pursue.

### **Conclusion**

How can federal managers dependent on other levels of government to accomplish their objectives manage most effectively to improve performance and enhance accountability? Answering this question is especially important since passage of the GPRA, which requires every federal agency to set outcome-focused goals and report progress annually to Congress. Goals, measurement, and incentives are powerful tools federal agencies can use to navigate the practical and political aspects of intergovernmental program delivery.

Federal agencies can exercise effective leadership by engaging experts and stakeholders in reviewing the evidence and then nominating goals needing national, state, and local attention, as the Department of Health and Human Services does with its decennial Healthy People reports, inspiring, activating, and orchestrating goal allies and stimulating healthy democratic discussion about priorities among the levels and branches of government. They can do this even without an explicit link to positive and negative incentives.

Federal agencies can promote learning in states and localities by encouraging measurement, using grants and regulations to stimulate common outcome-focused measurement as needed. They can give the data they collect

greater value by organizing, analyzing, disseminating, and using it to understand problems and identify solutions, as the NHTSA does routinely and as the Department of Education did when it normalized state and local performance data, assembled it into a single database others could analyze, and supported analysis by providing online data analysis tools.

Federal agencies should use grants to recruit, strengthen, and sustain goal allies and to stimulate the generation and sharing of standardized outcome-focused measurement, negotiating the content, timing, and technical aspects of data submission. If a federal agency lacks the authority to require outcome-focused data submission, it should make the case to Congress to obtain that authority. Federal agencies can even make this work with block grants, as the Health Resources and Services Administration did with the maternal and child health block grant, whereby each state chooses its own priorities but all states report on a common set of outcome indicators.

Federal agencies should use penalties with care to compel attention to federally set goals and obtain useful measurements. The most extreme penalties are more powerful as a threat than when imposed, when they can so upset those threatened by penalties that they provoke formation of an opposition coalition. Before penalizing another level of government because it has failed to achieve a target or to take a specific action, federal agencies should focus first on getting grantees and those threatened by penalties to adopt outcome-focused targets set by Congress (or at congressional instruction, by the agency) or set their own outcome-focused goals, measure progress toward goals using common indicators, develop cogent strategies based on the best available evidence, and implement the strategies. If local political views make it difficult even to get the goal adopted, federal agencies should help local goal allies make the case for change, using penalty threats incrementally to help them win local attention and allies.

More research is clearly needed to understand under what circumstances goals, measurement, grants, and penalties are likely to work in an intergovernmental context. It is hoped that this chapter inspires that research. It is also hoped that this discussion provides useful interim guidance to Congress in writing new laws and to federal agencies in managing programs that require them to depend on other levels of government to accomplish their objectives.

## Notes

1. *New State Ice Co. v. Liebmann*, 285 U.S. 262, 52 S.Ct. 371, 76 L.Ed. 747 (1932).
2. For an introductory discussion of knowledge management, see Robert M. Grant, "Toward a Knowledge-Based Theory of the Firm," special issue, *Strategic Management*

*Journal* 17 (Winter 1996): 109–12. For a discussion of leadership, see David M. Messick and Roderick M. Kramer, eds., *The Psychology of Leadership: New Perspectives and Research* (Mahwah, N.J.: Lawrence Erlbaum, 2005).

3. Martha Derthick, *The Influence of Federal Grants: Public Assistance in Massachusetts* (Harvard University Press, 1970); Timothy Conlan, “Grants Management in the Twenty-First Century: Three Innovative Policy Responses,” IBM Center for the Business of Government, 2005, pp. 10–12; quotation is from Eugene Bardach and Cara Lesser, “Accountability in Human Services Collaboratives: For What? And to Whom?” *Journal of Public Administration Research and Theory* 6 (April 1996): 197–224, p. 201.

4. Steven Kelman, “Improving Service Delivery in the United Kingdom: Organization Theory Perspectives on Central Intervention Strategies,” *Journal of Comparative Policy Analysis* 8 (December 2006): 393–419; David G. Frederickson, “The Potential of the Government Performance and Results Act as a Tool to Manage Third-Party Government,” PricewaterhouseCoopers Endowment for the Business of Government, August 2001; Shelley Metzenbaum, “Strategies for Using State Information: Measuring and Improving Program Performance,” PricewaterhouseCoopers Endowment for the Business of Government, December 2003 ([http://businessofgovernment.org/pdfs/Metzenbaum\\_Report.pdf](http://businessofgovernment.org/pdfs/Metzenbaum_Report.pdf)).

5. Paul E. Peterson, *The Price of Federalism* (Brookings, 1995).

6. The powers of the states and people are laid out in the Tenth Amendment to the U.S. Constitution. The system of checks and balances is the subject of James Madison, *Federalist* No. 51, in *Federalist Papers*, by Alexander Hamilton, James Madison, and John Jay (London: Penguin Classics, 1987).

7. Daniel J. Elazar, *American Federalism: A View from the States*, 2nd ed. (New York: Thomas Y. Crowell, 1972); Thomas R. Dye, *American Federalism: Competition among Governments* (Lexington, Mass.: Lexington Books, 1990).

8. Conlan, “Grants Management in the Twenty-First Century,” p. 13.

9. Peterson, *Price of Federalism*, p. 13.

10. An outcome-focused goal is one that focuses on the real-world conditions that a program seeks to change, such as water quality, percentage of people living in substandard housing, and commute times, rather than on agency activities, such as the number of permits issued, inspections conducted, or penalties collected.

11. James Fossett and others, “Federalism and Performance Management: Health Insurance, Food Stamps, and the Take-Up Challenge,” in *Quicker, Better, Cheaper: Managing Performance in American Government*, edited by Dall W. Forsythe (Albany: Rockefeller Institute Press, 2001), pp. 207–44.

12. Radin and others see the GPRA and the Program Assessment Rating Tool as two distinct and incompatible entities. According to the Office of Management and Budget, PART was developed “to give true effect to the spirit as well as the letter of the law [GPRA]” (Office of Management and Budget, “Rating the Performance of Federal Programs,” in *The Budget for Fiscal Year 2004* [2003] [[www.gpoaccess.gov/usbudget/fy04/pdf/budget/performance.pdf](http://www.gpoaccess.gov/usbudget/fy04/pdf/budget/performance.pdf)]). Thus PART can be seen as the guidance developed by the federal budget office to advance GPRA adoption in federal agencies, not distinct from the GPRA. For a discussion of the specific weaknesses of PART, see Shelley Metzenbaum,

“Performance Accountability: The Five Building Blocks and Six Essential Practices,” IBM Center for the Business of Government (2006).

13. Edwin A. Locke and Gary P. Latham, “Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey,” *American Psychologist* 57 (September 2002): 706–07; Alvin Zander, *Motives and Goals in Groups* (New Brunswick, N.J.: Transaction Books, 1996).

14. Locke and Latham, “Building a Practically Useful Theory.”

15. Gary P. Latham, “Motivate Employee Performance through Goal-Setting,” in *Handbook of Principles of Organizational Behavior*, edited by Edwin A. Locke (Malden, Mass.: Blackwell, 2004), pp. 107–19.

16. Janet A. Weiss and Sandy Kristin Piderit, “The Value of Mission Statements in Public Agencies,” *Journal of Public Administration Research and Theory* 9 (April 1999): 193–223.

17. U.S. Department of Health and Human Services, Public Health Service, *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*, GPO 017-001-00474-0 (Government Printing Office, 1991); U.S. Department of Health and Human Services, Public Health Service, “Promoting Health/Preventing Disease: Objectives for the Nation,” Fall 1980 ([www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/2e/23/e0.pdf](http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2e/23/e0.pdf)); U.S. Department of Health and Human Services, Public Health Service, *Healthy People: The Surgeon General’s Report on Health Promotion and Disease Prevention* (U.S. Department of Health, Education, and Welfare, Office of the Assistant Secretary for Health and Surgeon General, 1979 [<http://profiles.nlm.nih.gov/NN/B/B/G/K/segments.html>]); U.S. Department of Public Health, Public Health Service, *Promoting Health/Preventing Disease: Objectives for the Nation* ([http://odphp.osophs.dhhs.gov/pubs/hp2000/pdf/midcours/appdx\\_b.pdf](http://odphp.osophs.dhhs.gov/pubs/hp2000/pdf/midcours/appdx_b.pdf)); U.S. Department of Health and Human Services, *Healthy People 2010: Understanding and Improving Health*, 2nd ed. (GPO, November 2000 [[www.healthypeople.gov/Document/pdf/uih/2010uih.pdf](http://www.healthypeople.gov/Document/pdf/uih/2010uih.pdf)]).

18. Thomas R. Friedan, M.D., New York City Department of Health and Mental Hygiene, “Mayor’s Management Report,” September 2007 ([www.nyc.gov/html/ops/downloads/pdf/2007\\_mmr/0907\\_mmr.pdf](http://www.nyc.gov/html/ops/downloads/pdf/2007_mmr/0907_mmr.pdf)); Deborah R. Maiese and Claude Earl Fox, “Laying the Foundation for Healthy People 2010,” *Public Health Reports* 113 (January–February 1998): 92–95 ([www.phf.org/hp2010/ac1/dm\\_cf.htm](http://www.phf.org/hp2010/ac1/dm_cf.htm)). Also see “How Will the Objectives Be Used?” at the Healthy People website ([www.healthypeople.gov/About/objused.htm](http://www.healthypeople.gov/About/objused.htm)).

19. *Healthy People 2000 Midcourse Review*, appendix B, “History of the Objectives Development and the Midcourse Review Process” ([http://odphp.osophs.dhhs.gov/pubs/hp2000/pdf/midcours/appdx\\_b.pdf](http://odphp.osophs.dhhs.gov/pubs/hp2000/pdf/midcours/appdx_b.pdf)).

20. Maiese and Fox, “Laying the Foundation for Healthy People 2010.”

21. Public Health Service, *Healthy People: The Surgeon General’s Report on Health Promotion and Disease Prevention*, p. 1–7. This report builds on several earlier surgeon general’s reports identifying smoking as a problem.

22. American Lung Association Epidemiology and Statistics Unit, “Trends in Tobacco

Use,” November 2004 ([www.lungusa.org/atf/cf/%7B7A8D42C2-FCCA-4604-8ADE-7F5D5E762256%7D/SMK1.PDF](http://www.lungusa.org/atf/cf/%7B7A8D42C2-FCCA-4604-8ADE-7F5D5E762256%7D/SMK1.PDF)).

23. It is often suggested that those who set their own goals will have a stronger sense of goal “ownership” and therefore a stronger drive to achieve the goal. Research on goal setting, however, finds that goals set by third parties can be as powerful as self-set goals provided the externally set goals are plausible and the external goal setter possesses sufficient authority to be accepted as a goal setter. The five Healthy People 1979 goals targeting the major life stages were set by a small group in the office of the assistant secretary for health. A broader group of experts convened from around the country set the fifteen priority areas in the 1980 document. Subsequent Healthy People goals and objectives were informed by engagement of the broader public. This suggests that an iterative goal-setting process can work well, with the federal government proposing specific outcome-focused targets based on analysis of the evidence and states and others proposing adjusted targets based on more precise evidence about local conditions and expected effects on known treatments.

24. U.S. General Accounting Office, *Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*, GAO-04-38 (March 10, 2004).

25. Metzenbaum, “Strategies for Using State Information,” p. 31.

26. Lester M. Salamon, “The New Governance and the Tools of Public Action,” in *The Tools of Government: A Guide to the New Governance*, edited by Lester M. Salamon (Oxford University Press, 2002), pp. 1–47.

27. Metzenbaum, “Performance Accountability.”

28. Robert B. Cialdini, *Influence: Science and Practice*, 4th ed. (Boston: Allyn and Bacon, 2001), pp. 157–58.

29. Charles M. Tiebout, “A Pure Theory of Local Expenditures,” *Journal of Political Economy* 64 (October 1956): 416–24; Ginger Z. Jin and Phillip Leslie, “The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *Quarterly Journal of Economics* 118 (May 2003): 409–51.

30. Charles A. Keisler, *The Psychology of Commitment* (Orlando: Academic Press, 1971), cited in Zander, *Motives and Goals in Groups*; Locke and Latham, “Building a Practically Useful Theory.”

31. See Metzenbaum, “Performance Accountability,” pp. 22–25, for a discussion of the literature on peer comparison as a motivational tool.

32. Except where noted, the discussion of the federal roads and traffic safety office comes from Metzenbaum, “Strategies for Using State Information,” pp. 22–33, 57–61.

33. The Federal Highway Administration’s “Highway Statistics” for the year 2000 is described in its preface the “56th of an annual series” (U.S. Department of Transportation, Federal Highway Administration [[www.fhwa.dot.gov/ohim/hs00/preface.htm](http://www.fhwa.dot.gov/ohim/hs00/preface.htm)]).

34. “Editor’s Notes,” *Public Roads* 64 (November–December 2000) ([www.tfhrc.gov/ohim/pubrds/nov00/ednotes.htm](http://www.tfhrc.gov/ohim/pubrds/nov00/ednotes.htm)).

35. Quoted in Metzenbaum, “Strategies for Using State Information,” p. 23.

36. *Highway Safety Act*, P.L. 89-564.

37. California was the first adopter of “stop-and-check” laws, with police having primary enforcement authority, and the NHTSA was quickly able to see how the changes affected outcomes in the state (U.S. Department of Transportation, *National Highway Traffic Safety Facts 2000: Occupant Protection*, DOT HS 809 327 [www-nrd.nhtsa.dot.gov/Pubs/2000occfacts.pdf]).

38. See, for example, U.S. Department of Transportation, *Strong Legislation: The Key to Saving Lives*, DOT 809 823 (www.nhtsa.dot.gov/people/injury/research/6thAnnualBUARreport/pages/StrongLegislation.htm).

39. See National Highway Traffic Safety Administration, “Click It or Ticket Planner 2006” (www.buckleupamerica.org/ciot-planner/planner/index.cfm). Even before the Internet, the NHTSA and the Federal Highway Administration made their data and research findings available to the public in annual compendiums and newsletters.

40. Metzenbaum, “Strategies for Using State Information,” pp. 32–33.

41. Mark G. Solomon and others, *Evaluation of Click It or Ticket Model Programs*, Technical Report DOT HS 809 498, National Highway Traffic and Safety Administration, September 2002 (www.nhtsa.dot.gov/people/injury/airbags/clickit\_ticket/clickit\_composite/clickit\_composite.pdf).

42. Fatalities fell from more than 3.0 per 100 million vehicle miles traveled to less than 1.5 (National Highway Traffic Safety Administration, National Center for Statistics and Analysis, *Motor Vehicle Traffic Crash Fatality Counts and Estimates of People Injured*, August 2006, updated December 13, 2006, DOT HS 810 639 [www-nrd.nhtsa.dot.gov/pdf/nrd-30/ncsa/ppt/2006/810639.pdf]); on safety belt use, see National Highway Safety Administration, *Traffic Safety Facts 2006 Data: Occupant Protection*, DOT HS 810 807 (www.nhtsa.gov/portal/site/nhtsa/menuitem.6a6eaf83cf719ad24ec86e10dba046a0/).

43. Judith deNeufville, “Federal Statistics in Local Governments,” in *The Politics of Numbers*, edited by William Alonso and Paul Starr (New York: Russell Sage Foundation, 1987), pp. 343–62, 349. The study looked at urban renewal, public housing, area redevelopment, and hospital construction categorical grants and community development block grants.

44. Janet A. Weiss and Judith E. Gruber, “The Managed Irrelevance of Federal Statistics,” in Alonso and Starr, *Politics of Numbers*, pp. 363–91.

45. Ibid.

46. Ibid., p. 368.

47. Ibid.; Metzenbaum, “Strategies for Using State Information.”

48. National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform* (GPO, 1983 [www.ed.gov/pubs/NatAtRisk]).

49. David Grissmer and others, *Improving Student Achievement: What State NAEP Test Scores Tell Us* (Santa Monica, Calif.: Rand, 2000), p. 21.

50. One of the governors leading the 1984 charge, Lamar Alexander, became Bush’s secretary of education in 1991. Another, Bill Clinton, was elected president in 1992 and appointed a third governor who had been involved in the 1984 effort, Richard Riley, to be his education secretary. All three urged federal support to help states measure their performance, and all three urged the adoption of voluntary national standards. The

Improving America's Schools Act (P.L. 103-382), reauthorizing the 1965 Elementary and Secondary Education Act, linked Title I funding to the creation and content of state performance standards in the 1994 reauthorization of federal education law. For a more detailed discussion of the governors' role, see Maris A. Vinoskis, "The Road to Charlottesville: The 1989 Education Summit," National Educational Goals Panel, September 1999 (<http://govinfo.library.unt.edu/negp/reports/negp30.pdf>).

51. Andrew Rudalevige, "No Child Left Behind: Forging a Congressional Compromise," in *No Child Left Behind? The Politics and Practice of School Accountability*, edited by Paul E. Peterson and Martin R. West (Brookings, 2003), pp. 23–54.

52. As Rudalevige ("No Child Left Behind," pp. 34–42) explains, by the time the 1994 education law was up for reauthorization in 2000, educational performance had become a priority for yet another governor, George W. Bush of Texas. In Texas, Bush built on his predecessor's policies to test all students in reading and math in grades 3 to 8 and require that they pass state-issued tests to graduate. Educational accountability through measurement and other means became a campaign theme when Bush ran for president, one described in a campaign position paper that adopted as a goal the mission of the Children's Defense Fund, "to leave no child behind." The new president made action on the education campaign theme among his first postelection legislative moves.

53. In August 2005, the attorney general of the state of Connecticut sued the U.S. government for violating the Unfunded Mandates Reform Act, arguing that the No Child Left Behind law imposed large additional testing costs in violation of federal law (Sam Dillon, "Connecticut Sues the U.S. over School Testing," *New York Times*, August 23, 2005 [[www.nytimes.com/2005/08/22/nyregion/22cnd-child.html?st=cse&tsq=Blumenthal+education&scp=5](http://www.nytimes.com/2005/08/22/nyregion/22cnd-child.html?st=cse&tsq=Blumenthal+education&scp=5)]).

54. Regarding dropout rates, see "Governors Sign Compact on High School Graduation Rate at Annual Meeting," news release, National Governors Association, July 17, 2005. Regarding the governors' support for continuing strong federal accountability authority in the NCLB, see Alyson Klein, "Governors Enter Fray over NCLB; State Chiefs, Boards Join Plan for Revisions to Law," *Education Week*, April 10, 2007. Klein quotes Mary Kusler, assistant director of government relations for the American Association of School Administrators representing district superintendents, as observing, "It seems strange to me that three state organizations did not harp on the whole federalism issue."

55. Metzenbaum, "Strategies for Using State Information," pp. 34–42. See also the Education Trust website ([www2.edtrust.org/edtrust](http://www2.edtrust.org/edtrust)). As of 2006, the Department of Education's database of school-level performance scores was maintained under contract to the Department of Education by the American Institutes for Research.

56. See, for example, Institute of Educational Sciences, "What Works Clearinghouse" (<http://ies.ed.gov/ncee/wwc/>) and "State Comparisons" (<http://nces.ed.gov/nationsreportcard/nde/statecomp/>).

57. See, for example, *Personal Responsibility and Work Opportunity Reconciliation Act*, P.L. 104-193, sec. 103, amending Part A of title IV, with sec. 417, limiting federal authority to regulate the states except as expressly required in the new law.

58. See U.S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau, *Child Health USA 2006*, 2006 ([ftp://ftp.hrsa.gov/mchb/chusa\\_06/c06.pdf](ftp://ftp.hrsa.gov/mchb/chusa_06/c06.pdf)).

59. William T. Gormley and David L. Weimar, *Organizational Report Cards* (Harvard University Press, 1999).

60. Salamon, "New Governance and the Tools of Public Action," p. 17.

61. "Analytical Perspectives," in *Budget of the United States Government Fiscal Year 2007*, p. 108, table 8-3.

62. Paul L. Posner, "Federal Grant Design: What Washington Should Know and Why It Should Know It," paper prepared for the Ninety-Fourth Annual Meeting of the American Political Science Association, Boston, September 3–6, 1998, p. 4–5.

63. DeNeufville, "Federal Statistics in Local Governments," p. 344.

64. Derthick, *Influence of Federal Grants*, pp. 7, 201; Helen Ingram, "Policy Implementation through Bargaining: The Case of Federal Grants-in-Aid," *Public Policy* 25 (Fall 1977): 499–526; Paul E. Peterson, Barry G. Rabe, and Kenneth K. Wong, *When Federalism Works* (Brookings, 1986).

65. Derthick, *Influence of Federal Grants*; Peterson, Rabe, and Wong, *When Federalism Works*, p. 90; Richard P. Nathan, "State and Local Governments under Federal Grants: Toward a Predictive Theory," *Political Science Quarterly* 98 (Spring 1983): 47–57.

66. Derthick, *Influence of Federal Grants*, pp. 201–02; John W. Kingdon, *Agendas, Alternatives, and Public Policy*, 2nd ed. (New York: Longman, 1995).

67. See, for example, Peterson, Rabe, and Wong, *When Federalism Works*; Derthick, *Influence of Federal Grants*.

68. Peterson, Rabe, and Wong, *When Federalism Works*; see especially the discussion of Baltimore's use of community development block grant funds; also see Derthick, *Influence of Federal Grants*. If the block grant has a large matching requirement that increases the overview of local legislators or is in a state where the legislature must approve the receipt of a grant, it would garner considerably less local support.

69. Metzenbaum, "Strategies for Using State Information," p. 24.

70. Michael Hardy, "SAS Undergoes Identity Shift," *Federal Computer Week*, March 20, 2006.

71. Peterson, Rabe, and Wong, *When Federalism Works*.

72. Derthick (*Influence of Federal Grants*, p. 7) suggests that grants work in two ways: to "induce other levels of government to engage in [a specified] function or if they are already doing it, to do more of it" and by attaching "conditions that accompany the grants."

73. Congress adds some conditions to all grants, such as promotion of small business, to advance wholly new objectives. Other conditions seek to prevent the recurrence of problems that previously occurred. Local goal allies often pursue numerous federal grants to fund single programs or projects and must meet the conditions, often incompatible, set for each of the grants. Block grants are created to address this problem. As organized advocates for the narrower objectives funded by categorical grants lose funds locally and as

problems gain attention on the political action agenda, conditions get added to popular block grants to compel state and local attention to their concerns.

74. Derthick, *Influence of Federal Grants*, p. 195. For a good discussion of the reasons for grant conditions and interpretations of them, see Conlan, "Grants Management in the Twenty-First Century," and Derthick, *Influence of Federal Grants*, pp. 193–218.

75. See discussion of vocational education in Peterson, Rabe, and Wong, *When Federalism Works*.

76. For example, in its early GPRA reports, the EPA reported the number of completed water plans but paid more attention to confirming data submission than using the data it gathered. It did not report water quality conditions, despite having collected biennial water quality–monitoring data from every state. Nor did it identify which water bodies were getting cleaner and which were getting dirtier, despite the fact that states collected the data and the law sets a clear goal of zero discharge to the navigable waters of the United States. That analysis might have illuminated unknown problems and possible solutions or shown trends in pollutants discharged. Following the GPRA and PART, and with development in information technology, the EPA has begun the transition to measuring water quality conditions and discharge trends.

77. Federal laws and the federal budget office like to make sure grantees do not use federal money as a substitute for existing local spending (Posner, "Federal Grant Design"; Conlan, "Grants Management in the Twenty-First Century").

78. Peterson, Rabe, and Wong, *When Federalism Works*, pp. 140–42.

79. Shelley H. Metzenbaum, "Making Measurement Matter: The Challenge and Promise of Building a Performance-Focused Environmental Protection System," CPM 98-2 (Brookings, October 1998), p. 29.

80. On problems in mature grant programs, see David R. Beam and Timothy J. Conlan, "Grant," in Salamon, *Tools of Government*, p. 351.

81. Harry P. Hatry and others, "How Federal Programs Use Outcome Information: Opportunities for Federal Managers," in *Managing for Results 2005*, edited by John M. Kamensky and Albert Morales (Lanham, Md.: Rowman and Littlefield, 2005), pp.197–274.

82. Posner, "Federal Grant Design," p. 6.

83. Before 1992, federal law required states receiving federal alcohol, mental health, and substance abuse funding to prohibit the sale of tobacco to anyone under the age of eighteen, but few states enforced the law. The 1992 Synar amendment to the Public Health Service Act requires states to conduct random unannounced inspections to measure vendor compliance with the law and report compliance rates annually to the federal government. A state that fails to measure according to federal requirements can lose up to 40 percent of its mental health and substance abuse grant; the Center Substance Abuse and Mental Health Services Administration keeps track of compliance rates by state ([www.prevention.samhsa.gov/tobacco/01synartable.aspx](http://www.prevention.samhsa.gov/tobacco/01synartable.aspx)). See also Jim Hood, Mississippi attorney general, "Cigarette Sales in the U.S. Reach Historic 55-Year Low," press release, March 20, 2006 ([www.ago.state.ms.us/pressreleases/cigsalesdown.pdf](http://www.ago.state.ms.us/pressreleases/cigsalesdown.pdf)).

84. See “Cancer Trends Progress Report,” National Cancer Institute, January 7, 2008 ([http://progressreport.cancer.gov/doc\\_detail.asp?pid=1&did=2007&chid=71&coid=702&mid=#trends](http://progressreport.cancer.gov/doc_detail.asp?pid=1&did=2007&chid=71&coid=702&mid=#trends)). Tobacco sales control is one of several intervention strategies the government has encouraged to slow smoking among minors.

85. Metzenbaum, “Strategies for Using State Information,” p. 25.

86. *Ibid.*, p. 29.

87. *Ibid.*

88. Insurance Institute for Highway Safety, “Helmet Use Laws: March 2008” ([www.iihs.org/laws/state\\_laws/helmet\\_use.html](http://www.iihs.org/laws/state_laws/helmet_use.html)).

89. National Highway Traffic Safety Administration, *Motor Vehicle Traffic Crash Fatality Counts*.

90. Metzenbaum, “Strategies for Using State Information,” p. 29.

91. Ian Ayres and John Braithwaite, *Responsive Regulation: Transcending the Deregulation Debate* (Oxford University Press, 1992).

92. Harry and others, “How Federal Programs Use Outcome Information.”

93. This accountability principle is explored more fully in Metzenbaum, “Performance Accountability.” The quotation was attributed to Bratton in a slide presentation made by a police chief during a performance management training session the author ran. Bratton has since been contacted about this phrase being attributed to him and had no objection to it.

94. James Q. Wilson discusses this political phenomenon in *Political Organization* (Basic Books, 1973), pp. 327–46. See also Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Harvard University Press, 1965).

95. Metzenbaum, “Performance Accountability.”

96. On resistance to outcome-focused measurement, see Weiss and Gruber, “Managed Irrelevance of Federal Statistics,” pp. 374–76. On “cream skimming,” see Carolyn J. Heinrich, “Setting Performance Targets: Lessons from the Workforce Investment Act System,” in Kamensky and Morales, *Managing for Results 2005*, pp. 351–74, 370.