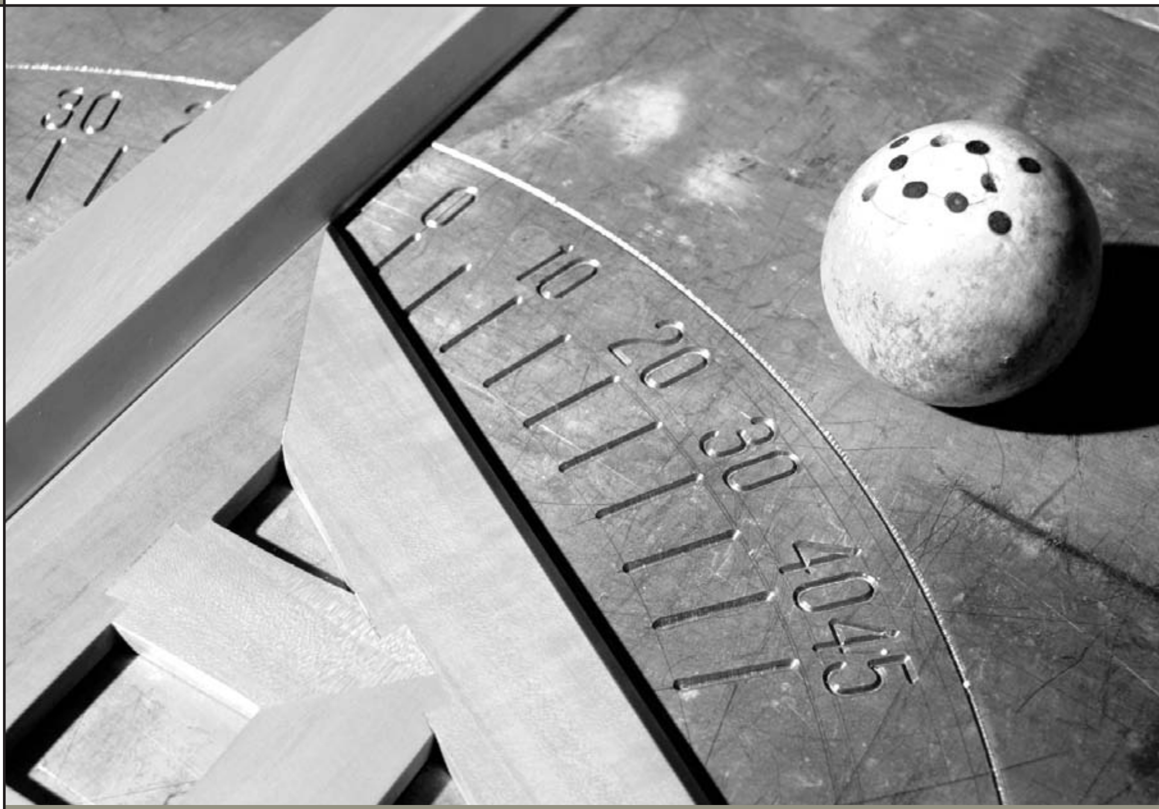
2006

Performance Accountability:
The Five Building Blocks and Six Essential Practices

Managing for Performance and Results Series



Shelley H. Metzenbaum
Visiting Professor
School of Public Policy
University of Maryland

IBM Center for
The Business of Government

MANAGING FOR PERFORMANCE AND RESULTS SERIES

Performance Accountability: The Five Building Blocks and Six Essential Practices

Shelley H. Metzenbaum
Visiting Professor
School of Public Policy
University of Maryland

IBM Center for
The Business
of Government

TABLE OF CONTENTS

FOREWORD

On behalf of the IBM Center for The Business of Government, we are pleased to present this report, "Performance Accountability: The Five Building Blocks and Six Essential Practices," by Shelley Metzenbaum.

Albert Morales

Performance measures are powerful management tools for creating both good and bad behaviors. Dr. Metzenbaum explores each of these dimensions and offers insights—based on real examples—on how to increase the good behavior and minimize the bad. For example, she describes how an overemphasis on the use of performance measures to enforce accountability in budgeting and pay systems can lead to weaker measures and reduced performance.

She does not, however, recommend not setting goals or measuring progress toward organizational goals. "Quite the contrary!" she says. "Goals and measures are among the most important tools public sector organizations can use to enhance both performance and accountability."

Debra Cammer Hines

Metzenbaum sets forth five building blocks—tools and techniques for constructing a good measurement system for an organization. And she describes six practices that leaders need to use to make appropriately designed systems work properly.

This report will be of interest to policy makers as well as managers. Policy makers need to understand the implications of the systems they design, such as performance-based pay tied to organizational targets. Managers need to understand that what they choose to do or not to do also has implications for the successful use of performance measures. Metzenbaum emphasizes the importance of constructive feedback and the use of constant back-and-forth discussions in improving performance based on metrics. "What is needed," she says, "is a performance management approach that is outcome focused, measurement rich, and inquisitive but not punitive." A simple statement for a tough task.

We hope that this report will be a useful resource for both policy makers and public managers at all levels of government across the nation as they continue the transformation of government to become more results oriented.

Albert Morales
Managing Partner
IBM Center for The Business of Government
albert.morales@us.ibm.com

Debra Cammer Hines
Partner and Vice President
Public Sector Financial Management
IBM Global Business Services
debra.cammer@us.ibm.com

EXECUTIVE SUMMARY

Governments around the world are increasingly adopting goals and measures as a way to improve societal outcomes and enhance accountability. To improve societal outcomes, governments seek to "grow the good"—the health, safety, well-being, and general quality of life. At the same time, they aim to "slow the bad"—harmful or unhealthy events, risk-raising causal factors, unnecessary costs, wasted time, fraud, corruption, and incivility. To strengthen government accountability, governments seek to meet four distinct categories of accountability expectations: fiscal, ethical, democratic, and performance.

Four Categories of Accountability

Fiscal accountability: Government spends its money as authorized, with as little waste as possible.

Ethical accountability: Government agencies operate honestly, without conflict of interest, self-dealing, other forms of fraud, or abuse of the power of governmental authority.

Democratic accountability: Government agencies do what their citizens want and need, engaging citizens and their elected representatives in understanding trade-offs and making well-informed choices among competing priorities. Government agencies treat people civilly and courteously, unless there are strong justifications not to, so people do not resent or resist government because it has acted in a rude, slow, or inappropriate manner.

Performance accountability: Government agencies and their employees work intelligently and diligently to deliver effective and cost-efficient government programs.

Experience has shown that goals and measures can be remarkably effective performance-driving, accountability-enhancing tools. But experience has also shown that they can provoke self-protective responses that interfere with performance and accountability gains: timid targets, measurement manipulation, measurement elimination, outcome avoidance (resulting in an affinity for output targets), and claim games where some rush to claim credit for accomplishments while others run from it, fearful of provoking resentment among their peers.

These performance-dampening responses most often arise when goals and measures are explicitly linked to externally provided (extrinsic) incentives, whether positive incentives in the form of rewards or negative ones in the form of penalties or poor public scores. When incentives are inappropriately linked to measurement of performance results or the wrong kinds of incentives are chosen, performance management systems tend to backfire, discouraging workers and even motivating them to cheat.

Indeed, an overwhelming body of research and experience suggests that promising rewards to individuals in government agencies seldom works when the rewards are linked to attainment of specific targets, progress relative to peers, progress relative to the past, or per unit of product. Somewhat surprisingly, in most cases government should resist the temptation to offer explicit rewards to individual employees for meeting or even making progress toward specific targets, except when employees are free to opt in and opt out of contests to attain new performance levels without fear of penalty or embarrassment for non-participation.

Similar caution needs to be exercised when using performance incentives for organizations. Extrinsic incentives can work, but they can also backfire, depressing both performance and accountability. They can rob goals and measures of their ability to stimulate the kind of effort and innovation that results in continual, sometimes dramatic, improvements in societal conditions. And, they easily provoke unproductive fear that interferes with improvement efforts, especially when accountability expectations are left vague.

This is not to suggest that organizations, their managers, and those who work for them should not set goals or measure progress toward the goals. Quite the contrary! Goals and measures are among the most important tools public sector organizations can use to enhance both performance and accountability.

What it does suggest is that less attention should be paid to incentives and far more to ensuring the active and effective use of outcome-focused goals and measures. Instincts to link rewards (and penalties) explicitly to goal attainment or comparative standing, whether through performance-based pay, budgets, or other extrinsic bounty, should be resisted. Linking rewards and penalties to goal attainment is not only an ineffective motivator but also an unfair and infuriating one when individuals or organizations lack the skills, resources, or authority to meet (or make progress toward) their targets and the means to secure those inputs. Indeed, some missed targets are inevitable in healthy, discovery-provoking, risk-tolerating enterprises. If individuals and organizations met all of their targets all the time, it would suggest that they had chosen timid targets and missed the performance-driving power of a stretch target.

What it also suggests is a need for public organizations to clarify accountability expectations both with those being held accountable and with those holding them to account, including supervisors, legislators, budget offices, grant-giving organizations, delivery partners, and the public. Specifically, public organizations and their managers should be held accountable for six essential practices:

- Emphasizing outcomes, using specific targets
- Measurement mastery
- Delivering feedback
- Assuring an ongoing venue for interactive inquiry
- Cogent strategies
- Implementation

To achieve both accountability and performance gains, public organizations and their managers need to tap the power and respect the limits of the five basic building blocks of performance management:

- Clear, measurable goals
- Measurement to motivate, illuminate, and communicate
- Verbal feedback to unleash the power of goals and measures
- Interactive inquiry
- Cautious use of externally provided incentives

The first four of the six essential practices characterize how the first four of the five basic building blocks should be used. The fifth and sixth essential practices, cogent strategies and implementation, are critical to converting the first four practices into improved societal outcomes.

Cogent strategies and implementation can also include the use of the fifth building block, incentives. Great caution must be exercised, however, when extrinsic incentives are employed as part of an implementation strategy lest they drive out intrinsic motivators. It is often preferable to let goals, measures, feedback, and interactive inquiry work on their own to drive performance goals without burdening them with explicit links to the promise of reward or the threat of penalty.

The Five Building Blocks of Performance Accountability

Clear, Measurable Goals

Goals motivate and drive performance improvement because they focus, energize, encourage persistence, and stimulate discovery not only for individuals but also for organizations. They have this effect even without an explicit link to incentives. To drive significant performance gains, goals need to be specific and challenging. To improve societal outcomes, goals also need to be outcome focused wherever possible, except in instances where resources or skills are inadequate to meet the goal, in which case near-term learning and capacity-building goals are appropriate. Where the positive effect of agency activities on outcomes has been clearly established, short-term output targets can be established but should be accompanied by longer-term outcome goals. A few specific goals should be challenging. At the same time, it can be helpful to set other targets to communicate areas where continual progress is sought, steady state is acceptable, and slippage will be tolerated. Group goals should be set when cooperation is needed. Goals do not need to be linked to incentives to be effective. They do, however, need to used, which is why feedback is so essential. Also, goals work best if used in a constructive, not a critical, manner.

Those holding government to account should expect government agencies to adopt specific targets that make clear to the public what government is trying to do (including adoption of a few stretch goals that are challenging but realistic) and what it is not trying to do. They should not blame or penalize agencies for missed targets, provided experience-informed, cogent strategies have been developed and implemented to meet the goals.

Measurement to Motivate, Illuminate, and Communicate

Measurement motivates, illuminates, communicates, and informs choice. Even without an incentive, measurement motivates both organizations and people because people enjoy a sense of accomplishment and take pride in a job well done. Measurement illuminates by identifying those with better performance who provide a path forward for those wise enough to learn from the experience of others. Measurement communicates what works and what doesn't, speeding both the uptake of effective practices and the discard of ineffective ones. Measurement also informs both electoral and consumption choices, serving as a shorthand language facilitating communication between government and both its citizens and consumers of government services.

Studying measurement patterns, changes, anomalies, and relationships reveals problems needing attention and program successes worthy of replication. Many useful measurements are not indicators of agency performance, per se. They simply serve to trigger focused follow-up questions that lead to the detection of the underlying causes of problems and progress. Unfortunately, measurement is often primarily used to answer the compliance-oriented question: "Did an agency or program meet its target?" It is far more constructive and effective when measurement is used primarily to answer performance-improving questions, such as: "What works and is worth replicating?" and "What does not, that needs attention?"

Verbal Feedback to Unleash the Power of Goals and Measures

Essential to effective performance management is not only quantitative feedback in the form of measurement but also verbal feedback. Well-delivered verbal feedback boosts confidence that a goal can be met, stimulates ideas and specific plans about how to meet it, and reinforces the importance of specific goals.

Interactive Inquiry

When cooperation among many parties is needed to meet a goal or when relevant expertise exists outside an organizational unit, a forum that facilitates frequent interactive inquiry enriches the performance-improving power of goals and measures. These meetings provide an efficient and effective forum for easing cross-organizational cooperation, delivering feedback when more than one person has information relevant to goal attainment, and obtaining quick answers and advice from managers, peers, and others in the organization. For these meetings to work, without deteriorating into show-and-tell sessions or stifling open and honest assessments, they should lead with questions and not with answers, avoid blame, and push for better understanding of conditions and causes, as well as specific action plans.

Cautious Use of Externally Provided Incentives

Extrinsic incentives in the form of externally promised rewards and threatened penalties can motivate, but they can also discourage and introduce unhealthy fears that compromise discoveries that lead to performance gains. Therefore, extrinsic incentives should be used only in a limited number of circumstances and applied with great care. A reluctant approach to the use of extrinsic incentives is advisable especially for individuals but also for organizations.

The Six Essential Practices for Improving Performance and Strengthening Accountability

The enormous potential problems that can arise when extrinsic incentives are inappropriately used underscore the importance of clarifying performance accountability expectations. Even when incentives are not explicitly linked to goal attainment or relative performance, problems often arise because those being measured fear such links will eventually be made. To use goals and measures in a way that improves societal results and strengthens democratic accountability, what is therefore needed in every public organization is a clear articulation and understanding of performance accountability expectations. Specifically, government organizations and their managers should be held accountable for six essential practices:

- **Emphasizing outcomes, using specific targets:** using specific outcome-focused goals or targets, a few of which are challenging; establishing specific targets when they have not been externally set; and communicating targets to the public.
- **Measurement mastery:** using measurements to gauge progress toward the targets; communicating measurements to those who can influence progress as well as to the public; and discovering what the measurements reveal by organizing and studying them to look for patterns, anomalies, changes, and relationships in the search for what works, what does not, causal connections, and areas where more understanding is needed.
- **Delivering feedback:** helping others in the organization to set outcome-focused targets (including a few ambitious ones), to believe in their own abilities to reach specific targets, to find specific ideas and practices that enable them to reach the targets, and to obtain needed skills, resources, and authority to meet targets or revise them to account for implementation obstacles.
- **Assuring an ongoing venue for interactive inquiry:** ensuring the existence of a venue that regularly engages others with expertise and resources relevant to the attainment of specific targets in providing feedback, stimulating synergistic thinking, sharing experience, planning and coordinating actions, assessing implementation efforts, and updating targets and action plans based on the best available evidence.
- **Cogent strategies:** developing cogent long-term strategies and shorter-term action plans based on the best available evidence and ideas.
- **Implementation:** implementing strategic and action plans and ensuring that insights from experience are fed back into the development of targets, strategies, and activity selection on a timely basis.

Will this prescription for performance management accountability expectations, informed by research findings, work in practice? Evidence from numerous government agencies suggest it can not only work, but it can work in a powerful way with great outcome and accountability returns. It is evolving in New York City, Baltimore, the state of Washington, the educational system in Tennessee, numerous United States federal agencies, and in the United Kingdom and New Zealand.

There are promising developments on the political front, as well. A small but increasing number of elected executives—for example, the mayors of Baltimore, Washington, D.C., and New York City—have boldly announced specific outcome-focused targets with deadlines, openly reported on progress and problems, and won re-election despite missed targets. And while most legislative decision makers at the federal level have ignored formal documents related to the Government Performance and Results Act (GPRA) and the Program Assessment Rating Tool (PART), many have in fact paid attention to agency goals and outcome measures when goals and measures are delivered in a format that is relevant to them.

What is needed is a performance management approach that is outcome focused, measurement

PART and Performance Accountability Expectations

The Program Assessment Rating Tool (PART) of the U.S. Office of Management and Budget appropriately gives federal agencies credit for many of the practices recommended here, including setting specific and challenging goals, emphasizing outcomes, and mastering measurement. Two aspects of the PART should be revised, however, to lessen the likelihood of dysfunctional agency responses:

- OMB should not score programs low for not meeting targets when programs have set challenging, outcome-focused targets, measured progress, and implemented what seemed like a sensible strategy at the time the target was set.
- OMB should not score programs low for not meeting targets they cannot control because of legislative barriers, even when an agency has proposed corrective legislation to the White House.

rich, and inquisitive but not punitive. When agencies adopt the six essential practices described here, outcomes improve and accountability increases. Outcomes rise because goals focus and motivate, measurement reveals what works and what doesn't, and feedback and interactive inquiry inspire, inform, and engage. Accountability (and democracy) increase because public articulation of specific goals clarifies what each organization will do and what it will not do, allowing citizens and their elected representatives to determine if the organization is doing what they want it to do and inviting them to use their electoral and administrative process voices to respond if they disagree. Outcomes and accountability also rise because goals, measurement, strategy transparency, and interactive inquiry encourage intelligent, honest, and diligent efforts.

Introduction

Governments in the United States and all over the world have adopted laws and directives requiring agencies to set goals and measure performance. These directives have at least two purposes. They seek to improve societal outcomes—growing the good (health, safety, well-being, and general quality of life) while slowing the bad (harmful or unhealthy events, risk-raising causal factors, unnecessary costs, wasted time, fraud, corruption, and incivility). They also seek to strengthen government accountability.

There is general agreement about what it means to improve societal outcomes, although differences in values and factual uncertainty often spur debate about what constitutes a better outcome. In contrast, a general notion of what it means to improve accountability is less well understood.

What Is Accountability?

What exactly is this concept of accountability? What does it mean to "hold someone accountable"? What does it mean for government agencies and employees to be answerable to someone and for what and to whom do they need to account? Presumably, it means in part the desire of citizens and their elected officials to be able to identify who is responsible for an organization's outputs or outcomes and for its successes and failures. But then what? When people talk about holding someone or some organization accountable, what happens? One public management expert has proffered a somewhat tongue-in-cheek answer:

> I know of no definitive answer, either theoretical or empirical. But I bet I know what the managers who are to be held accountable think. I bet they believe, from their own empirical experience, that "holding people accountable" means that they when they fail they are punished and that when they succeed nothing significant happens.[1]

Unclear accountability expectations—who is accountable to whom for what and what consequences arise when accountability expectations are not met—are problematic because they introduce fear into performance management, which is the use of goals and measures to manage. That fear, in turn, creates problems such as measurement manipulation, timid targets, outcome avoidance (resulting in an affinity for output targets), and claim games where some rush to claim credit for accomplishments while others run from it, fearful of provoking resentment among their peers. Occasionally, measurement systems even implode, seemingly overburdened by their own weight.

These problems arise for three primary reasons: vague accountability expectations, inadequate feedback and inquiry to probe the insights revealed by performance measures, and misguided notions of how and when to use incentives. Past experience and research suggests that many of these problems can be averted, performance improved, and accountability strengthened, but only if agencies and their watchdogs adopt an inquisitive, non-punitive approach to performance management.

Goals and measures are among the most powerful performance-improving and accountability-enhancing tools government has at its disposal. Even without an explicit link to incentives, goals and measures drive behavioral change both in individuals and in organizations. Goals do this by serving as a focusing point and by influencing attitudes, effort, and creativity. Measures do this by reinforcing goals, by guiding the search for more effective intervention approaches, and by informing choices.

Goals and measures cannot serve this powerful function, however, unless used in an atmosphere not overwhelmed by fear of penalties or even unfair rewards. Such an atmosphere necessitates an attitude change in the way government uses goals and measures—one that is active, not passive; one that is constructive, not critical; one that recognizes goals and measures as a robust resource, not simply an obligation to be generated in response to mandates for performance plans and reports. When goals and measures are simply placed on paper but never used, they are useless. Feedback is essential to unleashing their power. Feedback can be useful when provided one-on-one or delivered collectively. When delivered in a way that stimulates ongoing exploration of measurement implications among those with expertise and the potential to contribute to outcome gains, it illuminates, invigorates, enlists, and ultimately improves societal outcomes.

The necessary attitude change starts with questions such as: "What is working and merits replication?" and "What is not working that needs attention?" These questions contrast starkly with the initial question more frequently asked by agency leaders and influential observers such as budget analysts and appropriators: "Did you make your target?"

The instinct to hold people and organizations accountable for meeting a target is strong, among both practitioners and academics. In 2001, for example, congressional appropriators cut bonuses for the National Highway Traffic Safety Administration when it failed to meet some of its goals and threatened to cut bonuses for the Federal Aviation Administration (FAA) and the Internal Revenue Service.[2] Even leading public management experts suggest that bonuses should be withheld when targets are not met:

> If agencies give top managers bonuses even when they don't meet GPRA targets, they are saying that other goals are more important than GPRA. With these decisions, the appropriations [panel] is saying otherwise—that GPRA is a key measure of agency success, and there will be no bonuses for managers that fail to meet GPRA targets.[3]

Goals and Targets

The terms *goals* and *targets* are used interchangeably in this report. Goals are sometimes seen as more general and targets as more specific. That distinction is not used consistently or rigorously here, in part because that distinction does not consistently appear in the English language and because those seeking to establish a common nomenclature for performance management (for example, the Governmental Accounting Standards Board), do not make that distinction. Individual organizations may want to agree on a common terminology, but the objective in this report is to get key concepts across, not to force a precision of language.

For goals and measures to realize their performance-improving, accountability-enhancing potential, though, neither individuals nor government organizations should be help strictly accountable for meeting all targets. What is needed instead is a performance management system that anticipates missed targets and occasional performance slippage. What is needed is a system that recognizes that missed targets arise when programs set the sort of ambitious targets that most effectively motivate performance gain and that factors beyond an agency's control sometimes drive performance downward.

This kind of approach can sustain a high level of accountability despite its tolerance for missed targets. It does this not by holding managers and their organizations accountable for target attainment or even for steady performance climbs, but by continually holding them accountable for six essential performance management practices:

- A relentless focus on improving outcomes with clear, outcome-focused targets, a few of which are challenging
- Measurement mastery
- Frequent feedback
- Interactive and ongoing inquiry to find what works and what doesn't
- Cogent strategies and action plans
- Implementation

When agencies adopt this sort of inquisitive, non-punitive approach to performance management,

> **Accountability Expectations**
>
> A statement attributed to former New York City Police Department Commissioner Bill Bratton captures this notion of accountability expectations: "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 for dealing with it."

outcomes improve, accountability rises, and democracy is strengthened.

Organization of This Report

This report explores and explains how agencies should use goals and measures to engage the intelligence, interest, and commitment of those in government, not terrify or discourage them. The report identifies five building blocks integral to sustainable, effective, accountable performance management:

- Clear, measurable goals
- Measurement to motivate, illuminate, and communicate
- Verbal feedback to unleash the power of goals and measures
- Interactive inquiry
- Cautious use of externally provided incentives

The report explores what theory and experience suggest about the best ways to use these five building blocks, distinguishing between practices likely to advance outcomes and accountability and those more likely to trigger performance-dampening, dysfunctional behavior. This report offers a new notion of performance accountability, one that is inquisitive and expects persistent questioning to find program successes worth replicating and program and societal failures needing attention. It is an accountability that leads with goal clarity, information, and analysis—and that encourages with insights and the opportunity to make a difference. It is also a non-punitive accountability, employing sanctions only in limited circumstances and as a last resort, recognizing that penalties and even rewards often do not affect motivation in the ways intended. When goals are specified inaccurately, the wrong kinds of incentives chosen, or incentives inappropriately linked to who and what is measured, the motivational intent of incentives tends to backfire. Extrinsic incentives can discourage workers and even prompt cheating. In limited circumstances, inducements, penalties, and rewards can strengthen the performance-driving, accountability-enhancing effect of goals and measures. It is often preferable, however, to let goals and measures work on their own without an explicit link to incentives, relying instead on intensified attention to feedback and interactive inquiry.

Before examining the five building blocks more carefully, the report begins with a brief discussion of the oft-stated but seldom-defined term "accountability" to clarify assumptions used in the report regarding accountability expectations.

Changing Expectations of Accountability

In 2004, the comptroller of what was then the U.S. General Accounting Office (GAO) changed the organization's name to the Government *Accountability* Office. The name change captured evolving accountability expectations taking place in governments across the United States and the world, expectations that include "better performance" and "strengthened democracy" along with fiscal and ethical integrity.

Historically, U.S. accountability laws, rules, and policies have focused on fiscal and ethical accountability, following the trail of government money to prevent "the politics of personal favoritism and gain from meddling in the administrative decisions about personnel, procurement, finance, and service delivery."[4] In changing the name, U.S. Comptroller General David Walker sought to emphasize a shift in organizational emphasis from fiscal and ethical issues to performance and democratic ones:

> After 83 years, the General Accounting Office has changed its name to the Government Accountability Office. Some might wonder why GAO felt a need to tinker with an institutional identity so strongly associated with government economy, efficiency, and effectiveness. But our old name, as familiar and reassuring as it was, had not kept pace with GAO's evolving role in government. The truth is that "accounting" has never been our chief mission.

> Stereotypes, however, can be hard to shake. Some college students we were trying to recruit mistakenly assumed that you needed an accounting degree to work at GAO. New members of Congress, cabinet-level officials, and prominent journalists have, because of our name, thought that GAO's main job was to keep the government's books. In fact, a recent crossword puzzle in the *Washington Post* asked for a three-letter term describing a GAO employee; the answer was "CPA."[5]

Acknowledging that "GAO primarily scrutinized government vouchers and receipts in its early years," Walker observed that by the millennium, fiscal audits constituted only about 15 percent of the GAO workload.[6] The name change at GAO was intended to reflect a shift that had already occurred, one focused on how effective government agencies were in running programs (performance accountability) and whether government programs addressed the needs of society (democratic accountability):

> Today, most GAO blue-cover reports go beyond the question of whether federal funds are being spent appropriately to ask whether federal programs and policies are meeting their objectives and the needs of society. GAO looks at the results that departments and agencies are getting with the taxpayer dollars they receive.[7]

That is not to suggest that performance and democratic accountability substitute for fiscal and ethical accountability. Instead, new accountability expectations have been added to the existing ones.

The transition captured by GAO's name change suggests that four distinct accountability expectations currently exist for modern government agencies in democratic systems:

- **Fiscal accountability:** Government agencies will spend money as authorized, with as little waste and as efficiently as possible.
- **Ethical accountability:** Government agencies will operate without conflict of interest, self-dealing, other forms of fraud, or abuse of the power of governmental authority.
- **Democratic accountability:** Government agencies will do what its citizens want and need, engaging citizens and their elected representatives in understanding trade-offs and making well-informed choices among competing priorities. Government agencies will also treat people civilly and courteously, unless there are strong justifications not to, so that people do not resent or resist government because it has acted in a rude, slow, or inappropriate manner.
- **Performance accountability:** Government agencies and their employees will work intelligently and diligently to deliver effective and cost-effective government programs.

These four distinct accountability expectations are the ones used in this report to explore how agencies can use goals and measures to improve outcomes and enhance accountability.

Building Block 1: Clear, Measurable Goals

Goals, cognitive scientists have confirmed, play a remarkably powerful, performance-driving function. President John F. Kennedy intuitively understood this when he invited the nation in 1961 to accept the goal of putting a man on the moon in a decade: "I believe that this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the moon and returning him safely to the earth."[8] Kennedy did not threaten penalties, nor did he promise incentives when announcing this ambitious objective. Still, by proposing the goal and asking Congress to fund it, he inspired and challenged the nation to meet it.

Simply stated, goals focus, energize, encourage persistence, and stimulate discovery. Setting a goal releases a remarkable intrinsic motivational force in people. Researchers have determined that goals affect performance through four mechanisms: a directive function, an energizing function, persistence, and indirectly by leading to the arousal, discovery, and/or use of task-relevant knowledge and strategies.[9] This finding has been demonstrated in numerous and repeated experiments across countries and even in work with brain-damaged patients.[10]

Goals function this way not only for individuals but also for organizations. Whether set by an executive, a legislative body, or the people of the organization, goals also drive performance improvement in groups.[11] It is often suggested that goals are more effective performance drivers when the team, group, or organization picks its own goals. In fact, group goal setting is not a necessity and in some cases can be a costly enterprise. Both controlled experiments and retrospective studies indicate that assigned goals can be just as effective as self-set goals, but only if those doing the assigning have the appropriate power and authority.[12]

Of course, goals do not *always* drive performance gains. To drive significant performance gains, goals need to be *specific* and *challenging,* not general or easy.[13] The performance-driving power of a goal can be further enhanced by making it public. The driving power of a goal is reduced, however, when goals are overly complex or when individual goals are set rather than group goals when cooperation is needed. Also, goals have little value if they are not routinely articulated and discussed by organizational leaders. Finally, the power of challenging goals is unlikely to be tapped when people fear they will be penalized for not meeting them. Instead, goal setters will tend to select less challenging goals, which have a lower performance-driving effect.

Specific, Challenging, Outcome-Focused Goals Drive Performance Improvement

Goal specificity can be achieved by selecting quantitative and qualitative characteristics that refine a broad goal into a specific target. Goals are often refined into specific targets by indicating characteristics such as quantity (whether absolute or relative), time, place, population, or industrial sector. Embarrassed by the Russians beating the U.S. into space with its Sputnik launch, Kennedy responded in 1961 not just by saying the United States needed to build a strong space program. He selected a target that specified time (decade), place (moon), demographic characteristic (man), and quantity (a man). Target quantification can be absolute (10,000 currently obese children in California will achieve and maintain a healthy weight for their size and age, 99.9 percent of people served by community water systems will be served by systems with no water-quality problems), relative to a standard (air quality in every U.S. community will meet National Ambient Air Quality Standards), relative to the past (seat belt use will be 20 percent higher nationwide than the baseline established 20 years earlier), or relative to others (our school will perform in the top quartile of the state). Specific targets energize more than general ones, because they not only direct attention but also narrow goals, making them feel more attainable.

Challenging goals drive performance improvement because, perhaps counter to intuition, goal commitment is strengthened when goals are difficult and, more intuitively, because people (and organizations) work harder when they are strongly committed to a goal.[14] Goal commitment is further strengthened when commitments are made public.[15] But what constitutes a challenging goal, a stretch target? Edwin Locke, one of academia's leading "goal experts," has suggested that a stretch target is one with only a 10 percent chance of attainment.[16] Consider the "failure rate" implications of that statement: Nine out of 10 times, effective challenging targets will not be met! An incentive system that penalizes individuals and organizations for failure to attain all of their targets will therefore primarily be punitive and undoubtedly detract from the quality of the work environment.

Clearly, for ambitious targets to drive performance improvement, they must be allowed to operate in a climate that anticipates missed targets. Missed targets must be viewed as an indication of a healthy, discovery-provoking, risk-tolerating enterprise. If individuals and organizations meet all of their targets all the time, it suggests the targets are too easy and therefore not likely to tap the performance-driving power of stretch targets.

More challenging goals trigger highly effective work attitudes for groups as well as for individuals, stimulating cooperation and innovation. Researchers in lab experiments have concluded: "Subjects in the groups with the tougher task, compared to ones with the easier assignment, put out more effort, took more time to discuss how they would work together and what procedures they would use, changed their individual and group plans more often, and showed more concern for the quality of their work."[17] Setting specific stretch goals unleashes the workforce's reservoir of skills, knowledge, and instincts, and propels sustained, deliberate, and recurring strategy refinement to meet a goal. A specific, challenging target encourages delivery agents to "draw from a repertoire of skills that they have used previously in related contexts, and … apply them to the present situation … and … engage in deliberate planning to develop strategies that will enable them to attain their goals…."[18] In contrast, timid or easily met targets encourage a culture of compliance and control rather than one of creativity, learning, and rising performance.

Specific, challenging targets drive performance whether they focus on individual tasks, outputs, or outcomes. Agencies frequently achieve significant improvements in processing times, for example, when they set maximum allowable or average

Goals

Goals drive performance improvement and enhance accountability even without a link to extrinsic incentives when they are:

- **Challenging,** but only for a few priority goals.
- **Specific,** to focus and communicate where breakthrough progress is desired, continual progress is sought, steady state is acceptable, and slippage will be tolerated.
- **Outcome-focused** wherever possible—except in instances where resources or skills are inadequate to meet the goal, in which case learning and capacity-building targets are appropriate, or where the outcome goal is unitary in nature, in which case milestones are appropriate.
- **Constructively used** by managers and others in the organization in a routine manner.

Group, rather than individual, goals should be set when cooperation is needed.

Case Study: National Highway Traffic Safety Administration Sets Specific, Challenging Outcome-Specific Focused Goals

After passage of the Government Performance and Results Act of 1993, the National Highway Traffic Safety Administration (NHTSA) and its parent organization, the U.S. Department of Transportation (DOT), set an ambitious goal of raising seat belt use to 85 percent by 2000.[19] It chose the goal because strong evidence indicated that increased belt use significantly reduces fatalities and accident severity. DOT recognized that to attain its target, it would need cooperation from states, localities, and automobile drivers, all of whom it could influence but none of whom it directly controlled. Yet because of the potentially high safety return on each federal dollar invested, DOT took the risk of pursuing this ambitious target dependent on so many uncontrollable factors. By 2000, the agency had reached a 71 percent use rate, up from 67 percent in 1967. In 2001, it attained a 73 percent use rate.[20] By FY 2005, a nationwide 80 percent seat belt use rate had been achieved.[21] It missed the 85 percent target it had set, but nonetheless achieved an unprecedented seat belt use rate.[22]

The increase in seat belt use was obviously achieved not just by setting the goal, which focused and encouraged persistence, but by prompting a variety of non-traditional agency actions to meet it. It triggered funding shifts across DOT programs, for example. In FY 2003, the Federal Highway Administration directed funds from its budget to be used for state incentive grants to help NHTSA woo state and local cooperation, since the federal government does not directly enforce seat belt laws; states and localities do.[23] Winning state and local cooperation was not enough to change outcomes, however. NHTSA also needed to assist its service delivery partners in understanding how to change local decisions and behavior. It tested and evaluated a seat belt use marketing campaign, then promoted its uptake, acting almost as a franchisor.

Case Study: Ohio Environmental Protection Agency Sets Specific, Challenging Response Time Goals

The Ohio Environmental Protection Agency found goal setting very valuable when it set a goal in 2000 to clear out all administrative enforcement actions that had been in the docket more than two years. By the end of the year, Ohio EPA reduced the number of backlogged cases, some of which had been in the system three to five years, from 110 to 29. By the end of the second year, even with new cases, the number of cases exceeding the two-year threshold dropped to five. At the same time, the average age of all cases on the administrative docket dropped from 475 days to 325 days.[24]

Case Study: Federal Emergency Management Agency Encounters Problems by Focusing on Tasks, Not Outcomes

Reluctance to deal with outcomes, the problems people face, proved a key weakness in the Federal Emergency Management Administration (FEMA) response to Hurricane Katrina. Even when it was clear that the state and local governments were overwhelmed by the magnitude of the problems caused by Katrina, FEMA leaders in the hurricane-ravaged area resisted whatever instincts they must have felt to fix the problems they saw. FEMA is organized into discrete task groups, so the agency instead responded only to specific task requests made by the states and localities. The "not on my task list" orientation resulted in absurd situations. For example, when the state requested cots for firefighters, according to one senior state official who went to help his colleagues in Louisiana, "We got the 500 cots, but that is all we got. We didn't get any pillows or blankets."[25]

response time targets for a variety of transactions, including correspondence, complaints, general assistance calls to 311 or 800 numbers, permit applications, and even enforcement follow-up.

While specific, challenging targets drive progress toward any kind of goal, at some stage government agencies need to adopt *outcome-focused targets*. When they fail to do so, they inevitably get caught on a treadmill managing what they *do,* rather than

Case Study: New York City Measurement Efforts Thrive When Attention Shifts to Outcomes

New York City was an early adopter of performance measurement, begun after the city's fiscal crisis of 1977. Some city agencies—such as the Sanitation Department, which hired observers to rate the cleanliness of each neighborhood—focused on outcomes. Most, however, did not. In 1994, when the New York City Police Department adopted an outcome-focused goal to reduce crime, and then successfully used fresh and frequent measurement to drive crime down 25 percent in two years, it demonstrated the power of managing with outcome-focused goals and measures (Kennedy School of Government, 2001). This approach, known as CompStat, quickly spread from police to other agencies, with a similar positive effect on outcomes. Today, all city departments are expected to use outcome-focused measurement and management, now known as CAPStat, which stands for Citywide Accountability Statistics (Smith and Grinker, 2003).

what they are trying to *accomplish*. Outcomes tend to fade from attention despite the fact that government's obsession with outputs or processes, rather than with outcomes, was a primary catalyst that led to the reform efforts requiring goal setting and performance measurement.

Goals Communicate

The focusing function of goals serves not just to energize, but to communicate. This is an especially beneficial attribute when cooperation across large organizations or organizational boundaries is needed to meet a goal. It is also useful as a means to engage the public in the value-based decision making that goal setting inevitably is. Specific goals serve as a form of shorthand that inexpensively and concisely communicate where the people of an organization should concentrate their efforts and intelligence.[26] A goal serves as a "guide for directing members' actions and integrating their moves."[27] It also serves as a form of shorthand for communicating with those beyond organizational boundaries—to enlist assistance and expertise, compel cooperation, and invite those in a democracy to exercise their voice when they disagree with targets that have been chosen.[28]

The shorthand communication value of goals does not help an organization, however, if it has too many of them. Government agencies often find themselves overwhelmed because they are expected to do more than they can reasonably be expected to deliver. Successive legislative bodies and executives often articulate their priorities through law and policy directives, paying more attention to adopting something new than reconciling with or eliminating what was previously adopted. Even a single piece of legislation may teem with multiple objectives, the result of political compromise that fails to recognize (or chooses to ignore) the challenge it will present to implementing agencies. An abundance of goals, even if not inconsistent, overwhelm organizations, defeating the power of specific, challenging goals.

That is not to suggest that an organization cannot have multiple goals.[29] Effective organizations often have multiple goals in a single time period; they just cannot have too many *challenging* goals. All targets cannot be stretch targets.

Government organizations can consciously use targets as a shorthand communication tool to manage the problem of unreasonable expectations by broadcasting relative priorities. Agencies can specify stretch targets for priorities and where innovation and experimentation are sought, areas where steady progress is sought, areas where performance can remain level, and areas where some slippage will be tolerated. Agencies may want to set less ambitious targets, for example, when evidence suggests legal requirements no longer have a significant effect on improving outcomes but legislative bodies have not yet eliminated the requirement.

Goal specification in strategic and annual plans can, in other words, be used as a mechanism for managing the seemingly unending demand on government agencies to do everything with limited resources.[30] By inclusion and omission, the goals that an agency includes in its strategic and annual plans communicate to the people of an organization and to the public what an agency will do and what it will not do, strengthening democratic accountability.

At the same time, if the targets in strategic and annual plans are inconsistent, the communication value of goals is diminished. Thus, when the U.S. Department of Health and Human Services set a target in its FY 2004–2008 strategic plan of reducing the proportion of Americans who are obese by 50 percent in 10 years (from 30.9 percent in 1999–2000 to 15 percent by 2010)[31] but failed to include a specific annual target for obesity in its annual plan, it sent a confusing signal to the organization and missed an opportunity to drive performance improvement.[32] Similarly, if specific, challenging targets do not align with the topics being regularly discussed by management, the workforce tends to lose sight of the goals. Instead, it shifts attention to concerns managers are voicing.

Outcome-focused goals often increase the need for cross-organizational collaboration. The need to communicate the goals to the public with a plan to meet them often necessitates new inter-organizational arrangements. When the U.S. Department of Transportation set a specific target of cutting railroad crossings accidents in half within six years, for example, it spurred cooperation between the Federal Railway Administration and the Federal Highway Administration (FHWA) because FHWA controlled the incentives that could get states and localities to make the problem a priority.

Establishment and communication of a specific, challenging goal can also enlist goal allies and keep a problem that concerns the public on government's action agenda. The Kyoto Protocol, an amendment to a United Nations treaty on climate change negotiated (but not ratified) in 1997, established a specific target for each developed country to cut overall emissions of greenhouse gases by at least 5 percent below 1990 levels (by 2008 to 2012.) The Protocol illustrates the performance-driving power of a specific, challenging goal. Even without being legally enforceable, the goal's continued mention in the media, by other countries, and by organized political players is keeping the issue on the policy agenda, demanding attention even in countries that have chosen not to ratify the treaty (most notably the United States and Australia).

Goals Strengthen Democracy and Accountability

Agencies can use specific targets to inform the public about what they intend to do and what they intend not to do. This is an especially valuable tool when an agency believes it has more to do than its budget allows. Of course, the public and its elected representatives may not always agree with government's choice of specific targets. Numerous democratic mechanisms exist, such as legislative hearings and elections, which citizens and their elected representatives can use to express support for or displeasure with specific goals and suggest changes. Legislative hearings do not always afford the most conducive environment for discussing goal appropriateness or refinement, however, and agencies may find it hard to sort out political grandstanding from genuine concern at such highly staged events.

Unquestionably, those in government seeking to tap the power of goals and measures to improve outcomes fare better if they have thick political hides and anticipate the legislative, media, and advocacy predilection for criticism over praise.[33] Government agencies may be able to do more than develop a thick hide, however. An increasing body of evidence suggests that agencies can lessen beatings and enhance accountability by considering the interests of, informing, and engaging specific audiences. Few legislators, journalists, or advocates have either the experience or education in management methods likely to make them familiar with arguments for setting specific, challenging organizational goals ambitious enough to result in many missed

Case Study: Missed Targets Not a Problem for EPA on the Lower Charles River

The U.S. Environmental Protection Agency failed to meet an ambitious local goal that the Lower Charles River in Massachusetts would be swimmable in 10 years, a goal it set in 1995. Yet it encountered little criticism from the local community and media when it failed to make its 2005 target, most likely because of its sustained 10-year commitment to meeting the goal and the way it informed the community about its strategies, whether or not they worked, and its updated action plans.[34]

Case Study: Missed Targets a Problem for NHTSA Under PART

The National Highway Traffic Safety Administration got "dinged" by the Office of Management and Budget (scoring "Small Extent" rather than "Yes") in its Program Assessment Rating Tool, or PART, process for not making consistent progress toward its goal and for not meeting every annual target:

> Over the long-term, the program has demonstrated progress toward achieving its long-term goals, as the highway fatality rate has decreased from 1.75 fatalities per 100 million VMT [vehicle miles traveled] in 1992 to 1.50 in 2003. However, NHTSA has not shown significant progress and did not meet the targets over the past three years. For the past two years, the rate has stayed the same—1.5 highway fatalities per 100 million VMT.[35]

targets. Arguably, then, agencies could benefit if they try to discuss with legislators the motivational value of setting challenging goals and thus the implicit need to tolerate goal non-attainment in venues other than televised hearings.

Agencies may also want to create explicit mechanisms, such as open meetings, to discuss goal selection. Mayor Anthony Williams of Washington, D.C., launched the D.C. Scorecards, his goal-setting and performance-measurement effort, with a citywide forum of citizens to discuss city priorities. Used this way, specific goals not only drive performance improvement, but strengthen democracy and accountability as well.

One factor that threatens to tamp down U.S. federal agency adoption of ambitious targets is one aspect of OMB's annual Program Assessment Rating Tool (PART) scoring process, an integral part of the budget process. Most aspects of PART have had a positive effect, pushing agencies to adopt ambitious targets, including rewarding agencies that adopt ambitious long-term and annual targets with a "yes" score on two questions. PART penalizes the same agencies, however, when they do not attain all of those targets, inevitably frustrating agencies who genuinely appreciate the value of setting challenging, realistic targets. If not adjusted, the expectation that agencies will set ambitious targets and meet all of them or get penalized in the PART scoring process will undoubtedly prove confusing and frustrating to many agencies.

Cautionary Considerations

Specific, challenging goals are powerful tools that can drive performance improvement and enhance accountability, but only if used and only if individuals and agencies are not penalized so seriously for missed targets that they refuse to set ambitious, outcome-focused ones. Specific targets also have limited value if they are conflicting, confusing, or so numerous and unconnected that they are ignored. Researchers have identified several additional cautionary conditions that can interfere with the performance-driving potential of goal setting:

- A **lack of supervisory support** for goals undermines employee interest in them.[36] If organizations treat goals merely as words on paper, used in strategic and annual plans but never mentioned by managers, few in the organization will pay attention to them.
- Assigned goals can be as effective as self-selected goals, but not **if assigned tersely without explanation of purpose or rationale.**[37] Even for outcome-focused goals where the goal itself reveals the purpose, a rationale explaining why a specific target was chosen over others can enhance commitment to the goal.
- If a task is **overly ambitious and so complex that it makes people anxious and confused,** it is better to break the goal down into component parts or to set specific, challenging learning goals instead of outcome-focused ones.
- If a government program **lacks data to measure outcomes,** staff to analyze data, or contracts to hire needed expertise, then it can be more appropriate to set challenging initial goals to acquire those capacities than near-term outcome targets.
- While a few goals should be challenging, they must also be **realistic,** taking into consideration available resources, skills, and authority.[38]
- **When cooperation is needed** to accomplish a task, group goals are far more effective than

individual goals. Indeed, when cooperation is needed to do the job, individual goals prove inferior not just to group goals, but to no specific goal at all. Individual goals can work when combined with group goals, however, provided the objective is simple enough that group goals can accurately be broken down into individual goals.[39]

- If **goal commitment conflicts with workers' beliefs and prior behavioral practices,** cognitive dissonance sets in. Workers will try to restore consistency by changing either their beliefs and behavior or their goal commitment. Strong goal commitments can shift beliefs and behavioral practices.[40]

Summary and Implications

In sum, setting specific, challenging goals is a powerful, performance-driving tool except in instances where resources or skills are inadequate to meet the goal, in which case it is better to break the challenging goal into more manageable milestones, including near-term learning and capacity-building goals. In setting targets, care must be taken to provide clarity of rationale and purpose; avoid complexity; select targets that are reasonable in the context of available resources, skills, and authority; and resist individual goals when group cooperation is needed.

Specific goals are also an effective mechanism for enhancing democratic accountability. They start a conversation with citizens about whether or not an agency has adopted the right priorities.

Goals do not need to be linked to any incentive—rewards or punishments—to be effective. Indeed, linking incentives to extrinsic rewards or punishments can interfere with the performance-driving effects of a goal, a topic explored more fully in the section "Building Block 5: Cautious Use of Externally Provided Incentives." Specifically, if those setting targets fear the consequences of missing them, goal setters are more likely to select targets they can easily attain, forgoing the performance-driving potential of challenging goals.

To be useful, goals do, however, need to be used. And they need to be used in a constructive, not a critical, manner. Moreover, unused goals mislead the public and compromise one aspect of democratic accountability—that the public knows what government is trying to do. Therefore, managers should intentionally employ goals as a tool to stimulate performance gain. And legislators, budget offices, oversight agencies, and the media, in their quest for accountability, need to allow agencies to set and fail to meet specific, challenging, outcome-focused goals. If they do not and instead continually criticize agencies when goals are not met, it will interfere with the potential for goals to drive constructive change.

Specifically, to bring goals alive, they must be accompanied by feedback.[41] Managers need to ensure that workers are provided that feedback.

> For goals to be effective, people need summary feedback that reveals progress in relation to their goals. If they do not know how they are doing, it is difficult or impossible for them to adjust the level or direction of their effort or to adjust their performance strategies to match what the goal requires. If the goal is to cut down 30 trees in a day, people have no way to tell if they are on target unless they know how many trees have been cut. When people find they are below target, they normally increase their effort or try a new strategy.[42]

Without feedback, people and groups tend to believe past practices were more effective than they in fact were, prompting them to set overly ambitious goals and then invest in wasteful strategies to meet them.[43] Feedback comes in two forms which tend to be complementary: measurement and verbal feedback communication discussing progress toward a goal. These are the next two building blocks to which we turn our attention.

Building Block 2: Measurement to Motivate, Illuminate, and Communicate

Measurement is a quantitative form of feedback. Conversations about progress encourage people to pay attention to collected measurements and reinforce commitment to the goal. Measurement enables individuals and organizations to assess progress in relation to goals and adjust the level or direction of their effort as well as their performance strategies to match what the goal requires. Measurement is strengthened when reinforced by discussions exploring specific ways to advance toward a goal. Some goals such as landing a man on the moon in a decade may be so unitary in nature that measurement of progress toward the goal may not be appropriate. In those situations, milestone tracking, analysis of precursor and other contributory measurements, and discussions of progress toward the goal encourages persistence. Both individuals and organizations need feedback to tap the performance-driving power of goals. This section discusses how measurement improves outcomes and strengthens accountability; the next explores verbal feedback.

Measurement functions in four distinct ways to drive performance and enhance accountability. Measurement:

- Motivates
- Illuminates
- Communicates
- Informs individual choices

Measurement Motivates

People and organizations like to do well. Therefore, because measurement enables people to see how well they are doing and adjust their actions accordingly, measurement motivates them to work harder to *achieve specific goals* even without the explicit promise of reward or threat of punishment.

Measurement also motivates because measurement relative to the past and relative to peers can serve a goal-setting function. *Besting the prior year's accomplishments* can be seen as a dynamic mechanism that continually updates specific goals (although not always challenging ones) and therefore continually motivates because of the natural instinct to want to do better. In other words, even without a specific target, trend measurement sometimes plays a performance-driving role because past performance functions as a *de facto* goal. For both individuals and organizations, past performance motivates continual improvement for any trend that is both measured and discussed.[44] If a performance trend is measured but not discussed, however, and is one of so many measured outcomes that attention is not naturally directed to it, past performance is not likely to motivate performance improvement.

The motivational force of *peer performance* is somewhat more complicated. Sometimes, peer performance functions in a manner similar to past performance, with best performers setting de facto goals in the absence of explicit targets.[45] Other times, however, psychological reactions to comparison, competition, and relative position interfere with performance-improving aspects of measurement. Therefore, whether or not peer performance serves as a motivator is situational.[46] It depends on how those being compared react to comparison, the explicitness and competitiveness of the comparison, the culture of the organization, the political strength of those who fear faring poorly, and the fairness of the comparison. It also depends on links to rewards and penalties, a topic explored more fully in the incentive section of this report, starting on page 39. Therefore, great care must be exercised when using comparative measurement to motivate performance improvement.

The remainder of this discussion on the motivational value of measurement explores the numerous nuances that need to be understood to use comparative performance as an effective motivator both for individuals and for organizations.

The performance of individual leaders essentially sets a specific, challenging performance goal in an ongoing, dynamic manner. Goals set by peer leaders motivate some people, but peer comparisons discourage others.[47]

Even when comparative information is kept private (each person knows his or her individual score and performance relative to others, but individual information is not shared even among those being compared), comparison can depress the inclination of some individuals to do well.[48] Individual performance can actually decline with comparison.[49] To ensure a positive effect from comparison, the net effect of comparison would need to be determined for each situation.[50]

Another way to accommodate differential reactions to comparison on individual motivation (without having to calculate the net effect for each situation) would be to create opt-in comparative opportunities, letting people who are motivated by comparison self-select into it. A specific form of comparison, overt competition, not only helps some individuals thrive, it tends to lead to better performance by winners.[51] As in athletics, competition establishes higher and higher challenging targets for everyone, although only those who seek to be front-runners treat the very highest performance records as relevant.[52] Opt-in competitions would therefore, arguably, create a "trickle-up" effect.

People often opt into comparative situations through career selection. Many who select sales as a profession, for example, thrive on comparative environments.[53] In contrast, people who choose government as a career are less likely to be stimulated positively through comparison.[54] To engage people in government who thrive on comparison and to achieve the breakthrough performance levels competition can spur, government should experiment with opt-in opportunities for comparison and competition. This would avoid discouraging the many people in government deterred by comparison but capture the "trickle-up" effect of those motivated by it.

One form of comparison, *ranking,* can be particularly irksome. Rankings list compared parties in order, from best to worst or visa versa, rather than displaying results alphabetically or in some other manner. Ranking not only intensifies the distastefulness of comparison for those discouraged by it, it also introduces another problem because of the large number of individuals who overrate their own performance. People instinctively compare themselves with one another,[55] tend to overrate their own performance,[56] and get discouraged when given feedback that reveals their performance was worse than they believed. "Survey evidence consistently shows … that some 80 percent of us think we are better-than-average drivers, and that even more of us think of ourselves as more productive than the average worker."[57] In other words, people in everyone's hometown are like those in Garrison Keillor's Lake Wobegon, "where all the children are above average." Overestimating one's abilities is not a problem, per se, because people tend to perform better when they have confidence in their own abilities.[58] It becomes a problem if people get discouraged when they underperform their own self-expectations.[59] Since only about 50 percent of people can perform above average (or, more precisely, above median), a non-trivial number of people who consider themselves "above average" (about 30 percent of drivers, for instance) are likely to be surprised and disappointed by their poor relative ranking. And since disappointed people tend to reduce their work effort,[60] ranking everyone is likely to cause many workers to perform worse than they otherwise would have had they not been ranked.

To lessen the potential for discouragement, ranking is likely to be more effective if used primarily to identify the performance levels of leaders. Leaders can be chosen among those in a similar "pond" or "league," so those being measured see the leader's performance as relevant to their own.[61] At the same time, in choosing relevant leaders, caution needs to be exercised not to choose those too close, because negative reactions tend to intensify when people compare themselves to those who are more proximate.[62]

Another way to lessen the intensity and potentially discouraging effects of comparison on individuals while not losing its goal-setting, performance-driving value for everyone is to inform individuals privately about their own performance levels (but not their rank), group averages, and the performance levels of the leaders. This approach is currently being used in public schools in Tennessee. The state organizes and analyzes performance data for four discrete performance units: district, building (school), teacher, and student. Each teacher and his or her supervisor are provided data comparing the teacher's performance to the district and state average. To protect the privacy of other teachers, they are not given their performance relative to the school average. By law, no one but the teacher and the direct supervisor may obtain teacher performance information, although some teachers voluntarily share it.[63]

Organizational peer comparisons function somewhat as they do for individuals, but with some important distinctions because individual psyches are not so much at risk. As with individuals, some organizations see themselves as leaders and are spurred to higher heights because of comparison. Kennedy was spurred to set his "man on the moon in a decade" goal because Russia's leadership in space had embarrassed the United States. Similarly, comparative international statistics showing the relatively poor standing of the United States in international math and science tests relative to other developed countries successfully prodded policy and program changes.[64] Some localities, such as San Diego County in California, Fairfax County in Virginia, Montgomery County in Maryland, and the City of Charlotte in North Carolina, similarly see themselves as leaders and strive for national leadership recognition.

Whether or not organizations have a "drop-back," performance-depressing effect from comparison similar to that evident in individuals is less clear. It is hard to imagine a private, for-profit organization overtly resisting comparison, because comparison and competition is the conceptual premise of a free market economy. It is less difficult to imagine an aversion to comparison among not-for-profit or governmental organizations, which see themselves as noncompetitive almost by definition. Certainly, state and local agencies frequently espouse that aversion, except in the economic development arena.

What is less obvious is whether the espoused resistance to organizational comparison is accompanied by the sort of performance-dampening effect evident in individuals. Reportedly, in some cases, poor relative performance triggers the opposite reaction: self-generated, performance-improving efforts. Legislators in states ranked at the "back of the pack" in *Governing* magazine's Government Performance Project have used their low scores as ammunition to push legislative changes in their states.[65] Similarly, external advocates routinely use rankings, such as biggest polluters, highest teen pregnancy levels, or worst education scores, to press for (and sometimes win) needed changes in specific areas.[66] Whether or not some states get so discouraged that they drop out, as is the case with some individuals, is less clear.

Multidimensional ranking, indexing, or grading systems can be problematic motivators for both individuals and organizations if performance for each of the underlying factors that constitute the basis for comparison are not distinctly reported and easily accessed. Without that specificity, multidimensional rankings send confusing signals about how those being measured can improve their performance. Moreover, when made public, they simply embarrass poorly ranked parties without providing a clear path forward for improvement.

Unless the factors used in an index are based on research or compelling logic identifying the precise combination of past practices that contributed to successful outcomes, multidimensional ranking or

Case Study: NHTSA and the Federal Highway Administration Organize, Analyze, and Disseminate Data

The National Highway Traffic Safety Administration and the Federal Highway Administration have long gathered and published, on an annual basis, reams of tables showing *comparative state information*. The tables provide state data on different outcomes, such as accident rates involving bicycles and other pedalcycles, as well as important characteristics believed to affect outcomes, such as population levels and the number of vehicle miles traveled. The tables neither rank states nor try to combine multiple outcomes into a single outcome indicator. They do, however, support performance improvements thanks to an organizational emphasis on the illumination power, rather than motivational value, of measurement, a subject explored more fully in the next section.

grading systems are simply subjective. They will not necessarily improve performance even if they successfully stimulate copycat behavior by those seeking to fare better the next time around. Multidimensional comparison can be more effective and fair if rated individuals or organizations can learn precisely why they scored as they did, with comparative information on each factor that made up the combined indicator.

In sum, measurement can motivate even without incentives. When a goal exists, measurement functions as feedback that can inform specific future actions to reach the goal. When goals do not exist, measurement can also motivate by setting *de facto* “better than the past” goals. Comparative measurement can motivate as well, but can also have offsetting effects among those discouraged by or resentful of comparison. To lessen the prospects from problems associated with comparison, several specific alternatives have been suggested:

- Opt-in public comparative opportunities, especially for individuals, may motivate more performance gains than involuntary public comparison, which can trigger self-protective, unproductive responses.
- Ranking is useful for identifying individual performance leaders as benchmarks for study, for illuminating a path forward for others, and for identifying those needing the most assistance. Hiding the identity of those being ranked, even for those who rank highest except when they have agreed to have their identity revealed, lessens comparative tensions.
- Multidimensional comparisons of individuals and organizations should reveal, privately for individuals, how each performer fared on each dimension of performance used as the basis for comparison. This is essential for the multidimensional comparisons to motivate improvement.

Comparative measurement, ranking, and grading can create other problems, as well, and caution must be taken to avoid them. Comparative measurements irritate those being compared when they feel the process and criteria used as the basis for comparison are not fair.[67] For everyone, whether individuals or organizations, comparative measurement is irritating rather than performance-improving when measurements chosen as the basis for comparison are inaccurate. Great care needs to be taken to use accurate information not just for public comparisons, but to guide decisions about which programs (or providers) to continue and which to curtail. For goals and measures to improve performance and enhance accountability, government agencies need to make data accuracy a goal, and then invest in, measure, and publicly report on progress toward the data accuracy goal. When setting the goal, however, 100 percent accurate data is seldom needed. The data accuracy goal should support the intended use of the data. Data for illumination purposes, used to trigger focused follow-up, does not need to be as accurate (and therefore as costly to generate) as data used as the basis for reward or penalty.

Measurement Illuminates

As has been suggested in the previous section, measurement drives performance improvement not just by awakening intrinsic instincts to do better, but by illuminating a performance-improving path forward, shining a spotlight on problems needing attention, and sparking insights or questions. In many cases, comparative aspects of collected data—relative position, clusters or patterns where one group has significantly different measurement values than others, changes in measurement values relative to the past, outliers where a measurement value is extremely different from others, anomalies where the measurement value is different from what was expected, direction changes, and sudden speed-ups or slowdowns—guide the search for problems and progress.[68] Of course, analysis that finds or confirms relationships among measurements is also important. Comparative information sometimes reveals solutions worth replicating or problems needing attention. Other times, comparative measurement triggers focused follow-up questions, narrowing the search for causes of problems and more precise, effective, and cost-effective problem corrections.

Performance leaders not only set a bar for others to meet, they also provide experience others can study and replicate, providing a path to better performance.

Performance laggards, in contrast, signal a need for attention, whether the performance lag is relative to people, organizations, conditions, or events. That the United States ranks only 19th among countries in life expectancy for women and 25th for men, for example,[69]

Case Study: NHTSA Searches for Successes

The U.S. National Highway Traffic Safety Administration (NHTSA) continually scans accident data gathered from states to discover successful models. For example, through its ongoing scanning efforts, NHTSA found a program in the state of North Carolina (borrowed from Canada) that increased seat belt use, a practice prior measurement of accidents had revealed to be effective in cutting the severity of accident consequences. After detecting the successful North Carolina program, NHTSA then needed to test its applicability in other communities. To do that, it enlisted states to participate in a controlled study. Some states fully implemented the adapted North Carolina approach, which NHTSA dubbed the "Click It or Ticket" program; some tested their own method to increase seat belt use; and a control group did nothing. All states agreed to use a common method to measure seat belt use: Observers stood at busy intersections before the campaign was started and afterwards and counted seat belt wearers as a percentage of all car passengers. The controlled experiment showed a significant increase in seat belt use among states using the North Carolina–adapted NHTSA approach compared to states using other approaches and the control group. Scanning performance data from the states detected a possibly successful program. A controlled, measured experiment confirmed it. Once confirmed, a compelling argument could be made for all states to change their practices.

despite having the 4th highest gross national income per capita[70] suggests a problem needing attention. The U.S Department of Health and Human Services recognizes the problem in its Healthy People 2010 publication. It is now beginning to "drill down" to understand the causes, looking for variations in life expectancy rates by income, race, and gender that might explain the causes of the United States' comparatively low international status.

Data *anomalies,* unexpected measurement readings, can also direct the search for problems and progress.[71]

Measurement outliers can often be mere statistical artifacts, but they should not be completely disregarded because outliers can also be revealing. Agencies that measure the timeliness of responses to applications, permits, correspondence, and other queries have discovered the value of studying situations associated with the longest response times to understand why they happened. This understanding then informs the design of preventive strategies to lessen repeat occurrences.

Clusters of problems, where one place, kind of person, or kind of organization has a higher frequency of unwanted events than others, can also lead to detection of problems and solutions.

Measurements gathered for short-term or what some have characterized as street-level decision making

Case Study: NHTSA Searches for Problems and Factors That Affect the Magnitude of Consequences

The U.S. National Highway Traffic Safety Administration (NHTSA) also scans comparative data to find societal and program problems needing attention. It analyzes annual traffic accident statistics, for example, to identify the most common causes of traffic accidents, especially high-cost accidents. Its scanning efforts brought its attention to motorcycle accidents, especially accidents involving motorcycle riders not wearing helmets, as a priority traffic safety problem. Understanding this, NHTSA wanted to understand how laws affect helmet use and accident consequences, so it began tracking the number and characteristics of helmet laws in each state. When two states, Kentucky and Louisiana, repealed their motorcycle helmet laws in a time period when other states' helmet laws remained unchanged, it provided NHTSA the opportunity to compare changes in fatality and injury rates associated with motorcycle helmet laws. NHTSA commissioned a study to compare the fatality and injury rates in the two states to the rate of change in other states. The study found that Kentucky and Louisiana had significantly higher injury and fatality rates than other states, which experienced only a slight increase in fatality rates and a slight decrease in injury rates during the same period. By routinely tracking state laws, NHTSA was able to learn from a natural experiment that often arises when state laws change.[72] With this sort of comparative analysis to find problems and safer practices, NHTSA realizes the enormous value gained when federal agencies step forward to fill the role of scientist studying state experiments, realizing Brandeis's vision of the states as "laboratories of democracy." Too often, despite the much proclaimed value of states as laboratories, the lessons are lost because the scientists are missing.

Case Study: EPA Learns Lessons from Anomalies

An anomaly in collected data helped the U.S. Environmental Protection Agency (EPA) detect a previously unknown problem, find its cause, and then craft an effective solution that not only improved water quality at the site where the original measurement occurred but at other similar sites. A local watershed association for a river in Massachusetts started collecting monthly water-quality readings for 37 points along the 80-mile stretch of the river and posting it on its website. This made it possible for the regional EPA office to obtain and study frequent, up-to-date water-quality data. When EPA detected an anomalous reading, a downstream reading worse than an upstream one that could not be explained by permitted wastewater discharge between the two monitoring points, it triggered a simple but focused follow-up question: Why?

EPA and city officials walked the pipes of the sewer systems between the two monitoring points, and discovered a facility illegally hooked up to the storm sewer instead of the sanitary sewer, causing raw waste to discharge into the river. In other words, measurement helped the agency detect a problem and triggered focused follow-up to understand its cause. Someone in EPA then suggested lifting storm sewer drain caps on dry days to look for running water, leading to the discovery of numerous other illegal hook-ups. Measurement revealed a problem, triggered focused follow-up, and stimulated creative thinking about additional problem detection methods. Using enforcement warnings, enforcement actions, and other tools, EPA then compelled the illegal dischargers to eliminate illegal connections. Measurement also revealed that the solutions EPA tried were effective. Water quality increased dramatically in a five-year period.[73]

can also support longer-term *evaluations* such as retrospective regression analyses to confirm working assumptions, provided the collected measurements are well organized and disseminated so researchers can readily access them.

In sum, measurement can illuminate as well as motivate. Studying measurements provides a powerful tool for detecting problems, causes of problems, and technically effective and efficient solutions. While measurement relative to a goal is motivational, it is measurement relative to the past, peers, and other situations that reveals the greatest performance-improving and accountability-enhancing insights. The challenge with measurement, then, is to create a culture that encourages comparative measurement to illuminate solutions, detect problems, and positively motivate, while discouraging dysfunctional uses of and responses to measurement.

Measurement Communicates

Measurement communication can also be a powerful performance-driving, accountability-enhancing tool. Measurement communication enhances accountability by reporting progress to the public and revealing problems competing for public attention and resources. Communicating measures improves performance by promoting successes, alerting people to problems, and providing a shorthand language to support multi-party outcome improvement efforts. It can also improve performance by focusing supervisory attention on ways to improve performance rather than on debates about whether or not problems exist.

Public reporting of measurements indicating progress relative to targets enhances performance accountability by communicating how well government has accomplished what it set out to do and by identifying where it encountered problems.

Case Study: Coast Guard Detects and Diagnoses Clusters

The U.S. Coast Guard marine environmental protection program counts oil spills, chemical spills, debris, and nuisance (water) species. The Texas field office studied characteristics of oil spill patterns and noticed that most of the spills occurred at night. Since all inspections occurred during the day, this finding immediately suggested a program change—shifting some inspections to the evening hours. When the Coast Guard tried this, the number of oil spills dropped (a fact the Coast Guard could detect because it routinely counts oil spills). By studying the data, the Coast Guard found a problem. The data also suggested that a solution it tested was effective and worth trying in other areas where similar nighttime spill patterns are found.

Public reporting of measurements also helps governments select targets. Detecting a problem does not imply that the problem should be fixed, however, nor does detection of a solution necessarily warrant replication. Those decisions depend on other demands on organizational resources. Along with communicating measurements to the public, government should then communicate which problems it intends to address by the targets it sets and, by omission, which it will not.

Measurement transparency lets citizens and their elected officials determine whether they agree with the priorities government agencies have chosen. Even if an agency has not adopted a specific target for problems such as obesity, water quality, or the crime rate, reporting trends in those areas provides the community with better information about problems needing attention and the priority-setting trade-offs government is making, giving citizens the opportunity to dissent if they disagree.

It is often asserted that what gets measured gets managed. That is clearly not always the case. What is clear is this: That which is measured and publicly reported is more likely to ripen the policy stream and appear as an action item on the policy agenda than that which is not.[74] Simply stated, that which is not measured is hard to address in an effective manner. Even when it has not set explicit targets, government communicates to the public the societal and program issues it is monitoring by sharing the data it collects. Measurement communication is therefore a key aspect of democratic accountability. Without publicly reported measurement information in areas of concern, public policy and management attention will naturally drift to areas that are measured.

What If Government Fails to Measure a Problem or Communicate Its Findings?

Using the legislative and other public comment processes, the public should press for well-communicated data in areas of concern, even when no specific targets have been set. When government refuses to measure a problem, advocates should (and often do) secure the resources to make the case for government attention to the issue.

Measurement communication also supports performance improvement by promoting problem finding and problem fixing. Measurement communication can bolster efforts to find common patterns of problems needing attention. When the Texas field office of the U.S. Coast Guard marine environmental protection program found that most spills occurred at night, it could then alert other field offices to look for similar nighttime spill patterns. Measurement communication can also promote performance improvement, especially when it is integrated into well-packaged, measurement-rich material such as NHTSA's “Buckle Up America,” “Click It or Ticket,” and “Operation ABC—America Buckles Up Children” campaigns to promote the uptake of more effective, cost-effective program management practices.[75]

Measurement communication also supports performance improvement by supporting networks that find problems, solutions, and opportunities. Several multi-party watershed projects use measurement as a precise, shorthand language to provide updates on shared goals and to engage the efforts and enlist the expertise of a network of goal-sharing parties.[76] The Bureau of Labor Statistics and the U.S. Census Bureau have long shared data in ways that engage the expertise of thousands, stimulating studies that help communities and businesses advance their own outcomes. More recently, the U.S. Department of Education (ED) discovered how measurement publication stimulates non-governmental, value-adding analysis. ED wisely decided to fund the compilation of state educational outcome data into a single, nationwide database that normalizes school scores across states by reporting scores relative to each state's own median. This enabled a not-for-profit organization concerned about low-income students, the Education Trust, to create its own web-based database to facilitate the search for high-performing, high-poverty, high-minority schools, helping researchers study these schools to identify factors contributing to their success.[77]

Inside an organization, measurement communication can often defuse what, without evidence, becomes testy supervisory conversations about performance improvement. By letting the numbers reveal problem areas, well-constructed measurement systems relieve supervisors of the need to debate the existence of problems in discussions with their direct reports. This allows supervisory discussions to shift instead to considering specific ways to improve performance. Measurements "reduced the chances for the development of conflict and antagonism; they made possible a more cordial and cooperative relationship between subordinates and their supervisors."[78]

In sum, measurement communication enhances accountability by reporting to the public whether or not government achieved the goal it set out to meet, where problems remain, and what other problems are competing for attention. Measurement communication boosts performance by promoting successful interventions, alerting other places to possible problems needing attention, and supporting multi-party outcome-improving efforts. Measurement communication plays one other critical performance-improving, accountability-enhancing role: It informs consumer and electoral choice, the topic of the next section.

Measurement Informs Choices

Measurement communicates options to those faced with consumer, business, or electoral choices, enabling a better match between individual preferences and the services government provides. Although there is a tendency to think of government agencies as a monopoly, agencies can sometimes structure service delivery (and even some regulatory) programs in ways that drive performance improvement through healthy competition, demand shifting, and better matches to consumer tastes or client needs. The key is communicating operational and outcome measurement in a way that encourages these dynamics, transforming data through analysis in a way that serves the needs of specific audiences and then delivering the data to users so they can find and use it when they need it. The explosive growth in wireless communication technologies and handheld equipment makes this aspect of measurement communication an increasingly important performance-improving, accountability-enhancing opportunity for government.

Visitors to motor vehicle offices in some states benefit from online data about wait times at each office, just as parents have long made home-purchasing decisions based on the best available data about schools. Unfortunately, most school performance data tend to reveal more about parental income, highly correlated with test scores, than about the skills of the schools' teachers. If states reported school test scores in a manner that controlled for key background variables, parents could see the net "value" each school adds. This would undoubtedly lead to different home-purchasing patterns. Moreover, measurements could be analyzed and communicated in a manner that facilitates a better match between students and schools. Parents of slow learners, for example, may prefer a school that does particularly well teaching the lowest quartile of students, while those of average learners may want a school that is particularly adept at teaching those in the middle two quartiles.

An indexing process that compiles various performance indicators into a single ranking or grade to inform consumer choice inevitably involves

Case Study: Massachusetts Registry of Motor Vehicles Packages Measurement Information to Inform Consumer Choice

Consider how data analysis and audience-focused dissemination could transform state motor vehicle offices, often highly detested because of their long wait times. The Massachusetts Registry of Motor Vehicles (RMV) posts real-time (current) wait times for each office on its website.[79] This real-time posting is a noteworthy step in the right direction. Before leaving their home or office, visitors can check the website to find wait times at the closest registry offices and decide which one to visit. Imagine how much more powerful a bit more analysis and its dissemination would be if average wait time by type of transaction as well as averages for different hours of each day (time of day), day of the week, and week of the year were also posted for each office. Registry visitors could plan ahead to visit the nearest office on a slower day, also checking the real-time web post before departure to avoid unusual days. This slight bit of data dissemination would shift peak load pressures and lessen stress both on Registry visitors and Registry staff. It might also instill a bit of healthy competition among Registry offices to keep average wait times down. In addition, it could reduce program costs and provide users of government services with better options. If program budgets are not cut to offset the gains in cost-effectiveness, users of government services would most likely also receive faster service.

subjective judgments about the relative import of indicators included in (as well as those excluded from) the index. This can not only result in ill-informed consumer decisions, but also motivate ill-advised behavioral changes or measurement manipulations by those being measured. Some colleges and universities have reportedly engaged in denominator padding, for example, by changing their definition of an application to fare more favorably on *U.S. News & World Report's* annual report card.[80]

Measurement communication can also function as an effective regulatory tool driving performance improvement when government-reported data influences consumer choice. In Los Angeles County, health inspectors grade restaurant hygiene based on compliance with local health laws. Not surprisingly, the grades influence patron decisions about where to eat. To avoid losing customers, restaurants have upgraded their health practices to earn high grades.[81] Where grading is relatively simple, based on a few well-chosen performance indicators, and where it informs consumer purchasing decisions as is the case with restaurant grades, measurement communication can motivate improved performance. Even more complicated comparative measurement systems such as those associated with financial investments appear to motivate improved performance or at least reduce risks when consumers are sophisticated.[82]

Government agencies should consider more carefully if and how they can support the consumption choices of those whom they serve through better measurement communication. This will lead to healthy competition that motivates supplier effort and innovation and shifts in demand patterns, ultimately achieving a better match between demander preferences and provider products when the needs (or tastes) of consumers of government services vary.

Use Measurement to Inform Choice

Government agencies should consider if and how they can support the choices of those they serve through enhanced measurement communication. They should look for opportunities to stimulate healthy competition, shifts in demand patterns, and better matches between demander preferences and provider products with audience-focused characterization and dissemination of measurement information.

In addition, to enhance democratic accountability, government agencies should consider more carefully how they can better support the electoral choices of citizens by revealing not only specific goals, but also accomplishments and problems relative to goals, the past, and their peers. If the public does not agree with the bundle of goal and measurement choices that its elected officials have made, voters can select candidates who promise a preferable bundle of goals. Although most elected officials have shown little interest in setting goals and measuring performance toward them, presumably wary of inviting attack should they fail to meet their targets, several elected executives have embraced goal setting and performance measurement in recent years. Those who have done so include New York City Mayors Rudy Giuliani and Michael Bloomberg, District of Columbia Mayor Anthony Williams, Baltimore Mayor Martin O'Malley, and Washington State Governors Norman Locke and Christine Gregoire. All but Gregoire, who is in her first term, readily won re-election. Although this does not prove that goals and measures are a good political tool, it does suggest that they are not as politically poisonous as often feared.

In sum, measurement informs consumer and electoral choices. Governments can and should manage their measurement communication in a way that supports the choices of those they serve, supporting healthy competition that motivates effort, innovation, and better matches between consumers and providers when consumer (or citizen) needs and preferences vary.

Cautionary Considerations

As with goals, care must be exercised to tap the potential of performance measurement without unleashing fears that provoke dysfunctional behaviors, most notably measurement manipulation and measurement elimination. When measures inform decisions, care must also be exercised to ensure an adequate level of measurement accuracy.

Measurement manipulation is a serious issue because it provides organizations with bad information for decision making. Manipulated measures send bad signals about which interventions are working and which are not, resulting in misguided decisions about practices to expand and practices to discontinue.

Case Study: Cheating in Texas Schools

Reporters for the *Dallas Morning News* have detected measurement anomalies that suggest serious cheating problems in at least 400 schools out of 7,700 statewide; test scores for students in those 400 schools were three standard deviations away from test scores for the same students the previous year. One teacher told reporters, "You're expected to cheat...." Fellow teachers, she said, instructed her how to help students during the tests.[83] Teachers annually complain to the head of the Houston Federation of Teachers that principals urge them to cheat but will not file formal complaints for fear of retribution.[84]

Yet despite signs of serious cheating, the state of Texas has declined to follow up on the findings with all but a handful of schools, preferring to follow up only on verified complaints that can be prosecuted.[85] Rather than using measurements to find problems and work with the schools to fix them in a performance-improving learning approach, the state has apparently opted for a narrow compliance-focused accountability approach.

While some people respond to comparative measurement by shirking their work efforts, others so fear a poor performance rating that they manipulate the measures.

Measurement elimination, the dismantling of measurement systems, is obviously a problem, because for all but the most simple targets, the absence of relevant measures makes it impossible to determine whether or not progress is being made. Organizations sometimes so fear the consequences of measurements that they organize to dismantle or prevent the establishment of measurement systems. Data system destruction tends to occur most frequently when a poor relative standing threatens those being compared and the political power of those threatened by comparison exceeds that of those using comparative information. This has frequently been a problem plaguing recurring efforts to introduce measurement into hospitals, education, and the environment.[86]

For comparative organizational measurement systems to survive, the political strength of the audience using the measurements needs to exceed that of those at risk by being compared.[88] Agencies will be well served if they think about how to strengthen the political power of measurement consumers as a strategic issue so that measurement can survive and effectively motivate, illuminate, communicate, and inform choices.

Measurements are infuriating rather than performance-improving when inaccurate information is used to guide consumer choices, regulatory action, or program or provider funding decisions. As noted earlier, for goals and measures to improve performance and enhance accountability, government agencies need to make data accuracy a goal, and then invest in, measure, and publicly report on progress toward the data accuracy goal, setting the target level of the data accuracy goal commensurate with the intended use of the data.

Case Study: Fears Lead to Watered-Down Measures

During the debate over the 1996 amendments to the Safe Drinking Water Act in the U.S. Congress, representatives of local government lobbied hard to water down mandatory water-quality reporting requirements. They feared that the public would overreact to the presence of trace amounts of certain elements in drinking water.[87]

Rather than trying to figure out how to inform the public about inevitable risks, many localities preferred to limit the disclosure of performance information.

Case Study: Strong Demand of Military Drives Educational Performance Measures

Even the efforts of the illustrious Florence Nightingale and Horace Greeley to establish public performance reporting on the performance of hospitals and schools did not prevail.[89] It was only when the military needed comparative educational performance data to meet its hiring needs that a mechanism to compare student educational readiness survived. That's because a politically powerful constituency, the military—and eventually other well-organized political voices such as colleges and universities—deemed the data useful.[90]

Summary and Implications

In sum, measurement can motivate, illuminate, communicate, and inform choices, and, in so doing, strengthen accountability and improve outcomes. People and organizations like to do well and work harder when they see they are making progress relative to a goal or to the past. Comparison motivates some people and organizations. Other people, however, get discouraged by peer comparison. Measurement illuminates by revealing problems needing attention and program successes worthy of replication, and by triggering focused follow-up questions. Measurement also serves as a useful shorthand language for communication among goal sharers, citizens, and consumers of government services. Care must be exercised in using measurement, however, to tap its potential without unleashing fears that provoke dysfunctional behavior, such as measurement manipulation, measurement elimination, outcome avoidance, and timid targets. The problematic aspects of comparative measurement can be lessened if more emphasis is placed on its illumination value rather than its motivational or controlling value.

Building Block 3: Verbal Feedback to Unleash the Power of Goals and Measures

Goals and measures are potentially powerful performance-driving tools, but they have little performance-driving effect if simply communicated via written words in a plan or report. They have little value unless someone pays attention to them.[91] To be useful, they must be used. A leader (at any level of an organization, but ideally starting at the top) must not only frequently articulate a consistent set of goals so the organization stays focused on them, but must also encourage others in the organization to articulate them.[92] Managers need to reinforce the priority of specific, challenging goals, or they will be forgotten, pushed aside by other topics of discussion. More than simply articulating the goals, though, managers also need to pay attention to measurement of progress toward the goal and provide verbal feedback.

Feedback is the essential ingredient that unleashes the power of goals and measures. Years ago, researchers discovered the powerful effect of attention on worker productivity, even without goals. This is often referred to as the Hawthorne effect, named after the manufacturing plant where the surprising research finding was discovered. Researchers sought to test the effect of lighting changes on worker productivity. To their surprise, worker productivity improved whether lighting was increased or decreased. It did not, however, improve among those not in the experiment. The magic motivator proved to be attention, not lighting. One simple but often underutilized method managers can use to motivate workers is to discuss their work with them.[93]

Coaches in the Workplace

"The primary job of a coach is to improve performance rather than focus on blame. This is done through increasing the person's sense of control regarding the attainment of their goals. It is done by helping people to realize the outcomes they can expect from engaging in specific actions…. Setting a goal focuses attention on discovering solutions to its attainment." (Locke, *The Blackwell Handbook of Principles of Organizational Behavior,* p. 110).

How managers or team leaders provide feedback directly affects performance levels. Key to performance improvement is winning worker commitment to a goal. This can be achieved by helping people in an organization focus on specific, challenging goals and by building their self-confidence that they can get the job done. Formal recognition devices are not necessary. Effective feedback can be provided informally, simply expressing appreciation for work well done. The commendation must be specific and genuine, however. It cannot be general or phony.[94] Managers build confidence by helping workers set goals, stay focused on them, ensure adequate training so they acquire needed skills and knowledge, and maintain confidence that they can reach their goals.

In addition to building goal commitment, managers can "coach" improved performance. They can help employees sequence tasks so that prospects for early success are enhanced. Managers are most helpful if they discuss and provide advice about specific actions employees can try. Feedback using specific illustrations is more useful than generalized feedback about performance levels. "Research studies have demonstrated that the most useful feedback tends to be behaviorally specific, identifying those high-leverage behaviors that can be improved and providing guidance on how to do so."[95] Finding and presenting relevant benchmarking models to workers is also helpful, provided the models are not perceived as out of reach. Overly ambitious benchmarks tend to be daunting rather than inspiring. In addition, successful coaches find someone the worker respects to communicate the belief that the worker can achieve the goal. The influential person can be the supervisor, but can also be a peer or other persuasive associate. By framing measurement and feedback in a constructive manner, recognizing successes and showing an achievable path forward, managers help workers tap into the performance-enhancing aspect of measurement.[96]

Constructive feedback is far more motivating than critical or condemnatory feedback. "Positive performance feedback enhances intrinsic interest when provided in an informational (as opposed to a controlling) manner."[97] In contrast, negative supervisory feedback, especially that which does not align with what performance measurements reveal, can so enrage even enthusiastic workers that previously strong performance not only declines but remains weak even after a change in supervisors.[98] People tend to be more motivated when they believe they are doing well and disheartened when they sense they are doing poorly.

Managers can help workers boost performance by training them to carry on positive, rather than negative, internal conversations in their own heads. This works because people can function as their own internal coaches, encouraging and discouraging themselves with internal feedback.[99]

New technologies can also provide performance-improving feedback when used in constructive, non-critical ways. A recent study of 12 young drivers showed that when cameras were installed in their cars for subsequent viewing by both the teen and the parents, risky driving behavior by teens dropped 75 percent. One of the study participants recognized the value of specific feedback for learning. "That is why this was good, because people can tell you what to do all they want, but when you actually see it, that is when you learn."[100]

More frequent feedback is better than less feedback, in part because it helps workers recalibrate their goals based on experience,[101] leading to a higher rate of successful goal attainment and strategy adjustment. That is not to suggest frequent goal changes are essential, but that realistic adjustments are important.

Feedback enhances both individual and group performance. In addition to the feedback that benefits individuals, group feedback should clarify who has lead accountability for specific targets. That is, managers or the group should clearly identify a manager for each goal. Group feedback works best when it not only clarifies roles but at the same time builds a sense of organizational unity. In addition, it can be used to reinforce organizational ethical norms and behavioral boundaries, the "thou shalt nots" that reinforce fiscally and ethically accountable behavior. Clear delineation of the boundaries of unallowable behavior frees people up to innovate within those boundaries so long as they do not trespass beyond them.[102]

Specific functions managers should encourage as part of constructive group feedback are:

- Help members set challenging but realistic goals, adjusting them as needed.
- Make sure each member understands what his or her contribution is to the final product.
- Ensure group participants understand how their membership in the group is helpful.
- Emphasize the unity of the group.
- Change goals that are too difficult.
- Encourage talk in meetings about how performance can be improved.
- Avoid fear of failure.
- Help members feel responsible for the group's fate.
- Give members assignments that suit their abilities.[103]
- Clearly define boundaries of unacceptable practice.[104]

If direct supervisors fail to provide the needed feedback without prompting, these findings argue that goal managers charged with meeting a specific, challenging goal should initiate discussion of their progress toward the goal with the supervisor or with others whom they respect. The discussions should break the goal into manageable pieces, examine concrete illustrations of specific behaviors and strategies that might help advance progress, and reinforce confidence that the job can be done. The forum for the feedback is not as important as the way feedback is delivered.

In sum, feedback is the essential ingredient that unleashes the power of goals and measures. Leaders, supervisors, and sometimes even colleagues can use feedback to improve performance by helping people and organizations stay focused on specific, challenging goals; building confidence they can get the job done; ensuring adequate training and resources; sequencing tasks and adjusting goals when they prove overly ambitious; and providing advice about and illustrations of specific actions to try. In addition, feedback helps in building group unity and appreciation of individual contributions to the group; encouraging interactive exploration of performance improvement options; accepting failure as a by-product of challenging goals; clearly assigning roles; and clearly delineating boundaries of unacceptable practice.

Building Block 4: Interactive Inquiry

Feedback is a powerful tool both for individuals and for groups. It can be even more powerful for groups when it is delivered in a manner that goes beyond simple supervisory observations to stimulate ideas, encourage interactive inquiry, and tap resources and expertise beyond a goal manager's direct control.

More feedback is better than less, so group feedback can enrich individual feedback provided it is delivered in a well-directed, non-random, and constructive manner. Meetings afford an opportunity to invite non-supervisors to provide specific, relevant, and compelling examples to suggest a path forward and build confidence that specific targets can be met. This argues for engaging others who might be familiar with relevant benchmarks and whose views goal seekers respect in feedback, strategy development, and operational brainstorming.

In recent years, several government and nonprofit agencies have established frequent, regularly scheduled, goal-focused, measurement-rich meetings where "management discourse … is by the numbers."[105] Perhaps the most well-known of these efforts is CompStat (short for computerized statistics), an approach developed by the New York City Police Department (NYPD) to drive the crime rate down. The success of CompStat has spread across city agencies in New York, moving from the police department to the parks, corrections, and welfare departments. It has also been replicated in police departments across the country, and applied more broadly to other social issues in city (CitiStat in Baltimore and Pittsburgh) and recently in state government (Washington State's Government Management, Accountability, and Performance—GMAP). When inter-organizational cooperation is needed to accomplish a goal or when opportunities for cross-organizational learning exist, regularly scheduled, goal-focused, measurement-rich meetings provide a time- and cost-efficient forum to identify problems, refine goals, plan cooperative action, discuss progress, and improve results.[106] Meetings need to focus on goals, measurements that provide feedback relative to the goals, analysis and discussion to enhance measurement interpretation, long-term strategies, and shorter-term action plans informed by the evidence.

The NYPD convenes CompStat meetings for each police precinct or district at least once each month to focus on outcomes and the activities undertaken to influence them, and to engage the expertise and commitments of those in the organization with relevant information or resources.

> CompStat meetings … should not include discussions of policies, budgets, internal discipline, politics, organization structure, or any other item not directly related to crime reduction.…[107]
>
> Each session focuses on the activities in one or two geographic areas. Those commanders who are going to be involved in the discussion about trends and strategies in their districts may invite other members of their staffs who might play a role in, or add to, the discussion. Another group that should attend includes commanders from specialized units whose resources might be needed to adequately respond to crime patterns or other trends. Patrol commanders from adjacent areas or other patrol commanders who would benefit from exposure and experience may be directed to attend also.[108]

One of the oft-mentioned benefits of CompStat-like meetings is the presence of key personnel, including the legal, contracts, financial, human resource, and information departments. Much of the time, these staff office leaders ask questions of the goal managers. Other times, they are available to provide quick answers to key questions, speeding response times that otherwise delay a project. The diversity of experienced people in the room allows a stimulating mix of ideas informed by varied professional perspectives and experiences. It can also broaden consideration of program approaches and speed decision making and implementation. One key to the success of interactive inquiry meetings is that the person running the meeting "must have sufficient authority to order the coordination of services … and to implement strategies."[109]

The format of an interactive inquiry meeting is not as important as is its frequency, focus, attendees, and atmosphere. Meetings need not be structured in the formalized format of CompStat to be effective. Even in New York City, while some departments have adopted the CompStat approach, others integrate goals and measurements into less formally structured meetings.[110]

Interactive Inquiry

Frequent goal-focused, measurement-informed meetings reinforce the power of goals and measures when they are used to:

- Refine goals.
- Discuss progress.
- Identify problems.
- Challenge assumptions.
- Clarify roles and goal manager.
- Get feedback from different organizational units and professional perspectives.
- Stimulate ideas.
- Brainstorm implementation options.
- Devise action plans aligned with organizational priorities.
- Press for follow-through.
- Coordinate and encourage cooperation.
- Obtain quick approval or disapproval for new practices.
- Share information and insights broadly.
- Build unity of purpose and appreciation of individual contributions to the group.
- Reinforce boundary conditions.

Frequent interactive inquiry meetings are also useful for tackling three especially tricky operational problems for performance-focused organizations—mid-year goal adjustment, mid-year budgetary changes, and mid-year staff adjustment:

- *Goals, targets, or commitments may need mid-year adjustments* when targets are accomplished sooner than expected or new evidence suggests they are too ambitious. Frequent meetings provide a forum for goal managers to propose target adjustment and get quick approval or negotiate alternatives.
- *Mid-year budget adjustments* are complicated because budgets, in the best of circumstances, are batch processed, yet new evidence may sometimes suggest a need for a strategy revision that requires additional resources. To reap the performance-improving power of goals and measures, task adjustment must be nimble, something the government fund requisition process is not. Frequent goal-focused meetings provide a ready forum where goal managers can request additional funds for specific purposes, and get fast answers about whether or not additional resources might be available.
- Similarly, when evidence suggests that past activities are not working and a new course of action is needed, a goal manager's team may lack the new skills and knowledge required. They may need a mid-course *skill or staffing adjustment.* Frequent goal-focused meetings provide a venue for exploring options to deal with this challenge in the short term, such as borrowing staff resources from other divisions, contracting out, or hiring temporary employees.

Goal-focused, measurement-driven meetings among multiple organizations can also reduce turf battles. They allow cross-unit cooperation to occur naturally, even in organizations where previous efforts to encourage it never materialized.[111]

Key to the success of interactive inquiry meetings is that they be viewed by participants as constructive, not condemnatory. Nor can they be seen simply as a forum for merely "reporting out," with limited exchange. It is not enough to convene meetings; it is essential to get the tone right. Collins, in his study of successful corporations with breakthrough performance gains, identified four stylistic approaches used by successful companies for constructive feedback:

- Lead with questions, not answers.
- Engage in dialogue and debate, not coercion.
- Conduct autopsies, without blame.
- Build "red flag" mechanisms.[112]

Collins found that an ongoing, interactive process helps an organization realize the full performance-improving power of goal setting and measurement, but only if done in an honest, open, and accepting atmosphere. Meetings should invite healthy contention, but not in an accusative or threatening manner.

In sum, interactive inquiry enriches the performance-improving power of goals and measures in groups and organizations. To achieve these effects, though, it is essential that interactive inquiry meetings be run in a way that is constructive, not critical or controlling.

Building Block 5: Cautious Use of Externally Provided Incentives

The discussion so far has described the enormously powerful performance-driving, accountability-enhancing effect goals and measures have when combined with feedback and interactive inquiry, even without positive or negative incentives. What happens when incentives are combined with the other four building blocks of performance management?

Unquestionably, incentives have the power to change behavior that can improve outcomes. Consider the effect of one common form of negative incentive, the speeding ticket, intended to save lives by encouraging driving speeds no higher than a minimum standard, the speed limit. Incentives undoubtedly motivate progress toward a safety goal in this case. Drivers are less like likely to fly down the highway at breakneck speed when they think they will be caught and ticketed for speeding, especially if there is a possibility that the ticket will increase their insurance costs or cost them their license. Extrinsic incentives can motivate behavioral changes that lead to performance improvements.

But incentives can also easily discourage people and trigger dysfunctional individual and organizational responses that interfere with performance and accountability gains. Therefore, incentives should be employed with extreme caution. Indeed, a reluctant approach to the use of incentives is advisable. This is particularly true for individuals, but should also apply to organizations. It is often far more judicious to tap the "un-incented" power of goals and measures as

Use of Performance-Based Pay Expanding at Federal, State, and Local Levels

The U.S. Government Accountability Office (GAO) began exploring performance-based pay in the 1980s. The arrival in late 1998 of a new comptroller general brought a renewed effort to expand performance-based pay arrangements at GAO, for which Congress granted new authorities in 2004.[113]

Since then, the U.S. Congress has given other federal agencies—including the Departments of Homeland Security and Defense, which represent over half the U.S. federal workforce—legal authority to depart from the U.S. Office of Personnel Management (OPM) standards and adopt performance-based pay systems. In addition, OPM has issued regulations requiring "all future Senior Executive Service pay increases (beginning in January 2004) be based exclusively on individual and organizational performance."[114]

Interest in pay-for-performance has been a key component of many education reform efforts at the federal, state, and local levels over the last decade, as well. The basic idea is to reward teachers and school administrators whose students have attained certain performance standards, have high performance relative to peers, or have the greatest student performance improvement levels.[115] Punishment for non-performance is a prominent feature of the U.S. federal education No Child Left Behind law.[116]

Despite numerous efforts to introduce performance-based pay systems in government, few last beyond a few years. Moreover, there is little evidence that they achieve their intended motivational effect. Systems that have survived tend to reward a way of working or cost savings.

motivators, reinforced by feedback and interactive inquiry, than to introduce extrinsic incentives with their significant potential for deleterious side effects.

Whether or not to use incentives is not simply an interesting intellectual debate. Proposals to link individual or organizational performance levels to financial incentives are very popular and have accompanied many of the public management reform efforts around the world that require governments to set goals and then measure and report performance. The idea of motivating better performance through some sort of performance-based pay scheme, performance budgeting, performance contracting, or performance-linked grants is frequently discussed and has often been legislated or required by executive fiat.

What is noteworthy is that, despite all the hoopla, most government performance-based incentive systems that try to link positive incentives (rewards) to the performance of individuals tend not to last. One follow-up study on school systems with merit pay arrangements, for example, found that only 2 out of 10 were still operating 10 years later.[117] A performance-based pay system established in 1984 (for U.S. government employees just below the senior management level) was terminated by 1993.[118] Organizational reward and penalty systems seem to last longer and work better, but they too encounter difficulties that can lead to their demise. As discussed on page 31, multiple organizational measurement systems have imploded when those at risk of serious loss have had the political strength to bring them down.[119]

The poor track record of performance-based incentive systems has not, however, lessened interest in them. To garner the gains of incentives without releasing their woes, it is therefore important to delve into the lessons of experience and research to determine if, when, and how government should structure performance-based incentives systems for individuals and organizations so they are effective and sustainable, and so they do not overwhelm the untethered, performance-driving, accountability-enhancing power of goals, measures, feedback, and interactive inquiry. The remainder of this section examines these lessons.

Common Problems: Outcome Avoidance, Measurement Manipulation, and Timid Targets

Positive and negative extrinsic incentives sometimes have a positive motivational effect, leading to performance improvement, but they can also distort goal setting and measurement decisions. Specifically, they can provoke outcome avoidance, timid targets, and measurement manipulation, interfering with performance and accountability gains.

New Zealand's much touted public reform effort, launched in 1988, illustrates the problems incentives can introduce. The initial New Zealand performance management system called for a "purchase agreement" between each departmental manager and the relevant minister (a political official). The initial incentive arrangement put a manager at risk of losing his or her job if a department failed to meet its performance targets.[120] A 10-year retrospective analysis of the New Zealand effort faults the performance contract for creating a managerial aversion to outcome targets and an affinity for output goals, defeating key objectives of the reform:

> In fact, the 1988 reforms, with their accent on contract, sent the State in the other direction, tying outside agencies down so tightly in contracts that they are almost back to pre-1988 rules-based administration, far removed from imaginative management. The Plunket Society, for example, used to be funded to produce generalised outcomes; now it is contracted for highly specified outputs, satisfying central government monitors but stultifying initiative and perhaps delivering less for the money. Some lower-level managers within departments are complaining that their contracts have become rather like the old rules the 1988 reforms were supposed to have swept away.[121]

When incentives are introduced, government managers will naturally and very sensibly choose output targets if they can, because they can more directly control output levels. Linking a reward directly to goal attainment is also likely to lead to easier targets, even though it is the challenging ones that stimulate the discovery and persistence that result in significant performance gains.

Government managers and organizations do not always have the flexibility to choose their own targets. Sometimes, targets are set for them. Again, introducing incentives into the picture can create problems. Measurement manipulation and timid targets have proven a recurring problem when organizations have tried to link incentives and measurement levels. In a study of the Job Training Partnership Act (JTPA), for example, Marschke concluded, "[W]hen performance measures are compensated, bureaucrats respond by finding the least-cost strategies of boosting those performance measures"[122] without a compensating boost in outcomes.

Heinrich found similar measurement manipulation and timid target problems in a study of the successor program to JTPA, the Workforce Investment Act, which sanctions states that fail to meet 80 percent of negotiated performance levels for 17 measures. Some states skimmed the clientele, limiting their client base to avoid harder-to-serve clients.[123] Heinrich also found states adopting timid targets, in some cases setting performance expectations below the state's own baseline level.[124]

The problems are not limited to one kind of government program. Jacob found measurement machinations in a study of Chicago schools. While "math and reading achievement increased sharply following the introduction of the accountability policy … driven largely by increases in test-specific skills and student effort, … the teachers responded strategically to the incentives along a variety of dimensions—by increasing special education placements, preemptively retaining students, and substituting away from low-stakes subjects like science and social studies."[125]

Marschke also found evidence that, at least in some circumstances, incentives do not even have the intended influence on outcome improvements: "[O]nly the studies of programs where performance is uncompensated show statistically significant correlation between JTPA-style performance measures and impacts."[126] His findings are surprising, but perhaps they should not be. Similar experience was found in the private sector, where extrinsic rewards are presumed to work so well. In the book *Good to Great,* an exhaustive study comparing companies that went through a transition from average to outstanding performance with similar companies whose performance stayed steady or declined, researcher Jim Collins found that financial incentives for corporate executives were not correlated with corporate performance levels:

> Most importantly, when we analyzed executive compensation patterns *relative to comparison companies, we found no systematic differences* on the use of stock (or not), high salaries (or not), bonus incentives (or not), or long-term compensation (or not.) The only significant difference we found was that the good-to-great executives received slightly *less* total cash compensation 10 years after the transition than their counterparts at the still mediocre comparison companies.[127]

Outcome avoidance, measurement manipulation, and timid targets all directly interfere with the performance-increasing, accountability-improving power of goals and measures. Why do these problems occur? Do extrinsic incentives prompt these unproductive responses? To begin to answer that question, we turn our attention to understanding the effect of incentives on individuals and organizations, looking first at the structure of the incentive, and then considering whom the incentive is intended to influence and what it seeks to reward or penalize.

How the Structure of External Incentives Affects Motivation

The characteristics of an extrinsic incentive affect how people and organizations respond to it. Incentives can be extrinsic (externally provided) or intrinsic (internally motivated). Extrinsic incentives may or may not be monetary, and they can reward or penalize. The structure of an incentive can affect how an incentive influences behavior and, ultimately, outcomes.

Extrinsic Incentives Drive Out Intrinsic Motivators

Perhaps the greatest problem that arises with extrinsic incentives, especially monetary bonuses, is that they tend to drive out intrinsic ones. Researchers have found that introducing financial incentives into a work situation where people had previously been intrinsically motivated dampens intrinsic instincts to do a good job.[128] In some situations, those offered financial rewards actually underperform those not offered any reward. In others, performance improvements induced by financial reward are not sustained after rewards are terminated.

The problem of extrinsic motivators driving out intrinsic ones is likely to be a greater concern in the public sector than in the private sector, because people attracted to the public sector are more likely to be intrinsically motivated by their job and less likely to be motivated by money. This is especially the case for people at the higher grades of government service, who have often chosen their occupations because of their intrinsic motivation to serve or their belief in advancing a specific agency mission.[129]

Financial Rewards Work in Limited Situations

Financial incentives such as the promise of a bonus, raise, or promotion can drive improved performance when they are large enough. "Monetary incentives are one practical outcome that can be used to enhance goal commitment. … more money gains more commitment."[130] Monetary incentives work best, however, for jobs that are single-tasked, clearly measurable, and linked to individual workers.[131] In those cases, one sort of incentive that works well is piece-rate pay offered as a supplement to a base "safety net" hourly wage. Some government jobs, such as pothole repair, are single-tasked, easily measured, and can be linked to individual workers.

Government has a problem using piece-rate payments that the private sector does not, however. Few governmental organizations enjoy a rising revenue stream funded by the improved productivity of piece-rate incentives that can be shared with those generating the gains. Government does not, for example, enjoy a revenue flow that could fund a pay-per-pothole-filled or pay-per-park-toilet-cleaned payment scheme.

> **In the Right Setting, the Right Incentives Can Work**
>
> In one private sector example, a windshield replacement company, both workers and the company profited after a piece-rate payment system was introduced that supplemented the base hourly wage (maintaining a "safety net" for the worker).[132] The company was able to pay high-performing workers more because workers increased their effort, productivity, and ultimately output; weaker workers left; and the company attracted new workers who expected to do well under a piece-rate system. This, in turn, generated increased net revenues that could be used to pay the workers while also raising company profits.

Therefore, except for those government organizations that can charge a fee for services, it has to cap piece-rate payments based on fixed revenues.

One area where government has been able to link incentive payments to a revenue stream is with cost savings, essentially a negative expenditure stream. Maricopa County, Arizona, has adopted what it calls a "gain-sharing" program that distributes half of annual cost savings as a bonus among all employees.[133] The experience in Maricopa County suggests that incentives to encourage cost savings in government may work effectively. The goal, cost savings, is simple and easily measurable. It should be noted that this incentive structure only drives performance improvement, though, if cost reductions are accomplished without a compromise in performance. Otherwise, cost-cutting incentives merely result in service reductions.

It is noteworthy that most federal programs that have participated in performance-based pay demonstration projects appear to be revenue-generating units of federal agencies. It is also noteworthy that what may be the longest-lasting federal performance-based pay system, that of the Naval Air Warfare Center Weapons Division in China Lake, California, does not reward attainment of specific performance targets but rather a way of working.[134]

Rewarding Complex Goals Confounds Outcome Improvement

When goals are too difficult or complex, the motivating potential of a monetary reward linked to the goal is mitigated. Rewarding goal attainment tends not to work for difficult goals unless the goal can be broken down into individual pieces and payment made for each task completed. In most cases, monetary incentives large enough to induce behavioral changes overemphasize that which can be measured. Unfortunately, they overlook that which is more difficult to measure but which better serves the interests of the organization and, in the case of government agencies, of society.

Temporal Disparity Between Reward Cycle and Outcome Causes Problems

One specific problem that frequently arises with annual monetary incentives (as opposed to longer-term rewards such as stock distributions) is that they create an unwanted bias toward short-term returns.

The collapse of Enron can be attributed, in part, to a poorly structured monetary reward system that directly linked huge end-of-year bonuses to overly simplistic performance metrics.[135]

Comparison Problems Exacerbated by Monetary Incentives

Monetary incentives also exacerbate the comparative measurement problems discussed in the section on measurement. People care not just about how their performance is ranked but how they are rewarded relative to others. Monetary rewards heighten comparative tension because money is easily compared, and people tend to judge the value of a reward by its relative rather than its absolute value.[136] What might seem like a nice-sized bonus to an award-giving manager may be judged an insult by the award-receiving employee if it is less than that received by a colleague the employee considers of lesser ability[137] or if it is less than what the employee expected based on past practice.

Given these findings, it is unclear why an agency would introduce financial rewards for public sector managers and professionals, who already tend to have strong intrinsic motivation to do their jobs well. Nor is it clear why they would introduce them for frontline workers, at least those with budgets not easily replenished by the fruits of their increased effort and ingenuity. In most government agencies, financial awards are necessarily limited in amount and number because of budget constraints. Non-financial and non-tangible rewards, such as praise, have fewer problems so long as they are genuinely and appropriately delivered. And if they are not scalable, the amount of a non-financial award cannot be easily compared to other awards, reducing problems that arise because awards are judged more by their relative than absolute value.

This does not suggest that all government agencies should precipitously eliminate individual financial bonuses if they already have them. Since workers treat rewards as relative rather than absolute, they may look at financial incentives relative not only to others but also to their own expectations based on past practices. Sudden changes from past bonus practices may lead to frustration and depress performance. Additional examination of the literature on and experience of organizations that have eliminated bonuses would be useful to inform agency decision making in this area. What is clear is that if financial incentives are introduced into agencies where they have not previously existed, as is frequently proposed for many government agencies, they are not likely to have the positive benefits anticipated and are likely to provoke numerous unconstructive responses.

Instances Where Positive Financial Incentives Have Promise

Financial rewards do hold promise in a few specific situations, especially for organizations. They can work, for example, in opt-in situations where individual or organizations are wooed to compete or pursue a target they might not otherwise have considered with potential prizes, grants, or contracts. The promise of financial reward has successfully been used to woo individuals, such as baseball and basketball players, and organizations, such as architecture firms, to put forth their best efforts or ideas. The watermark records they set in these contests may then have a trickle-up effect as more and more people strive for the equivalent of Roger Bannister's four-minute mile.

Monetary incentives have also proven effective in wooing organizations to adopt new outcome-focused goals, adjust existing priorities, or at least bring organizations with shared goals forward to cooperate by offering monetary incentive for doing so. U.S. federal domestic policy agencies have long used financial incentives to enlist state and local governments in the health, welfare, education, employment, environmental protection, housing, transportation, and economic development policy areas to advance goals identified by the U.S. Congress as needing attention.[138] States, localities, and nonprofits agree to goal adoption in their grant applications.

In many of these goal-adoption incentive situations, the applicants do not need financial incentives to want to adopt a specific goal. They are already motivated to meet it and just looking for funding for that purpose. In other words, they are not motivated by the money but the cause. They do not need a financial incentive to improve, just financial fuel to operate. In these cases, the challenge for government is not figuring out how to calibrate the right sort of goal-adoption incentive. Instead, the challenge for the incentive-offering government is figuring out how to find grantees (or contractors) that share the goal and have a strong capacity to deliver progress toward it, how to engage them in

Case Study: The Value of EPA Regulatory Sanctions as Motivators

The Clean Air Act effectively uses the threat of a serious penalty, the ability to stop federally funded highway development projects, to motivate states and localities to meet specific, challenging air-quality goals.[139] The law required the U.S. Environmental Protection Agency (EPA) to set six air-quality goals for the nation, the National Ambient Air Quality Standards. All major metropolitan areas were required to attain these minimum standards by specific dates. Air-quality monitors around the country monitor attainment. Initial failure to meet the standard triggers an initial sanction, that the state must write an implementation plan. In the plan, a state is expected to set forth a cogent, data-informed strategy for attaining national air-quality goals (standards) in non-attainment areas. The law required EPA to trigger the second penalty, loss of federal funds for highway development projects, if non-attainment persisted past a specified period and if EPA considered the state attainment plan unacceptable.[140] The system has successfully motivated not only attention to the problem, but air-quality improvement. In 1999, for example, 400 counties were designated as non-attainment areas for ozone. By 2002, 157 of those counties had moved from non-attainment to maintenance status.[141] Similar progress has been made for the other five initial criteria air-quality standards.[142]

In contrast, the U.S. Clean Water Act, which also set a specific, challenging goal of zero discharge to the navigable waters of the United States, lacked a similarly powerful automatic trigger to compel attention to and measure progress toward the goal. As a result, despite evident cleanups in some water bodies, EPA cannot similarly report where and by how much water quality has improved across the nation. That information gap is finally changing as EPA responds to the requirements of the Government Performance and Results Act and a spate of local lawsuits requiring EPA or states to set targets for the amount of discharge specific water bodies can tolerate.[143]

interactive inquiry to identify more effective and cost-efficient approaches, and how to share what they learn broadly with others.

Government agencies have also used grants successfully to entice states, localities, and sometimes nonprofit organizations to feed data into federal systems. Grant-induced data collection efforts are often beset with debate about data standards and data quality, especially when a state already had its own system in place before applying for a federal grant. Nonetheless, especially when incentive-offering agencies require submission of outcome data, common data about performance and other program characteristics strengthen analyses to detect successes and problems, contributing to performance and accountability increases.

Instances Where Penalties Are Effective

So far, the discussion has concentrated on positive extrinsic incentives, those offered as a reward. Incentives do not always come in a positive form. Extrinsic incentives can also be structured as punishment. Penalties such as the withholding of promised payments or public criticism can effectively deter certain performance-reducing practices. Agreement with and understanding of the need to avoid a practice or outcome makes penalty threats more effective, tapping into and boosting intrinsic performance drivers instead of driving them out. Researchers have found that penalties work most effectively when the "penalty threatener" has high prestige and those at risk of being penalized know clearly what is expected of them. The threat of a penalty can also work well as a way to compel commitment from others, whether an individual or organization, to a specific goal.

Penalties have proven an effective mechanism for ensuring that for-profit firms meet goals (minimum standards) established by government when granting a permission to operate (regulatory permit) as well.[144] Sometimes, especially when stretch goals are set, meeting the goal requires private sector experimentation. Penalty threats have also proven effective for compelling private sector measurement and public reporting.

Instances Where Penalties Create Problems

Penalties create problems when used without explanation of why something should be prevented. They do not work well when a large proportion of those at risk of being penalized disagree with what they are being asked to do. Even in those circumstances, however, the threat of a penalty can bring about compliant responses. The motivation to comply tends not to last, however, beyond the penalty threat.

Among colleagues, it is more effective, at least initially, to encourage adoption of certain behaviors and avoidance of bad ones than to threaten negative consequences. Collegial pressure on members of a group works better if it is non-coercive, because coercive pressures make the person threatened too concerned about the consequences of failure and how to avoid them than about how to improve his or her performance.[145]

When penalty threats play a more prominent role than constructive feedback, they can ignite a confrontational relationship, impeding what might otherwise have been a performance-improving collaboration. One way to mitigate this possibility is to delineate clearly the conditions or behaviors that will be penalized and complement that information with constructive feedback and interactive inquiry. When this is done, penalties function more as a border guard, signaling the location of the "do not cross" line. Used this way, penalties play a useful backstop role to constructive feedback.

Who Gets Rewarded or Penalized?

Performance-driving, accountability-enhancing incentives require good decisions not just about the nature of the incentive, but about the intended incentive recipient. Should incentives go to individuals, teams, or whole organizations? Should everyone share in rewards or just those that have the best performance? Answering the "who" questions well, pertaining to cooperation, comparison, and the risk orientation needed for different kinds of work, can lessen incentive problems.

Reward the Group, Not the Individual, When Cooperation Is Needed

When cooperation is needed, rewards given to individual people or organizations rather than to the whole contributing team tend to irk the unrewarded. Consider, for example, a government agency whose mission is reducing homelessness. To be successful, the agency needs assistance from the housing and social service agencies, as well as the general counsel and contracting offices. Yet if the homeless agency is the only unit rewarded for reducing homelessness, the spirit of cooperation is not likely to last long. When cooperation is needed, group rewards make sense.

Rewarding individuals or single organizations can exacerbate unhealthy credit-claiming problems, especially when attribution is difficult. If people know that their individual contribution must be recognized to be rewarded, they must devote energy to attention getting and credit claiming, or risk not being fairly rewarded for their own work. This diverts intellect and time from improving outcomes to improving personal position. Group rewards are superior to individual rewards when cooperation is needed.

Compare with Care

The measurement section explored how comparison can illuminate, identifying benchmarking models and problem areas and providing valuable performance-improving insights for both organizations and individuals. At the same time, comparison can make people feel less effective than they previously believed themselves to be, curbing rather than inspiring their inclination to do better. When used as the basis for rewards, comparative measurement problems are exacerbated. If people feel their relative ranking or their comparative compensation is too low relative to others they deem not as effective, they often adjust their work effort downward to match compensation received.[146]

Discouragement also sets in if the comparisons used as the basis for rewards are deemed unfair in terms of distributive and procedural justice. Worker satisfaction with jobs and pay levels depends on a sense that distributive justice (the fair allocation of rewards) has been met. The feeling that the process used to make reward allocation decisions (procedural justice) is fair is also important. Workers' attachment to the organization and their feelings about their bosses is lowered when they feel procedural justice has been violated.[147]

Reward Quotas Cause Many Problems

Reward quotas limiting the number of people in an organization who can receive bonuses can further discourage cooperation. Rewarding only a few performers based on relative ranking interferes with needed cooperation because one team member's gain is another's loss. An incentive rewarded to one person or organization comes at the expense of others and limits the inclination to be mutually supportive. Moreover, organizations with reward quotas are not likely to be the kind of workplace where the "best and brightest" want to work, because it artificially limits the number who can be recognized for their good work, even when most workers are outstanding.[148]

This argues strongly against setting quotas limiting the percentage of high performers who can earn the best rewards. Research suggests that these systems will also frustrate the large numbers of workers who are above average but not quite good enough to make the cutoff for the best performers.[149] Reward quotas may work well for organizations that need a few superstars, where everyone else can be mediocre. They do not work well, however, for organizations that want to attract and keep large numbers of strong performers. And even in star-seeking systems (such as those wooing the best athletes and investment managers), very high rewards for the very few can result in wasteful competition if too many seek star status, absorbing motivational and training resources more productively directed to other endeavors.[150]

Rank-and-Yank Not for Permanent Government Workers

Ranking to penalize those at the bottom has been advocated as a way to identify individuals with the worst performance for employment termination. This "rank and yank" approach, advocated and named by former General Electric CEO Jack Welch, calls for identification and then termination of the bottom 10 percent of the workforce each year, replacing that 10 percent with the best available applicants. A simulation model of this approach, which does not incorporate likely morale and productivity effects, concludes that this system will work in the private sector, but even there only for a short period.[151] Given the greater difficulty of terminating employees in the public sector, even testing a rank-and-yank approach would make sense in government only for probationary and temporary employees.

Default Position: Limit Use of Individual Incentives

The evidence leads to a surprising, non-intuitive conclusion. Individual incentives are seldom likely to work in government. Except in a few specific types of situations, it is far preferable to reward everyone or to reward no one beyond basic pay, especially when cooperation is needed among individuals in an organization. It is far too difficult to measure individual contributions and compare performance, and far too costly in terms of worker discouragement, to make performance-based incentives for individuals a sound practice. Where breakthrough performance and risk taking is sought, opt-in contests can be used to attract those motivated by comparison and competition. When cooperation is needed among those being compared or when a non-trivial proportion of the workforce is likely to be discouraged by comparison, individual rewards and those based on relative position should be avoided.

Rewarding no one or everyone does not imply, it is important to note, that individual feedback should not be provided. As noted earlier, individual and organizational feedback, delivered constructively, directly contributes to performance improvement, infusing with life the motivation, illumination, and communication power of goals and measures.

Design Extrinsic Incentives to Match the Job Requirements

Which situations warrant performance-linked incentives for individuals? One strategy for private sector performance-based pay schemes for individuals that may hold promise for the public sector is "reward the extremes, muddle the middle."[152] Baron and Kreps, in a textbook on private sector human resource management, offer three conceptual categories of jobs worth distinguishing for incentive purposes: *stars, guardians,* and *foot soldiers*. *Stars* hold jobs with significant upside potential and limited downside risk: "where a bad performance isn't too bad, but a good performance is very good for the firm." Venture capitalists and researchers fall in this category. *Guardians* hold the opposite sorts of jobs, where "a bad performance is a disaster but a good performance is only slightly better for the firm than an average performance." Airline pilots fall in this category, and, arguably, so do federal workers who review products for safety before allowing them to be sold on the market. Those covering the territory in between are labeled *foot soldiers*.[153]

Baron and Kreps advise different recruitment, training, evaluation, and incentive systems for these three types of jobs. They urge, for example, that organizations structure job assignments that involve high-cost risks in ways that minimize the risks of failure, such as by introducing redundancy into the system when more people doing the same tasks can

Incentives

Extrinsic incentives drive out intrinsic ones and therefore need to be employed with great caution and care. Because they so often backfire when used inappropriately, they should be used only in very limited circumstances.

When Extrinsic Incentives Work
Extrinsic rewards linked to attainment of a goal or specific measurement level work when rewards are large enough and:

- Goals are simple, easily measured, and clearly associated with an individual.
- Goals are simple, easily measured, and the rewards are shared fairly among all who contributed to goal attainment.
- The revenue or cost-savings stream used to fund rewards is not finite, but replenished by outcome or productivity gains.
- They are used to attract people and organizations into contests to establish breakthrough performance watermarks, raising the bar on what constitutes a challenging goal for other high achievers and creating a "trickle up" effect on others' performance.
- They are used to attract others to adopt specific outcome targets and to report measures to the public or a central repository.

Penalties, when they are large enough, work to encourage:

- Goal adoption.
- Measurement generation.
- Reporting to the public or a central repository.
- Adoption of effective goal-attainment practices when effective attainment methods are known.
- Experimentation when effective attainment methods are not known.

When Extrinsic Incentives Do Not Work
Extrinsic rewards linked to attainment of a specific goal or measurement level tend to discourage workers and often trigger dysfunctional, self-protective responses when:

- Rewards are given to individuals rather than everyone who contributed, when cooperation is needed.
- The number of rewards is limited
  - and therefore interferes with cooperation because people need to compete with each other to earn a reward, and
  - necessitates comparison across multiple dimensions of performance, which in turn requires subjective rather than evidence-based judgments about the relative import of each dimension of performance
- The value of rewards can be easily compared (as is the case with financial rewards) and self-perceptions of performance exceed the perceptions of those determining who gets rewarded.
- Differences of opinion exist about the relative value of others' contributions.
- Goals are complicated.
- Goal attainment is highly uncertain and depends on factors that are difficult to influence through agency action.
- Goal attainment necessitates experimentation because knowledge about effective intervention is so limited.
- A long lag time exists between agency action and societal outcomes.
- An agency goal involves the prevention of rare events or other difficult-to-measure outcomes.

Penalty effectiveness is compromised by:

- A lack of understanding of the importance of the goal (especially minimum standards that trigger penalties).
- A penalty-threatening entity that is weak, distrusted, or otherwise not respected.

reduce the risk of a high-cost problem. For stars, they suggest that "performance evaluations should recognize outstanding performance and treat failure as being close to on par with mediocrity."[154] They note, even further, "If it can be demonstrated that 'failure' entailed some chance of success, failure might even be evaluated ahead of conservative mediocrity, to encourage risk taking."[155] In contrast, measurement systems for guardian jobs should primarily focus on unacceptably low performance, taking care not to associate top ranking with risky behaviors.[156] They also warn against combining star and guardian jobs, where possible, because of the difficulty of structuring effective incentives.[157]

Organizational Rewards Not as Sensitive to Comparison, but Caution Still Required

Many organizations, most notably businesses, often respond favorably to monetary incentives linked to comparative performance, striving both to win customers and to survive by demonstrating past performance levels better than their competitors. An increasing number of public sector organizations, as customers, are also successfully using financial incentives to reward contractors for strong past performance, delivering the incentive in the form of new contract awards.[158] Whether or not rewards delivered, after the fact, to top-performing government organizations can similarly motivate performance improvement is less clear from the evidence, because no examples of this sort of arrangement were found during the preparation of this report. Rewards delivered before the fact, in the form of a grant, can be an effective tool, however, which one organization can use to attract other organizations to adopt new outcome-focused goals or cooperate to meet a shared goal.

However, problems can arise when incentives are linked to organizational performance comparisons, as noted earlier in the discussion of the way measurements inform customer or electoral choice. Those threatened by measurement systems (for example, by the loss of customers or an election) often try to dismantle them. For incentives to drive improved performance in these situations, those who want comparative performance measurements need to assemble and maintain greater, or at least better organized, political power than those being measured.

What Gets Rewarded or Penalized?

One of the biggest challenges in establishing a performance-linked incentive system is deciding what to reward. Should incentives be linked to outcome targets, output targets, or something else? And should they reward outcome attainment, progress toward the target, comparative position, or something else? It turns out that the smartest approach for most government agencies, perhaps surprisingly, is not linking incentives to attainment of or progress toward either outcome or output targets. Instead, as the arguments presented below suggest, the most effective and fair system may be one that holds organizations and the people in them accountable for setting outcome-focused goals, measuring progress toward them, and using goals and measures to manage.

Target Attainment, Progress Toward Target, or Something Else?

Linking rewards and penalties to goal attainment is unfair and infuriating when individuals or organizations lack the skills, resources, or authority to meet (or make progress toward) their targets and the means to secure those inputs. Indeed, some missed targets are inevitable in healthy, discovery-provoking, risk-tolerating enterprises. If individuals and organizations meet all their targets all the time, it suggests that they chose timid targets and missed the performance-driving power of challenging goals.

Outcome Targets: Adoption Essential, but Attainment Should Not Be Tied to Rewards or Penalties

Many circumstances make it infeasible and unfair to reward (or penalize) people or organizations for meeting (or failing to meet) outcome targets. These include the inability to control external factors that significantly affect outcomes; long lag times between outcome attainment and action; outcomes that are hard to measure because they occur as infrequent, unwanted events; and attribution problems associated with production interdependencies.

External factors can be influenced, but not controlled. Government cannot control all factors influencing outcomes, and therefore linking incentives to outcomes necessarily includes a significant gambling component. Rewarding those that attain specific outcome levels favors not just those whose

Advance Rewards for Delayed Outcomes

Research on private sector software developers concludes that for outcomes with long lag times and attribution problems, linking rewards to outcomes does not work. Instead, the advice is to invest in extensive screening to find and hire the right people for the job, pay them well to show management confidence in their work, provide stock shares in the company so they share in long-term returns, and let them do their work without trying to measure individual performance.[159]

actions contributed to performance gains but also those who were just plain lucky. Conversely, penalizing those that fail to meet their targets unfairly hurts the hapless, even if they worked very hard.

Those that set stretch targets, the sort that stimulate innovation, are at a special disadvantage if incentive systems penalize the non-attainment of a target. In addition, rewarding outcomes can be a problem when outcomes vary significantly from year to year because of uncontrollable factors such as a group of particularly high- or low-achieving students in a class one year or weather conditions. (Year-to-year variance problems can be mitigated to some extent by linking incentives to multi-year averages of outcomes.[160])

Long lag times preclude timely incentives. It is also difficult to reward outcome attainment when a long lag time occurs between the societal outcomes sought, such as reducing the incidence of a disease whose causal factors are not yet understood, and the actions taken to influence them. In such situations, payouts or penalties would occur long after they are needed to have a motivational effect. Long lag times characterize many government programs, including those supporting basic research, software development, efforts to address problems with long gestational periods, economic development, and arguably even education programs (if self-sufficiency is seen as the ultimate objective). Publicly traded companies accommodate this problem with stock and options distributions; public organizations have no similar way to deliver a share in future returns.

Unwanted events are hard to count. Incentives linked to performance outcomes can similarly be difficult when the societal objective is to prevent unwanted events, such as terrorist attacks, fires, outbreaks of insect-linked diseases, or traffic accidents. How do you measure what did not happen? Some programs have learned how to measure common unwanted events, test different risk-reduction approaches, measure their effectiveness, and implement increasingly cost-effective approaches. These programs measure outcomes as reductions in unwanted events. When these types of programs mature, and succeed in managing unwanted events down to a low incidence, maintaining a low incidence level constitutes program success. While it is feasible to reward attainment of a steady performance level every year, steady-state accomplishments tend to be undervalued when compared to areas with more significant performance gains. This becomes problematic when rewards are limited in number and their allocation is competitively determined.

Attribution tensions interfere with cooperation. Attribution problems arise when incentives are offered to individuals or organizational units for improving outcomes that depend upon multiple parties for success. Two types of attribution problems are common, credit claiming and credit avoidance, both of which interfere with the cooperation needed to improve outcomes.

Much of government's work requires cross-office and often cross-organizational cooperation to improve outcomes. Increasing seat belt use, for example, requires cooperation among school principals, the local police, state legislators, governors' safety advisors, Department of Transportation regional offices, and NHTSA staff. If incentives are offered for raising seat belt use but not shared among everyone involved, how much of the performance increment credit should each party claim? Efforts to parse out credit for accomplishments delivered through a suite of partners quickly lead to artificial measurement machinations, with no discernible performance or accountability gain. And once an assessment system is seen as artificial or unfair, it discourages the very parties the system seeks to motivate. It can also interfere with a dynamic learning system where all parties readily share the information they collect, so strategies can be quickly revised based on the latest evidence.

Ironically, while some try to claim credit, others seek to avoid it. Some top performers, faced with

Fear of Turf Wars Leads to Outcome Avoidance and Affinity for Output Targets

In the author's work with state environmental agencies, very bright, experienced, and committed senior compliance and enforcement officials have been reluctant to adopt environmental outcome targets, such as a water-quality goal, despite evidence that this approach worked surprisingly well in another environmental agency, revealing previously unknown violators. Their reluctance stemmed not from their inability to control all factors affecting water quality, but from a concern that others in their agencies, including the permit and rule writers, would view their adoption of an outcome-focused goal as turf invasion.

One consequence of this stovepipe mentality is that, historically, few in environmental agencies pay attention to whether or not they are improving water quality. Operating units with the most staff resources instead traditionally focus on specific activities assigned to them, such as permitting, inspections, rule writing, and enforcement. Agency planning and science offices measure environmental outcomes such as water quality and human health, but, in the words of a local environmental regulatory official when asked about using his state's health tracking data, "There is little connection between that and the work we do."

comparative rewards, lessen their efforts lest they alienate co-workers with whom they work on a regular basis. Some are embarrassed by higher bonuses. Others are reluctant to consider issues beyond those clearly considered their own territory, fearful they will be seen as turf encroachers. The result is disjointed production lines that lack the glue needed to meld component parts into a coherent, functional whole that improves outcomes. Interactive inquiry meetings, with outcome-focused goal leaders and team members, can address these concerns, while incentives tend to exacerbate them.

Outputs Targets Not a Useful Substitute

Because of all the problems that make it difficult to link incentives to outcome targets, many agencies have done what New Zealand originally did. They have adopted output targets, such as clients served, requests processed, permits issued, and trainings offered, instead of focusing on outcomes. Yet anything other than outcome targets tends to distort behavior in ways that do not improve either societal outcomes or accountability.

A classic piece of organizational behavior research conducted by Blau in 1963 documented how output indicators interfere with outcome gains.[161] Blau studied an employment office with a mission of serving "workers seeking employments and employers seeking workers."[162] The employment office initially evaluated its own workers' performance by counting the easily tallied number of interviews conducted. Blau found the output-oriented assessment triggered serious dysfunctional distortions:

> The interviewer's interest in a good rating demanded that he maximize the number of interviews and therefore prohibited spending much time on locating jobs for clients. This rudimentary statistical record interfered with the agency's objective of finding jobs for clients in a period of job scarcity.[163]

The employment office subsequently adjusted its program performance indicators, substituting a set of eight indicators for the one, including an outcome indicator, job placement. The shift to outcome measurement successfully motivated desired changes in worker behavior, resulting in higher placement rates, although it also created some measurement manipulation problems. A subsequent effort to add many more indicators confused workers because it lacked specific targets to communicate priorities.

Blau's research and many additional studies since then have documented the exceptional difficulty of picking the right outputs to measure without triggering behavioral changes that compromise the pursuit of the outcomes government agencies seek to achieve. When situations vary and workers or organizations need to select from among a combination of interventions to improve outcomes, linking incentives to a small set of output indicators tends to distort choices to favor actions that are measured, thereby interfering with outcome gains.

Output indicators can work in some situations, most notably when a clear link has been established between a particular activity, such as the adoption of motorcycle helmet use laws, and outcomes. They can also be useful when the outputs themselves are a component part of the outcomes, such as pothole repairs. Even with pothole targets, however, failure to measure the outcome attributes of the repairs, such as durability, along with the output indicators is likely to bias activity choices toward what is being measured. Milestone output targets are also useful for infrastructure and other singular objectives with long lead times, such as letting a contract for or completing a section of a major construction project. In addition, output targets can be valuable for certain essential measurement infrastructure activities that support outcome measurement and ultimately the identification of what works and what doesn't, such as building a database and training people how to use measurements.

A Solid Alternative: Link Incentives to the Use of Outcome-Focused Goals and Measurement Mastery

Incentives linked to outcome targets create numerous performance-dampening, accountability-cramping problems. These do not get fixed by linking incentives to output targets, most of which divert attention and effort away from improving outcomes.

There is an alternative, suggested by CompStat and other successful performance management efforts. It requires a change in accountability expectations. Instead of linking incentives to meeting or making progress toward outcome targets, it suggests the use of incentives to motivate attention to outcome-focused targets, including the establishment of targets when they have not already been set, and to the mastery of the full array of measurements that can illuminate progress and problems. The CompStat experience suggests how powerful and motivational outcome-focused goals and measures can be without any explicit link between target attainment and reward or punishment. To reiterate the message attributed to former New York City Police Commissioner William Bratton, managers did not get in trouble if the crime rate went up in their precinct, but were taken to task if they did not know *why* it had gone up and did not have a plan to deal with it.

Bratton clarified his accountability expectations for all his managers. He did not expect target attainment. He did expect the use of specific objectives, measurement mastery (timely and accurate intelligence), effective strategies and tactics, rapid deployment of personnel and resources, and relentless follow-up and assessment.[164] And through Compstat meetings, managers provided feedback and stimulated interactive inquiry.

Summary and Implications

In sum, extrinsic incentives in the form of externally promised financial rewards and threatened penalties can motivate, but they can also discourage and trigger dysfunctional responses, including performance-dampening, accountability-reducing responses such as outcome avoidance, measurement manipulation, and timid targets. Extrinsic incentives drive out people's intrinsic motivators and complicate interpersonal comparison problems. They can work for individuals in a limited number of circumstances, most notably for simple tasks where the motivated performance gain generates revenues that fund the rewards, but these circumstances occur infrequently in the public sector. Extrinsic incentives can be more effective for organizations, especially when used to win attention to a goal, participation in a measurement system, and participation in performance-improving experiments. Even with organizations, however, extrinsic incentives have to be used with care, because they can motivate measurement games instead of performance improvement.

When extrinsic incentives are used, rewards should be shared among all who contributed to progress to encourage cooperation. Comparative performance measurement can be a powerful tool for detecting successes and problems, but using it as the basis for reward or penalty is likely to lessen its accuracy and thereby compromise its illumination value. Therefore, government agencies should adopt a reluctant approach to the use of incentives both for individuals and for organizations, especially since goals and measures, accompanied by feedback and interactive inquiry, are such powerful motivators and performance drivers on their own.

Six Essential Practices for Using the Building Blocks Effectively

Extrinsic incentives can work, but they can also backfire, depressing both performance and accountability. They can rob goals and measures of their ability to stimulate the kind of effort and innovation that results in continual, sometimes dramatic, improvements in societal conditions. And, they easily provoke unproductive fears that interfere with improvement efforts, especially when accountability expectations are left vague. What this suggests is that less attention should be paid to incentives and far more to ensuring the active and effective use of outcome-focused goals and measures. What it also suggests is a need for public organizations to clarify accountability expectations both with those being held accountable and with those holding them to account, including supervisors, legislators, budget offices, grant-giving organizations, delivery partners, and the public. Specifically, government organizations and their managers should be held accountable for six essential practices:

- **Emphasizing outcomes, using specific targets:** Focus on specific outcome-focused goals or targets, a few of which are challenging; set targets when they have not been externally set; and use targets to communicate relative priorities in all areas.
- **Measurement mastery:** Measure progress and other factors affecting progress and accountability, communicate it broadly, and discover what the measurements reveal. This is accomplished by organizing and studying the data to look for patterns, anomalies, changes, and relationships to find what works, what doesn't, causal connections, and where more understanding is needed.
- **Delivering feedback:** Provide feedback to those trying to reach targets so they stay focused on them, believe in their own abilities to meet them, and think about specific ideas and practices that will help them reach their targets.
- **Assuring an ongoing venue for interactive inquiry:** Encourage interactive inquiry to engage others with expertise and resources in delivering feedback, action planning, and implementation; stimulate synergistic thinking; and facilitate coordination and collaboration, usually through regularly scheduled meetings.
- **Cogent strategies:** Develop cogent long-term strategies and shorter-term action plans (not necessarily written plans) based on the best available evidence and ideas.
- **Implementation:** Implement the strategy and action plans, with ongoing revision based on frequent and timely review of experience.

With this sort of approach, incentives are best used to set new benchmarks, enlist goal allies, and recruit measurement contributors, while penalties are most appropriate for those who do not stay focused on their goals, measure and analyze progress, develop cogent strategies and action plans based on available evidence, implement the plans and revise plans and actions as needed. If those being held accountable in government, whether individual managers or organizations such as grant recipients, try what seemed like a sound strategy when adopted and it subsequently fails, penalties are not needed or appropriate unless the goal is abandoned and cogent strategies neither adopted nor executed. Penalties may be appropriate in some cases—for example, with for-profit firms regulated by government if minimum standards agreed to as a condition of operation are not met, especially if the means for meeting the goal are well known or if promised measurements have not been delivered. If the means are not known, however, it may even be appropriate to hold off penalizing private firms that fail to meet their targets. This would be the case so long as the firms stayed focused on their goals, frequently measured and reported progress, analyzed and shared lessons revealed by the measurement, engaged in frequent interactive inquiry not only in-house but with the government and neighbors, developed cogent strategies based on the best available information, executed the strategies, and revised them as suggested by the evidence.

Will this prescription for performance management work in practice? Evidence from numerous government agencies, some of it included in this report, suggests it can not only work, but work in a powerful way with great outcome and accountability returns. It is evolving in the police, corrections, parks, health, homeless, and welfare departments of New York City, as well as in the U.S. Coast Guard and the National Highway Traffic Safety Administration. Recent changes in New Zealand suggest its public management reform efforts are evolving in this direction as well, with a new emphasis on managing for outcomes (MFO) that no longer links pay directly to specific targets.[165] The Tennessee Value-Added Assessment System suggests a possible way to measure performance without unduly threatening those being measured by placing the emphasis on feedback rather than on punishment. Promising developments are by no means limited to those mentioned in the body of this report. They are emerging at the Indian Housing program at the Department of Housing and Urban Development[166] and at the Maternal and Child Health Bureau in the Department of Health and Human Services, which requires all grantees to report on 18 outcome-focused performance indicators but places primary emphasis on data sharing, feedback, and problem solving.

Promising developments are also evident in the Program Assessment Rating Tool, or PART, developed by the Office of Management and Budget. It appropriately gives federal agencies credit for many of the practices recommended here, including setting specific and challenging goals, emphasizing outcomes, and mastering measurement. Two aspects of the PART need to be revised, however, to lessen the likelihood of dysfunctional responses: (1) scoring agencies low for not meeting targets when programs have set challenging, outcome-focused targets, measured progress, and implemented what seemed like a sensible strategy at the time the target was set; and (2) scoring agencies low for not meeting targets they cannot control because of legislative barriers, even when an agency has proposed corrective legislation to the White House.

There are promising developments on the political front, as well. A small but increasing number of elected executives have boldly announced outcome-focused targets with specific quantities and deadlines, openly reported progress and problems, and won re-election despite missed targets. And while most legislative decision makers at the federal level have ignored formal GPRA and PART documents, many have paid attention to agency goals and outcome measures when delivered in a format that is relevant to them.

With a performance management approach that is outcome focused, measurement rich, and inquisitive but non-punitive, outcomes improve and accountability rises. Outcomes rise because goals focus and motivate, measures reveal what works and what doesn't, and feedback and interactive inquiry inspire, inform, and engage. Accountability (and democracy) increases because public articulation of specific goals clarifies what each organization will do and what it will not do, allowing citizens and their elected representatives to determine if the organization is doing what they want it to do and inviting them to use their electoral and administrative process voices to respond if they disagree. Outcomes and accountability also rise because goals, measurement, strategy transparency, and interactive inquiry encourage intelligent, honest, and diligent efforts.

Acknowledgments

I would like to thank Mark Abramson and John Kamensky of the IBM Center for The Business of Government for their advice and guidance throughout the writing of this report.

Endnotes

1. Robert D. Behn, "Linking Measurement and Motivation: A Challenge for Education," *Advances in Educational Administration,* Vol. 5 (1997), pp. 15–58, in *Rethinking Democratic Accountability* (Brookings, 2001). Behn offers an answer to the question, "What does it mean to be held accountable?" I suggest, in this report, a different answer.

2. Brian Friel, "No Results, No Bonus" (*GovExec.com,* August 2, 2001) and correspondence with Friel, January 2006.

3. Donald Kettl, University of Pennsylvania, quoted in Friel.

4. Robert D. Behn, "The New Public-Management Paradigm and the Search for Democratic Accountability," *International Public Management Journal,* 1(2), 1999, pp. 131–165.

5. David M. Walker, "GAO Answers the Question: What's in a Name?" Roll Call, July 19, 2004. http://www.gao.gov/about/rollcall07192004.pdf (accessed December 26, 2005).

6. Walker.

7. Walker.

8. President John F. Kennedy, "Special Message to the Congress on Urgent National Needs," May 25, 1961.

9. Edwin A. Locke and Gary P. Latham, "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey," *American Psychologist,* pp. 705–717.

10. Gary P. Latham, "Motivate Employee Performance through Goal-setting," in *The Blackwell Handbook of Principles of Organizational Behavior,* ed. Edwin A. Locke (Blackwell, 2004), p. 108.

11. Alvin Zander, *Motives and Goals in Groups* (New Brunswick: Transaction, 1996), p. xvi.

12. Zander, pp. xvi, xxvi.

13. Locke and Latham, p. 706; Zander, p. xxiv.

14. Locke and Latham, pp. 707–708.

15. Locke and Latham, pp. 706–707.

16. E-mail exchange with Edwin A. Locke, March 26, 2006.

17. Zander, p.xxiv.

18. Locke and Latham, pp.706–707.

19. U.S. Department of Transportation Performance Plan (FY 2001) and Report (FY 1999).

20. U.S. Department of Transportation FY 2003 Performance Plan (May 2002). http://www.nhtsa.dot.gov/nhtsa/whatis/planning/perf-plans/2003/Index.html (accessed December 6, 2005).

21. U.S. Department of Transportation Performance and Accountability Report. http://www.dot.gov/perfacc2005/parpart1.htm#safety (accessed March 11, 2006).

22. U.S. Department of Transportation FY2004 Performance Plan. http://www.dot.gov/PerfPlan2004/safety_highway.html (accessed March 11, 2006).

23. U.S. Department of Transportation FY 2003 Performance Plan.

24. Ohio Environmental Protection Agency, Summary of 2001 Enforcement Performance (Columbus, Ohio: April 2002).

25. *Governing* blog posted by Elizabeth Daigneau, October 11, 2005. http://governing.typepad.com/13thfloor/2005/10/index.html (accessed December 30, 2005). Supplemented by author's notes from conference.

26. Robert S. Kaplan and David P. Norton, *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment* (Harvard Business School Press, 2001), p. 10.

27. Zander, p. xii.

28. For example, Mark T. Imperial, "Collaboration and Performance Management in Network Settings: Lessons from Three Watershed Governance Efforts," in *Managing for Results 2005,* ed. John M. Kamensky and Albert Morales (Rowman and Littlefield, 2005), p. 404.

29. Robert S. Kaplan and David P. Norton, *The Balanced Scorecard: Translating Strategy into Action* (Harvard Business School Press, 1996).

30. Harvard Business School strategy expert Michael Porter regularly asserts that strategy necessitates decisions not only about what an organization will do, but what it

will not do. "Sustainable competitive advantage comes from being different. What the strategic mind focuses on is: 'How can we, by limiting what we do, by not seeking to serve all customers, by not seeking to offer every product, by not seeking to enter every geography, how can we, by limiting ourselves, be unique?' That is the essential strategic question." Quoted in interview with Steve Klinkerman, "The Power of Saying No," (BAI Banking Strategies. http://www.bai.org/bankingstrategies/1998-jan-feb/power/ (accessed March 12, 2006).

In *Good to Great,* Jim Collins found that "good-to-great companies did not focus principally on what to do to become great; they focused equally on what *not* to do and what to *stop* doing." Jim Collins, *Good to Great: Why Some Companies Make the Leap and Others Don't* (New York: Harper Collins, 2001), p. 11.

31. http://aspe.hhs.gov/hhsplan/2004/hhsplan2004.pdf (accessed December 7, 2005).

32. http://www.hhs.gov/budget/06budget/FY2006AnnualPlan.pdf (accessed December 7, 2005).

33. See, for example, on media preference for negative stories, Thomas Patterson, *Out of Order* (Vintage Books, 1993).

34. Shelley H. Metzenbaum, "Measurement that Matters: Cleaning Up the Charles River," in *Environmental Governance,* ed. Donald F. Kettl (Brookings, 2002), pp. 58–117. For an update, see Shelley H. Metzenbaum, "Performance Management at the Local Level: The Case of the Charles River," in *Water Resources Impact* (Washington, DC: American Water Resources Association), July 2005, pp. 8–12.

35. U.S. Office of Management and Budget 2006 Budget, Department of Transportation PART review, NHTSA Operations and Research.

36. Latham, p. 117.

37. Locke and Latham, pp. 707–708.

38. Latham, p. 113.

39. Terence R. Mitchell and William S. Silver, "Individual and Group Goals When Workers Are Interdependent: Effects on Task Strategies and Performance," *Journal of Applied Psychology,* 75-2 (1990): pp. 185–193.

40. Zander, p. xvii.

41. Albert Bandura and Daniel Cervone, "Self-Evaluative and Self-Efficacy Mechanisms Governing the Motivational Effects of Goal Systems," *Journal of Personality and Social Psychology,* 45-5 (1983): pp. 1017–1028.

42. Locke and Latham.

43. C. Keisler. *The Psychology of Commitment* (Orlando Academic Press, 1971) in Zander, p. xi.

44. Peter M. Blau, *The Dynamics of Bureaucracy: A Study of Interpersonal Relationships in Two Government Agencies* (University of Chicago Press, 1963), p. 38.

45. Bandura and Cervone.

46. This is sometimes referred to as "contingency theory." See, for example, Carol Sansone, and Judith M. Harackiewicz, " 'Reality' Is Complicated," *American Psychologist,* Vol. 53- 6 (June 1998): p. 673. "This approach suggests that to understand and predict an event's effect on motivational outcomes, one needs to account for how the event affects individuals' initial orientation to the task (e.g., their goals, what they attend to, what they will feel is important). Just as critically, one must account for how this orientation subsequently affects individuals' actions (e.g., strategies), cognitions (e.g., cognitive interference), and affect (e.g., positive mood) while performing."

47. John M. Tauer and Judith M. Harackiewicz, "Winning Isn't Everything: Competition, Achievement Orientation, and Intrinsic Motivation," *Journal of Experimental Social Psychology* 35 (1999): pp. 209–238.

48. Tauer and Harackiewicz.

49. E. L. Deci, et al., "When Trying to Win: Competition and Intrinsic Motivation," *Personality and Social Psychology Bulletin,* 7 (1981): pp. 79–83.

50. Robert H. Frank and Philip J. Cook more fully explore some aspects of this question regarding the net value of competition in *The Winner-Take-All Society* (New York: The Free Press, 1995).

51. Uri Gneezy, Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics,* CXVIII, August 2003, pp. 1049–1074. This study finds that competition where the winner takes all leads to better performance than using piece-rate or random payments, but it also finds that women fare significantly worse than men when men and women compete against one another but not when women compete only against women. Thus, performance competitions in the workplace could lead to gender-biased outcomes.

52. Frank and Cook.

53. James N. Baron and David M. Kreps, *Strategic Human Resources: Frameworks for General Managers* (Wiley & Sons, 1999).

54. Hal G. Rainey, *Understanding and Managing Public Organizations* (Jossey-Bass, 2003), pp. 229–247.

55. Robert H. Frank, *Choosing the Right Pond: Human Behavior and the Quest for Status* (Oxford, 1985), p. 8.

56. Max H. Bazerman, *Judgment in Managerial Decision-Making* (John Wiley and Sons, 2002), pp. 65–69; David C. Matz and Verlin B. Hinsz, "Social Comparison in the Setting of Goals for Own and Others' Performance," *Journal of Business and Psychology,* 14-4 (Summer 2000): pp.563–572 (564).

57. Frank and Cook, p. 8, citing Thomas Gilovich, *How We Know What Isn't So* (New York: Free Press, 1991), chapter 5.

58. Albert Bandura, "Cultivate Self-Efficacy for Personal and Organizational Effectiveness," in Locke, 2002, pp. 121–136.

59. Judith M. Harackiewicz, Carol Sansone, and George Manderlink, "Competence, Achievement Orientation, and Intrinsic Motivation: A Process Analysis," *Journal of Personality and Social Psychology* 48-2 (1985): pp. 493–508.

60. Bandura and Cervone. The effect of self-perception on effort and performance varies depending both on how much actual performance fell short of the standard and on individual variations in the sense of self-efficacy. For example, just missing the performance levels set by others can make some people feel even more challenged, while significantly or consistently missing it by a long shot can make them feel greatly discouraged.

61. Frank.

62. Baron and Kreps, pp. 102–104.

63. William T. Gormley and David L. Weimar, *Organizational Report Cards* (Harvard, 1999); interview with Dr. William L. Sanders, September 26, 2005.

64. Shelley H. Metzenbaum, "Strategies for Using State Information: Measuring and Improving Program Performance" (IBM Center for The Business of Government, Dec. 2003).

65. Statements made at *Governing's* "Managing Performance 2005" conference by state legislators.

66. See, for example, Friends of the Earth and Environmental Working Group, "Clean Water Report Card," March 2000.

67. Jerald Greenberg, "Promote Procedural Justice to Enhance Acceptance of Work Outcomes," in Locke, 2004, pp. 181–195.

68. Shelley H. Metzenbaum, "Measuring to Comply, Measuring to Perform," a Cognos white paper, 2006.

69. U.S. Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health*. 2nd ed. (Washington, DC: U.S. Government Printing Office, November 2000). http://www.healthypeople.gov/ (accessed March 13, 2006).

70. http://www.finfacts.com/biz10/globalworldincome percapita.htm (accessed December 12, 2005).

71. George Stalk, Rob Lachenauer, and John Butman, *Hardball: Are You Playing to Play or Playing to Win* (Cambridge: Harvard Business School, 2004), pp. 39–54.

72. R. G. Ulmer and D. F. Preusser, "Evaluation of Repeal of Motorcycle Helmet Laws in Kentucky and Louisiana." Technical summary done for the U.S. National Highway Traffic Safety Administration, October 2002. http://www.nhtsa.dot.gov/People/injury/pedbimot/motorcycle/kentuky-la03/TechSumm.html (accessed March 13, 2006).

73. Metzenbaum (2002).

74. John W. Kingdon, *Agendas, Alternatives, and Public Policies* (Longman, 1995).

75. http://www.buckleupamerica.org/nmay05/nmay05.htm (accessed December 7, 2005).

76. Imperial.

77. Metzenbaum (2003).

78. Blau, p. 43.

79. See, for example, http://www.mass.gov/qrmv/boston.shtm (accessed March 13, 2006).

80. Gormley and Weimar.

81. Ginger Jin and Philip Leslie, "The Effects of Disclosure Regulation: Evidence from Restaurants," October 2001. http://www.complianceconsortium.org/StoredDocuments/Academic/HygieneDisclosureRestaurants.pdf.

82. Archon Fung, et al., "The Political Economy of Transparency: What Makes Disclosure Policies Effective?" (Harvard Kennedy School Ash Institute for Democratic Governance, OP-03-04, December 2004), pp. 19–20.

83. Joshua Benton and Holly K. Hacker, "Celebrated School Accused of Cheating; Exclusive: TAKS Results Too Good to be True at Houston Elementaries, *Dallas Morning News,* March 31, 2005.

84. Joshua Benton, "Monitors will Police HISD Tests," *Dallas Morning News,* January 12, 2005.

85. Interview with Joshua Benton, *Dallas Morning News* reporter.

86. Gormley and Weimar.

87. Based on the author's experience as Associate Administrator for Regional Operations and State/Local Relations at the U.S. Environmental Protection Agency at the time of the debate.

88. Fung et al.

89. Gormley and Weimar.

90. Gormley and Weimar, p. 44.

91. Dennis C. Smith and William J. Grinker, "The Promise and Pitfalls of Performance-Based Contracting" (New York: Seedco, June 2004), p. 14. Prepared for presentation at the 25th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM), Washington, DC, November 5–8, 2003. http://www.seedco.org/publications/publications/promise_and_pitfalls.pdf (accessed January 2, 2006).

92. Zander, p. xviii.

93. This experiment gave rise to what is known as the Hawthorne effect, where studied objects change their behavior because they are studied, because the experiment was conducted at the Hawthorne plant. George Elton Mayo, *The Human Problems of an Industrial Civilisation* (New York: Macmillan, 1933).

94. Fred Luthans and Alexander D. Stajkovic, "Provide Recognition for Performance Improvement," in Locke (2004).

95. Baron and Kreps, p. 218.

96. Zander, p. xviii.

97. Harackiewicz, Sansone, and Manderlink.

98. Blau, pp. 44–45.

99. Latham, pp. 109–113.

100. Lisa Stark, "Putting the Brakes on Reckless Teen Driving; Study Uses Cameras in Car to Show What Teens Are Really Doing," Manakato, MN, *World News Tonight,* ABC News Internet Ventures, Dec. 5, 2005.

101. Zander, p. xviii.

102. Robert Simons, "Control in an Age of Empowerment," *Harvard Business Review,* March-April 1995, pp. 80–88.

103. Zander, p. xviii.

104. Simons.

105. Dennis C. Smith and William J. Grinker, "The Transformation of Social Services Management in New York City: 'CompStating' Welfare" (New York: Seedco, 2005). Paper presented at the 26th Annual Research Conference of the Association for Public Policy Analysis and Management in Atlanta, Georgia, October 28–30, 2004. http://www.seedco.org/publications/publications/compstating_welfare.pdf (accessed January 2, 2006).

106. Harry P. Hatry, et al., "How Federal Programs Use Outcome Information: Opportunities for Federal Managers," and Imperial in Kamensky and Morales (2005).

107. Phyllis Parshall McDonald, *Managing Police Operations: Implementing the New York Crime Control Model—CompStat* (Wadsworth/Thomson Learning, 2002), p. 33.

108. McDonald, p. 34.

109. McDonald, p. 134.

110. New York City presentation at *Governing* magazine's Managing Performance 2005.

111. Author conversation with San Diego County officials.

112. Collins, p. 74–80.

113. Jonathan Walters and Charles Thompson, "The Transformation of the Government Accountability Office," IBM Center for The Business of Government, 2005.

114. Kay Coles James, Director, Office of Personnel Management, Memorandum for Heads of Departments and Agencies, December 17, 2003. http://www.opm.gov/oca/compmemo/2003/2003-19.asp (accessed December 29, 2005). The National Defense Authorization Act for Fiscal Year 2004 (Public Law 108-136, November 24, 2003) established a new performance-based pay system for members of the Senior Executive Service effective on the first day of the first applicable pay period beginning on or after January 1, 2004 (January 11, 2004, for most employees).

115. Harry P. Hatry, et al., "Issues and Case Studies in Teacher Incentive Plans," Urban Institute, 1994. More recently, see Maria Sechetti and Tracy Jan, "Romney Wants Teacher Merit Pay, *Boston Globe,* September 22, 2005, p. A1.

116. No Child Left Behind Act of 2002. If a school is deemed non-performing after two years, it must allow parents to transfer their children to other public schools. It also must develop a school improvement plan and spend 10 percent of its Title I allocation on teacher professional development. After three years, schools must provide eligible students with "supplemental services," which generally means tutoring. After four years, schools must take "corrective action," such as replacing school staff, adopting new curriculum, or extending the school day. After five years, schools face takeover by the state or a contracted private education firm. See http://www.stateline.org/live/ViewPage.action?siteNodeId=136&languageId=1&contentId=41610 (accessed March 13, 2006).

117. Hatry, et al. (1994), p. 232.

118. John Kamensky, "Performance-Based Pay in the Federal Government: An Historical Overview," in Howard Risher, "Pay for Performance: A Guide for Federal Managers," IBM Center for The Business of Government, November 2004, pp. 8–9.

119. Gormley and Weimar; Fung et al.

120. Graham Scott, Ian Ball, and Tony Dale, "New Zealand's Public Sector Management Reform: Implications for the United States," *Journal of Policy Analysis and Management,* Vol. 16 (Summer 1997): pp. 357–381.

121. Colin James, "The State Ten Years On from the Reforms," Occasional Paper No. 6, State Services Commission, October 1998. http://www.ssc.govt.nz/display/document.asp?NavID=82&DocID=2834 (10 of 11) (accessed December 29, 2005).

122. Gerald Marschke, "The Economic of Performance Incentives in Government with Evidence from a Federal Job Training Program," in *Quicker, Better, Cheaper: Managing Performance in American Government,* ed. Dall W. Forsythe (Albany: Rockefeller Institute Press, 2001), p. 83.

123. Carolyn J. Heinrich, "Setting Performance Targets: Lessons from the Workforce Investment Act System," in *Managing for Results 2005,* ed. John M. Kamensky and Albert Morales (Lanham, MD: Rowman & Litttlefield, 2005), p. 370.

124. Heinrich, pp. 360–361.

125. Brian A. Jacob, "Accountability, Incentives, and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools," NBER Working Paper 8968 (Cambridge: National Bureau of Economic Research, June 2002).

126. Marschke, p. 83.

127. Collins, p. 49.

128. E. Deci, “Effects of Externally Mediated Rewards on Intrinsic Motivation,” *Journal of Personality and Social Psychology,* 18 (1971): pp. 105–115.

129. Rainey, pp. 234–236.

130. Locke and Latham, p. 708.

131. Iris Bohnet and Susan C. Eaton. “Does Performance Pay Perform? Conditions for Success in the Public Sector,” in *For the People: Can We Fix Public Service?* ed. John D. Donahue and Joseph S. Nye, Jr. (Brookings, 2003), pp. 238–254.

132. Baron and Kreps, pp. 243–245. When the company moved to the piece rate, the rate was set so that an installer completing an average number of jobs in a week would be better off under, and paid according to, the old hourly wage formula.

133. Material from David R. Smith, Mike Schaiberger, and Diane Golat, Maricopa County. All employees (excluding temporary and initial probationary employees) who have been employed by MCDOT for at least six consecutive months share in the savings.

134. United States General Accounting Office Report to Congressional Requesters, January 2004, “Human Capital: Implementing Pay for Performance at Selected Personnel Demonstration Projects,” 04-83, p. 33.

135. Kurt Eichenwald, *Conspiracy of Fools: A True Story* (Broadway Books, 2005).

136. Frank.

137. Baron and Kreps, p. 103. “The existence of wide reward differentials has been shown to act as an impediment to developing close, interdependent working relations.”

138. Timothy J. Conlan, “Grants Management in the 21st Century: Three Innovative Policy Responses,” IBM Center for The Business of Government, October 2005.

139. The Clean Air Act (CAA), Section 179 (42 United States Code 7401-7626).

140. The Clean Air Act.

141. Hatry et al. (2005), p. 231.

142. http://www.epa.gov/oar/airtrends/2005/econ-emissions.html

143. U.S Bureau of Land Management, *Water Quality Law Summary,* http://www.blm.gov/nstc/WaterLaws/Chap4.html#Anchor-7638 (accessed March 14, 2006).

144. See, for example, Wayne B. Gray and John T. Scholz, “Does Regulatory Enforcement Work? A Panel Analysis of OSHA Enforcement,” *Law and Society Review,* 27-1 (1993), pp. 177–213.

145. Zander, p. xxv.

146. D. Campbell and R. Pritchard, “Motivation Theory in Industrial and Organizational Psychology,” in *Handbook of Industrial and Organizational Psychology,* ed. M. D. Dunnette (Rand McNally, 1976), pp. 63–130. Cited in Nancy Katz, “Incentives and Performance Management in the Public Sector,” note done for Kennedy School Executive Session on Public Sector Performance Management, June 2000. http://www.ksg.harvard.edu/visions/performance_management/katz_incentives.htm (accessed March 14, 2006).

147. Baron and Kreps, pp. 107–108.

148. Frank.

149. Baron and Kreps, p. 232.

150. Frank and Cook.

151. Steven E. Scullen et al., “Forced Distribution Rating Systems and the Improvement of Workforce Potential: A Baseline Simulation,” *Personnel Psychology* 58 (2005): pp. 1–32.

152. Baron and Kreps, pp. 220–238.

153. Baron and Kreps, p. 27.

154. Baron and Kreps, p. 220.

155. Baron and Kreps, p. 241.

156. Baron and Kreps, p. 220.

157. Baron and Kreps, p. 29.

158. FCS Group, “Best Practices and Trends in Performance Based Contracting,” paper prepared for the State of Washington Office of Financial Management (Redmond, WA: FCS, 2005).

159. Robert D. Austin and P. D. Larkey. “Performance-based Incentives in Knowledge Work: Are Agency Models Relevant?” *International Journal of Business Performance Management* (Spring 2000).

160. Thomas J. Kane and Douglas O. Staiger, “Volatility in School Test Scores: Implications for Test-Based Accountability Systems,” in *Brookings Papers on Education Policy 2002* (Brookings, 2002), pp. 235–283.

161. Blau, p. 38.

162. Blau, p. 38.

163. Blau, p. 38.

164. McDonald, p. 8.

165. http://www.ssc.govt.nz/display/document.asp?NavID=253 (accessed January 2, 2006).

166. Hatry et al., (2005).

Bibliography

Austin, Robert D. and P. D. Larkey. "Performance-based Incentives in Knowledge Work: Are Agency Models Relevant?" *International Journal of Business Performance Management,* Spring 2000.

Bandura, Albert and Daniel Cervone. "Self-Evaluative and Self-Efficacy Mechanisms Governing the Motivational Effects of Goal Systems." *Journal of Personality and Social Psychology,* 45-5 (1983), pp. 1017–1028.

Bandura, Albert. "Cultivate Self-Efficacy for Personal and Organizational Effectiveness." In Locke, 2004, pp. 120–136.

Baron, James N. and David M. Kreps. *Strategic Human Resources: Frameworks for General Managers.* John Wiley & Sons, 1999.

Bazerman, Max H. *Judgment in Managerial Decision-Making.* John Wiley & Sons, 2002.

Behn, Robert D. "Linking Measurement and Motivation: A Challenge for Education." *Advances in Educational Administration,* Vol. 5 (1997), pp. 15–58.

Behn, Robert D. *Rethinking Democratic Accountability.* Brookings, 2001.

Behn, Robert D. "The New Public-Management Paradigm and the Search for Democratic Accountability." *International Public Management Journal* 1(2), 1999, pp. 131–165.

Benton, Joshua. "Monitors will Police HISD Tests." *Dallas Morning News,* January 12, 2005.

Benton, Joshua, and Holly K. Hacker. "Celebrated School Accused of Cheating; Exclusive: TAKS Results Too Good to be True at Houston Elementaries." *Dallas Morning News,* March 31, 2005.

Blau, Peter M. *The Dynamics of Bureaucracy: A Study of Interpersonal Relationships in Two Government Agencies.* University of Chicago Press, 1963.

Bohnet, Iris and Susan C. Eaton. "Does Performance Pay Perform? Conditions for Success in the Public Sector." In *For the People: Can We Fix Public Service?* edited by John D. Donahue and Joseph S. Nye, Jr., Brookings, 2003, pp. 238–254.

Broadnax Walter D. and Kevin J. Conway. "The Social Security Administration and Performance Management. In Forsythe, pp. 143–176.

Campbell, D. and R. Pritchard. "Motivation Theory in Industrial and Organizational Psychology." In *Handbook of Industrial and Organizational Psychology,* edited by M. D. Dunnette. Rand McNally, 1976, pp. 63–130.

Collins, Jim. *Good to Great: Why Some Companies Make the Leap and Others Don't.* New York: Harper Collins, 2001.

Conlan, Timothy J. "Grants Management in the 21st Century: Three Innovative Policy Responses." Washington, DC: IBM Center for The Business of Government, October 2005.

Deci, E. L. "Effects of Externally Mediated Rewards on Intrinsic Motivation." *Journal of Personality and Social Psychology* 18 (1971), pp. 105–115.

Deci, E. L. et al. "When Trying to Win: Competition and Intrinsic Motivation." *Personality and Social Psychology Bulletin* 7 (1981), pp. 79–83.

Eichenwald, Kurt. *Conspiracy of Fools: A True Story*. Broadway Books, 2005.

FCS Group. "Best Practices and Trends in Performance Based Contracting." Paper prepared for the State of Washington Office of Financial Management. Redmond, WA: FCS, 2005.

Forsythe, Dall W., ed. *Quicker, Better, Cheaper: Managing Performance in American Government*. Albany, NY: Rockefeller Institute Press, 2001.

Frank, Robert H. *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford, 1985.

Frank, Robert H. and Philip J. Cook. *The Winner-Take-All Society*. New York: The Free Press, 1995.

Friel, Brian. "No Results, No Bonus." *GovExec.com,* August 2, 2001.

Fung, Archon et al. "The Political Economy of Transparency: What Makes Disclosure Policies Effective?" Harvard Kennedy School Ash Institute for Democratic Governance, OP-03-04, December 2004.

Gilovich, Thomas. *How We Know What Isn't So.* New York: Free Press, 1991.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, CXVIII, August 2003, 1049–1074.

Gormley, William T. and David L. Weimar. *Organizational Report Cards*. Harvard, 1999.

Gray, Wayne B. and John T. Scholz. "Does Regulatory Enforcement Work? A Panel Analysis of OSHA Enforcement." *Law and Society Review*, 27-1 (1993), pp. 177–213.

Greenberg, Jerald. "Promote Procedural Justice to Enhance Acceptance of Work Outcomes." In Locke, 2004.

Harackiewicz, Judith M., Carol Sansone, and George Manderlink. "Competence, Achievement Orientation, and Intrinsic Motivation: A Process Analysis." *Journal of Personality and Social Psychology* 48-2 (1985), pp. 493–508.

Hatry, Harry P. et al. "How Federal Programs Use Outcome Information: Opportunities for Federal Managers." In Kamensky and Morales, 2005.

Hatry, Harry P. et al. "Issues and Case Studies in Teacher Incentive Plans." Urban Institute, 1994.

Heinrich, Carolyn J. "Setting Performance Targets: Lessons from the Workforce Investment Act System." In Kamensky and Morales, 2005.

Imperial, Mark T. "Collaboration and Performance Management in Network Settings: Lessons from Three Watershed Governance Efforts." In Kamensky and Morales, 2005.

Jacob, Brian A. "Accountability, Incentives, and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." NBER Working Paper 8968. Cambridge: National Bureau of Economic Research, June 2002.

James, Colin. "The State Ten Years On from the Reforms." Occasional Paper No. 6. State Services Commission, October 1998. http://www.ssc.govt.nz/display/document.asp?NavID=82&DocID=2834 (10 of 11) (accessed December 29, 2005).

James, Kay Coles. Director, Office of Personnel Management, Memorandum for Heads of Departments and Agencies, December 17, 2003. http://www.opm.gov/oca/compmemo/2003/2003-19.asp (accessed December 29, 2005).

Jin, Ginger and Philip Leslie. "The Effects of Disclosure Regulation: Evidence from Restaurants." October, 2001. Available at: http://www.complianceconsortium.org/StoredDocuments/Academic/HygieneDisclosureRestaurants.pdf

Kamensky, John. "Performance-Based Pay in the Federal Government: An Historical Overview." In Howard Risher, "Pay for Performance: A Guide for Federal Managers." Washington, DC: IBM Center for The Business of Government, November 2004, pp. 8–9.

Kamensky, John M. and Albert Morales, ed. *Managing for Results 2005*. Lanham, MD: Rowman & Littlefield, 2005.

Kane, Thomas J. and Douglas O. Staiger. "Volatility in School Test Scores: Implications for Test-Based Accountability Systems." In *Brookings Papers on Education Policy 2002*. Brookings, 2002, pp. 235–283.

Kaplan, Robert S. and David P. Norton. *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business School Press, 1996.

Kaplan, Robert S. and David P. Norton. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Harvard Business School Press, 2001.

Katz, Nancy. "Incentives and Performance Management in the Public Sector." Note done for Kennedy School Executive Session on Public Sector Performance Management, June 2000. http://www.ksg.harvard.edu/visions/performance_management/katz_incentives.htm (accessed March 14, 2006).

Kelman, Steven. *Unleashing Change: A Study of Organizational Renewal in Government*. Brookings, 2005.

Kennedy School of Government. Executive Session on Public Sector Performance Management. *Get Results Through Performance Management: An Open Memorandum to Government Executives*. Harvard University, 2001.

Kingdon, John W. *Agendas, Alternatives, and Public Policies*. Longman, 1995.

Latham, Gary P., "Motivate Employee Performance through Goal-setting." In *The Blackwell Handbook of Principles of Organizational Behavior,* edited by Edwin A. Locke. Blackwell, 2004, p. 108.

Locke, Edwin A. *The Blackwell Handbook of Principles of Organizational Behavior*. Blackwell, 2004.

Locke, Edwin A. and Gary P. Latham. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey." *American Psychologist* 57-9 (September 2002), pp. 705–717.

Luthans, Fred and Alexander D. Stajkovic. "Provide Recognition for Performance Improvement." In Locke, 2004.

Marschke, Gerald. "The Economic of Performance Incentives in Government with Evidence from a Federal Job Training Program." In Forsythe, 2001.

Matz, David C. and Verlin B. Hinsz. "Social Comparison in the Setting of Goals for Own and Others' Performance." *Journal of Business and Psychology,* 14-4 (Summer 2000), pp. 563–572.

Mayo, George Elton. *The Human Problems of an Industrial Civilisation*. New York: Macmillan, 1933.

McDonald, Phyllis Parshall. *Managing Police Operations: Implementing the New York Crime Control Model—CompStat*. Wadsworth/Thomson Learning, 2002.

Metzenbaum, Shelley H. "Measurement that Matters: Cleaning Up the Charles River." In *Environmental Governance,* edited by Donald F. Kettl. Brookings, 2002, pp. 58–117.

Metzenbaum, Shelley H. "Measuring to Comply, Measuring to Perform." A Cognos white paper, 2006.

Metzenbaum, Shelley H. "Performance Management at the Local Level: The Case of the Charles River." In *Water Resources Impact*. Washington, DC: American Water Resources Association, July 2005, pp. 8–12.

Metzenbaum, Shelley H. "Strategies for Using State Information: Measuring and Improving Program Performance." In Kamensky and Morales, 2005.

Mitchell, Terence R. and William S. Silver. "Individual and Group Goals When Workers Are Interdependent: Effects on Task Strategies and Performance." *Journal of Applied Psychology*, 75-2 (1990), pp. 185–193.

The National Defense Authorization Act for Fiscal Year 2004 (Public Law 108-136, November 24, 2003).

Ohio Environmental Protection Agency. Summary of 2001 Enforcement Performance. Columbus, Ohio: April 2002.

Patterson, Thomas. *Out of Order*. Vintage Books, 1993.

Rainey, Hal G. *Understanding and Managing Public Organizations*. Jossey-Bass, 2003.

Sansone, Carol and Judith M. Har ackiewicz. " 'Reality' Is Complicated." *American Psychologist*, Vol. 53-6 (June 1998).

Scott, Graham, Ian Ball, and Tony Dale. "New Zealand's Public Sector Management Reform: Implications for the United States." *Journal of Policy Analysis and Management,* Vol. 16 (Summer 1997), pp. 357–381.

Scullen, Steven E. et al. "Forced Distribution Rating Systems and the Improvement of Workforce Potential: A Baseline Simulation." *Personnel Psychology* 58 (2005), pp. 1–32.

Sechetti, Maria, and Tracy Jan. "Romney Wants Teacher Merit Pay." *Boston Globe,* September 22, 2005, p. A1.

Simons, Robert. "Control in an Age of Empowerment." *Harvard Business Review*, March-April 1995, pp. 80–88.

Smith, Dennis C. and William J. Grinker. "The Promise and Pitfalls of Performance-Based Contracting." New York: Seedco, June 2004. Prepared for presentation at the 25th Annual Research Conference of the Association for Public Policy Analysis and Management (APPAM), Washington, DC, November 5–8, 2003. http://www.seedco.org/publications/publications/promise_and_pitfalls.pdf (accessed January 2, 2006).

Smith, Dennis C. and William J. Grinker. "The Transformation of Social Services Management in New York City: 'CompStating' Welfare." New York: Seedco, 2005. Paper presented at the 26th Annual Research Conference of the Association for Public Policy Analysis and Management in Atlanta, Georgia, October 28-30, 2004. http://www.seedco.org/publications/publications/compstating_welfare.pdf (accessed January 2, 2006).

Stalk, George, Rob Lachenauer, and John Butman. *Hardball: Are You Playing to Play or Playing to Win.* Cambridge: Harvard Business School, 2004.

Stark, Lisa. "Putting the Brakes on Reckless Teen Driving; Study Uses Cameras in Car to Show What Teens Are Really Doing." Manakato, MN, *World News Tonight,* ABC News Internet Ventures, December 5, 2005.

Tauer, John M. and Judith M. Harackiewicz. "Winning Isn't Everything: Competition, Achievement Orientation, and Intrinsic Motivation." *Journal of Experimental Social Psychology* 35 (1999), pp. 209–238.

Ulmer, R. G. and D. F. Preusser. "Evaluation of Repeal of Motorcycle Helmet Laws in Kentucky and Louisiana." Technical summary done for the U.S. National Highway Traffic Safety Administration, October 2002. http://www.nhtsa.dot.gov/People/injury/pedbimot/motorcycle/kentuky-la03/TechSumm.html (accessed March 13, 2006).

U.S. Bureau of Land Management. *Water Quality Law Summary*. http://www.blm.gov/nstc/WaterLaws/Chap4.html#Anchor-7638 (accessed March 14, 2006).

U.S. Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health*. 2nd ed. Washington, DC: U.S. Government Printing Office, November 2000. http://www.healthypeople.gov/ (accessed March 13, 2006).

U.S. Department of Transportation. FY 2003 Performance Plan (May 2002). http://www.nhtsa.dot.gov/nhtsa/whatis/planning/perf-plans/2003/Index.html (accessed December 6, 2005).

U.S. Department of Transportation. FY 2004 Performance Plan. http://www.dot.gov/PerfPlan2004/safety_highway.html (accessed March 11, 2006).

U.S. Department of Transportation. Performance and Accountability Report. http://www.dot.gov/perfacc2005/parpart1.htm#safety (accessed March 11, 2006).

U.S. Department of Transportation. Performance Plan (FY 2001) and Report (FY 1999).

U.S. General Accounting Office. Report to Congressional Requesters, January 2004. "Human Capital: Implementing Pay for Performance at Selected Personnel Demonstration Projects," 04-83.

U.S. Office of Management and Budget. FY 2006 Budget, Department of Transportation PART review, NHTSA Operations and Research.

Walker, David M. "GAO Answers the Question: What's in a Name?" *Roll Call*, July 19, 2004. http://www.gao.gov/about/rollcall07192004.pdf (accessed December 26, 2005).

Walters, Jonathan and Charles Thompson. "The Transformation of the Government Accountability Office." Washington, DC: IBM Center for The Business of Government, 2005.

Zander, Alvin. *Motives and Goals in Groups*. New Brunswick: Transaction, 1996.

ABOUT THE AUTHOR

Shelley H. Metzenbaum is a Visiting Professor at the University of Maryland School of Public Policy and Executive Director of the Environmental Compliance Consortium, a collaborative effort of state environmental protection agencies to improve the measurement and management of their programs.

Professor Metzenbaum is the faculty chair of the University of Maryland School of Public Policy's executive education program for the U.S. Food and Drug Administration. Prior to that, she served as adjunct faculty at the Brookings Institution, running the "Science and Technology Policy" and "Managing for Results" programs. From 1998 to 2000, she headed the executive session on public sector performance management at Harvard's Kennedy School of Government. She currently serves as senior consultant to the Facility Reporting Project, a multi-stakeholder project to develop generally accepted, credible, and comparable sustainability metrics for regulated facilities.

From 1993 to 1997, Metzenbaum served as associate administrator of the U.S. Environmental Protection Agency for Regional Operations and State/Local Relations. As associate administrator, she managed the design and implementation of the National Environmental Performance Partnership System (NEPPS) and the Sustainable Development Challenge Grant program. Previously, she served as undersecretary of the Massachusetts Executive Office of Environmental Affairs (EOEA.) Before serving as undersecretary, she was director of capital budgeting for the Commonwealth of Massachusetts.

Metzenbaum's teaching, writing, and consulting emphasize ways governments can make performance measurement more useful, with a special emphasis on ways information can improve the effectiveness and efficiency of the environmental protection system. Her first report for the IBM Center, "Strategies for Using State Information: Measuring and Improving Program Performance," was published in 2003. She is a fellow of the National Academy of Public Administration and a principal of the Council for Excellence in Government. Metzenbaum holds a Ph.D. in public policy from the Kennedy School of Government, Harvard University.

KEY CONTACT INFORMATION

To contact the author:

Shelley H. Metzenbaum
Visiting Professor
Maryland School of Public Policy
University of Maryland
2101 Van Munching Hall
College Park, MD 20742
(978) 371-3099

e-mail: smetzenbaum@verizon.net

CENTER REPORTS AVAILABLE

COMPETITION, CHOICE, AND INCENTIVES

Determining a Level Playing Field for Public-Private Competition (1999)
Lawrence L. Martin

Implementing State Contracts for Social Services: An Assessment of the Kansas Experience (2000)
Jocelyn M. Johnston and Barbara S. Romzek

A Vision of the Government as a World-Class Buyer: Major Procurement Issues for the Coming Decade (2002)
Jacques S. Gansler

Contracting for the 21st Century: A Partnership Model (2002)
Wendell C. Lawther

Franchise Funds in the Federal Government: Ending the Monopoly in Service Provision (2002)
John J. Callahan

Making Performance-Based Contracting Perform: What the Federal Government Can Learn from State and Local Governments (2002, 2nd ed.)
Lawrence L. Martin

Moving to Public-Private Partnerships: Learning from Experience around the World (2003)
Trefor P. Williams

IT Outsourcing: A Primer for Public Managers (2003)
Yu-Che Chen and James Perry

The Procurement Partnership Model: Moving to a Team-Based Approach (2003)
Kathryn G. Denhardt

Moving Toward Market-Based Government: The Changing Role of Government as the Provider (2004, 2nd ed.)
Jacques S. Gansler

Transborder Service Systems: Pathways for Innovation or Threats to Accountability? (2004)
Alasdair Roberts

Competitive Sourcing: What Happens to Federal Employees? (2004)
Jacques S. Gansler and William Lucyshyn

Implementing Alternative Sourcing Strategies: Four Case Studies (2004)
Edited by Jacques S. Gansler and William Lucyshyn

Designing Competitive Bidding for Medicare (2004)
John Cawley and Andrew B. Whitford

International Experience Using Outsourcing, Public-Private Partnerships, and Vouchers (2005)
Jón R. Blöndal

Effectively Managing Professional Services Contracts: 12 Best Practices (2006)
Sandra L. Fisher, Michael E. Wasserman, and Paige P. Wolf

E-GOVERNMENT

Supercharging the Employment Agency: An Investigation of the Use of Information and Communication Technology to Improve the Service of State Employment Agencies (2000)
Anthony M. Townsend

Assessing a State's Readiness for Global Electronic Commerce: Lessons from the Ohio Experience (2001)
J. Pari Sabety and Steven I. Gordon

Privacy Strategies for Electronic Government (2001)
Janine S. Hiller and France Bélanger

Commerce Comes to Government on the Desktop: E-Commerce Applications in the Public Sector (2001)
Genie N. L. Stowers

The Use of the Internet in Government Service Delivery (2001)
Steven Cohen and William Eimicke

State Web Portals: Delivering and Financing E-Service (2002)
Diana Burley Gant, Jon P. Gant, and Craig L. Johnson

Internet Voting: Bringing Elections to the Desktop (2002)
Robert S. Done

Leveraging Technology in the Service of Diplomacy: Innovation in the Department of State (2002)
Barry Fulton

Federal Intranet Work Sites: An Interim Assessment (2002)
Julianne G. Mahler and Priscilla M. Regan

The State of Federal Websites: The Pursuit of Excellence (2002)
Genie N. L. Stowers

State Government E-Procurement in the Information Age: Issues, Practices, and Trends (2002)
M. Jae Moon

Preparing for Wireless and Mobile Technologies in Government (2002)
Ai-Mei Chang and P. K. Kannan

Public-Sector Information Security: A Call to Action for Public-Sector CIOs (2002, 2nd ed.)
Don Heiman

The Auction Model: How the Public Sector Can Leverage the Power of E-Commerce Through Dynamic Pricing (2002, 2nd ed.)
David C. Wyld

The Promise of E-Learning in Africa: The Potential for Public-Private Partnerships (2003)
Norman LaRocque and Michael Latham

Digitally Integrating the Government Supply Chain: E-Procurement, E-Finance, and E-Logistics (2003)
Jacques S. Gansler, William Lucyshyn, and Kimberly M. Ross

Using Technology to Increase Citizen Participation in Government: The Use of Models and Simulation (2003)
John O'Looney

Seaport: Charting a New Course for Professional Services Acquisition for America's Navy (2003)
David C. Wyld

E-Reporting: Strengthening Democratic Accountability (2004)
Mordecai Lee

Understanding Electronic Signatures: The Key to E-Government (2004)
Stephen H. Holden

Measuring the Performance of E-Government (2004)
Genie N. L. Stowers

Restoring Trust in Government: The Potential of Digital Citizen Participation (2004)
Marc Holzer, James Melitski, Seung-Yong Rho, and Richard Schwester

From E-Government to M-Government? Emerging Practices in the Use of Mobile Technology by State Governments (2004)
M. Jae Moon

Government Garage Sales: Online Auctions as Tools for Asset Management (2004)
David C. Wyld

Innovation in E-Procurement: The Italian Experience (2004)
Mita Marra

Computerisation and E-Government in Social Security: A Comparative International Study (2005)
Michael Adler and Paul Henman

The Next Big Election Challenge: Developing Electronic Data Transaction Standards for Election Administration (2005)
R. Michael Alvarez and Thad E. Hall

RFID: The Right Frequency for Government (2005)
David C. Wyld

FINANCIAL MANAGEMENT

Credit Scoring and Loan Scoring: Tools for Improved Management of Federal Credit Programs (1999)
Thomas H. Stanton

Using Activity-Based Costing to Manage More Effectively (2000)
Michael H. Granof, David E. Platt, and Igor Vaysman

Audited Financial Statements: Getting and Sustaining "Clean" Opinions (2001)
Douglas A. Brook

An Introduction to Financial Risk Management in Government (2001)
Richard J. Buttimer, Jr.

Understanding Federal Asset Management: An Agenda for Reform (2003)
Thomas H. Stanton

Efficiency Counts: Developing the Capacity to Manage Costs at Air Force Materiel Command (2003)
Michael Barzelay and Fred Thompson

Federal Credit Programs: Managing Risk in the Information Age (2005)
Thomas H. Stanton

Grants Management in the 21st Century: Three Innovative Policy Responses (2005)
Timothy J. Conlan

Performance Budgeting: How NASA and SBA Link Costs and Performance (2006)
Lloyd A. Blanchard

HUMAN CAPITAL MANAGEMENT

Profiles in Excellence: Conversations with the Best of America's Career Executive Service (1999)
Mark W. Huddleston

Reflections on Mobility: Case Studies of Six Federal Executives (2000)
Michael D. Serlin

Managing Telecommuting in the Federal Government: An Interim Report (2000)
Gina Vega and Louis Brennan

Using Virtual Teams to Manage Complex Projects: A Case Study of the Radioactive Waste Management Project (2000)
Samuel M. DeMarie

A Learning-Based Approach to Leading Change (2000)
Barry Sugarman

Labor-Management Partnerships: A New Approach to Collaborative Management (2001)
Barry Rubin and Richard Rubin

Winning the Best and Brightest: Increasing the Attraction of Public Service (2001)
Carol Chetkovich

A Weapon in the War for Talent: Using Special Authorities to Recruit Crucial Personnel (2001)
Hal G. Rainey

A Changing Workforce: Understanding Diversity Programs in the Federal Government (2001)
Katherine C. Naff and J. Edward Kellough

Life after Civil Service Reform: The Texas, Georgia, and Florida Experiences (2002)
Jonathan Walters

The Defense Leadership and Management Program: Taking Career Development Seriously (2002)
Joseph A. Ferrara and Mark C. Rom

The Influence of Organizational Commitment on Officer Retention: A 12-Year Study of U.S. Army Officers (2002)
Stephanie C. Payne, Ann H. Huffman, and Trueman R. Tremble, Jr.

Human Capital Reform: 21st Century Requirements for the United States Agency for International Development (2003)
Anthony C. E. Quainton and Amanda M. Fulmer

Modernizing Human Resource Management in the Federal Government: The IRS Model (2003)
James R. Thompson and Hal G. Rainey

Mediation at Work: Transforming Workplace Conflict at the United States Postal Service (2003)
Lisa B. Bingham

Growing Leaders for Public Service (2004, 2nd ed.)
Ray Blunt

Pay for Performance: A Guide for Federal Managers (2004)
Howard Risher

The Blended Workforce: Maximizing Agility Through Nonstandard Work Arrangements (2005)
James R. Thompson and Sharon H. Mastracci

The Transformation of the Government Accountability Office: Using Human Capital to Drive Change (2005)
Jonathan Walters and Charles Thompson

INNOVATION

Managing Workfare: The Case of the Work Experience Program in the New York City Parks Department (1999)
Steven Cohen

New Tools for Improving Government Regulation: An Assessment of Emissions Trading and Other Market-Based Regulatory Tools (1999)
Gary C. Bryner

Religious Organizations, Anti-Poverty Relief, and Charitable Choice: A Feasibility Study of Faith-Based Welfare Reform in Mississippi (1999)
John P. Bartkowski and Helen A. Regis

Business Improvement Districts and Innovative Service Delivery (1999)
Jerry Mitchell

An Assessment of Brownfield Redevelopment Policies: The Michigan Experience (1999)
Richard C. Hula

San Diego County's Innovation Program: Using Competition and a Whole Lot More to Improve Public Services (2000)
William B. Eimicke

Innovation in the Administration of Public Airports (2000)
Scott E. Tarry

Entrepreneurial Government: Bureaucrats as Businesspeople (2000)
Anne Laurent

Rethinking U.S. Environmental Protection Policy: Management Challenges for a New Administration (2000)
Dennis A. Rondinelli

The Challenge of Innovating in Government (2001)
Sandford Borins

Understanding Innovation: What Inspires It? What Makes It Successful? (2001)
Jonathan Walters

Government Management of Information Mega-Technology: Lessons from the Internal Revenue Service's Tax Systems Modernization (2002)
Barry Bozeman

Advancing High End Computing: Linking to National Goals (2003)
Juan D. Rogers and Barry Bozeman

MANAGING FOR PERFORMANCE AND RESULTS

Corporate Strategic Planning in Government: Lessons from the United States Air Force (2000)
Colin Campbell

Using Evaluation to Support Performance Management: A Guide for Federal Executives (2001)
Kathryn Newcomer and Mary Ann Scheirer

Managing for Outcomes: Milestone Contracting in Oklahoma (2001)
Peter Frumkin

The Challenge of Developing Cross-Agency Measures: A Case Study of the Office of National Drug Control Policy (2001)
Patrick J. Murphy and John Carnevale

The Potential of the Government Performance and Results Act as a Tool to Manage Third-Party Government (2001)
David G. Frederickson

Using Performance Data for Accountability: The New York City Police Department's CompStat Model of Police Management (2001)
Paul E. O'Connell

Moving Toward More Capable Government: A Guide to Organizational Design (2002)
Thomas H. Stanton

The Baltimore CitiStat Program: Performance and Accountability (2003)
Lenneal J. Henderson

Strategies for Using State Information: Measuring and Improving Program Performance (2003)
Shelley H. Metzenbaum

Linking Performance and Budgeting: Opportunities in the Federal Budget Process (2004, 2nd ed.)
Philip G. Joyce

How Federal Programs Use Outcome Information: Opportunities for Federal Managers (2004, 2nd ed.)
Harry P. Hatry, Elaine Morley, Shelli B. Rossman, and Joseph S. Wholey

Performance Management for Career Executives: A "Start Where You Are, Use What You Have" Guide (2004, 2nd ed.)
Chris Wye

Staying the Course: The Use of Performance Measurement in State Governments (2004)
Julia Melkers and Katherine Willoughby

Moving from Outputs to Outcomes: Practical Advice from Governments Around the World (2006)
Burt Perrin

Using the Balanced Scorecard: Lessons Learned from the U.S. Postal Service and the Defense Finance and Accounting Service (2006)
Nicholas J. Mathys and Kenneth R. Thompson

Performance Leadership: 11 Better Practices That Can Ratchet Up Performance (2006, 2nd ed.)
Robert D. Behn

Performance Accountability: The Five Building Blocks and Six Essential Practices (2006)
Shelley H. Metzenbaum

NETWORKS AND PARTNERSHIPS

Leveraging Networks to Meet National Goals: FEMA and the Safe Construction Networks (2002)
William L. Waugh, Jr.

Applying 21-Century Government to the Challenge of Homeland Security (2002)
Elaine C. Kamarck

Assessing Partnerships: New Forms of Collaboration (2003)
Robert Klitgaard and Gregory F. Treverton

Leveraging Networks: A Guide for Public Managers Working across Organizations (2003)
Robert Agranoff

Extraordinary Results on National Goals: Networks and Partnerships in the Bureau of Primary Health Care's 100%/0 Campaign (2003)
John Scanlon

Public-Private Strategic Partnerships: The U.S. Postal Service-Federal Express Alliance (2003)
Oded Shenkar

The Challenge of Coordinating "Big Science" (2003)
W. Henry Lambright

Communities of Practice: A New Tool for Government Managers (2003)
William M. Snyder and Xavier de Souza Briggs

Collaboration and Performance Management in Network Settings: Lessons from Three Watershed Governance Efforts (2004)
Mark T. Imperial

The Quest to Become "One": An Approach to Internal Collaboration (2005)
Russ Linden

Cooperation Between Social Security and Tax Agencies in Europe (2005)
Bernhard Zaglmayer, Paul Schoukens, and Danny Pieters

Leveraging Collaborative Networks in Infrequent Emergency Situations (2005)
Donald P. Moynihan

Public Deliberation: A Manager's Guide to Citizen Engagement (2006)
Carolyn J. Lukensmeyer and Lars Hasselblad Torres

A Manager's Guide to Choosing and Using Collaborative Networks (2006)
H. Brinton Milward and Keith G. Provan

TRANSFORMATION

The Importance of Leadership: The Role of School Principals (1999)
Paul Teske and Mark Schneider

Leadership for Change: Case Studies in American Local Government (1999)
Robert B. Denhardt and Janet Vinzant Denhardt

Managing Decentralized Departments: The Case of the U.S. Department of Health and Human Services (1999)
Beryl A. Radin

Transforming Government: The Renewal and Revitalization of the Federal Emergency Management Agency (2000)
R. Steven Daniels and Carolyn L. Clark-Daniels

Transforming Government: Creating the New Defense Procurement System (2000)
Kimberly A. Harokopus

Trans-Atlantic Experiences in Health Reform: The United Kingdom's National Health Service and the United States Veterans Health Administration (2000)
Marilyn A. DeLuca

Transforming Government: The Revitalization of the Veterans Health Administration (2000)
Gary J. Young

The Challenge of Managing Across Boundaries: The Case of the Office of the Secretary in the U.S. Department of Health and Human Services (2000)
Beryl A. Radin

Creating a Culture of Innovation: 10 Lessons from America's Best Run City (2001)
Janet Vinzant Denhardt and Robert B. Denhardt

Transforming Government: Dan Goldin and the Remaking of NASA (2001)
W. Henry Lambright

Managing Across Boundaries: A Case Study of Dr. Helene Gayle and the AIDS Epidemic (2002)
Norma M. Riccucci

Managing "Big Science": A Case Study of the Human Genome Project (2002)
W. Henry Lambright

The Power of Frontline Workers in Transforming Government: The Upstate New York Veterans Healthcare Network (2003)
Timothy J. Hoff

Making Public Sector Mergers Work: Lessons Learned (2003)
Peter Frumkin

Efficiency Counts: Developing the Capacity to Manage Costs at Air Force Materiel Command (2003)
Michael Barzelay and Fred Thompson

Managing the New Multipurpose, Multidiscipline University Research Centers: Institutional Innovation in the Academic Community (2003)
Barry Bozeman and P. Craig Boardman

The Transformation of the Government Accountability Office: Using Human Capital to Drive Change (2005)
Jonathan Walters and Charles Thompson

Transforming the Intelligence Community: Improving the Collection and Management of Information (2005)
Elaine C. Kamarck

Executive Response to Changing Fortune: Sean O'Keefe as NASA Administrator (2005)
W. Henry Lambright

Ramping Up Large, Non-Routine Projects: Lessons for Federal Managers from the Successful 2000 Census (2005)
Nancy A. Potok and William G. Barron, Jr.

The *Next* Government of the United States: Challenges for Performance in the 21st Century (2005)
Donald F. Kettl

Reforming the Federal Aviation Administration: Lessons from Canada and the United Kingdom (2006)
Clinton V. Oster, Jr.

PRESIDENTIAL TRANSITION

Government Reorganization: Strategies and Tools to Get It Done (2004)
Hannah Sistare

Performance Management for Political Executives: A "Start Where You Are, Use What You Have" Guide (2004)
Chris Wye

Becoming an Effective Political Executive: 7 Lessons from Experienced Appointees (2005, 2nd ed.)
Judith E. Michaels

Getting to Know You: Rules of Engagement for Political Appointees and Career Executives (2005)
Joseph A. Ferrara and Lynn C. Ross

SPECIAL REPORTS

Enhancing Security Throughout the Supply Chain (2004)
David J. Closs and Edmund F. McGarrell

Assessing the Impact of IT-Driven Education in K–12 Schools (2005)
Ganesh D. Bhatt

Investing in Supply Chain Security: Collateral Benefits (2005, 2nd ed.)
James B. Rice, Jr., and Philip W. Spayd

Six Trends Transforming Government (2006)
Mark A. Abramson, Jonathan D. Breul, and John M. Kamensky

BOOKS*

Collaboration: Using Networks and Partnerships
(Rowman & Littlefield Publishers, Inc., 2004)
John M. Kamensky and Thomas J. Burlin, editors

Competition, Choice, and Incentives in Government Programs
(Rowman & Littlefield Publishers, Inc., 2006)
John M. Kamensky and Albert Morales, editors

E-Government 2001
(Rowman & Littlefield Publishers, Inc., 2001)
Mark A. Abramson and Grady E. Means, editors

E-Government 2003
(Rowman & Littlefield Publishers, Inc., 2002)
Mark A. Abramson and Therese L. Morin, editors

Human Capital 2002
(Rowman & Littlefield Publishers, Inc., 2002)
Mark A. Abramson and Nicole Willenz Gardner, editors

Human Capital 2004
(Rowman & Littlefield Publishers, Inc., 2004)
Jonathan D. Breul and Nicole Willenz Gardner, editors

Innovation
(Rowman & Littlefield Publishers, Inc., 2002)
Mark A. Abramson and Ian Littman, editors

Leaders
(Rowman & Littlefield Publishers, Inc., 2002)
Mark A. Abramson and Kevin M. Bacon, editors

Learning the Ropes: Insights for Political Appointees
(Rowman & Littlefield Publishers, Inc., 2005)
Mark A. Abramson and Paul R. Lawrence, editors

Managing for Results 2002
(Rowman & Littlefield Publishers, Inc., 2001)
Mark A. Abramson and John M. Kamensky, editors

Managing for Results 2005
(Rowman & Littlefield Publishers, Inc., 2004)
John M. Kamensky and Albert Morales, editors

Memos to the President: Management Advice from the Nation's Top Public Administrators
(Rowman & Littlefield Publishers, Inc., 2001)
Mark A. Abramson, editor

New Ways of Doing Business
(Rowman & Littlefield Publishers, Inc., 2003)
Mark A. Abramson and Ann M. Kieffaber, editors

The Procurement Revolution
(Rowman & Littlefield Publishers, Inc., 2003)
Mark A. Abramson and Roland S. Harris III, editors

Transforming Government Supply Chain Management
(Rowman & Littlefield Publishers, Inc., 2003)
Jacques S. Gansler and Robert E. Luby, Jr., editors

Transforming Organizations
(Rowman & Littlefield Publishers, Inc., 2001)
Mark A. Abramson and Paul R. Lawrence, editors

* Available at bookstores, online booksellers, and from the publisher (www.rowmanlittlefield.com or 800-462-6420).

About the IBM Center for The Business of Government

Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion of new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

The Center is one of the ways that IBM seeks to advance knowledge on how to improve public sector effectiveness. The IBM Center focuses on the future of the operation and management of the public sector.

About IBM Global Business Services

With consultants and professional staff in more than 160 countries globally, IBM Global Business Services is the world's largest consulting services organization. IBM Global Business Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build and run those solutions in a way that delivers bottom-line business value. For more information visit www.ibm.com.

For additional information, contact:

Mark A. Abramson
Executive Director
IBM Center for The Business of Government
1301 K Street, NW
Fourth Floor, West Tower
Washington, DC 20005
(202) 515-4504, fax: (202) 515-4375

e-mail: businessofgovernment@us.ibm.com
website: www.businessofgovernment.org