

**MANAGING FOR SOCIAL CHANGE: IMPROVING LABOR  
DEPARTMENT PERFORMANCE IN A PARTISAN ERA**

*Seth D. Harris\**

I.	INTRODUCTION.....	988
II.	COMPLIANCE WITH THE LEGAL FRAMEWORK VERSUS GOOD PERFORMANCE.....	991
	<i>A. The Legal Infrastructure for Executive Branch Management Planning.....</i>	992
	1. Government Performance and Results Act of 1993 (GPRA) ..	992
	2. GPRA Modernization Act of 2010 (GPRAMA) .....	995
	<i>B. The Labor Department’s Serious Performance Problems Despite GPRA Compliance.....</i>	998
	1. An Overview of the Labor Department .....	998
	2. Labor Department Performance Before 2009.....	1000
	<i>i. Wage &amp; Hour Division (Wage &amp; Hour) .....</i>	1001
	<i>ii. Office of Federal Contract Compliance Programs (OFCCP) .....</i>	1002
	<i>iii. Mine Safety and Health Administration (MSHA).....</i>	1003
	<i>C. GPRA’s Failings and the GPRAMA’s Failure To Fix Them.....</i>	1005
	1. GPRA’s Failings.....	1005
	<i>i. No Congressional Accountability.....</i>	1005
	<i>ii. No Clear Definition of Success .....</i>	1007
	<i>iii. No Assurance that Agencies Use Data in Their Decision-Making .....</i>	1007
	<i>iv. Failure To Require Evidence of Relationships Between Outputs and Outcomes .....</i>	1009
	<i>v. Complexity of the Mandate.....</i>	1010

---

\* © Seth D. Harris. Distinguished Scholar, Cornell University School of Industrial & Labor Relations; Deputy U.S. Secretary of Labor, May 2009 through January 2014; Acting U.S. Secretary of Labor, January 2013 through July 2013. I presented an earlier version of this paper at West Virginia University College of Law’s “Zealous Advocacy for Social Change Labor Law Conference” on October 23, 2014. I am grateful to Michael Kerr, Natalie Palugyai, Demetra Nightingale, Braye Cloud, and Holly Donnelly for their comments and Associate Dean Anne Marie Lofaso for her invitation to the conference and encouragement in concluding this Article. Benjamin Hawkins provided essential, patient, and efficient research and editorial assistance. I appreciate Cornell University’s Industrial & Labor Relations School’s support for my research. All errors are mine.

	2. GPRAMA’s Failure To Fix GPRA’s Failings .....	1010
III.	TURNING AROUND THE LABOR DEPARTMENT’S PERFORMANCE.....	1015
	A. <i>The Turnaround Strategy</i> .....	1016
	B. <i>The Results</i> .....	1025
IV.	A PROGRAM FOR REFORM .....	1034
	A. <i>Congress</i> .....	1035
	B. <i>Stakeholders</i> .....	1040
V.	CONCLUSION .....	1045

## I. INTRODUCTION

For decades, labor and employment policy at the federal level has been debated through a partisan lens. Unions and other worker advocates generally support Democrats who, in turn, pursue greater government intervention in the workplace and labor market both to achieve certain social and economic goals and to disrupt unfair hierarchies. The business community and small-government conservatives generally support Republicans who, in turn, typically oppose government intervention in the workplace and labor market, preferring that markets and private decision-makers assign resources and distribute economic product.<sup>1</sup>

Hidden beneath this heatedly ideological and interest-driven partisan divide is a surprising and rarely discussed consensus regarding the management of labor and employment policies and programs. Both Democrats and Republicans want the United States Department of Labor, for example, to be run efficiently and cost-effectively, and to achieve meaningful outcomes for the people and entities it serves. While there is debate about ends and some means, there is broad consensus that money should be spent wisely and taxpayers should expect a meaningful return on their investments.

In truth, “consensus” may overstate the case. Government management is rarely debated in Washington, even in the partisan world of labor and employment policy. It draws little attention or interest. One quick anecdote will support this assertion. During my tenure as Deputy Secretary and Chief Operating Officer of the Labor Department, the department’s performance management staff created an “Annual Performance Index” that collected data showing how well or poorly the department and its constituent agencies had done their jobs during the preceding fiscal year.<sup>2</sup> The index offered valuable

<sup>1</sup> See, e.g., Chris Opfer, *Chamber, Labor Spend Big in Senate Races with Different Strategies for Reaching Voters*, BLOOMBERG BNA (Oct. 18, 2014), <http://www.bna.com/chamber-labor-spend-n17179897271/>; Lydia Saad, *Democrats More Confident in Labor; Republicans, in Business*, GALLUP (June 20, 2012), <http://www.gallup.com/poll/155267/democrats-confident-labor-republicans-business.aspx>.

<sup>2</sup> See *infra* notes 104–05 and accompanying text.

insights into the management of the department's programs and agencies, how the agencies' leadership teams had prioritized their work, and the trajectory of the agencies' performance as compared with their own past performance. The index was the culmination of a multi-year effort to make dramatic improvements in the department's performance.<sup>3</sup> Beginning in fall 2013, the Labor Department's congressional affairs office sought to organize a meeting with congressional staff at which I would brief them on the latest index and discuss the department's performance management program. How many congressional staff agreed to attend the briefing? Zero. Based on prior experience, I considered that response to be an accurate representation of their bosses' interest in the topic.

When asked, both Republicans and Democrats would agree they want the Labor Department to be managed effectively. But neither party campaigns on management issues, or even raises them frequently in congressional hearings or speeches on the floors of the House of Representatives or Senate.<sup>4</sup> These issues hold little appeal to the parties' political bases, and they generate much less heat than a fight over, for example, the Employee Free Choice Act.<sup>5</sup> A cynic might suggest it "goes without saying" that both parties support effective management of labor and employment policy.

Yet, meaningful social change can result from the effective management of the resources available to the Labor Department and other government agencies administering labor and employment laws, as well as other federal agencies. Good management should not be viewed as an end in itself, or a means to achieve the ill-defined goal of "good government." Rather,

---

<sup>3</sup> Seth D. Harris, *Op-Ed: Labor's Moneyball Approach to Boosting Performance*, NEXTGOV (Jan. 14, 2014), <http://www.nextgov.com/big-data/2014/01/op-ed-labors-moneyball-approach-boosting-performance/76781/?oref=ng-dropdown>.

<sup>4</sup> We searched for hearings relating to the Labor Department's compliance with or performance under the Government Performance and Results Act or the GPRA Modernization Act held by the Senate Health, Education, Labor and Pensions Committee or the House Education and Workforce Committee—the Labor Department's principal "authorizing committees"—and the Senate and House Appropriations Committees' sub-committees with jurisdiction over the Labor Department since January 2009. We found none. *See infra* note 32.

<sup>5</sup> Employee Free Choice Act of 2009, H.R. 1409, 111th Cong. (2009); Employee Free Choice Act of 2009, S. 560, 111th Cong. (2009). The Employee Free Choice Act would have amended the National Labor Relations Act to better facilitate union organizing. Most important, it would have permitted unions to be established in workplaces through a "card check" process involving the collection of authorization cards from a majority of workers rather than an election run by the National Labor Relations Board. This proposed legislation proved to be a partisan lightning rod. *See* Steven Greenhouse, *Bill Easing Unionizing Is Under Heavy Attack*, N.Y. TIMES, Jan. 9, 2009, at A12, available at [http://www.nytimes.com/2009/01/09/us/09labor.html?\\_r=0](http://www.nytimes.com/2009/01/09/us/09labor.html?_r=0); Steven Greenhouse, *Democrats Drop Key Part of Bill To Assist Unions*, N.Y. TIMES, July 17, 2009, at A1, available at <http://www.nytimes.com/2009/07/17/business/17union.html>; Deborah L. Cohen, *Now Playing: The Employee Free Choice Act*, REUTERS (Apr. 29, 2009, 1:32 PM), <http://www.reuters.com/article/2009/04/29/us-column-cohen-efca-idUSTRE53S67T20090429>.

effective management, particularly performance management, can mean that more workers' lives are saved in the workplace, more stolen wages are recovered for the workers who earned them, more workplace civil rights violations are remedied, more workers receive better skills training to prepare for higher paying and long lasting jobs, and more workers' pension and health plans are secure.<sup>6</sup> These results are achieved through improving the administration of existing laws and programs rather than new legislation or regulations.

In a partisan era, improved management may be the only noncontroversial pathway to achieving these results. Congress's general acquiescence in most executive branch planning and management choices, and the general sense that better government management is a good thing for the country, creates political space for executive branch leaders to dramatically improve the performance of their departments and agencies. At the same time, congressional inattention and failure to hold the executive branch accountable for excellent management makes such improvement much less likely. For this reason, those seeking social change should be more actively engaged in ensuring that the Labor Department and other federal and state labor agencies (e.g., the National Labor Relations Board, the Equal Employment Opportunity Commission, state labor departments) are managed effectively to significantly improve the lives of America's working families. Simply, without meaningful reform and outside pressure, Congress will not do it.

This Article tells the story of the successful efforts to turn around the Labor Department's performance during the first five years of the Obama Administration and deliver better outcomes for working families and others served by the department. At its core, it is a story of how improved performance management produced the results progressive advocates seek. But it is also a tale of serious flaws in the legal infrastructure that governs federal executive branch performance management and measurement, and Congress abandoning the field after establishing this legal infrastructure. The Labor Department's success came in spite of the congressionally enacted legal system, not because of it, particularly in light of Congress's neglect of that system. As a result, this Article argues that there is a great deal more to be done to guarantee the continuation of the Labor Department's success and, perhaps, to proliferate that success to other government agencies. Reforming the laws that guide executive branch performance management and measurement is one necessary step; however, this Article also argues that stakeholders—particularly progressive stakeholders who care deeply about government's role—must engage aggressively in policing government performance.

Part II of this Article describes the legal framework for executive branch performance management and how the Labor Department, although

---

<sup>6</sup> See *infra* notes 94–96 and accompanying text.

complying with the formal requirements of that legal framework, was not well managed and did not perform well, in some regards, before President Obama took office. Part III describes the Labor Department's performance management turnaround efforts in the first five years of the Obama Administration. These efforts far exceeded those required by law, but were consistent with the larger government reform goals that President Obama articulated and Congress sought to codify. Part III also describes the real-world results these performance management efforts produced for working families and others served by the Labor Department. Finally, Part IV proposes a reform agenda to improve the way the Labor Department, other labor agencies, and other executive branch entities manage their work.

## II. COMPLIANCE WITH THE LEGAL FRAMEWORK VERSUS GOOD PERFORMANCE

There is good reason to question Congress's constitutional authority to direct how the Cabinet departments and other non-legislative agencies conduct performance management and measurement.<sup>7</sup> The Constitution's Article II, Section 1 vests the President with the federal government's "executive" powers.<sup>8</sup> A plethora of statutes delegate authority to the President and his Cabinet and sub-Cabinet officers to execute the functions established by those laws. Thus, there is a nontrivial argument that the Constitution expects that Congress will leave executive branch performance to the President and his agency heads. On the other hand, Congress is vested with near plenary authority over federal spending under Article I, Sections 8 and 9, including the power to condition the use and receipt of funds on compliance with specific congressional dictates.<sup>9</sup> In an exercise of its responsibility to ensure that the funds it appropriates are spent properly, Congress could argue that it should be deeply and directly involved in measuring and assessing whether the executive branch is performing well.

Ultimately, the constitutional argument is not the most important issue. The larger problems are congressional competence and will to oversee executive branch performance. Over the last 20 or so years, Congress has enacted two statutes—the Government Performance and Results Act of 1993 (GPRA) and the GPRA Modernization Act of 2010 (GPRAMA)—that were intended to give structure and direction to the executive branch's strategic and

---

<sup>7</sup> For the reader's convenience, I am going to use "agency" or "agencies" to describe both Cabinet departments, like the Labor Department, and non-Cabinet executive branch agencies, like the Office of Personnel Management. For the purposes of this Article, the difference between a department and agency is not important. I will use "department" on occasion to refer only to the Labor Department.

<sup>8</sup> U.S. CONST. art. II, § 1, cl. 1.

<sup>9</sup> U.S. CONST. art. I, §§ 8–9.

management planning. The first section will describe the system established by these two statutes, but it will also argue that Congress has not taken the requisite steps to ensure effective implementation of these laws. In particular, Congress has neither used its oversight function to cajole the executive branch into better performance nor exercised its appropriations power to direct it. To illustrate this point, the second section will describe seriously problematic performance in the Labor Department before 2009 and the law's failure to force better performance and better outcomes. The third section will specify the lessons to be learned from the Labor Department's pre-2009 example—that is, how and why specific flaws in the performance management legal infrastructure, combined with congressional neglect, facilitated problematic performance in the Labor Department before 2009.

*A. The Legal Infrastructure for Executive Branch Management Planning*

This section will explain and critique the two laws that constitute the legal infrastructure for performance management and measurement in the federal executive branch: GPRA<sup>10</sup> and GPRAMA.<sup>11</sup>

1. Government Performance and Results Act of 1993 (GPRA)

GPRA, enacted with support from President Bill Clinton,<sup>12</sup> established a reasonable and well-structured strategic and operating planning process for the federal government's executive branch. Each executive branch agency was directed to develop and publish a five-year strategic plan containing the agency's mission statement, goals and objectives (commonly called "GPRA goals"), and strategies to achieve the GPRA goals.<sup>13</sup> Agencies were also directed to develop annual performance plans consisting of quantifiable and measurable performance targets and a means for comparing actual performance to performance targets.<sup>14</sup> Agencies were required to report annually on actual performance compared with annual targets and to evaluate their success in achieving targets while explaining any performance failures.<sup>15</sup>

---

<sup>10</sup> Government Performance and Results Act of 1993, Pub. L. No. 103-62, 107 Stat. 285 (codified as amended in scattered sections of 5 U.S.C. and 31 U.S.C.).

<sup>11</sup> GPRA Modernization Act of 2010, Pub. L. No. 111-352, 124 Stat. 3866 (2011) (codified as amended in scattered sections of 5 U.S.C. and 31 U.S.C.).

<sup>12</sup> See Remarks on Signing the Government Performance and Results Act of 1993 and an Exchange with Reporters, 29 WEEKLY COMP. PRES. DOC. 1540, 1540 (Aug. 3, 1993) ("The Vice President and I were both enthusiastic about this bill, and I am very, very pleased that it has passed so rapidly.").

<sup>13</sup> Government Performance and Results Act of 1993 § 3.

<sup>14</sup> *Id.* § 4.

<sup>15</sup> *Id.*

With good reason, Congress expected that clearer articulation of management goals and strategies, with regular reporting to Congress on actual performance, would bring change. Two assumptions about human behavior lie at GPRA's core. First, planning for success increases the likelihood of success. Second, disclosure of plans and activities will cause the planner-actor to behave more responsibly and effectively. GPRA requires both planning and disclosure, so Congress expected that its enactment would cause performance to improve, accountability to strengthen, the confidence of the American people in their government to grow, and Congress to be better able to engage in meaningful oversight of the executive branch.<sup>16</sup> Even given the loftiness of these goals, the GPRA structure made a great deal of sense as a means of achieving those goals. Federal agencies were required to define clearly their purpose for existing by answering, what is the agency's mission? Even more important, each agency was forced to state plainly how it would achieve its mission (strategies) and tell Congress how it could know that the mission had been accomplished or that meaningful progress had been made (goals and objectives). At the conceptual level, GPRA was good legislation.

Perhaps GPRA's greatest contribution was its requirement that agencies publicly articulate their "outcome measures"—their GPRA goals.<sup>17</sup> These are metrics that count the real-world impacts of government programs that serve their agency's mission. Of course, this is the ultimate measure of the government's success: whether it is improving American society to the extent promised. One example that shows the importance and occasional complexity of GPRA outcome goals comes from the Labor Department's Mine Safety and Health Administration (MSHA). MSHA defines its mission to be to "prevent death, disease, and injury from mining and promote safe and healthful workplaces for the Nation's miners."<sup>18</sup> Thus, one measure of MSHA's effectiveness is whether miners are dying in workplace accidents. The common-sense goal should be to reduce the number of fatalities to the lowest level possible.

MSHA does not simply measure the absolute number of miners killed at work each year. Instead, it measures a five-year rolling average of fatal injuries per 200,000 hours worked.<sup>19</sup> Why use such a complicated measure to communicate whether MSHA is succeeding at one of its most important jobs? Is MSHA obfuscating? Interestingly, MSHA publicly reports the number of miners killed in coal and "metal/nonmetal" mines each year on its website.<sup>20</sup>

---

<sup>16</sup> *Id.* § 2; 139 CONG. REC. S3078–01 (daily ed. Mar. 17, 1993) (statement of Sen. William S. Cohen).

<sup>17</sup> See *infra* Part II.C.1.iv for further discussion of "outcome" and "output" measures.

<sup>18</sup> U.S. DEP'T OF LABOR, STRATEGIC PLAN: FISCAL YEARS 2014–2018, at 35 (2014).

<sup>19</sup> *Id.* at 37.

<sup>20</sup> *MSHA Fatality Statistics*, MINE SAFETY & HEALTH ADMIN., <http://www.msha.gov/stats/charts/chartshome.htm> (last visited Feb. 27, 2015). The Federal Mine Safety and Health Act

So, the agency is not hiding alternative fatality measures from Congress and the public. To the contrary, the more complicated measure provides a better assessment of MSHA's success.

MSHA divides the absolute number of fatal injuries by a fixed number of hours worked to control for the size of the mining workforce. Simply, 10 deaths among 100 miners would be a qualitatively different outcome from 10 deaths among 100,000 miners. The first result would represent a catastrophic failure by MSHA. The second result must be considered a greater success, even acknowledging that any workplace fatality is unacceptable. This normative difference is unavoidable even though the same number of workers have died.

Smoothing the results with a five-year average is a little more controversial, but it has its own logic. As one former MSHA Assistant Secretary once queried me in a private conversation, should the agency be held accountable for the different outcomes if a miner lights a cigarette, setting off an explosion in an empty mine between shifts, thereby killing only himself, and another lights a cigarette in the middle of a crowded day shift and kills 50 co-workers? The best answer is "no." Judgments about agency effectiveness should depend upon agency performance, not serendipity. Smoothing corrects for serendipity, among other things. But regardless of which side of this argument has the upper hand, MSHA's fatality measure offers the public and Congress a great deal of insight into its performance with respect to one of the most important mine health and safety outcomes. MSHA has been charged with reducing the number of miners who die in mines. MSHA's measure tells us whether it is accomplishing this part of its mission, arguably in a nuanced way. Thus, Congress and the American people have a way of knowing if the government, at least this part of the government, is working well. This is vital information in a democracy.

The hope expressed in several "pilot project" provisions of GPRA was that this kind of information could support more sophisticated decision-making about federal spending. GPRA held out the promise that the federal government would eventually move to a performance-based budgeting strategy—that is, funding agencies based on their ability to produce, and the extent to which they have produced, the outcomes promised by their GPRA goals.<sup>21</sup> More simply, GPRA hinted that Congress would fund programs based on performance data.

In theory, performance-based budgeting is a worthwhile idea. Let us return to the MSHA example. Unlike every other employment law enforcement agency in the Labor Department (and perhaps outside, as well), MSHA is statutorily mandated to investigate every mine, sand and gravel pit, and other covered workplaces either twice (for above ground operations) or four times

---

requires mine operators to report fatalities to MSHA within 15 minutes. Federal Mine Safety and Health Act, 30 U.S.C. § 813(j) (2012).

<sup>21</sup> Government Performance and Results Act of 1993 § 2(b)(2).

(for below ground operations) each year.<sup>22</sup> As a result, MSHA has rich data about every violation and workplace fatality, injury, or illness. MSHA has compiled a record, over time, of every investigation and other type of intervention in the workplace, including its exercise of exceptional enforcement powers like orders closing portions of mines until violations are remedied.<sup>23</sup> MSHA knows how long every investigation took, how many violations were found, and whether the same employer violated the same provisions of the law or any others after the initial investigation. As a result, MSHA has a good understanding of which enforcement and compliance strategies work well to reduce violations, including in different circumstances and different types and sizes of mines.

In a world of rational, outcome-focused congressional decision-making, congressional appropriators should fully fund every successful MSHA strategy at the level required to achieve the maximum reduction in workplace fatalities, injuries, and illnesses. If this world were entirely populated by agencies as data-rich and evidence-based as MSHA, then Congress could go further and undertake the morally difficult task of balancing investments in saving miners' lives against health and safety outcomes achieved by other agencies (e.g., the Centers for Disease Control reducing the incidence of infectious disease and the U.S. Agriculture Department guarding against foodborne illness). Yet, it has not happened. As of this writing, and despite GPRA's two-decades-old hints that it would, Congress has not embraced performance-based budgeting. This does not lead to the inevitable conclusion that Congress is irrational. But it supports the argument offered below that Congress failed to play its oversight role under GPRA.

## 2. GPRA Modernization Act of 2010 (GPRAMA)

The GPRAMA refined the performance management process established by GPRA. For example, agencies are now required to produce strategic plans every four years, rather than five, so that the plans' schedule coincides with presidential terms.<sup>24</sup> But there are three particularly meaningful reforms of GPRA included in the GPRAMA that demonstrate real learning about government performance management.

---

<sup>22</sup> Federal Mine Safety and Health Act § 813(a).

<sup>23</sup> Another example is that MSHA takes more aggressive enforcement and investigative action against mine operators with "patterns of violations." See 30 C.F.R. §§ 104.1–104.4 (2013) (MSHA's newly revised pattern of violations regulation). MSHA's regulation sets standards for determining whether a mine operator has exhibited a pattern of violations and what remedial actions MSHA may take. See *generally* MINE SAFETY & HEALTH ADMIN., U.S. DEP'T OF LABOR, PATTERN OF VIOLATIONS (POV) PROCEDURES SUMMARY (n.d.).

<sup>24</sup> GPRA Modernization Act of 2010, Pub. L. No. 111-352, § 2, 124 Stat. 3866, 3867 (2011) (codified at 5 U.S.C. § 306(b) (2013)).

With its first reform, Congress acknowledged that many outcomes expected from government require the involvement of more than one agency. Under the GPRAMA, the Director of the Office of Management Budget (OMB) must assemble a government-wide performance plan consisting of priority performance goals that cut across agencies (“cross-cutting goals”) with quarterly targets.<sup>25</sup> Little more needs to be said about this acknowledgement of a truism of governmental life: Congress often assigns the same, overlapping, or closely related responsibilities to more than one agency and has only limited means of holding multiple agencies accountable. With more agencies involved, responsibility gets diffused and conflicts can arise. Seeking ways of holding all of the involved agencies accountable to a single measure was a sound and uncontroversial reform.

Reading between the lines, however, the two other notable reforms were more important and signaled that the conceptual soundness of GPRA had not been matched by effective implementation. GPRA implicitly promised that Congress would engage regularly and deeply on executive branch performance. In its second notable GPRAMA reform, Congress implicitly pleaded for help in carrying out its oversight role by requiring public disclosure of both performance plans and performance results. The GPRAMA directs agencies to develop and publish their annual performance plans on a public website.<sup>26</sup> The agencies are also required to report annually on their progress in achieving their targets and, where performance fails to meet targets for two years, to explain their failures to Congress and offer remediation plans, possibly including requests for additional funding.<sup>27</sup> In essence, since Congress tacitly acknowledged that it was not paying attention to executive branch performance, it hoped that requiring agencies to disclose their performance to the public would embarrass them into (or perhaps “nudge” them toward)<sup>28</sup> better performance. Perhaps it also hoped that the public, or particular

---

<sup>25</sup> *Id.* § 3; see also *Cross-Agency Priority Goals*, PERFORMANCE.GOV, <http://www.performance.gov/cap-goals-list?view=public> (last visited Feb. 27, 2015).

<sup>26</sup> GPRA Modernization Act § 2(a). There is no commonality or consistency in the structure of budgets in the federal government, or even within a single agency. For example, the Labor Department’s Mine Safety and Health Administration has eight budget activities, while the department’s Wage & Hour Division has only one. Compare OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, BUDGET OF THE U.S. GOVERNMENT: FISCAL YEAR 2015 APPENDIX at 810 (Mine Safety and Health Administration budget explained and itemized), with *id.* at 805 (Wage & Hour Division budget explained and itemized).

<sup>27</sup> GPRA Modernization Act § 4(h).

<sup>28</sup> See RICHARD H. THALER & CASS R. SUNSTEIN, *NUDGE* 53–71 (Penguin Books 2009). But see Peter M. Gollitzer et al., *When Intentions Go Public: Does Social Reality Widen the Intention-Behavior Gap?*, 20 PSYCHOL. SCI. 612, 616 (2009) (“When other people take notice of one’s identity-relevant behavioral intentions, one’s performance of the intended behaviors is compromised.”).

agencies' stakeholders, would bring pressure directly on agencies to improve their performance without waiting for congressional oversight.

The third reform fits this interpretation, as well. The GPRAMA sought to strengthen accountability mechanisms within the executive branch by promoting the "one neck to choke" theory of management. With respect to agencies' priority goals, Congress required each agency's deputy secretary, newly designated the "Chief Operating Officer," to hold quarterly meetings about the agency's priority goals.<sup>29</sup> Presumably, the Chief Operating Officer would hold the agency's staff accountable in these meetings and, in turn, the Secretary, the OMB Director, and Congress could wrap their collective hands around the Chief Operating Officer's neck to ensure accountability and performance. The GPRAMA also required agencies to publicly identify a leader responsible for achieving success on each goal—again, allowing Congress and the public to know which neck to choke without waiting for Congress to perform the choking on their behalf.<sup>30</sup> Congress further directed the OMB Director to hold quarterly meetings with agencies and goal leaders regarding their priority goals.<sup>31</sup>

Again, the best interpretation of this reform was that Congress knew it had failed in its oversight role and needed someone else to review regularly each agency's performance. This responsibility was effectively delegated to the deputy secretaries and the OMB Director. Congress's tacit acknowledgement of its failure is consistent with the lived experience of at least one Cabinet agency. During my nearly five years as the Deputy Secretary and Chief Operating Officer of the Labor Department in the Obama Administration, I was never called to testify before the department's congressional authorizing committees (the Senate Health, Education, Labor and Pensions Committee and the Senate Finance Committee, and the House Education and Workforce Committee and the House Ways & Means Committee) or its appropriations subcommittees, regarding the Labor Department's performance, or even its compliance with GPRAMA and the GPRAMA.<sup>32</sup> Further, I was never invited to meet with those committees' and subcommittees' staffs to discuss the Labor Department's

---

<sup>29</sup> GPRAMA Modernization Act § 8.

<sup>30</sup> *Id.* § 5(b)(1)(C).

<sup>31</sup> *Id.* § 6.

<sup>32</sup> I conducted an informal survey of a half-dozen other deputy secretaries of domestic Cabinet agencies whose service in the Obama Administration overlapped with mine. All of my colleagues reported that they, too, had never been called to testify about GPRAMA, the GPRAMA, or comprehensive agency performance by their authorizing committees or appropriations subcommittees. To check my memory of the Labor Department's experience, we searched in ProQuest's Congressional Database and found no hearings or reports relating to GPRAMA or the GPRAMA held by the Labor Department's authorizing or the appropriations committees or its sub-committees during my tenure.

compliance with GPRA and the GPRAMA.<sup>33</sup> And, as my vignette rendered in the Introduction discloses, congressional staff *en masse* failed to attend a meeting that the Labor Department's congressional affairs staff sought to organize for the purpose of disclosing and discussing the department's performance. Simply, Congress played no role in overseeing the Labor Department's compliance with either GPRA or the GPRAMA during President Obama's first five years in office.

*B. The Labor Department's Serious Performance Problems Despite GPRA Compliance*

When I returned to the Labor Department in early 2009, the department was largely compliant with GPRA's requirements. The GPRAMA was still one year or so removed, but the department had a five-year strategic plan with a mission statement, goals, objectives, and strategies as required by GPRA.<sup>34</sup> It published annual performance plans that, on one occasion, won an award.<sup>35</sup> The department reported annually on its actual performance compared with annual targets,<sup>36</sup> but the department's technical compliance with GPRA did not result in effective performance across its agencies. In fact, in several agencies (although certainly not all), performance was quite poor in 2009. After a brief overview of the Labor Department's complex and diverse organizational structure, this section offers some illustrative stories about Labor Department performance in 2009 and explains how its performance management system, such as it was, had failed.

1. An Overview of the Labor Department

The Labor Department is a federated agency; in essence, it is a collection of smaller agencies with a roughly common mission. Depending upon how you count, the department has around 23 operating units. There are 13 "line agencies" serving workers, employers, or others either directly or through state, local, or not-for-profit grantees or contractors. Two agencies—

---

<sup>33</sup> Again, other deputy secretaries reported that they were similarly ignored, at least on this topic.

<sup>34</sup> See U.S. DEP'T OF LABOR, STRATEGIC PLAN: FISCAL YEARS 2011–2016 (2011), available at [http://www.dol.gov/\\_sec/stratplan/StrategicPlan.pdf](http://www.dol.gov/_sec/stratplan/StrategicPlan.pdf).

<sup>35</sup> See *Performance Planning and Results*, U.S. DEP'T OF LABOR, <http://www.dol.gov/oasam/PMC/PMC-ppr.htm> (last visited Feb. 27, 2015); *U.S. Labor Department Earns Number One Ranking of Federal Agencies for Program and Financial Performance Reporting*, U.S. DEP'T OF LABOR (May 5, 2009), <http://www.dol.gov/opa/media/press/oasam/OASAM20090505.htm>; Henry Wray et al., *Mercatus Center Unveils 10th Annual Performance Scorecard*, MERCATUS CENTER (May 5, 2009), [http://mercatus.org/media\\_clipping/mercatus-center-unveils-10th-annual-performance-scorecard](http://mercatus.org/media_clipping/mercatus-center-unveils-10th-annual-performance-scorecard).

<sup>36</sup> *Performance Planning and Results*, *supra* note 35.

the Employment and Training Administration (ETA) and the Veterans Employment and Training Service (VETS)—provide or fund workforce development and job training services. Six worker protection agencies enforce employment laws focused on various aspects of the employment relationship and workplace. The Office of Workers' Compensation Programs (OWCP) manages four diverse benefits programs for different groups of workers injured or made ill on the job, and survivors of those killed. Three agencies are principally engaged in developing and advocating policies relating to workers with disabilities, women in the workplace, and international labor affairs. The Bureau of Labor Statistics (BLS) is the federal government's labor and price statistics resource.

Each of these line agencies has multiple product lines. For example, the Occupational Safety and Health Administration (OSHA), one of the department's best known worker protection agencies, engages in law enforcement, worker education, compliance assistance to employers, regulation, advisory services, and a host of other activities designed to make the workplaces over which it has jurisdiction safer and healthier for workers.<sup>37</sup> ETA funnels billions of dollars to hundreds of local workforce investment boards to fund thousands of American Job Centers that provide job search and job training services to workers. It also runs (through contractors) 125 Job Corps residential and non-residential centers, providing education and job training services to disadvantaged youth, and oversees states' unemployment insurance systems, among many other activities.<sup>38</sup> BLS supplies a long list of data series and statistical reports, including monthly releases on the previous month's unemployment rate, wage trends, and job creation.<sup>39</sup> In sum, the Labor Department's agencies produce dozens of different products and services to serve a variety of goals and to achieve an array of outcomes.

Six agencies support the work of the Labor Department's 13 line agencies with staffs of lawyers, lobbyists, public affairs specialists, procurement and budget experts, human resources personnel, information technologists, and financial managers and accountants. The Labor Department even houses four sets of judges who adjudicate certain categories of cases produced by the department's line agencies. Overall, the department has approximately 17,000 employees distributed throughout the United States and a

---

<sup>37</sup> See Occupational Safety and Health Act, 29 U.S.C. § 651 (2012).

<sup>38</sup> See *Program Administration*, JOB CORPS, [http://www.jobcorps.gov/AboutJobCorps/program\\_admin.aspx](http://www.jobcorps.gov/AboutJobCorps/program_admin.aspx) (last visited Feb. 27, 2015); *The "Plain English" Version of the Workforce Investment Act of 1998*, U.S. DEP'T OF LABOR, <http://www.doleta.gov/usworkforce/wia/Runningtext.cfm> (last visited Feb. 27, 2015); *Training: One-Stop Career Centers*, U.S. DEP'T OF LABOR, <http://www.dol.gov/dol/topic/training/onestop.htm#doltopic> (last visited Feb. 27, 2015).

<sup>39</sup> *Databases, Tables & Calculators by Subject*, BUREAU OF LABOR STAT., <http://www.bls.gov/data/> (last visited Feb. 27, 2015).

discretionary budget in the vicinity of \$12 billion each year.<sup>40</sup> It is a large, complex, and diverse organization. Most important, it serves workers, employers, and its other constituents through component agencies that are semi-autonomous but operate under the direction of the department's Secretary and Deputy Secretary.

## 2. Labor Department Performance Before 2009

Perhaps the most telling fact about the Labor Department in 2009 was that, despite its compliance with GPRA, it did not have a system in place that permitted an accurate assessment of the department's overall performance and, in some cases, individual agencies' performance. The department and some of its agencies did not measure many of the agencies' activities. One employment law enforcement agency with a budget of approximately \$100 million had only three metrics in 2009, even though it had many more than three product lines. Further, many of the department's metrics did not measure anything meaningful. Some focused on process rather than outputs or outcomes. Others seemed to be designed to ensure that the agency being measured would always receive an "A" on its report card, regardless of the quality of its performance. But this was not universally true. Some agencies, including BLS and OWCP, had longstanding and comprehensive performance management systems that pre-dated GPRA and had extensive sets of measures designed to capture all of their performance. These systems had long been closely monitored by these agencies' senior leaders. In fact, these well-managed agencies contributed a disproportionate share of the department's total stock of performance measures in 2009.

President Bush's Administration had instituted a system called the Program Assessment Rating Tool (PART) that gave "[e]ffective," "[m]oderately [e]ffective," "[a]dequate," or "[i]neffective" ratings to many programs across the federal government.<sup>41</sup> In private conversations, career staff in the Labor Department admitted to me that the PART results reported by the department to OMB were not an honest reflection of performance in the department. Results were reported with the goal of achieving an "effective" rating rather than to disclose the truth of an agency's performance management and how it might improve. The same was true of the department's annual performance report under GPRA. In a meeting early in 2009 with the department's central performance staff, I asked if the glossy, aesthetically pleasing, award-winning report was a "report to report" or a report that was used to manage. Sheepishly, the staff admitted that it was a "report to report"

---

<sup>40</sup> U.S. DEP'T OF LABOR, FY 2016 BUDGET IN BRIEF 6 (2015), available at <http://www.dol.gov/dol/budget/2016/PDF/FY2016BIB.pdf>.

<sup>41</sup> *The Program Assessment Rating Tool (PART)*, EXPECTMORE.GOV, <http://georgewbush-whitehouse.archives.gov/omb/expectmore/part.html> (last visited Feb. 27, 2015).

2015]

MANAGING SOCIAL CHANGE

1001

and that managers did not use it to assess or improve their agency's performance.<sup>42</sup>

Later in my tenure, as we began to institute a new performance management system at the Labor Department, I learned that two-thirds of the department's operating units—including some of its largest and highest profile agencies—did not have annual operating plans. This meant that the agencies had not clearly articulated to the department's leadership, OMB, Congress, or the public what they expected to produce and how much they expected to produce from month-to-month and quarter-to-quarter, or how they made choices about what and how much they would produce. In many cases, there was no explanation of the relationship between agencies' day-to-day activities and the outcomes they promised in the department's strategic plan. In fairness, while GPRA required an annual departmental performance plan, it apparently did not require performance or operating plans from the department's constituent agencies.<sup>43</sup> However, a federated agency, like the Labor Department, necessarily operates through its constituent parts. It would be very difficult, if not impossible, to develop a genuine departmental performance plan without building on agency operating plans. In a world of responsible management, the departmental performance plan would amalgamate and summarize its agencies' performance plans. But that seemingly did not happen in the Labor Department prior to 2009. There was nothing to amalgamate. In other words, the department had achieved seeming GPRA compliance without actually managing well and responsibly throughout the department.

Three brief anecdotes about the condition of the Labor Department's performance in 2009 will demonstrate how these weaknesses in the department's performance management and measurement system were associated with bad outcomes for workers.

*i. Wage & Hour Division (Wage & Hour)*

Wage & Hour administers the nation's minimum wage, overtime, and child labor law—the Fair Labor Standards Act<sup>44</sup>—as well as the Family and Medical Leave Act<sup>45</sup> and two prevailing wage laws for federal contractors. One

---

<sup>42</sup> The staff's sheepishness may have been the product of their accurate reading of my body language that I wanted meaningful management reports rather than glossy reports-to-report. In their defense, it must have been difficult to dedicate themselves to producing a report that Congress, OMB, and the public, not to mention the agency's leadership, would not read. Further, OMB and GPRA included an extensive number of requirements for these reports such that even the best intentioned performance staff would have found it difficult to produce an accessible, rational, and worthwhile report.

<sup>43</sup> See *supra* Part II.A.1 for a discussion of GPRA's requirements.

<sup>44</sup> Fair Labor Standards Act of 1938, 29 U.S.C. § 201 (2013).

<sup>45</sup> Family and Medical Leave Act of 1993, 29 U.S.C. § 2601 (2013).

of the principal means by which Wage & Hour enforces these laws is to accept complaints from workers who believe they have not received pay or leave to which they are entitled. In addition to investigations it initiates on its own, Wage & Hour investigates seemingly valid complaints. If violations are found, Wage & Hour collects the workers' pay and returns it to them. In appropriate cases, it might seek to impose liquidated damages or civil money penalties on the law-breaking employer.<sup>46</sup> The expectation is not merely that Wage & Hour will remedy the instant violation, but that requiring the payment of unpaid wages and the potential imposition of penalties will deter future violations.

In 2009, the Government Accountability Office (GAO), an independent auditing arm of Congress, issued a scathing report essentially finding that Wage & Hour was performing this basic part of its job poorly.<sup>47</sup> Operating undercover, GAO filed 10 typical worker complaints with Wage & Hour's local offices across the country. Wage & Hour staff deterred these "workers" from filing complaints by encouraging them to resolve the issue themselves, directing most calls to voicemail, not returning phone calls to both employees and employers, and providing conflicting or misleading information about how to file a complaint. Wage & Hour policies required investigators to enter all reasonable complaints into Wage & Hour's database. Half of GAO's fictitious complaints were not so recorded. Looking beyond its undercover cases, GAO identified 20 real-life cases affecting at least 1,160 employees whose complaints were inadequately investigated by Wage & Hour. Five of the cases were closed based on false information provided by the employer that could have been verified by a search of public records to which Wage & Hour staff did not have access. GAO also found that Wage & Hour's investigations were often delayed by months or years, partly because backlogs in some offices prevented investigators from initiating cases within six months.<sup>48</sup> In sum, whether due to a lack of resources, poor management, inappropriate performance measures, ineffective systems, or underqualified or uninterested employees, Wage & Hour had failed to collect and act upon information and complaints from vulnerable workers who believed their rights had been violated.

*ii. Office of Federal Contract Compliance Programs (OFCCP)*

OFCCP enforces workplace anti-discrimination and affirmative action laws that apply specially to federal contractors and their subcontractors. Its

---

<sup>46</sup> Fair Labor Standards Act § 216(b).

<sup>47</sup> U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-09-458T, DEP'T OF LABOR: WAGE AND HOUR DIVISION'S COMPLAINT INTAKE AND INVESTIGATIVE PROCESSES LEAVE LOW WAGE WORKERS VULNERABLE TO WAGE THEFT (2009), available at <http://www.gao.gov/assets/130/122107.pdf>.

<sup>48</sup> *Id.* at 23.

performance problem was not a failure of competence. Rather, OFCCP's leadership under the Bush Administration had adopted management policies designed to inhibit finding and remedying workplace discrimination.

In most enforcement agencies, investigators go to the site of the alleged violation, review evidence, and interview potential witnesses, among other things, to determine the underlying facts. Prior to 2010, OFCCP investigators were directed not to conduct on-site reviews of workplaces. Instead, they were permitted to conduct only "desk audits," which essentially consisted of reviewing employer-provided documents at their desks. Desk audits that revealed compelling evidence of systemic discrimination would lead not to an on-site investigation, but to a further desk audit. But a "notice of violation" could be issued only after an on-site review. In the Orwellian world of poor management of government resources, this was called the "Active Case Management" system.<sup>49</sup>

The result was a system that only very rarely found workplace discrimination and often turned away cases that might have involved discrimination. After we instituted a new performance management regime in OFCCP, we learned that the agency's investigators made technical mistakes in more than 40% of their cases and substantive errors in about 25% of their cases. So, even in the rare instances when OFCCP stumbled upon workplace discrimination, there was an unacceptably high risk that OFCCP investigators would miss or ignore it.

### *iii. Mine Safety and Health Administration (MSHA)*

As noted above, MSHA protects the health and safety of miners and people working in sand and gravel pits, among other workplaces. In 2009, MSHA had been without an Assistant Secretary confirmed by the U.S. Senate for five years.<sup>50</sup> The employee who was in charge of this more-than-2,000-person organization when I arrived at the Labor Department in 2009 was a dedicated, but temporary, career executive whose most recent experience was running the Birmingham, Alabama, district office.

During this time, MSHA suffered employee turnover so severe that a large majority of its investigators and front-line supervisors had less than five years of experience.<sup>51</sup> Like most professional jobs, greater experience

---

<sup>49</sup> OFFICE OF FED. CONTRACT COMPLIANCE PROGRAMS, U.S. DEP'T OF LABOR, ACTIVE CASE MANAGEMENT PROCESS & PROCEDURES NO. 285 (2008).

<sup>50</sup> Kathy Snyder, *Stickler's Arithmetic*, MINESAFETYWATCH (Mar. 31, 2008, 1:01 PM), <http://minesafetywatch.blogspot.com/2008/03/sticklers-arithmetic.html>.

<sup>51</sup> MINE SAFETY & HEALTH ADMIN., U.S. DEP'T OF LABOR, INTERNAL REVIEW OF MSHA'S ACTIONS AT THE UPPER BIG BRANCH MINE-SOUTH 4 (2012) [hereinafter UPPER BIG BRANCH REVIEW], available at <http://www.msha.gov/performancecoal/UBBInternalReview/UBBInternalReviewReportNoappx.pdf>.

contributes significantly to greater effectiveness, more skills, and broader knowledge for mine safety and health inspectors. Only a few months after President Obama nominated, and the Senate confirmed, a new Assistant Secretary for MSHA, the Upper Big Branch mine in West Virginia exploded and 29 miners were killed—the worst mine disaster in 40 years.<sup>52</sup> MSHA certainly did not cause that mine explosion and kill those miners. Massey Energy Company, the owner of Upper Big Branch, did.<sup>53</sup> But MSHA's own internal review said that the inexperience of the investigators and supervisors who inspected Upper Big Branch resulted in Massey Energy Company getting away with violations of law that should not have been permitted.<sup>54</sup> Perhaps worse, Upper Big Branch was not the first mine disaster under the Bush Administration's watch. Two serious mine accidents in 2006, Sago and Aracoma, also caused the deaths of multiple miners. The Crandall Canyon mine collapse in 2007 not only killed miners working in the mine, but also produced an ill-considered and poorly managed MSHA effort to rescue miners that caused additional deaths.<sup>55</sup>

In sum, while the Labor Department largely complied with GPRA's technical requirements, it did not achieve high-quality performance in several of its largest and most important agencies. I have offered three examples, but there are others. While some parts of the Labor Department were well-managed and high performing leading up to 2009, GPRA could not claim credit for these results, as these agencies' systems pre-dated the statute. GPRA certainly did not avoid bad outcomes and bad performance in the remaining agencies. The next section posits reasons for GPRA's failures across large parts of the Labor Department.

---

<sup>52</sup> Ian Urbina, *No Survivors Found at Site of W.Va. Mine Disaster*, N.Y. TIMES, Apr. 10, 2010, at A1, available at [http://www.nytimes.com/2010/04/10/us/10westvirginia.html?page\\_wanted=all&\\_r=0](http://www.nytimes.com/2010/04/10/us/10westvirginia.html?page_wanted=all&_r=0).

<sup>53</sup> UPPER BIG BRANCH REVIEW, *supra* note 51, at 1.

<sup>54</sup> *Id.* at 26.

<sup>55</sup> See ELLIOT P. LEWIS, OFFICE OF INSPECTOR GEN., MSHA ACCOUNTABILITY PROGRAM COAL MINE SAFETY AND HEALTH MANAGEMENT LETTER NO. 05-06-007-06-001 (2006), available at <http://www.oig.dol.gov/public/reports/oa/2006/05-06-007-06-001.pdf>; MINE SAFETY & HEALTH ADMIN., U.S. DEP'T OF LABOR, INTERNAL REVIEW OF MSHA'S ACTIONS AT THE ARACOMA ALMA MINE #1 (2007), available at <http://www.msha.gov/readroom/FOIA/2007InternalReviews/Aracoma%20Internal%20Review%20Report.pdf>; MINE SAFETY & HEALTH ADMIN., U.S. DEPARTMENT OF LABOR, INTERNAL REVIEW OF MSHA'S ACTIONS AT THE SAGO MINE (2007), available at <http://www.msha.gov/Readroom/FOIA/2007InternalReviews/Sago%20Internal%20Review%20Report.pdf>; see also Steven Greenhouse, *Report Cites Mine-Safety Agency Failures*, N.Y. TIMES, Nov. 18, 2007, at 124, available at <http://www.nytimes.com/2007/11/18/us/18labor.html>; Matthew Davis, *US Mining Safety Under Scrutiny*, BBC NEWS (Jan. 5, 2006, 9:52 PM), <http://news.bbc.co.uk/2/hi/americas/4585482.stm>; William M. Welch, *Rescuers Pulled from Mine Indefinitely After Cave-In Kills 3*, USA TODAY (Aug. 19, 2007), [http://usatoday30.usatoday.com/news/nation/2007-08-16-utah-mine\\_N.htm](http://usatoday30.usatoday.com/news/nation/2007-08-16-utah-mine_N.htm).

C. *GPRA's Failings and the GPRAMA's Failure To Fix Them*

The glib answer to the question, “why didn’t the Labor Department perform better under GPRA” would be to blame the ideology or leadership of Secretary Elaine Chao and the Bush Administration. But, like most glib answers, this approach would not take us very far. If the leadership of the Labor Department or any other executive branch agency is not carrying out its organization’s mission, then an effective performance management system should disclose this fact and create mechanisms for holding leaders accountable. This is the very heart of GPRA: disclosure to Congress, and to a lesser extent the public, so that performance will be a part of the congressional debate about the role of government, the effectiveness of particular government programs, and the competence of particular governmental leaders. Again, GPRA failed. The first section will discuss five important reasons why it failed. The second section will explain that the GPRAMA—legislation intended to reform GPRA—did not fix these failures.

1. GPRA’s Failings

i. *No Congressional Accountability*

Ironically, in a system that is intended to create greater governmental accountability, GPRA neglected to hold Congress accountable for overseeing the GPRA process. As a result, executive branch entities like the Labor Department were able to get away with poor performance and pro forma GPRA compliance because they knew Congress was not paying attention. GPRA required agencies to consult with Congress regarding their strategic plans and little else.<sup>56</sup> But Congress’s laissez-faire attitude was not inevitable. Congress could have expressly mandated in GPRA that each of its authorizing committees or appropriations sub-committees hold at least one annual hearing on agencies’ compliance with and performance under GPRA. In preparation for these hearings, or even without the hearings, Congress could have directed the GAO or each agency’s Inspector General to conduct annual audits of every agency’s compliance and performance, rather than merely directing the GAO to produce one government-wide report on GPRA compliance in 1997.<sup>57</sup> Instead, Congress chose to require nothing of itself.

---

<sup>56</sup> Government Performance and Results Act of 1993, Pub. L. No. 103-62, § 3(d), 107 Stat. 285, 286 (codified as amended in scattered sections of 5 U.S.C. and 31 U.S.C.).

<sup>57</sup> *Id.* § 8. GAO has issued a series of general reports about aspects of GPRA and the GPRMA, but none like those for which I am advocating. See *Managing for Results in Government*, U.S. GOV’T ACCOUNTABILITY OFFICE, [http://www.gao.gov/key\\_issues/managing\\_for\\_results\\_in\\_government/issue\\_summary#t=1](http://www.gao.gov/key_issues/managing_for_results_in_government/issue_summary#t=1) (last visited Feb. 27, 2015) (listing reports).

Closely related, there are no consequences for an agency's failure to comply or for its poor performance. Budgets are not cut, and programs are not eliminated. Appropriations decisions in Congress are driven more by ideology and constituency politics than evidence.<sup>58</sup> But even if performance data were to drive government spending decisions, it would be hard to argue that these sanctions are directed at the agency's decision-makers and managers rather than the programs' beneficiaries. Decision-makers remain unscathed in either event. There is no requirement that political appointees and senior career leaders suffer discipline for a failure to comply with GPRA or poor programmatic or departmental performance. In the absence of congressional oversight and third-party audits, there is not even a shaming sanction. Similarly, there are no rewards in GPRA for compliance or outstanding governmental performance.

The same appropriations decision-making process that sustains poor performing programs also starves the best performing programs when they do not have sufficient political support. Top managers are given few incentives to improve performance. Bonuses for political appointees are closely scrutinized and generally disfavored.<sup>59</sup> GPRA says nothing about rewarding senior career managers when their programs provide excellent service to the American people or save taxpayers money.<sup>60</sup> With a near total absence of rewards and punishments, it is easy—even rational—to achieve pro forma compliance, like the Labor Department had prior to 2009, without actually producing good-quality government performance.

---

<sup>58</sup> See generally John Bridgeland & Peter Orszag, *Can Government Play Moneyball?*, ATLANTIC (June 19, 2013, 10:05 PM), <http://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/> (discussing Congress's unwillingness to eliminate programs even when there is evidence demonstrating the programs' ineffectiveness).

<sup>59</sup> 5 U.S.C. § 4508 (2012) prohibits bonuses to political appointees from June 1 of a presidential election year through the President's inauguration. After a scandal involving bonuses during the President George H.W. Bush Administration, President Clinton's chief of staff issued a memorandum urging agencies to refrain from giving bonuses to political appointees, although a GAO report found that some Clinton Administration appointees received such bonuses. See Tanya N. Ballard, *Despite Prohibition, Clinton Appointees Got Bonuses*, GOV'T EXECUTIVE (Dec. 13, 2002), <http://www.govexec.com/federal-news/2002/12/despite-prohibition-clinton-appointees-got-bonuses/13111/>. President George W. Bush allowed political appointees to receive performance bonuses. See Eric Lichtblau, *Bush Restoring Cash Bonuses for Appointees*, N.Y. TIMES (Dec. 4, 2002), available at <http://www.nytimes.com/2002/12/04/us/bush-restoring-cash-bonuses-for-appointees.html>. President Obama froze all bonuses for political appointees. See Scott Wilson, *Obama Orders Freeze on Bonuses, Monetary Awards for Federal Political Appointees*, WASH. POST (Aug. 4, 2010), <http://www.washingtonpost.com/wp-dyn/content/article/2010/08/03/AR2010080306839.html>.

<sup>60</sup> See *infra* notes 64–65 and accompanying text.

*ii. No Clear Definition of Success*

GPRA is almost entirely silent about a definition of “success” for executive branch performance. GPRA’s “purposes” section uses phrases like “waste,” “inefficiency in Federal programs,” “vital public needs,” and “improve program efficiency and effectiveness.”<sup>61</sup> These phrases mean little to agency staff attempting to set performance goals. By contrast, GPRA is specific about its process requirements. As a result, it necessarily communicates to the executive branch agencies that, to the extent they are judged at all, judgments will be based on processes rather than substance or practice. More directly, GPRA ratifies pro forma compliance because it does not offer guidance regarding the substance of compliance.

Admittedly, defining “success” for many government programs is difficult. Organic statutes often offer broadly stated goals, but not specific means of measuring achievement of those goals or any targets on the way to achieving those goals. For example, the congressionally articulated purpose for the Occupational Safety and Health Act of 1970, which created the Labor Department’s OSHA, is “to assure so far as possible every working man and woman in the Nation safe and healthful working conditions and to preserve our human resources.”<sup>62</sup> How does OSHA know whether it has succeeded in satisfying this congressional purpose? Should OSHA benchmark its performance against other agencies, its own past performance, or some guidepost in the economy? Merely counting up the number of workers killed, injured, or ill in unsafe and unhealthy workplaces would fail to take into account the limits on OSHA’s jurisdiction and resources, as well as the ebbs and flows of different industries over time in the U.S. economy even though some jobs and some industries are demonstrably more dangerous and unhealthy than others. It would also fail to accommodate any measure of the cost effectiveness of OSHA’s programs. Should OSHA measure lives and limbs saved divided by appropriated dollars spent, for example? GPRA has nothing to say about these important and difficult questions, and does not create a forum for Congress, the executive branch, and the public to debate them.

*iii. No Assurance that Agencies Use Data in Their Decision-Making*

Recall the story told above about the Labor Department’s annual performance report being a “report to report” rather than a management document. While tacitly favoring pro forma compliance, GPRA also failed to establish a means of looking behind pro forma compliance to assess actual compliance and executive branch performance. Yet, such means were

---

<sup>61</sup> Government Performance and Results Act of 1993 § 2.

<sup>62</sup> 29 U.S.C. § 651(b) (2013).

available. First, every five years, GAO conducts a study to ask federal executive branch managers whether their agencies use evidence-based, data-driven decision-making—in essence, asking managers if their agencies implement strategic and operating plans and rely on the performance data produced under GPRA.<sup>63</sup> The survey asks a host of questions that directly and indirectly disclose whether the agency manages its operations using the performance data it reports, along with other evidence available in a fulsome performance management and measurement system. The first such survey was administered in 1996, that is, before the first GPRA strategic plans were published. This survey could have been used as a means of validating agencies' adoption of evidence-based, data-driven management practices, particularly if the survey were administered annually and its results were translated for Congress and the public.

Second, GPRA did not require the establishment of close connections between agency performance and strategic plans, on the one hand, and the individual performance plans of career managers and employees, on the other. Every federal employee must have an individual performance plan that details the elements of good performance in his or her job.<sup>64</sup> Employees' supervisors and managers are required to use these elements to review the employees' performance at least annually. Raises and bonuses, or demotions and discharges, can result from these performance reviews. Yet, when Congress enacted GPRA, it did not amend the longstanding statutory provisions establishing and regulating this performance review process to require a direct connection between individual and agency performance.

The Office of Personnel Management regulations that implement the individual performance review statute specifically require agencies to consider GPRA results in Senior Executive Services (SES) managers' performance review. However, the regulations treat GPRA results as only one part of the process. There are five criteria, including customer satisfaction; “[e]mployee perspectives”; “[t]he effectiveness, productivity, and performance quality of the employees for whom the senior executive is responsible”; and affirmative action, equal employment opportunity, and diversity goals.<sup>65</sup> In other words, it is conceivable under the literal terms of the regulations that a SES manager could get a good, or very good, performance rating without having helped her agency achieve good or very good performance. Equally important, there is no requirement in the regulations that front-line employees (so-called “GS” or “GM” employees) are judged based on agency performance.

---

<sup>63</sup> U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-13-519SP, 2013 FEDERAL MANAGERS SURVEY ON ORGANIZATIONAL PERFORMANCE AND MANAGEMENT ISSUES (2013), <http://www.gao.gov/special.pubs/gao-13-519sp/index.htm>.

<sup>64</sup> 5 U.S.C. § 4302 (2013).

<sup>65</sup> 5 C.F.R. § 430.307(a)(2) (2013).

*iv. Failure To Require Evidence of Relationships Between Outputs and Outcomes*

There are several different kinds of performance metrics employed by agencies. According to GPRA, outcome measures offer “an assessment of the results of a program activity compared to its intended purpose.”<sup>66</sup> In other words, like MSHA’s mining fatalities metric, they measure the real-world effects of government programs as compared with their intended effects. Output measures describe the level of activity or the number of activities produced by a government program over a specified period.<sup>67</sup> An agency’s outputs are supposed to produce its designated outcomes.

Yet, many agencies have great difficulty measuring outcomes. Even defining outcomes can be difficult. For example, OFCCP enforces workplace anti-discrimination and affirmative action laws that apply to federal contractors. Is OFCCP’s outcome supposed to be federal contractor workforces that more closely approximate the demographics of the labor markets from which they are hiring—the classic definition of affirmative action’s beneficial effects?<sup>68</sup> Or is OFCCP’s outcome supposed to be federal contractors’ compliance with the laws it administers? If it is the latter, how should OFCCP measure this outcome? The agency undertakes compliance evaluations of roughly 2% of the universe of federal contractors covered by the laws OFCCP administers. How can it know from its evaluation of this small, non-randomly selected subset of the contractor universe whether compliance among all federal contractors is increasing? The fact is that OFCCP, and those who oversee its performance, cannot know on a quarter-by-quarter or year-by-year basis whether compliance is increasing or decreasing. Further, how can we know whether it is OFCCP’s interventions, as opposed to exogenous factors, that changed the compliance rate in any given year or overall? From the available performance measures, we cannot.

GPRA acknowledges this fact by requiring that agencies’ strategic plans include a description of the program evaluations it has used to set goals and objectives, along with a schedule for future program evaluations.<sup>69</sup> Performance evaluations are essentially social science studies of the relationship between outputs and outcomes, among other things.<sup>70</sup> However, GPRA does not require program evaluations where outcome-focused performance measures are difficult or impossible. Rather, it requires only that, when agencies choose to undertake program evaluations, they report them in

---

<sup>66</sup> Government Performance and Results Act, 31 U.S.C. § 1115(h)(7) (1993).

<sup>67</sup> *Id.* § 1115(h)(8).

<sup>68</sup> *See, e.g.,* United Steelworkers v. Weber, 443 U.S. 193, 208 (1979).

<sup>69</sup> Government Performance and Results Act, 5 U.S.C. § 306(a)(8) (1993).

<sup>70</sup> Government Performance and Results Act, 31 U.S.C. § 1115(h)(7).

their multi-year strategic plans. Agencies were permitted, under the GPRA structure, to report output measures without any hypothesis or evidence about the connection between the output and its intended outcome.

v. *Complexity of the Mandate*

GPRA requires the creation of a performance plan for every “program activity” in each agency.<sup>71</sup> In essence, this requirement imposes an obligation to set performance targets for each line of the federal budget.<sup>72</sup> But the budget of the U.S. government’s executive branch agencies is dizzyingly diverse and eye-glazingly complex to the point of frustrating this requirement. Again, the Labor Department offers a helpful example. Wage & Hour, which spent around \$229 million in fiscal year 2013, has one budget activity. OFCCP, which spent more than \$100 million, also has only one budget activity. By contrast, OSHA has ten, MSHA has eight, and ETA has dozens.<sup>73</sup> So, OSHA and MSHA were technically required to produce a performance target for their “compliance assistance” activities (i.e., help for employers to bring them into compliance with employment laws) because they have budget line items funding that activity. Wage & Hour and OFCCP, technically speaking, were not required to establish a performance target for the same activity. ETA was expected to have dozens, even hundreds, of performance measures across its many programs.

Perhaps with this complex diversity in mind, GPRA allows agencies, in consultation with OMB, to offer alternative forms of measurement, or to “aggregate, disaggregate, or consolidate program activities,” as long as “any aggregation or consolidation may not omit or minimize the significance of any program activity constituting a major function or operation for the agency.”<sup>74</sup> In the case of the Labor Department, the budget’s complexity and the laxity in Congress’s direction and oversight allowed two-thirds of the department’s agencies to avoid creating annual operating/performance plans and effectively avoid establishing performance goals for many of their functions.<sup>75</sup>

## 2. GPRAMA’s Failure To Fix GPRA’s Failings

As its description in the earlier part of this section should make clear, the GPRAMA did not fix these five failings of GPRA. In particular, the

---

<sup>71</sup> *Id.* § 1115(b).

<sup>72</sup> U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-05-734SP, A GLOSSARY OF TERMS USED IN THE FEDERAL BUDGET PROCESS 77 (Sept. 2005), available at <http://www.gao.gov/products/GAO-0-734SP>.

<sup>73</sup> U.S. DEP’T OF LABOR, BUDGET IN BRIEF: FISCAL YEAR 2014 62, 65, 6–7 (2013), [www.dol.gov/dol/budget/2014/PDF/FY2014BIB.pdf](http://www.dol.gov/dol/budget/2014/PDF/FY2014BIB.pdf).

<sup>74</sup> Government Performance and Results Act, 31 U.S.C. § 1115(d).

<sup>75</sup> See *supra* Part II.B.2 for my earlier discussion of this point.

GPRAMA did not aggressively assert congressional oversight of the GPRA process. There is no mandate regarding oversight hearings or GAO or Inspector General audits of agencies' annual performance plans. There is no interpretation or analysis of agencies' self-reporting of their performance at all. The GPRAMA did include a slight ratcheting up of congressional involvement. While GPRA had only generally required agencies to consult with Congress when developing their strategic plans,<sup>76</sup> the GPRAMA required congressional consultations "when developing or making adjustments to a strategic plan" and specified "including majority and minority views from the appropriate authorizing, appropriations, and oversight committees . . ."<sup>77</sup> Thus, Congress sought engagement in agencies' mission definition and goal-setting, but continued to abstain from playing a significant role in holding agencies accountable for their performance against their goals.

One provision in the GPRAMA seemingly moved Congress closer to a meaningful role in a performance accountability system. The GPRAMA requires OMB to determine whether the agencies' program activities have met their performance goals and objectives outlined in the agency performance plan. When performance goals are not met, OMB is expected to submit a report to the Senate's Committee on Homeland Security and Governmental Affairs (HSGAC) and the House's Committee on Oversight and Governmental Reform (COGR), as well as GAO.<sup>78</sup> If underperformance persists for more than one fiscal year, the GPRAMA requires further reports from OMB describing remediation plans, requesting more money, seeking authority to move already budgeted money to those underperforming programs, or proposing statutory reforms or elimination of the program.<sup>79</sup>

Superficially, these requirements represent the kind of congressional oversight that should hold agencies accountable for compliance and improving performance. But any deeper inquiry discloses that these reforms were simply more of the same. For example, note Congress's passivity in its putative oversight role. The committees responsible for general government operations receive information and plans from OMB, but there is no obligation imposed on the committees to act or even investigate when they are notified about performance problems. There is no reason to believe that these reports will be used for any purpose other than dust collection on congressional bookshelves or, perhaps worse, as tools for partisan attacks on disfavored programs. This concern is heightened by the fact that the specified committees are not the right forums for congressional action on most programs' performance. Their staffs

---

<sup>76</sup> See Government Performance and Results Act, 5 U.S.C. § 306 (1993).

<sup>77</sup> GPRA Modernization Act, 5 U.S.C. § 306(d) (2013). The same consultation requirement is imposed on the OMB Director in the process of choosing priority goals. 31 U.S.C. § 1120 (2013).

<sup>78</sup> 31 U.S.C. § 1116(f).

<sup>79</sup> *Id.* § 1116(g).

and members do not have the requisite expertise to participate meaningfully in reforming the hundreds of government programs that are outside their jurisdiction. Should HSGAC and COGR debate how to improve America's job training system for dislocated workers or should the Senate HELP and House Education and Workforce Committees? What about mine safety and health or pensions? If HSGAC and COGR were to attempt to dive into these issues, or any others arising out of the Labor Department, these authorizing committees would certainly and justifiably complain about infringements on their jurisdiction. Worse than congressional infighting, HSGAC and COGR simply do not have the competence to undertake substantive reform of these programs. While it is conceivable that HSGAC and COGR could function like clearinghouses referring performance issues to the appropriate authorizing committees, the GPRAMA neither requires nor encourages them to do so.

Perhaps most starkly, HSGAC and COGR do not decide the budget or appropriations questions that the GPRAMA requires OMB to submit to them. Answering those questions is the responsibility of the Senate and House Budget and Appropriations Committees. If Congress were serious about performance management, and moving toward performance-based budgeting as GPRA suggested it would, then the budget and appropriations committees would be at the center of the discussion. They are not. The best interpretation is that the centrality of HSGAC and COGR to this purported oversight process is yet more evidence that GPRA and the GPRAMA elevate form over substance. The GPRAMA created a dead end for performance information, not a true path to congressional oversight of executive branch performance.

The GPRAMA's path to performance accountability did not lead to Capitol Hill. Instead, the GPRAMA delegated Congress's oversight role to others. The first group was the deputy secretaries of each department, newly dubbed "Chief Operating Officers" (COOs).<sup>80</sup> The second overseer is OMB. The third overseer is the public. But the delegation to the first two groups—or, more precisely, the obligation imposed—was only partial. The COOs were required to meet quarterly with the responsible officials in their agencies to review performance, but only with respect to the agency's "priority goals."<sup>81</sup> The priority goal concept was apparently Congress's effort to emulate the widely accepted practice in corporate America of focusing performance management on "key performance indicators" that capture an organization's achievement of the goals that drive its success.<sup>82</sup> But Congress did not require a priority goal for every program activity, every agency operating unit, or even every major function in an agency. It left the requirement vague. As a result,

---

<sup>80</sup> *Id.* § 1123.

<sup>81</sup> *Id.* § 1121(b)(1).

<sup>82</sup> *See, e.g.,* DAVID PARMENTER, KEY PERFORMANCE INDICATORS: DEVELOPING, IMPLEMENTING, AND USING WINNING KPIs (2010).

the Labor Department—with 23 operating units, a \$12 billion discretionary budget, and dozens and dozens of products—had only three priority goals after the GPRAMA.<sup>83</sup> If I had followed the GPRAMA-mandated quarterly review process during my time as the Labor Department's COO, I would not have addressed most of the performance issues in the Labor Department and I would never have met with most of the Labor Department's constituent agencies. The GPRAMA requires very little oversight by the COOs.<sup>84</sup>

The same problem limits the GPRAMA's direction to the OMB Director. The Director is supposed to work with the agencies in the development of the federal government's priority goals and to meet quarterly with the agency officials in charge of those goals.<sup>85</sup> During my time as COO and Deputy Secretary, I never met with any of President Obama's OMB Directors to discuss the Labor Department's priority goals. I was not the responsible official for any of those goals, but I expect I would have been invited to join those meetings as the department's COO and principal performance overseer.<sup>86</sup> Even if those meetings had taken place, only three goals would have been discussed and most of the Labor Department's performance would have been ignored.

This leaves the public. The GPRAMA requires agencies to publish their annual performance plans on a public web site every February and data regarding actual performance within 150 days after the end of the federal government's fiscal year (i.e., September 30th plus 150 days).<sup>87</sup> Agencies' annual performance plans are expected to address all of their performance metrics for every program activity, not merely the so-called priority goals. Thus, the public is provided with a reasonably complete picture of agency performance on an annual basis. Congress gets only the bad news of unmet performance goals.

The problem is that the public is being given too much unprocessed data rather than actionable information. In particular, the GPRAMA did not address GPRAMA's failure to define successful government performance so that the public could judge actual performance against that definition. As a result, it is difficult for the public to know whether a particular agency is performing well or poorly. The public is told by each agency whether the agency has met its targets, but the public has no way to know if these are the right targets or why the targets were chosen. For example, because agencies have limited

---

<sup>83</sup> U.S. DEP'T OF LABOR, FISCAL YEAR 2013 ANNUAL PERFORMANCE REPORT 20 (2014), available at [www.dol.gov/dol/aboutdol/](http://www.dol.gov/dol/aboutdol/).

<sup>84</sup> See 31 U.S.C. § 1120(b) (describing agency priority goals).

<sup>85</sup> *Id.* §§ 1120(a)(1), 1121(a)(1).

<sup>86</sup> In fairness to my former colleagues at OMB, my staff and I met annually with OMB senior and budget staff to review the Labor Department's annual performance index, discussed *infra* in text accompanying notes 104–05.

<sup>87</sup> GPRAMA Modernization Act, 5 U.S.C. § 306(a); 31 U.S.C. § 1116(b)(1).

resources, lower targets set for one activity may be necessitated by increased investments in another activity. In the Labor Department's employment law enforcement agencies during the Obama Administration, an effort was made (to the extent permitted by budget structures) to invest more resources in enforcement by reducing spending on employer-focused compliance assistance. Even if this strategy was apparent, the public would not have had a means of knowing the extent to which it dictated performance targets related to enforcement and compliance assistance.

Similarly, unless it performs the analysis itself, the public cannot know whether a particular agency is meeting all or most of its targets. It is therefore very difficult to know whether an agency is well managed, poorly managed, or mediocre. There is no comparison across agencies or ranking of agencies' performance. There is no requirement that agencies disclose trend lines in performance in a manner that will allow the public to understand whether an agency is improving. There is no summary tool that provides an assessment of an agency's overall performance. Given the impenetrable complexity of the federal budget, and particular agencies' budgets, the public cannot know whether an agency is complying with the requirement to have goals and targets for every program activity.

Of course, the public, like Congress, does not have access to tools that disclose whether agencies used GPRA-generated data in their decision-making. As noted above, the GAO survey of executive branch manager is produced only every five years. As a result, agency performance that meets its target in intervening years could be entirely the product of happenstance rather than sound management or shrewd resource allocation. If a goal of GPRA is to allow the public to determine whether its government is well-managed, the mountain of data that the GPRAMA requires agencies to disclose to the public does not necessarily provide the answer, at least not in the form presented.

Perhaps most problematic, the public cannot know or study the relationships between agencies' outputs and the outcomes defined in their strategic plans. For example, a sizable percentage of the data in the Labor Department's annual performance reports relates to outputs: the number of investigations undertaken by an employment law enforcement agency, the quality of those investigations, the speed with which benefits checks are distributed to workers compensation beneficiaries, the number of on-site reviews of Job Corps centers, and the percentage of BLS data series produced on time, to name just a few.<sup>88</sup> This is valuable information that tells the observer whether the Labor Department is doing its job and whether its staff is working hard. But these data do not disclose the real-world consequences of the Labor Department doing its job. Ultimately, this is the most important

---

<sup>88</sup> See generally U.S. DEP'T OF LABOR, FISCAL YEAR 2013 ANNUAL PERFORMANCE REPORT (2014), available at [www.dol.gov/dol/aboutdol/](http://www.dol.gov/dol/aboutdol/) (listing these metrics and others).

information: whether government is serving its customers and constituents. Performance data may not provide this information to the public.

With all of these failings in GPRA and the GPRAMA, the public does not have the tools it needs to serve as an effective overseer of the federal government's performance. In section three, I will propose reforms to the GPRAMA and actions that stakeholder groups can take that will remedy Congress's failure to provide a meaningful system for overseeing government performance. The next part will describe efforts by the Labor Department's new leadership team to turn around the organization's performance beginning in 2009. This story of an important performance management and measurement success in a large and complex federal Cabinet department will further inform and serve as something like a proof of concept for the reform agenda to be described in section three.

### III. TURNING AROUND THE LABOR DEPARTMENT'S PERFORMANCE

Despite compliance with GPRA, the Labor Department faced serious management and performance challenges when I returned in 2009, as the preceding section explained. The new Obama Administration leadership team faced the challenge of turning around performance where it was failing without interfering with good (even excellent) performance in other parts of the department. We could not simply scrap everything that had come before and start anew, as some new management teams feel compelled to do. Further, while GPRA offered a worthwhile starting place for a performance turnaround effort, the preceding section explained that it did not offer a complete answer. I knew that Congress, and even OMB, were not going to hold the Labor Department (or me) accountable for excellent performance. I also knew that we would not face close, unsparing scrutiny from GAO and our IG on department-wide performance despite the generally high quality of their work on particular management issues.

Pro forma compliance with GPRA was not the only available option. The Labor Department's new leadership team could choose to comply with the law and also dramatically improve performance, as long as it understood that achieving the former did not inevitably produce the latter. Better results for the Labor Department's constituents would require the department's leaders to hold themselves and each other accountable. Before describing the process by which the Labor Department's leadership team established this measurement and accountability system, it is worth noting that the team's members—like many agency leadership teams in administrations of both parties—did not bring substantial management experience to this task.

Many executive branch presidential appointees arrive at their jobs after service as congressional staff (or members, at the Cabinet level), academics, lawyers, political operatives, non-profit organization or foundation executives,

think-tank leaders, or advocates.<sup>89</sup> Few have managed large or mid-size organizations. Except those with prior executive branch or state or local government experience, almost none have managed organizations as complex and surrounded by politics as executive branch agencies.<sup>90</sup> As a result, these presidential appointees do not bring formal management training, management experience, or any particular interest in management to their roles as senior federal managers.<sup>91</sup> They are typically hired for their policy expertise, their political acumen, their appeal to important constituencies, or all three, rather than their ability to measure and drive performance in a large bureaucracy. Compounding this challenge, there is little management or performance measurement training available for presidential appointees at any point during their typically short tenures. It is easy for presidential appointees to delegate management and measurement issues to their senior career executives and concentrate solely on policy and political concerns. These facts magnify the importance of the accomplishments of the Obama Labor Department's leadership team recounted in this section. The team overcame not only a broken congressional mandate, but their own lack of management expertise, to turn around performance in their organization and better serve workers and other constituents.

The first section below will describe the turnaround strategy implemented by the Labor Department's leadership team beginning after 2009. The second section will offer data and narrative demonstrating the strategy's success through Fiscal Year (FY) 2013 and the real-world results it produced for working families and others served by the Labor Department.

#### A. *The Turnaround Strategy*

*Strategic Plan.* The department needed an operational framework to improve performance management. The starting place was the GPRA-mandated strategic plan. In 2009, the department's most recent plan, drafted under Secretary Chao, was set to expire. We chose to use this opportunity—along with a one-year delay to facilitate it—to re-think this governing document for the department's performance. I was struck that the Chao

---

<sup>89</sup> In some administrations, this list would include lobbyists; however, President Obama banned recent lobbyists from service in his administration except in very limited circumstances. *See* Exec. Order No. 13490, 74 Fed. Reg. 467 (Jan. 21, 2009).

<sup>90</sup> *See supra* Part II.B.1 for discussion of the Labor Department's complexity. There is a rich debate to be joined—although not in this Article—regarding whether government agencies are more difficult to manage and complex than even the largest private-sector businesses. My view is that government entities like the Labor Department are significantly more challenging, although I admit that this perspective is biased by the experiences recounted in this Article.

<sup>91</sup> I do not exempt myself from this observation. To the contrary, as discussed below, I may have exemplified it.

strategic plan differed little from the strategic plan that Secretary Alexis Herman had put in place during President Bill Clinton's second term with my help.<sup>92</sup> It seemed unlikely, to say the least, that a center-left Democrat and a conservative Republican would have the same vision and goals for an organization that finds itself at the center of so many partisan struggles and controversial issues. Certainly, the public perception was that Secretary Chao led the Labor Department in a very different direction from Secretary Herman.<sup>93</sup> The more likely explanation for the similarity in the two secretaries' strategic plans was that these strategic plans mattered very little or not at all to the day-to-day management of the department during Chao's tenure.

The strategic plan is a leadership document. It sets forth the leadership team's vision, or "mission" in GPRA terms, for the organization. We scrapped the four-part, overgeneralized framework in the existing departmental strategic plan for a simple and capacious vision statement: "Good Jobs for Everyone."<sup>94</sup> This vision statement gave us the opportunity to define "good jobs" as consisting of those things that the Labor Department's agencies existed to create: skills training and job placement assistance; safe and healthy workplaces; workplaces free of discrimination; jobs with pensions and health care; a safety net when people leave their jobs due to layoffs, illness or injury, or retirement; and a voice for workers in workplace decision-making, among others.<sup>95</sup> "For everyone" allowed us to focus on the department's efforts to ensure equal opportunity and fair treatment for workers of color, women workers, workers with disabilities, veterans, immigrant workers, and workers in other countries. This portion of our vision statement fit very well with Secretary Solis's enthusiastic embrace of stakeholder communities and her commitment to ensure that those who had not been included in policymaking decisions in the past would be included during her leadership of the Labor Department.

*Outcome Goals.* In the language of GPRA, we deduced the department's and agencies' outcome goals from the vision statement and laid them out in the strategic plan.<sup>96</sup> We conceived these outcome goals to be the

---

<sup>92</sup> Compare U.S. DEP'T OF LABOR, STRATEGIC PLAN: FISCAL YEARS 1999–2004 (2000), available at <http://babel.hathitrust.org/cgi/pt?id=coo.31924092386634>, with U.S. DEP'T OF LABOR, STRATEGIC PLAN: FISCAL YEARS 2006–2011 (2006) (on file with author). I served as Counselor to the Secretary during President Clinton's second term.

<sup>93</sup> See Steven Greenhouse, *Departing Secretary of Labor Fends Off Critics*, N.Y. TIMES, Jan. 9, 2009, at A12, available at [www.nytimes.com/2009/01/10/washington/10chao.html](http://www.nytimes.com/2009/01/10/washington/10chao.html); *Profile: Secretary of Labor Elaine Chao*, ABC NEWS (Nov. 1, 2005), <http://abcnews.go.com/Politics/Inauguration/story?id=122080>.

<sup>94</sup> U.S. DEP'T OF LABOR, STRATEGIC PLAN: FISCAL YEARS 2011–2016 at ii (2010), [www.dol.gov/\\_sec/stratplan/StrategicPlan.pdf](http://www.dol.gov/_sec/stratplan/StrategicPlan.pdf).

<sup>95</sup> *Id.*

<sup>96</sup> *Id.* at 9.

department's promises to the people and organizations it serves, as well as the public, Congress, and the President. We felt we needed a persuasive response to the question, how does the Labor Department make your life, American society, and the American economy better? What outcomes should it produce for you? Although we did not use this lexicon, these promises were the Labor Department's key performance indicators—the goals that could be said to define whether the department had succeeded in carrying out its mission.<sup>97</sup> However, we also acknowledged that we could not always connect our agencies' every-day activities to outcomes that our constituents would feel in their daily lives.

*Program Evaluation.* As noted above, the department's outputs do not always self-evidently result in outcomes that are experienced by workers, employers, and other stakeholders and regulated entities. For this reason, in 2010, we established the Labor Department's first Office of the Chief Evaluation Officer.<sup>98</sup> This office, led in succession by two of the nation's best known and best respected program evaluation experts,<sup>99</sup> was charged with helping each agency to develop a five-year "learning plan" largely dedicated to developing evidence of how each agency's strategies and outputs produced either the outcomes promised in the strategic plan or similar and equally important outcomes. We also dramatically expanded the departmental budget for program evaluation with additional funds available through a new budget set-aside authority Congress approved for the department. At its core, this program evaluation agenda existed to gather evidence useful to holding the department, its agencies, and its staff accountable to the promises in the strategic plan and statutory goals and objectives.

*Operating Plans.* Program evaluations can take years to complete. Responsible leaders of large organizations cannot wait two or three years to hold their staffs accountable for their performance. The solution to this challenge was to require almost every agency in the department to have an annual operating plan. The operating plan served two purposes. First, it was the place where agencies articulated the connection between their day-to-day strategies and their outputs and the outcomes they expected to produce, including those promised in the departmental strategic plan. We required every agency operating plan to include either evidence of the causal connection between the strategies in the plan and the desired outcomes, or a testable hypothesis of that connection. The program evaluations listed in the agency's learning plan would test those hypotheses. For example, if an agency leadership

---

<sup>97</sup> See discussion in text accompanying note 82.

<sup>98</sup> See Harris, *supra* note 3.

<sup>99</sup> The Labor Department's two Chief Evaluation Officers: Professor Jean Grossman of Princeton University, who established the office, and Demetra Nightingale of the Urban Institute, who continues to lead the office at the time of this writing.

2015]

*MANAGING SOCIAL CHANGE*

1019

team believed that teaching workers about worker protection laws would increase employers' compliance with the law (as some did), we required evidence of that causal relationship, or a testable hypothesis and a plan to test that hypothesis through program evaluation. We also built a maturity model to assess agencies' progress in this regard.

Second, the operating plan was our principal accountability tool. It catalogued all of the agency's activities and set both quarterly and annual targets for each activity with an output measure or, if possible, an outcome measure. We did not allow ourselves to get bogged down in GPRA's definitional challenges of what constituted a budgetary program activity. Every activity that mattered to carrying out the agency's mission had to be measured. Any program or activity that was not being measured, by definition, should be considered for elimination because it was not important, with the exception of a smaller number of administrative activities. Returning to the same example, if an agency believes that educating workers about worker protection laws produces more compliance among employers, then the agency propounding that hypothesis was required to report how many workers it had educated about worker protection laws in each quarter and over the course of a year. But it was also required to report how many workplace investigations it planned to conduct, and what percentage of its activities met a specified quality standard, among other things. Targets set in isolation for individual activities do not disclose an agency's priorities and its strategies for achieving desired outcomes. The operating plan had to provide a complete picture of performance, not a mere focus on a few highlighted activities.

Three important points about the operating plans should be highlighted. Point number one: I required every agency, except for the smallest service agencies and our four groups of judges,<sup>100</sup> to have an operating plan. They were the centerpieces of our performance management system and our best means of holding agency leaders and staff accountable. The problem that GPRA permitted—two-thirds of the Labor Department's operating units did not have an annual performance or operating plan in 2009—could not be allowed to continue. It did not. When I left the department, 17 of 19 operating units (excluding the adjudicative boards) had operating plans. If planning for performance improves performance, then everyone had to have a plan.

Point number two: agencies were required to plan for continuous improvement in their operating plans. The agencies were required to set an annual target for every measure that was above the average of the agency's last

---

<sup>100</sup> The Labor Department's adjudicative bodies are data-driven organizations with a firm grasp on their workloads, productivity, and the quality of their products. However, the nature of their work made prioritization and targeting very difficult, if not impossible. Nonetheless, I held quarterly meetings with the chief judges of the adjudicative bodies in which we discussed various strategies for addressing backlogs and other issues. But operating plans for this purpose would not have made sense.

five years of actual performance on that same measure (or however many years up to five for which the agency had data). We did not require year-over-year improvement because too many exogenous factors—for example, unexpected increases in workload, natural disasters like Hurricane Sandy, human-made disasters like the Deepwater Horizon oil spill or the Upper Big Branch mine explosion, or surprising actions by Congress or other institutions—could intervene in any given year. Instead, we required an improvement *trend* demonstrated by surpassing the average of past performance. However, we permitted one category of exceptions to this rule. If the agency’s leadership had decided to deemphasize an activity, then the agency could set a target below the average of past performance. In fact, we insisted on it. We wanted agency heads to prioritize, spend fewer resources on low priority activities, and invest more in high priority activities. For example, as noted above, the Obama Administration prioritized enforcement of employment laws and education of workers about employment laws above employer-focused compliance assistance activities. Accordingly, employment law enforcement agencies were permitted to reduce targets for employer-focused compliance activities. As will be discussed below, this prioritization process was particularly important to proper budget formulation, especially when it became apparent that most budgets in the department would remain flat or decrease during much of my post-2010 tenure.

Even before Congress enacted it in the GPRAMA, we rejected a definition of “success” focused on meeting or beating performance targets. Particularly in the form included in the GPRAMA, this approach creates the wrong incentives. Recall that the GPRAMA requires OMB to report to Congress about any agency that fails to meet its performance goals and objectives. Agencies repeatedly failing to meet particular performance goals must submit remediation, budget, and reform plans that OMB would share with Congress. No agency wants to be the subject of a report to Congress about failed performance. It is embarrassing for the agency’s leaders and, despite Congress’s passivity on performance management issues, drastically increases the risk of partisan attacks on programs’ budgets, activities, purposes, and leaders. So, agencies will certainly set their targets at low levels which their leaders know they can easily achieve to avoid being included in this “failure report” to Congress. This is entirely rational behavior, at least in a political environment, but it also hurts performance.

Most employees seek to perform in a manner that meets or slightly exceeds their employer’s expectations. Low targets will mean low levels of performance or, more precisely, performance below the level that could be achieved with a more rational and aggressive targeting system. There are times when it is appropriate for agency leaders to set “stretch goals” that urge their employees to work harder, reengineer business processes, or reallocate resources to produce more in the coming year than has ever been produced before. I used this approach relentlessly with the Labor Department’s

leadership team in the firm belief that pushing targets higher generally forces positive change inside an organization. Even if the stretch goal is not reached, performance is likely to improve, and improve dramatically. I often told our leadership team that I preferred that they receive an “A-” on a stretch goal than an “A+” for exceeding a goal that lacked ambition. Of course, better performance is the goal of this entire enterprise, but the stretch-goals strategy for achieving better performance would be punished under the GPRAMA with the shaming sanction of an OMB report to Congress. It is entirely irrational and self-defeating to expect federal managers to put themselves in the difficult position of trying to explain these kinds of results to OMB and Congress.

Point number three: we sought to construct a “balanced scorecard” for each agency over time.<sup>101</sup> In the corporate world, the balanced scorecard represented a move away from an exclusive focus on bottom-line financial measures toward a more holistic assessment of organizational performance. It focuses on four distinct categories of measures: financial, customer response, internal process and production, and innovation and learning. We sought to adapt this approach to the particularized needs of the Labor Department. For Fiscal Year 2011, we focused on the development of a comprehensive set of production measures in every agency that would capture the quantity and timeliness (where appropriate) of every product and process promised by the agency in its operating plan. This effort necessarily produced a mix of outcome, output, and process measures. Because of many agencies’ missions, the mix included diverse measures of policy outputs, law enforcement activities, outreach efforts, customer service operations, and others products.

For Fiscal Year 2012, we worried that an excessive focus on quantity would drive down quality. Accordingly, we encouraged every agency to pair each quantity and timeliness measure with a quality measure. Several agencies, including OFCCP and ETA’s Trade Adjustment Assistance program (TAA program), developed “protocol fidelity” quality measures. In essence, the agencies identified the necessary steps to generating one of their products (investigations for OFCCP, certifications of benefit applications for the TAA program) and then assessed the percentage of instances in which employees followed each step of the protocol. We also required agencies to include some measure of customer satisfaction in their operating plans and, in November 2013, established a Customer Service Program Office to support the agencies. For Fiscal Year 2013, we incorporated a comprehensive set of administrative measures into the operating plans: financial management, human resources, information technology, procurement services, and other business functions.

*Accountability System.* Where GPRA, and later the GPRAMA, did not impose a thoroughgoing accountability system, we did. Every quarter, in my role as Chief Operating Officer, I held people accountable. I convened

---

<sup>101</sup> See generally ROBERT S. KAPLAN & DAVID P. NORTON, *THE BALANCED SCORECARD: TRANSLATING STRATEGY INTO ACTION* (1996).

quarterly meetings with every agency head to review every measure in their operating plans, not just the very few “priority goals” contemplated by the GPRAMA. Agency heads would report on whether their agencies’ actual performance during the quarter and year-to-date met, missed, or exceeded their targets. But we would also have rich, detailed discussions about strategies, tactics, resources, and priorities. We would discuss whether measures should be scrapped or added, and we would discuss evaluations being conducted that related to various measures. I would push for ever higher performance and performance targets, and the agency heads and their leadership teams would push back—often very convincingly—by explaining the unintended consequences of higher targets or excessive focus in one area or another. These quarterly performance meetings were not lovefests. I was blunt and, within the constraints that exist in any hierarchical organization, so were the agency leadership teams. But I promised everyone involved that I would never impose the death penalty on anybody for missing a quarterly or annual target. The worst thing that happened was they had to have a further discussion with me, which some of my colleagues compared unfavorably to being subjected to the death penalty.

The importance of my deep involvement in the process, playing the role of Chief Operating Officer, cannot be overstated. It was not because I brought a brilliant management mind with uncanny insights into organizational transformation. I did not. I had been a law professor before becoming Deputy Secretary, and my only previous management experience had been leading the Labor Department’s small policy office and a few dozen employees during the Clinton Administration. I had no formal training in management, much less performance management. So, it was not me. It was the fact that the person holding the title of Deputy Secretary was involved, reading and annotating all of the briefing memos, reviewing and commenting upon the charts disclosing each agency’s performance, attending every meeting, asking a long list of questions, and opining about how we could all achieve the Labor Department’s mission. It was the Hawthorne effect—my observation of and engagement in the process helped to change behavior.<sup>102</sup>

Another powerful force was the deep involvement of the department’s agency heads and their willingness, whether voluntary or in response to the pressure I applied, to take personal responsibility for operational excellence.

---

<sup>102</sup> At the same time, it is essential not to excessively personalize this success. The systems described in this Article required a sizable undertaking that involved coordination by a dedicated central staff of performance and budget experts and performance managers and analysts in the department’s agencies, meaningful and diverse organizational changes throughout the department, and a large resource investment both in the time of leaders and employees throughout the department, including my office and the Office of the Assistant Secretary for Administration and Management, and in consultants and technology that facilitated this transformation.

2015]

*MANAGING SOCIAL CHANGE*

1023

Each assistant secretary or agency director became deeply knowledgeable about and invested in his or her agency's operations, their successes, and their failures. Agency heads evaluated how the senior career managers who reported to them presented their agencies' performance and responded to my questions in the quarterly review meetings. They valued their senior career managers' competence in managing to achieve planned results. As a result, word spread through the department that all of the agency heads and I were reviewing actual performance against every measure. Front-line supervisors and employees whose performance was being measured took notice, made their own assessments of whether their performance would pass muster, and changed their behavior accordingly.

But we did not rely on the Hawthorne effect alone. We employed five additional accountability tools. The first tool was the most important: requiring agencies to link all managers' individual performance standards to their agency's performance under the agency's annual operating plan. Each SES manager's individual performance plan expressly referenced the agency operating plan, and specific performance results were incorporated into the SES manager's performance review. For front-line supervisors, we began the process of ensuring that the agency strategies and targets for which they were responsible were incorporated into their individual performance plans. In this way, we could be assured that managers and supervisors would take performance data into account in the performance of their work.

The second tool was less important, but it is worth discussing because it discloses my worst failure as the Labor Department's Chief Operating Office. In its refined and successful form, this tool involved the Labor Department's central performance staff providing each agency head with a ranking of its performance as compared with the performance of the other agencies in the Labor Department. In essence, agencies were ranked based on how many targets they significantly exceeded, how many they met, and how many they did not meet. These rankings were not published. Each agency head's report showed only where his or her agency was ranked, but did not identify the agencies that ranked higher and lower. The central performance staff, their boss, my staff, the Secretary, and I were the only Labor Department employees who knew how every agency ranked.

These extensive precautions were taken because, in 2011, we attempted to use this ranking tool to help determine the performance ratings and bonuses of each agency's SES managers. Simply, this tool was not mature enough to be used for this purpose. As a result, the rankings, ratings, and bonuses were an unmitigated disaster. Feelings were unnecessarily hurt. Doubts were raised about our entire performance management enterprise. I had long promised that performance management was not going to be used for punitive purposes, but here I was using the system to label managers and take money out of their pockets. It was a poor decision that was my fault. Fortunately, we recovered.

The third accountability tool was deployed within my last few months as the Deputy Secretary: we liberated almost all of the performance data to every employee in the Labor Department. While employees did not get to see the extensive briefing memos I received every quarter from every agency, they were given access to a dashboard and portal to almost all of the data. They also were given access to visualization tools to conduct their own analyses. Just before I left, agencies began reporting not only nationwide results, but performance results broken out by regional office. For example, OFCCP divides the United States into five regions. Employees were able to see data disclosing how each of these regions performed. So, OFCCP employees in one region could compare their own performance against OFCCP employees in other regions. But they could also compare their own performance against the OSHA or Wage & Hour employees in their own region. Employees were permitted to see data from all agencies, not merely their own. Ultimately, if a successful performance management and measurement system is going to survive, the employees, their unions, and front-line supervisors must own it. Liberating our performance data was designed to facilitate that result.

The fourth accountability tool involved integrating performance data and the departmental budget. We required that any agency submitting a request for additional funding during the department's annual budget cycle specify, with performance data and projections for future performance, how Congress granting the request would improve performance. Predicting improvement required the agency to have a full understanding of its successes and failures to date, and a clear set of priorities. This proved to be a difficult requirement for the agencies and the department's budget staff, if only for reasons of timing. In the odd world of federal budgeting, agencies' budgets must be written more than one year before the start of the fiscal year in which they would take effect. For example, the Labor Department's portion of the President's Fiscal Year (FY) 2016 budget, which would take effect on October 1, 2015, if enacted by Congress, was originally submitted to OMB in August or September 2014, in part so that the President's proposed budget could be issued in February 2015.<sup>103</sup> While this timing may make sense for the President and OMB, it is difficult to reconcile with the department's annual performance planning cycle. Agencies that were finishing up implementation of their FY 2014 operating plans and refining their FY 2015 operating plans in August or September 2015 were expected to predict how they would perform in FY 2016 with added resources. This is a challenging thought exercise, but it required agencies' leadership teams to think deeply about performance, performance planning,

---

<sup>103</sup> See Memorandum from Brian C. Deese, Deputy Director of the Office of Mgmt. and Budget, to Heads of Departments and Agencies, Fiscal Year 2016 Budget Guidance (May 5, 2014), available at <https://www.whitehouse.gov/sites/default/files/omb/memoranda/2014/m-14-07.pdf>.

prioritizing of activities and outcomes, and the real value of their existing resources as well as any additional resources Congress might provide.

The fifth accountability tool involved the creation of an annual performance index that sought to show the entirety of the Labor Department's performance.<sup>104</sup> We sought to establish a yardstick for the department's overall performance management system. We wanted it to disclose not merely whether a particular agency or a particular function within an agency had improved, but whether we had produced improved performance across the department. In a sense, this was my report card as the Chief Operating Officer, and I used it as a performance report to my superiors. I met with Secretary Solis in 2012 and 2013, and Secretary Tom Perez in 2014, to report on the previous year's performance using the annual performance index. I instigated meetings with the Deputy Director of OMB for Management and the OMB Labor Branch<sup>105</sup> and its political boss each year for the same reason with the same tool in hand. As discussed earlier, I attempted to meet with congressional staff to brief them on the annual performance index, but to no avail. In the next section, I will explain the standards used in the annual performance index to define and judge whether the Labor Department and its leadership were successful in improving performance.

### *B. The Results*

There is no other enterprise in the American economy that performs all of the same functions performed by the Labor Department. The department is not like McDonald's which can compare, or "benchmark," its performance against the performance of Burger King, Wendy's, and Arby's. The Labor Department has no competitors and no comparators. As a result, we could only benchmark the Labor Department's performance against its own past performance to assess whether there had been improvement. In the process, we were forced to answer the question that GPRA and the GPRAMA did not answer: what is success in performance management?

We defined "success" by looking at the answers to two questions. The first question was, for every measure, did this year's performance exceed performance during all of the prior years for which we have data? In other words, was this year's performance the best performance yet? We did not literally mean "yet"; rather, we set a five-year horizon on our benchmarking to ensure a fair comparison. The second question was, for every measure, did our performance beat the average of the prior years' performance for that measure? Again, we used a five-year horizon, so the agencies would average the past five

---

<sup>104</sup> U.S. DEP'T OF LABOR, PERFORMANCE INDEX: 2013 [hereinafter LABOR PERFORMANCE INDEX] (on file with author).

<sup>105</sup> This is the division of OMB that oversees the Labor Department's budget and policy processes.

years of performance, or as many years of data as they had up to five years, and compare the result to the present year's actual performance.

The following results from FY 2013, which is the last year for which data were available at the time of this writing, demonstrate the success of the turnaround strategy. For all Labor Department performance measures for which prior years' data were available, FY 2013 was the best year of performance on 57% of all measures.<sup>106</sup> That was an increase from FY 2011 and FY 2012. In FY 2013, performance on three-quarters of all measures beat the average—again, an increase from FY 2011 and FY 2012, where the numbers were already high.<sup>107</sup> In sum, performance improved in a very large majority of the Labor Department's functions and activities.

These broad-based improvements in performance were partly the product of the rule for agency operating plans described above: every annual target for every measure had to be set above the average of the agency's last five years of performance. Agencies were required to plan and strive for our definition of "success," they did, and they achieved success in the large majority of their activities. But recall, also, that there was an exception: if an agency's leadership team sought to deemphasize an activity, then it could set a below-average target. To put the 75% above-average performance result in perspective, and to show how truly broad-based performance improvements were, in about half of the instances in which agencies failed to exceed the average of their prior performance, it was the product of a conscious management decision to deemphasize the activity and reduce actual performance.<sup>108</sup> It was not poor performance. In other words, performance improved when management wanted it to improve and declined when management wanted it to decline for more than 85% of measures.

For one-quarter of the remaining 15% of measures, performance failed to exceed the average of past performance within a variance of 5%.<sup>109</sup> In other words, the agencies just barely missed succeeding by our standard. In another small share of cases, surprises like unexpected increases in workload caused performance in FY 2013 to fall below the average of prior years.<sup>110</sup> In sum, genuinely problematic performance was the cause of below-average performance for only 5% of measured activities.<sup>111</sup> By any measure, this is an impressive result.

It is worth addressing several potential challenges to these broad-based improvements in performance. The starting place should be budgets. Many of

---

<sup>106</sup> LABOR PERFORMANCE INDEX, *supra* note 104, at 3.

<sup>107</sup> *Id.* at 8.

<sup>108</sup> *Id.* at 10.

<sup>109</sup> *Id.*

<sup>110</sup> *Id.*

<sup>111</sup> *Id.*

2015]

## MANAGING SOCIAL CHANGE

1027

the Labor Department's agencies, but not all, received increases in their FY 2010 appropriations, which was President Obama's first budget cycle. We conducted extensive discussions within the department regarding when the performance benefits of budget increases would and should be felt, but we never established a causal relationship or succeeded in quantifying or pinpointing the timing of its effects. Nonetheless, budget increases can have delayed effects that might have shown up in FY 2013. For example, many agencies used their additional funds to hire more employees. In most cases, these employees needed training and close supervision in their earliest days, so I was told not to expect meaningful performance improvements during their first year on the job. More training and more experience was expected to increase their productivity over time, and the reduced demand for close supervision should have freed more experienced staff to be more productive over time. But we did not have real-time evidence of these effects or their scope. They were hypotheses advanced by senior career managers based on their experience that I found plausible.

Nonetheless, the excellent performance results in FY 2013 cannot be entirely credited to additional budget resources received in FY 2010, and may not be attributable to larger budgets at all. These performance improvements were achieved in FY 2013, which was the fiscal year in which Congress "sequestered" 5% of the funds from almost every federal agency's budget and almost every budget program activity.<sup>112</sup> In essence, almost every Labor Department program activity had fewer resources in FY 2013 than in FY 2011 and FY 2012. Agencies were forced to cut back training, travel, overtime, procurement of technology applications and hardware, and other productivity enhancing activities. Performance should have suffered, and perhaps it did, but it did not decline. It improved. Further, even the agencies that did not receive meaningful budget increases in FY 2010 (e.g., OWCP and BLS), and one agency that suffered a real cut in its FY 2010 budget (i.e., the Office of Labor Management Standards) accomplished performance improvements in FY 2013 alongside their colleagues in agencies that received more resources. Money alone does not explain the Labor Department's performance improvement. But budgets cannot be entirely discounted when discussing performance. More resources certainly make better performance easier, although they do not guarantee it.

Challenges beyond budgets and their impacts should also be addressed. Skeptics reading the Labor Department's FY 2013 performance results might ask whether the leadership team, led by its Chief Operating Officer, somehow "juiced" the numbers to make overall performance look better than it actually was. For example, did we add a sizable number of new performance measures

---

<sup>112</sup> See U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-14-244, 2013 SEQUESTRATION: AGENCIES REDUCED SOME SERVICES AND INVESTMENTS, WHILE TAKING CERTAIN ACTIONS TO MITIGATE EFFECTS (2014) [hereinafter GAO-14-244], available at [www.gao.gov/assets/670/661444.pdf](http://www.gao.gov/assets/670/661444.pdf).

to our operating plans where we knew performance would inevitably improve? The following data should prove that the answer is “no.”<sup>113</sup> We tracked the results for all of the performance measures that had been instituted before 2009—that is, before the Obama-appointed Labor Department leadership team took over. In FY 2013, the department had its best year of performance yet on 58% of those measures, or roughly the same results as for all Labor Department measures.<sup>114</sup> Performance on 80% of these measures beat the average of our past performance in FY 2013, or slightly better than the department’s performance on all of its measures.<sup>115</sup> Again, these were improvements over FY 2011 and FY 2012. In sum, there is no support for the argument that the department cheated with respect to the kinds and quality of performance measures instituted. To the contrary, the new measures we instituted were a sincere effort to account for and improve performance.

Another reasonable challenge to the department’s excellent performance results would be to ask whether the leadership team or the agencies overloaded the department’s portfolio of performance measures with unimportant process or administrative measures to produce a more impressive aggregate result.<sup>116</sup> We tracked the department’s performance on its outcome

---

<sup>113</sup> In addition, I can testify that it is not true.

<sup>114</sup> LABOR PERFORMANCE INDEX, *supra* note 104, at 7.

<sup>115</sup> *Id.*

<sup>116</sup> I reject any diminution in the importance of measuring administrative functions like human resources, financial management, information technology management, and procurement services. These functions are often defined as “business services” that are outside the core mission and core competencies of government entities. This is a cramped view of these functions. They are essential to the success of the other functions that may be more easily defined as serving an agency’s mission.

For example, early in my tenure as Deputy Secretary, for a variety of reasons, the department’s financial managers were having difficulty paying the department’s bills on time. Under the Prompt Pay Act, 31 U.S.C. §§ 3901–3907 (2012), agencies that do not pay a bill within a fixed time frame must pay a specified interest penalty. Late payments, at one point, cost the department more than \$1 million in interest penalties. We introduced a performance measure that tracked whether bills were paid on time. Accrued penalties declined dramatically and ultimately fell below the federal government’s standard. Most important, the money saved was available for investment in additional personnel, improved information technology, or other productivity-enhancing expenditures.

Similarly, the speed with which we hired new employees is an important measure, and a measure employed across the federal government. Every vacant position reduces productivity in any agency that depends, like virtually all Labor Department agencies, on labor-intensive efforts for their success. Assessing the quality of hires, or at least hiring managers’ satisfaction with new employees, was another important measure. Quickly hiring unqualified people is not a path to improved performance. In the same vein, the speed and accuracy with which the department’s information technology “help desk” responded to employees’ inquiries is an important contributor to productivity given the ubiquity of technology in modern work life.

It is important to add emphatically that not all administrative measures are important or should survive in a mission-focused performance management system. For example, no one cares

measures—that is, the very important measures that assess real-world impacts on workers and the department’s other customers.<sup>117</sup> In 2013, performance on 72% of the department’s outcome measures was the best yet.<sup>118</sup> Performance on 89% of outcome measures beat the average of past performance.<sup>119</sup> Again, FY 2013 was better than FY 2011 and FY 2012.

There is an important caveat that should be added to these impressive results on outcome measures. Some sizable portion of the Labor Department’s outcome measures are associated with ETA’s job training programs, their success in placing workers into jobs, and whether those workers are retained in their jobs. Of course, the United States suffered a catastrophic finance-sector-led economic downturn beginning in 2007 with a jobs recession that continued long thereafter. After reaching its peak at around 10% in late 2009,<sup>120</sup> the unemployment rate began to improve, Americans slowly returned to work, and employers began haltingly to hire again in 2011, 2012, and 2013.<sup>121</sup> This improvement in the labor market was an exogenous factor that almost certainly effected some unknown increase in ETA’s “entered employment rate” (i.e., success in placing workers exiting job training programs into unsubsidized jobs) and its “retained employment rate” (i.e., success of these workers in keeping their jobs for several quarters) in all of the many job training programs it funds.<sup>122</sup>

While this fact does not minimize the improvement in the outcomes experienced by the customers of ETA-funded job training programs, it makes crediting ETA with the improvements more difficult. As a result, the stunning improvements in outcomes measures seen in FY 2013 may be somewhat less amazing. But not all outcome measures belonged to ETA and its job training programs. Some related to workplace fatalities and injuries overseen by MSHA and OSHA. Other agencies also contributed outcome measures. So, this caveat does not undermine the conclusion that departmental performance on outcome

---

or should care how many paper clips an employment law enforcement agency buys. But these kinds of discussions about which measures are meaningful and meaningless was a critical part of the Labor Department’s turnaround effort.

<sup>117</sup> LABOR PERFORMANCE INDEX, *supra* note 104, at 8.

<sup>118</sup> *Id.*

<sup>119</sup> *Id.*

<sup>120</sup> *Databases, Tables & Calculators by Subject*, U.S. DEP’T OF LABOR: BUREAU OF LAB. STAT., <http://data.bls.gov/timeseries/LNS14000000> (last visited Feb. 27, 2015).

<sup>121</sup> *Id.*; see also Lucia Mutikani, *Recession Fears Ease on Factory, Hiring Data*, REUTERS (Aug. 31, 2011), <http://www.reuters.com/article/2011/08/31/us-usa-economy-idUSTRE77U25D20110831>.

<sup>122</sup> These are two of four so-called “WIA common measures” required by the Workforce Investment Act of 1998, Pub. L. No. 105-220 § 136(b)(2), 112 Stat. 936, 999, and its successor, the Workforce Investment Opportunity Act, Pub. L. No. 113-128, § 116(b)(2), 128 Stat. 1425 (2014). The third and fourth measures are average (or median) earnings in unsubsidized employment and attainment of industry-recognized credentials.

measures improved significantly, although perhaps not as much as these gaudy numbers might suggest.

In fact, a look at some of the Labor Department's outcomes in FY 2013 shows the importance of this turnaround strategy to the lives of working families. In FY 2013, the department achieved the lowest fatality rate for miners in the United States, ever.<sup>123</sup> The number of miners who died in workplace accidents also was the lowest, ever. The number of workplace injuries in mines was the lowest, ever. The department achieved the lowest rate of fatalities in general industry, ever.<sup>124</sup> Over the five years ending in FY 2013, the department achieved the lowest fatality rate in the construction industry, ever.<sup>125</sup>

In FY 2013, the department conducted the largest number of whistleblower investigations, ever.<sup>126</sup> In FY 2013 alone, the department responded to more complaints from miners suffering retaliation for raising health and safety concerns than were responded to in the entire second term of the Bush Administration or the entire second term of the Clinton Administration. Over the first five years of the Obama Administration, the Labor Department returned more than \$1.1 billion dollars in wages lost through wage theft to the workers who earned them. The department also did the best job, ever, of targeting Wage & Hour investigations to the workplaces that had violations.

In FY 2013, the department conducted the largest number of pension and health plan civil investigations and criminal investigations, ever.<sup>127</sup> Over the five years ending in FY 2013, the department recovered more than \$1.3 billion in pension and health plan benefits for more than 710,000 participants and beneficiaries. In FY 2013, the department assured that the largest percentage, ever, of workers exiting Labor Department job training programs got industry-recognized credentials that would increase their likelihood of finding and keeping jobs.

In FY 2013, the department did the best job, ever, of finding the very small number of union officers and staff who embezzled funds or engaged in fraud. The department also concluded investigations of union elections faster than ever before with the fewest resources available in more than a decade. The department helped the largest percentage of federal employees who acquired disabilities and collected workers compensation to return to work, ever. The department also processed workers compensation claims for longshore workers, federal employees, and energy employees at the fastest clip, ever.

---

<sup>123</sup> LABOR PERFORMANCE INDEX, *supra* note 104, at 13.

<sup>124</sup> *Id.* at 12.

<sup>125</sup> *Id.*

<sup>126</sup> *Id.* at 15.

<sup>127</sup> *Id.* at 17.

2015]

## MANAGING SOCIAL CHANGE

1031

The department accomplished these outcomes, and many others, while securing clean financial audits for five consecutive years and successfully implementing two large-scale information technology procurements. In FY 2013, the department did the best job, ever, of paying its bills on time, and paid its small business contractors faster than ever before. The percentage of the Labor Department's contracts that were entered into with small businesses also was the highest, ever.

There was no conflict between achieving these progressive outcomes and significantly reforming employment law through regulations and other processes. While accomplishing this broad-based performance improvement, the department expanded minimum wage and overtime protections to nearly two million home health aides.<sup>128</sup> The department strengthened affirmative action for people with disabilities and veterans.<sup>129</sup> It significantly strengthened its regulations protecting miners from black lung<sup>130</sup> and proposed a new regulation to protect workers from exposure to respirable silica.<sup>131</sup> The department also promulgated almost two dozen rules in cooperation with the Treasury and Health and Human Services Departments implementing the President's Patient Protection and Affordable Care Act.<sup>132</sup> This is only a small sampling of the department's regulatory output during the first five years of the Obama Administration. And it does not include the department's involvement in the President's legislative, trade, and foreign policy agendas.

Of course, these accomplishments did not occur in a vacuum. They were achieved despite a very difficult environment for the federal government and federal employees, as well as the people and organizations served by the Labor Department. The department achieved these results while responding to the worst financial crisis in seven decades,<sup>133</sup> the largest economic stimulus

---

<sup>128</sup> See Application of the Fair Labor Standards Act to Domestic Service, 29 C.F.R. § 552 (2013).

<sup>129</sup> See Affirmative Action and Nondiscrimination Obligations of Contractors and Subcontractors Regarding Individuals with Disabilities, 78 Fed. Reg. 58681 (Sept. 24, 2013) (to be codified at 41 C.F.R. § 60-741); Affirmative Action and Nondiscrimination Obligations of Contractors and Subcontractors Regarding Special Disabled Veterans, Veterans of the Vietnam Era, Disabled Veterans, Recently Separated Veterans, Active Duty Wartime or Campaign Badge Veterans, and Armed Forces Service Medal Veterans, 78 Fed. Reg. 58613 (Sept. 24, 2013) (to be codified at 41 C.F.R. §§ 60-250, 60-300).

<sup>130</sup> See Lowering Miners' Exposure to Respirable Coal Mine Dust, Including Continuous Personal Dust Monitors, 79 Fed. Reg. 24813 (May 1, 2014).

<sup>131</sup> Occupational Exposure to Respirable Crystalline Silica, 78 Fed. Reg. 177 (proposed Sept. 12, 2013).

<sup>132</sup> See *Affordable Care Act*, U.S. DEP'T OF LAB., <http://www.dol.gov/ebsa/healthreform/> (last visited Feb. 27, 2015) (listing regulations and guidance).

<sup>133</sup> See *Recession Took Bigger Bite Than Estimated*, BLOOMBERGBUSINESS (July 29, 2011), <http://www.bloomberg.com/news/2011-07-29/recession-took-bigger-bite-out-of-u-s-economy-than-previously-estimated.html>.

legislation ever,<sup>134</sup> the worst oil spill ever,<sup>135</sup> the worst mine disaster in four decades,<sup>136</sup> two shut downs of the federal government and several threatened shut downs,<sup>137</sup> at least two potential defaults on the U.S. government's debts,<sup>138</sup> sequestration of the funds necessary to the accomplishment of the department's mission,<sup>139</sup> and a three-year pay freeze for federal employees.<sup>140</sup>

Outcome measures are not the only way of assessing the department's success in achieving its mission. Another method is to consider performance on the department's key performance indicators—the outcomes and outputs promised in the department's revamped strategic plan—including many of those discussed in the preceding paragraphs. The department's analysis showed that performance on those strategic plan measures in FY 2013 was essentially the same as performance on all Labor Department measures. The department had its best performance yet on 53% of strategic plan measures.<sup>141</sup> Performance on 71% of strategic plan measures beat the average of past performance.<sup>142</sup> So, the strategic plan indicators appear to be representative of departmental performance overall. They also demonstrate that the department performed well in the activities that the leadership team considered to be most important to the accomplishment of the Labor Department's mission.

---

<sup>134</sup> American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115 (2009).

<sup>135</sup> Campbell Robertson & Clifford Kraus, *BP May Be Fined up to \$18 Billion for Spill in Gulf*, N.Y. TIMES, Sept. 5, 2014, at A1, available at <http://www.nytimes.com/2014/09/05/business/bp-negligent-in-2010-oil-spill-us-judge-rules.html>.

<sup>136</sup> See *supra* Part II.B.2.iii.

<sup>137</sup> See Carl Hulse, *Budget Deal To Cut \$38 Billion Averts Shutdown*, N.Y. TIMES, Apr. 8, 2011, at A1, available at <http://www.nytimes.com/2011/04/09/us/politics/09fiscal.html>; Jonathan Weisman & Jeremy W. Peters, *U.S. Government Is Shutting Down in Fiscal Impasse*, N.Y. TIMES, Oct. 1, 2013, at A1, available at <http://www.nytimes.com/2013/10/01/us/politics/congress-shutdown-debate.html>; Ashley Parker & Jeremy W. Peters, *House Republicans Ready Plan To Avoid a Government Shutdown*, N.Y. TIMES, Dec. 3, 2014, at A18, available at <http://www.nytimes.com/2014/12/03/us/house-gop-weighs-symbolic-immigration-vote-in-plan-to-avoid-shutdown.html>.

<sup>138</sup> See Jackie Calmes & Carl Hulse, *Debt Ceiling Talks Collapse as Boehner Walks Out*, N.Y. TIMES (Apr. 8, 2011), <http://www.nytimes.com/2011/07/23/us/politics/23fiscal.html>; Ashley Parker & Annie Lowrey, *Boehner Pledges To Avoid Default, Republicans Say*, N.Y. TIMES, Oct. 4, 2013, at A1, available at <http://www.nytimes.com/2013/10/04/us/politics/debt-limit-impasse.html>.

<sup>139</sup> See GAO-14-244, *supra* note 112.

<sup>140</sup> Josh Hicks, *Bill Would Increase Federal Pay by 3.3 Percent Next Year*, WASH. POST (Mar. 26, 2014), <http://www.washingtonpost.com/blogs/federal-eye/wp/2014/03/26/bill-would-increase-federal-pay-by-3-3-percent-next-year/>.

<sup>141</sup> LABOR PERFORMANCE INDEX, *supra* note 104, at 9.

<sup>142</sup> *Id.*

The bottom line of all of these data analyses is that the department, by almost any reasonable measure, did not merely beat its past performance in FY 2013, and FY 2012, and FY 2011. The department got better at continuous improvement every year, with FY 2013 performance achieving the highest level in recent memory, and perhaps ever. Even more important, the workers and others who are served by the Labor Department experienced these improvements in their lives and were better off because the department improved its performance.

There is one more challenge to these conclusions that raises doubts about the basic premise that better performance was the result of improved management. It is fair to ask, was all of this mere happenstance? Do we know that Labor Department managers and supervisors actually used performance data to manage performance in 2013 and produce these dramatic and broad-based performance improvements? Can these results be attributed to evidence-based, data-driven management practices? Fortunately, the GAO survey, discussed in the first section of this Article, answers these questions.

In 2012, GAO surveyed managers in all of the departments that make up the President's Cabinet and ten other large executive branch agencies.<sup>143</sup> Three-quarters of managers answered the survey as part of a stratified sample. GAO asked these managers several dozen questions about whether their agencies used evidence-based, data-driven decision-making throughout the department. Did their supervisors use it? Did their employees use it? Did they use it? Were they held accountable for using it? Was every imaginable kind of measure used?<sup>144</sup> The executive branch managers answering the survey were given a spectrum of choices ranging from (paraphrased) "No, we didn't use that type of evidence and data-driven management at all" to "we used it to a very great extent"—the top rating.<sup>145</sup>

The Labor Department's Office of the Chief Evaluation Officer analyzed the results of the survey and found that the Labor Department received the top rating—that is, its managers said it used specific evidence-based, data-driven practices "to a very great extent"—more than any other executive branch agency.<sup>146</sup> The analysis also found that the Labor Department placed fourth out of twenty-four executive branch agencies, and ahead of all other Cabinet departments, with respect to the percentage of managers giving

---

<sup>143</sup> See U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-13-518, *MANAGING FOR RESULTS: EXECUTIVE BRANCH SHOULD MORE FULLY IMPLEMENT THE GPRA MODERNIZATION ACT TO ADDRESS PRESSING GOVERNANCE CHALLENGES* (2013).

<sup>144</sup> See *Managing for Results: 2013 Federal Managers Survey on Organizational Performance and Management Issues*, U.S. GOV'T ACCOUNTABILITY OFF. (2013), available at <http://www.gao.gov/special.pubs/gao-13-519sp/index.htm>.

<sup>145</sup> *Id.*

<sup>146</sup> U.S. DEP'T OF LABOR, DEPARTMENT OF LABOR ANALYSIS OF DATA FROM THE GAO FEDERAL MANAGER SURVEY 7 (2013) (on file with the author).

the top two ratings.<sup>147</sup> In September 2014, GAO issued its analysis of the survey results showing that the Labor Department was one of only two agencies with a statistically significant improvement in its average score from 2007 to 2012.<sup>148</sup> GAO also found that the Labor Department's average score on all questions was the third highest in the executive branch and the highest among all Cabinet departments.<sup>149</sup> This is powerful circumstantial evidence that the department's dramatic and broad-based improvement in performance was the result of a commitment to using performance data and analysis to manage and make decisions.

The next section offers a reform agenda that would increase the likelihood that these kinds of performance improvements will occur across the federal executive branch, including in its labor-related agencies.

#### IV. A PROGRAM FOR REFORM

The Labor Department's turnaround strategy succeeded in producing good, improving performance across functions and agencies. This is not to suggest that the department's improvement process was complete when I left in January 2014. I often joked with my colleagues just prior to my departure that we had completed only the fourth year of a ten-year process. There is a great deal more the Labor Department could accomplish to improve its performance with more sophisticated data analysis and predictive analytics, real-time display of performance data to Labor Department managers and employees along with other transparency measures, and the transformation of program evaluation results into performance data. One critical innovation implemented by Secretary Tom Perez and my successor Deputy Secretary Chris Lu was to review, including directly with front-line employees, how agencies' performance measures and the expectations that attend them are being communicated to front-line supervisors and by front-line supervisors to front-line employees.

Nonetheless, one reason the department achieved the performance improvements detailed in the preceding section is that the department's leadership team solved the problems left unsolved in GPRA and the GPRAMA. But those flaws remain. The success of a government performance system should not depend upon agency leadership teams working around the laws defining that system. These systems are fragile. Continuity, particularly in the collection of data, is critical. An election or an intraparty change in personnel can overturn the entire enterprise and result in a return to superficial GPRA

---

<sup>147</sup> *Id.* at 9.

<sup>148</sup> U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-14-747, MANAGING FOR RESULTS: AGENCIES' TRENDS IN THE USE OF PERFORMANCE INFORMATION TO MAKE DECISIONS 9 (2014), available at [www.gao.gov/assets/670/666187.pdf](http://www.gao.gov/assets/670/666187.pdf).

<sup>149</sup> *Id.* at 17–18.

compliance without the substance of good-quality performance. After taking office, Secretary Perez and Deputy Secretary Lu gave their inherited leadership team an opportunity to discontinue or dramatically scale back the Labor Department's performance management and measurement system. The Assistant Secretaries and other agency leaders, after four years of experience with the system and its successes, chose to keep the system in place with only minor modifications.

One superficially appealing solution short of law reform would be for presidents to appoint only skilled, well-trained professional managers with experience overseeing performance management systems to lead every executive branch agency. This Article should offer testimony that this solution is not needed. As noted above, the members of the Obama Labor Department's leadership team could not have been fairly described as experienced, formally-trained performance managers. Yet, they achieved dramatic performance improvements under very difficult circumstances in a complex organization. Further, the costs of trading policy expertise for management experience are difficult to calculate, but potentially threatening to a president's agenda. A more nuanced mix of senior, presidentially appointed personnel along with a performance management training program should be part of the next president's agenda. There is a critical role for both policy and management at the highest levels of the federal executive branch. The task is to maximize the results from both functions.

But the surest solution would come from improved systems. At a minimum, Congress should fix identifiable problems in GPRA and the GPRAMA. The greater challenge is to institutionalize high-quality government performance and systems that ensure continuous improvement across the United States government. This section will offer a reform agenda that has as its goal ensuring effective performance management and measurement throughout the federal executive branch without unnecessarily disrupting existing systems that may be entirely effective and essential to improved performance. Congress must make reforms, and the first section below elaborates on those reforms. But the following section will also suggest efforts that should be undertaken by stakeholder organizations that are the voice of government's customers and constituents. Since the principal problem with GPRA and the GPRAMA is that Congress failed to establish effective mechanisms for overseeing executive branch performance, one important remedy is to hold Congress accountable and establish a thoroughgoing system of citizen oversight of government performance.

#### *A. Congress*

Most of the reforms that Congress should undertake have been previewed in earlier sections of this Article, particularly in the discussion of the GPRAMA's failure to address problems with GPRA. In this section, I will make these reforms explicit and expand upon them.

The first and most important problem to be solved is the lack of a system whereby Congress holds the executive branch accountable for GPRA/GPRAMA compliance and, more importantly, improving performance. In essence, GPRA and the GPRAMA amounted to drive-by legislating. Congress passed the laws, but then walked away without taking the necessary steps to ensure they were implemented properly. The solution is simple: Congress must establish oversight systems and hold itself accountable for operating those systems successfully.

The House of Representatives and the Senate should change their rules to require that every authorizing committee and every appropriations sub-committee hold an annual hearing in which they review the GPRA compliance and performance of the agencies within their jurisdictions. Where there are overlapping jurisdictions, the Senate Majority Leader and the Speaker of the House should make assignments. HSGAC and COGR should remain responsible for oversight of the process and GPRA and the GPRAMA generally, and they should keep their jurisdiction over OMB's role in the GPRA compliance process. Yet, these committees should not be involved in the substance of performance management and measurement for agencies over which they do not have jurisdiction.

Congressional committees are thinly staffed and not necessarily prepared to take on this sizable new performance oversight role. They will need help. Congress should direct GAO or the agencies' inspectors general to audit each agency's performance in advance of the authorizing committees' and appropriations sub-committees' hearings. GAO's and the inspectors general's roles should be strictly limited to comparing actual performance to goals set in agency operating/performance plans, investigating the integrity of performance data used by the agencies, and assessing compliance with GPRA and the GPRAMA. Policy and strategic judgments should be left to a debate between Congress and agencies' leadership teams. Of course, both Congress's hearings and all audits should be accessible to the public.

Annual hearings and individual agency audits alone will not provide a full measure of the quality of an agency's performance. Federal agencies should be benchmarked against one another. Just as there is no entity in the American economy that performs all of the same functions as the Labor Department, there are no competitors or comparators for the U.S. Departments of Defense, Interior, State, Transportation, or Agriculture, or any other Cabinet department or executive branch department. There are only two possible benchmarks for these organizations: their own past performance and the performance of other large federal agencies. Congress should use both benchmarks as a means of putting each agency's annual performance results in context.

Congress should direct every agency to construct an annual performance index that offers a complete picture of the agency's actual performance compared with its past performance. Congress should also direct

GAO and OMB to develop a system for ranking agencies' performance on the basis of these annual performance indices. Plainly, it is not possible to compare the Labor Department's success in training workers with the Defense Department's success in defeating terrorism threats in foreign countries. This is the classic "apples-to-oranges" comparison. But the purpose of the rankings is not to offer this kind of comparison. Effective benchmarking depends upon an "apples-to-apples" comparison. The goal should be to compare each agency's success in sustaining or surpassing its own performance with every other agency's success on the same measure.

The closely related second problem is that GPRA and the GPRAMA do not define "success." Worse, they implicitly define "failure" to be actual performance that does not meet an agency's target for performance. As noted above, meeting or exceeding targets is a poor measure of successful performance. Congress should abandon targets as the measure of success or failure. In place of targets, Congress should shift its focus to past performance as agencies' principal benchmark. Agencies should be required, in some form, to compare their actual performance in the present year to performance in prior years.

Although we did not encounter such a circumstance in the Labor Department during President Obama's first five years, it is possible that this approach is not appropriate or optimal for every activity in every agency. Accordingly, agencies should be permitted to propose some substitute measure that discloses performance trends. For example, some agencies may have well-functioning cost-accounting systems. These kinds of systems can be valid alternatives. But the default should be a standard that an agency's actual performance in the present year is equal to or better than some measure of performance in prior years (e.g., the Labor Department's system of averaging a fixed number of prior years' data). Exceptions should be permitted for those activities where performance has been deemphasized, real (i.e., inflation-adjusted) budgets have been reduced, or some other legitimate excuse is articulated by the agency's leadership in the annual performance plan. The validity of these excuses can be debated in the committees' and subcommittees' oversight hearings. But the debate over which agency has been most successful would be facilitated by the ranking system to be developed by OMB and GAO.

Similarly, Congress must abandon its stated focus on, but actual indifference to, "priority goals." Instead, Congress and OMB should turn their attention to all of the goals articulated in an agency's strategic plan. Plainly, trying to assess performance using all of the metrics included in an agency's operating plan, or the operating plans of a federated department's (like the Labor Department) many constituent agencies, would be overwhelming for Congress and its committees. But the strategic plan measures should be an agency's key performance indicators. If constructed properly, they should reflect overall agency performance and emphasize the highest priority activities

and outcomes. The existing priority goals offer only a very small subset of this larger, more comprehensive, but still manageable universe. Congress should be concerned with the most important, best, and most comprehensive measures of whether an agency is carrying out its mission. Further, while the strategic plan measures should be reported as one important category of information for observers' consideration, the committees and sub-committees conducting oversight hearings should be free to inquire into any measure in any agency's performance plan. Agencies' choices regarding which measures are included in the strategic plan should not be a means of obfuscating performance in problematic areas. The result would be a strong incentive to strip selected, potentially important outcome goals out of agencies' strategic plans.

The third problem is that the GPRA/GPRAMA system offers no assurance that the data it generates are actually employed in agencies' management decision-making. I have already described the GAO survey of executive branch managers that broadly tests this question. Congress must require and fund GAO to conduct this survey and analyze its results every year rather than every five years. GAO's analysis of the survey's results should be expanded beyond average scores and distributed much more widely.

To further ensure that federal managers use performance data in their everyday decision-making, Congress also must amend the statute establishing the individual performance review system. While it should be permissible for agency leaders to take other factors like equal employment opportunity and efforts to increase workforce diversity into account in these managers' reviews, agency and individual performance outcomes and outputs must be preeminent. Binding managers' individual performance reviews—and their opportunities for promotion and bonuses, along with their risks of discipline—to agency performance will incentivize close attention to performance data and efforts to improve performance. Accordingly, Congress should amend the law to make unmistakably clear that the principal factors in review and bonus determinations are the success of the SES manager's agency in improving performance, the SES manager's contributions to improving performance and accomplishing her agency's performance goals and, most importantly, performance results for activities that are within the manager's scope of responsibility.

The fourth problem with GPRA and the GPRAMA is that connections between agencies' strategies, outputs, and outcomes are not always self-evident, and may be quite difficult for observers to understand. Congress should require agencies to include plain-English statements in their strategic plans regarding how their strategies and outputs produce the outcomes articulated in their strategic plans—that is, their hypotheses of the causal connections between outputs and outcomes. Congress should also require and fund program evaluations in any circumstance where an agency does not or cannot report on outcomes as part of its performance measurement system. Additional funding is critical, and should be a precondition for the imposition

2015]

*MANAGING SOCIAL CHANGE*

1039

of any such requirement. While evaluations of specific strategies or procedures within a program might be conducted at fairly low cost using administrative data, a properly conducted evaluation of a program as a whole can cost millions of dollars. Appropriations necessary to carrying out an agency's mission must not be re-directed to program evaluation simply because Congress is unwilling to make the requisite investment in studying how government best serves its customers and constituents. Degrading performance to measure performance is not good government.

In some cases, including the ETA and MSHA examples offered above, agencies' performance data include outcome reporting. Nothing further should be required to evaluate the program as a whole when this is true, although it is important to continue to evaluate strategies and components of programs to, for example, test how to improve outcomes. But in other circumstances, like OFCCP's program outcome quandary also discussed above, further study is required. In the simplest terms possible, Americans want to know that government is delivering results that improve their lives, their society, their communities, and their country. Agencies should be required to study and explain how every program they manage is doing exactly that, with the possible exception of administrative functions that support mission-focused activities. Agencies should also explain how they intend to transform program evaluation information into performance data over time.

The fifth problem is that Congress's mandate regarding when and for which activities performance plans are required is confusing, at best, and ineffective, at worst. Too many agency activities slip through this porous requirement. In the case of the Labor Department before 2009, a large percentage of constituent agencies were able to escape the requirement to create and implement operating/performance plans. Congress should require that every department, agency, and operating unit of the executive branch produce an operating plan that measures performance with respect to every activity or function that is relevant to carrying out that entity's mission. A "relevance" standard would likely lead agencies to include every activity and function in its operating plan out of fear that any implicit admission of a program's "irrelevance" could lead Congress to eliminate its funding. An alternative and more forgiving standard would be to require operating plans that cover every activity or function that is "important" to carrying out that entity's mission and how those parts of the program can be improved.<sup>150</sup> This lower standard increases the risk that agencies will not measure performance in certain functions; however, this is precisely the kind of issue that the GAO or inspector audits and congressional hearings could address.

---

<sup>150</sup> It might make sense to leave a safety valve by allowing agency heads, at their discretion, to exempt operating units that exclusively provide supportive services like congressional and public affairs.

It would be fair to criticize this focus on Congress as both unrealistic and potentially perilous to the independence of the executive branch. Congress has struggled to carry out its own fundamental constitutional duty to enact annual appropriations bills for the past several years.<sup>151</sup> Expecting Congress to oversee executive branch performance effectively when its own performance has been fodder for late-night television hosts' monologues and single-digit approval ratings in public polls may be too ambitious. But Congress is the only means of achieving the entirety of this reform agenda. The next president could direct his or her appointees in OMB and the executive branch agencies to implement pieces of the agenda, including defining "success," creating annual performance indices, explaining relationships between outputs and outcomes, articulating comprehensive program evaluation agendas, and requiring operating plans from every operating unit, for example. Other large and important pieces of the agenda, particularly those that would facilitate serious debate about the functioning of government, can only be accomplished through Congress. Like it or not, Congress must participate.

### *B. Stakeholders*

Almost every report required by GPRA, the GPRAMA, and the reforms I have advocated must be made public. Every strategic plan, most aspects of agencies' operating/performance plans, every program evaluation, every quarterly and annual performance measure, and actual performance from every quarter and year must be published on agencies' web sites. Every annual performance index, the GAO survey and its analysis and results, every GAO and Inspector General audit, and every ranking of the agencies' performance and commitment to evidence-based, data-driven decision-making should be collected and published on a central OMB web site. Every congressional hearing on government performance must be open to the public.

Only one aspect of agencies' annual operating/performance plans arguably should not be published. Certain strategic choices to be made by law enforcement agencies, including the Labor Department's employment law enforcement agencies, should be kept confidential as a means of avoiding advance notice to likely targets of enforcement activity. The police typically do not broadcast advance notice of targets for raids or sweeps because doing so would likely result in the targeted criminal activity being hidden or moved. The same is true for OSHA, Wage & Hour, MSHA, OFCCP, and the Labor Department's other law enforcement agencies, as well as the Department of Homeland Security's Immigration and Customs Enforcement and the Secret Service, the Justice Department's Federal Bureau of Investigation and Drug Enforcement Agency, and other law enforcement entities in the federal

---

<sup>151</sup> See *supra* notes 137–38.

2015]

MANAGING SOCIAL CHANGE

1041

executive branch. Of course, the same concern arises with respect to national security agencies like the Defense Department, the Central Intelligence Agency, the National Security Agency, and the Coast Guard. Agencies should be empowered to redact sections of their operating/performance plans that would disclose law enforcement, national security, or intelligence gathering plans if, and only if, disclosure would risk the accomplishment of the agencies' missions. Of course, these redacted sections should be subject to *in camera* review by congressional authorizing and appropriating committees and their sub-committees. With this limited exception, agencies' operating/performance plans should be made public every year.

The purpose of public disclosure is to empower non-governmental actors to hold both Congress and the executive branch agencies accountable. This is not to suggest that Congress will never play its appropriate oversight role. Whether or not Congress takes responsibility for overseeing GPRAMA compliance and executive branch performance, stakeholders must be actively engaged, although the need for stakeholder engagement is greater if Congress continues to abdicate its role. But stakeholder engagement will keep all of the players honest: Congress, OMB, and the executive branch agencies. Also, stakeholder engagement often leads to congressional engagement as members of Congress learn that involvement in an issue may win the favor of their constituents or other important interests. So, a virtuous cycle may result.

Some might argue that public disclosure alone will cause the agencies to perform better and pay closer attention to their performance and operating planning. But this argument is premised on the expectation that someone is watching. The shame of public embarrassment and the implicit threat that an outsider's review of disclosed information will lead to consequences together cause public disclosure to change behavior. But is anyone watching? Earlier, I used the phrase "the public" to describe one of the overseers of executive branch performance.<sup>152</sup> Of course, 300+ million Americans are not going to review the mountains of plans and data that agencies are required to produce by GPRAMA and the GPRAMA, much less those I would have Congress require. "The public" is not watching. Americans are simply too busy working, caring for their families, attending to their health and the health of friends and family, serving and otherwise participating in their communities, and planning their futures.

Instead, "the public" depends upon its elected representatives to carry out this function. They also depend upon stakeholders as a proxy army focusing on different aspects of the public interest and holding elected and appointed officials accountable. Unionized and many other workers rely on the AFL-CIO and its affiliates, the Service Employees International Union, and other independent unions and worker organizations. Environmentally conscious

---

<sup>152</sup> See *supra* text accompanying note 28.

voters rely on the League of Conservation Voters, the Sierra Club, the Natural Resources Defense Council, and other environmental organizations. Civil libertarians look to the American Civil Liberties Union and civil rights advocates depend upon the NAACP, the Leadership Conference on Civil Rights, the National Council of La Raza, the Human Rights Campaign, and other civil rights and advocacy groups. Small businesses rely on the National Federation of Independent Business and other business lobbying groups, including the U.S. Chamber of Commerce. A very long list of trade associations advocate for various domestic and foreign industries and occupations.

The critical opportunity in reforming and improving federal government performance is the engagement of well-informed stakeholders using accessible tools provided to them by the government to hold both Congress and the executive branch accountable. The tasks in which these stakeholder overseers should engage are easily enumerated. First, they should lobby Congress to enact the legislative and organizational reforms explained in this section. Second, they should review drafts of agencies' strategic plans so that they can engage actively with the agencies and Congress regarding the definition of agencies' missions and which outcomes and key performance indicators would demonstrate that the agency is carrying out its mission. Stakeholders should insist on extensive and public opportunities to meet with top agency leaders and performance staff during the strategic plan formulation process to discuss these topics. In an ideal world, these sessions would be live-streamed on agencies' web sites so that stakeholder groups' members, and stakeholder organizations unable to afford a trip to Washington, D.C., can participate remotely. At a minimum, and akin to the notice and comment process in the Administrative Procedure Act,<sup>153</sup> agencies should publish a summary of all such discussions with stakeholders along with any written comments submitted by stakeholder groups. Of course, stakeholder groups should submit detailed written comments, but they should also educate their members about the content of the strategic plans and their importance.

These discussions regarding agencies' strategic plans should include a deep exploration of the agencies' program evaluation plans. In my experience, discussions between agency officials and stakeholders about which strategies they should pursue are heavily infused with unproven, often ideologically driven assumptions about causal connections between outputs and outcomes. Agency leaders should be able to explain to stakeholders in clear terms how they intend to demonstrate that their activities and strategies will improve the lives of the stakeholder group's constituents and others who are similarly situated. Stakeholders should insist that assumptions about connections between outputs and outcomes are made explicit and evidence to support them

---

<sup>153</sup> Administrative Procedure Act, 5 U.S.C. § 553 (2013).

either disclosed or developed in the agency's strategic plan. It is essential to remember that, under the GPRAMA, agency strategic plans will remain in place through an entire presidential term. So, these stakeholder interactions with agencies about program evaluation, outcomes, and mission are very high stakes. They should be given a high priority by stakeholder organizations.

Third, stakeholders should engage aggressively with agencies every year as they develop their annual operating/performance plans. These plans will determine how agencies invest their resources on a day-to-day basis, so stakeholder organizations should insist on an opportunity to review draft plans and challenge agency leaders' strategic choices and target-setting decisions. The mere demand that agencies publicly release their operating plans will increase the likelihood that they will create such plans for all of their operating units. Nonetheless, this demand is going to be met with a great deal of resistance from agency officials conditioned to withhold this kind of information from the public as a means of guarding against public criticism and judgment. Stakeholders must insist.

The value of non-governmental actors' engagement in this context cannot be overstated. Outsiders are going to be hostile to arguments that are too common and too commonly accepted in the federal government: "because we've always done (or measured) it that way," "because this function requires staff who are rated a GS-12 and my staff members are all rated GS-13," and "because Congressman Smith is going to be upset if we close our office in Omaha to staff that priority initiative in Los Angeles." Breaking out of internal logic systems that rationalize inadequate performance and hamstringing agency leaders will be one important product of stakeholder engagement. Stakeholders will not quietly acquiesce in inertia and business-as-usual if their constituents' interests are at stake.

Again, stakeholders should insist on meeting with agency leaders and demand an opportunity to provide them with detailed written comments on their draft operating plans. Agency leaders should be required to explain the choices and resource investments included in their operating plans. Stakeholders should demand plain-English explanations of how agency leaders are going to ensure that their activities and strategies improve the lives of the stakeholders' constituents and others. Perhaps most important, the operating plans released to the public must contain quarterly and annual targets for every important activity in an agency. Stakeholders must have this information if they are going to hold agencies accountable for their performance in real time.

Fourth, stakeholders should demand public release of all quarterly and annual performance data. Along with public release of operating/performance plans, this is a necessary mechanism to allow stakeholders to compare actual performance with the targets in the agency's operating/performance plan. But stakeholders must not be passive recipients of agency data dumps. Stakeholders should offer their judgments about how agencies are performing using these data and, if Congress or the President requires them, the agencies' annual

performance indices. If neither Congress nor the GAO rank agencies' performance, then stakeholder groups should. If agencies do not produce annual performance indices, then stakeholders must. But stakeholder groups should not merely re-hash or disclose the numbers. They should offer detailed critiques of agency performance that they publish and share with Congress to inform any congressional hearings that may be held.

Annual performance indices would provide insight into agency-wide performance, but they would not look at cross-agency functional areas despite the GPRAMA's requirement of the creation of a few cross-agency priority goals. This is another role stakeholder groups can fill. For example, the Labor Department's OFCCP, the Justice Department's Civil Rights Division, and the Equal Employment Opportunity Commission share responsibility for the enforcement of anti-discrimination laws that protect workers.<sup>154</sup> Reviewing the Labor Department's annual performance index alone, or even side-by-side with the Justice Department and EEOC indices, would not give observers a clear picture of the government's performance on workplace civil rights. However, civil rights stakeholder groups might construct a "civil rights performance index" that looks at and compares performance measures drawn from all three agencies to assess how the incumbent administration is performing on the issues of greatest interest to those groups' constituents. Similar cross-agency indices would be valuable in many areas of federal government responsibility ranging from environmental protection to health care to worker training to immigration to consumer protection to national security.

Stakeholder organizations could organize themselves to accomplish all of these tasks and contribute to improving government performance in at least three ways. Each stakeholder group may choose to work alone, with each group focusing on the agencies and agency sub-units that serve or protect the group's members. Groups may form coalitions, either with other groups that have similar interests or with groups across several disciplines. For example, civil rights groups could band together only with each other or they could work with environmental and worker rights groups. Finally, groups across ideologies and interests could form a clearinghouse organization that lobbies on government performance issues and provides information, data, and analysis to all of its members without regard to ideology, as well as the public.

---

<sup>154</sup> For example, OFCCP enforces Executive Order No. 11246. Exec. Order No. 11246, 30 Fed. Reg. 12319 (1965). Executive Order No. 11246 applies only to federal contractors, but is substantively similar to Title VII of the Civil Rights Act of 1964, which is enforced or administered by the EEOC and the Justice Department. 2 U.S.C. § 1311 (2013). Similarly, OFCCP enforces Section 503 of the Rehabilitation Act of 1973, which also applies only to federal contractors. 29 U.S.C. § 793 (2013). The Rehabilitation Act of 1973 is substantively similar to the employment provisions of the Americans with Disabilities Act, which is enforced by the EEOC and the Justice Department. 42 U.S.C. § 12101 (2013).

Either the coalition approach or the clearinghouse approach seems most likely to succeed. Oversight of the federal government's performance and performance management and measurement, as well as the data analysis required to understand and influence government performance, require advanced skills and knowledge. Just as Congress needs help from the GAO and the inspectors general, stakeholder groups need help from their own experts with experience, deep knowledge, and sophisticated skills. Too many stakeholders cannot afford this level of investment. Banding together is the best method to achieve economies of scale, reduce each stakeholder group's costs, and increase the likelihood that stakeholders will play a meaningful role in improving government performance.

#### V. CONCLUSION

The fundamental premise of this Article is that government matters to the lives of the American people. Government can help to make Americans' lives better, facilitate greater and fairer opportunities, protect from the depredations of markets as well as natural and human-made disasters, provide for the common welfare, and make us safe from enemies foreign and domestic. For most Americans who seek progressive social change, government is one mechanism through which this change will be accomplished. While legislative and regulatory improvements to our laws and new laws can be important channels for accomplishing progressive outcomes, in a harshly polarized political environment like that found in Washington for much of the last quarter century, progressive law reform through Congress may not be possible. Even in the best political environment, law reform can be rendered meaningless or made fully effective depending upon the quality of government performance. Social change can also be realized in either event through a tighter and more aggressive focus on the day-to-day management of the government.

Progressives must not allow themselves to become knee-jerk defenders of government. Government is a tool to achieve a result, not an end in itself. Good intentions, flowery rhetoric, and even policies with which progressives agree should not be sufficient. Those who are committed to social change should demand evidence, in real time and readily understandable, demonstrating that government is effecting that change. When government fails, progressives should be the first to sound the alarm and demand better.

A second fundamental premise of this Article is that solutions to government performance failures begin with more democracy, more transparency, and more accountability for elected and appointed officials. The biggest challenge is not overcoming wrongheaded priorities or bad judgment or management, although there are some of each in the federal government. Rather, the challenge is to ensure that government's goals and actions are perennially grounded in the genuine needs and desires of the American people. This is the very essence of democracy. Only extensive engagement between government and governed, and open government information and data that

facilitates deeper understanding of government, can ensure government agencies that function in this way.

This Article sought to critique the laws that purport to structure and guide the federal government's performance planning and measurement while telling the story of how one agency, the Labor Department, succeeded in improving its performance, sometimes despite these laws. This Article also sought to offer a path forward to further improving federal government performance both through congressional reforms and activist stakeholder engagement. If government matters to progressive social change, and its performance can be improved, then government performance must be at the very center of any progressive agenda for the future.